15-17 November 2024

IV. INTERNATIONAL CONGRESS ON ARTIFICIAL INTELLIGENCE IN HEALTH

"Human Machine Interface"

PROCEEDINGS BOOK

TÜBİTAK

UNIVERSITY OF BAKIRÇAY

YAŞAR UNIVERSITY

TINAZTEPE UNIVERSITY

CONGRESS COMMITTEES

Congress President

Mustafa BERKTAŞ	İzmir Bakırçay University Rector

Congress Guests of Honor

Orhan AYDIN	Scientific and Technological Research Council of Turkey (TÜBİTAK)
Hasan MANDAL	İstanbul Technical University
Ümit KERVAN	Turkey Institutes of Health (TÜSEB)
Bayram YILMAZ	Dokuz Eylül University
Mustafa GÜVENÇER	İzmir Tınaztepe University
Levent KANDİLLER	Yaşar University
Evren YAŞAR	Bozok University
Mehmet DEMİRHAN	Turkish Orthopedics and Traumatology Association
Mustafa BERKTAŞ	International Association for Artificial Intelligence in Health

Congress Scientific Secretaries

Abdulkadir HIZIROĞLU	İzmir Bakırçay University
Görkem YAMAN	İzmir Bakırçay University
Mehmet HAKLIDIR	Scientific and Technological Research Council of Turkey (TÜBİTAK)
Songül VARLI	Turkey Institutes of Health (TÜSEB)
Caner ÇAVDAR	Dokuz Eylül University
Çağın ŞENTÜRK	İzmir Tınaztepe University
Yiğit KAZANÇOĞLU	Yaşar University
Yusuf HIDIR	Bozok University
Cemil YILDIZ	Turkish Orthopedics and Traumatology Association
Orhan ER	International Association for Artificial Intelligence in Health

Organizing Committee

Görkem ATAMAN	Yaşar University
Ahmet Emin ERBAYCU	İzmir Bakırçay University
Şahin BOZOK	İzmir Bakırçay University
M. Kemal GÜLLÜ	İzmir Bakırçay University
Orhan ER	İzmir Bakırçay University
Ruşen HALEPMOLLASI	Scientific and Technological Research Council of Turkey (TÜBİTAK)
Elif GÜLER KAZANCI	University of Health Sciences Bursa Faculty of Medicine
Halil İbrahim KAYRAL	Turkey Institutes of Health (TÜSEB)
Mehmet BAKIR	Bozok University
Murat KORKMAZ	Turkish Orthopedics and Traumatology Association
Dilek ORBATU	H.S.U. İzmir Dr. Behçet Uz Pediatrics and Surgery E.R.H.
Senem ALKAN ÖZDEMİR	H.S.U. İzmir Dr. Behçet Uz Pediatrics and Surgery E.R.H.
Ayşegül YURT	Dokuz Eylül University<
Esra ARUN ÖZER	İzmir Tınaztepe University<
Özge TÜZÜN ÖZMEN	İzmir Bakırçay University
Bülent TOĞRAM	İzmir Bakırçay University
Ahu PAKDEMİRLİ	International Association for Artificial Intelligence in Health

Organization Team

İbrahim Onur ALICI	İzmir Bakırçay University
Bayram ÇOLAK	İzmir Bakırçay University
Onur DOĞAN	İzmir Bakırçay University
Ümit Hüseyin KAYNAR	İzmir Bakırçay University
Seren DÜZENLİ ÖZTÜRK	İzmir Bakırçay University
Nurullah ÖZDEMİR	İzmir Bakırçay University
Onur UĞURLU	İzmir Bakırçay University
Tarık SEMİZ	İzmir Bakırçay University
Research Support and Infrastructure	İzmir Bakırçay University
Kader MERT	İzmir Bakırçay University
Gülnur ÖNSAL;	İzmir Bakırçay University
Yıldız ÖZKILIÇ	İzmir Bakırçay University
Gülbin ERGİN GEDİK	

Scientific Committee

Foreign Science Committee

Hadi SEYEDARABİ	University of Tabriz
Mohammad Ali BALAFAR	University of Tabriz
Ayraz BARADARANİ	University of Windsor
Swati Vijay SHINDE	Pimpri Chinchwad College of Engineering
Jafar TANHA	University of Tabriz
Narjes Nikzad KHASMAKHI	Cologne University of Applied Sciences
Abejide Ade-Ibijola	University of Johannesburg
Tankiso Mololi	University of Johannesburg
Jazz (Ijaz) Rasool	Oxford University
Mirbek Turduev	Kyrgyz-Turkish Manas University
Guneet Kaur	University of Stirling

Tech Team

Ali PİŞİRGEN	İzmir Bakırçay Üniversitesi
Emre TURAN	İzmir Bakırçay University
M. Furkan AKSU	İzmir Bakırçay University
Betül YÜRDEM	İzmir Bakırçay University
Ali Mert ERDOĞAN	İzmir Bakırçay University
Mehmet GENCER	İzmir Bakırçay University
Yasemin İNCİ	İzmir Bakırçay University
Derya ÇITAK	İzmir Bakırçay University
Mutlu BAÇARU	İzmir Bakırçay University
Duygu KOCABAŞ ATILGAN	İzmir Bakırçay University
Selçuk BAZARCI	İzmir Bakırçay University
Kaan ÖZKAN	İzmir Bakırçay University
Serdar DEMİRCİ	İzmir Bakırçay University
Okay KINIK	İzmir Bakırçay University
Sercan TURAN	İzmir Bakırçay University
Cahit Canberk ANDAÇ	İzmir Bakırçay University
Beyza ASLAN	İzmir Bakırçay University
Yeşim AYGÜL	İzmir Bakırçay University
Sağlık Yönetimi Topluluğu	İzmir Bakırçay University

PREFACE TO CONGRESS PROCEEDING BOOK

The IV. International Congress on Artificial Intelligence in Health, hosted by İzmir Bakırçay University between November 15-17, 2024, was successfully held with the theme of “Artificial Intelligence in Health and Human-Machine Interaction”. The congress provided an important platform where current developments in artificial intelligence and human-machine interaction in the field of health were discussed and national and international experts shared their knowledge and experiences.

Within the scope of the congress, R&D studies, areas of use and what awaits us in the future of health technologies and artificial intelligence-based applications were discussed through thematic panels, presentations and oral presentations. In addition, a one-day Preparation Camp for the Artificial Intelligence in Health Congress 2024 was organized with the theme of human-machine interface technologies, providing participants with practical training.

The statistics of our congress are as follows;

- Number of panel sessions held: 11
- Number of oral presentation sessions held: 11
- Number of papers presented: 83
- Number of participants: 310
- Number of institutions from which participants came: 43

The sessions held and papers presented during the congress increased awareness of artificial intelligence and human-machine interaction in health and encouraged the sharing of knowledge and experience among participants. In this way, it is thought that new breakthroughs can be made in developing human-machine interfaces with artificial intelligence tools in the field of health and that the presented studies can guide policy makers.

We would like to thank all participants who contributed to our congress and the supporting institutions and organizations. We hope that similar successes will continue in the congresses to be held in the coming years.

With our love and respect

Prof. Dr. Mustafa BERKTAŞ

Rector of Izmir Bakircay University

Congress President

TABLE OF CONTENTS

ADEQUACY OF CHATGPT-4O'S RESPONSES TO QUESTIONS OF RESIDENCY TRAINING DEVELOPMENT EXAM ON SPINAL SURGERY

Uğur Yüzügüldü[1]*, Anıl Özgür[2], Bilge Kağan Yılmaz[3], Ömer Erşen[2], Serkan Savaş[3]

[1] Department of Orthopedics, Balıkesir Atatürk State Hospital, Balıkesir, Türkiye

[2] Department of Orthopedics, Gülhane Education and Research Hospital, Ankara, Türkiye

[3] Department of Orthopedics, Afyonkarahisar Health Sciences University, Afyonkarahisar, Türkiye

** Corresponding author: uguryuzuguldu@gmail.com*

Introduction-Aim: Artificial intelligence chatbots gained popularity due to their ability to analyse substantial scientific data from machine learning techniques and generate human-like responses in medicine. Chat Generative Pre-Trained Transformer (ChatGPT) has become the most popular chatbot after its release in November 2022, and every new version of ChatGPT has increased its ability to respond appropriately [1].

Residency Training Development Exam (RTDE) aims to evaluate the theoretical training of assistants, consisting of a total of 200 multiple true-false questions. In this exam, there are three options for the resident to choose: true, false, or don't know [2]. Wrong answers decrease residents score to avoid guessing in the exam. Recent studies have shown that multiple false positive questions can provide similar assessment methods to multiple choice questions for the evaluation of students. Also, it has been shown that exams with multiple true-false questions can better demonstrate the student thinking system. This study aims to evaluate the accuracy of ChatGPT-4o responses of questions of RTDE on Spinal Surgery [3,4].

Materials-Methods: The questions on spinal surgery of four RTDE between 2021 and 2024 were asked and responses of ChatGPT-4o were recorded in June 2024. We used a new window in the internet browser to avoid the memory bias of artificial intelligence for each question. Authors graded the responses using a rating system similar to the one introduced by Mika et al. Grading system included four grades as follows: 1- Correct answer with an excellent explanation, 2-Correct answers with inadequate explanation, 3- Wrong answer due to incorrect understanding of the statement in the question by artificial intelligence, 4- Wrong answer with wrong explanation [5]. Since patient data was not used in this study, ethical approval was not required or obtained. We used Microsoft Excel (Microsoft Corp, Redmond, WA) for statistical analysis. Scoring and variables are presented as percentages and continuous variables as mean ± standard deviation (SD) or median (quartiles).

Results: A total of 83 questions on spinal surgery were asked of the residents in four exams. ChatGPT gave correct responses to 75% of the questions. Success rates were 55% in the 2021 exam, 85% in the 2022 exam, 87% in the 2023 exam, and 82% in the 2024 exam. Grading scores were 2.1±1.3 in 2021, 1.4±0.9 in 2022, 1.5±0.9 in 2023, and1.8±1.2 in 2024.

Since the distribution of questions according to the subtopics of the spine was different in different years, when the evaluation was made according to the topics, the topics of deformity and destructive spinal diseases were determined to be the areas in which ChatGPT achieved 81% (22/27 and 9/11 respectively) success [6]. The success rate was 73% (11/15) in basic sciences, including spinal anatomy, spinal imaging, and physical examination. The success rate on spinal trauma questions was 63% (11/19), while it was 54% (6/11) in the degenerative diseases subtopic [7].

The results of this study demonstrated that ChatGPT-4o had more difficulty in recognizing false statements as false. Although 78% of the statements that ChatGPT failed were false, it evaluated them as true. The most important reason for this was considered to be the inability of the artificial intelligence to fully perceive the nuances in the statements. Information containing only part of the true statement on was accepted as true by ChatGPT, but the real answer was false.

Discussion-Conclusion: Nowadays, when almost everyone can easily access artificial intelligence through their smartphones or mobile devices and artificial intelligence is rapidly entering medicine and health sciences, including education. Although the overall success rate of ChatGPT-4o was 75% in spinal surgery true-false questions of RTDE, it fell short of expectations.

Keywords: artificial intelligence, ChatGPT, multiple true false, question format

REFERENCE

[1] Ahmed W, Saturno M, Rajjoub R, Duey AH, Zaidat B, Hoang T, Restrepo Mejia M, Gallate ZS, Shrestha N, Tang J, Zapolsky I, Kim JS, Cho SK (2024). ChatGPT versus NASS clinical guidelines for degenerative spondylolisthesis: a comparative analysis. Eur Spine J. Mar 15

[2] Brassil C, Couch B (2019). Multiple-true-false questions reveal more thoroughly the complexity of student thinking than multiple-choice questions: a Bayesian item response model comparison. International Journal of STEM Education 6:16.

[3] Couch B, Hubbard JK, Brassil C (2018). Multiple–True–False Questions Reveal the Limits of the Multiple–Choice Format for Detecting Students with Incomplete Understandings. BioScience 68:6, 455–463.

[4] Johns WL, Kellish A, Farronato D, Ciccotti MG, Hammoud S (2024). ChatGPT Can Offer Satisfactory Responses to Common Patient Questions Regarding Elbow Ulnar Collateral Ligament Reconstruction. Arthrosc Sports Med Rehabil. 13;6(2):100893.

[5] Mika AP, Martin JR, Engstrom SM, Polkowski GG, Wilson JM (2023). Assessing ChatGPT Responses to Common Patient Questions Regarding Total Hip Arthroplasty. J Bone Joint Surg Am. 4;105(19):1519-1526.

[6] Schaap, L., Verkoeijen, P., Schmidt, H. (2014). Effects of different types of true–false questions on memory awareness and long-term retention. *Assessment & Evaluation in Higher Education, 39*(5), 625–640.

[7] Zhang S, Liau ZQG, Tan KLM, Chua WL (). Evaluating the accuracy and relevance of ChatGPT responses to frequently asked questions regardi2024ng total knee replacement. Knee Surg Relat Res. 2;36(1):15.

ANALYSIS OF THE VALIDITY AND RELIABILITY OF THE SMARTPHONE GONIOMETER APPLICATION IN PATIENTS WITH PATELLAFEMORAL PAIN BY COMPARING IT WITH UNIVERSAL GONIOMETER AND DIGITAL INCLINOMETER

Musa Çankaya[1]*, Aleyna Bekaroğlu[2]

[1] Department of Physiotherapy and Rehabilitation, Necmettin Erbakan University, Konya, Türkiye

** Corresponding author: mcankaya@erbakan.edu.tr*

Introduction-Aim: Patellofemoral pain (PFP) is one of the most common knee problems that impairs function and daily activities [1]. PFP is one of the most confusing and clinically challenging chronic conditions. Symptoms usually occur from the anterior aspect of the patella and commonly along the medial aspect of the knee [2]. Symptoms of PFP can develop slowly or suddenly, and pain tends to worsen with activities such as squatting, prolonged sitting, climbing stairs, jumping or running [3]. The development of PFP has a multifactorial nature with multiple contributing factors including proximal, local and distal factors. PFP affects 29% of active adolescents and 23% of both active and sedentary adults in the general population [4]. People with PFP have lower health-related quality of life and physical activity levels than their peers. The prognosis of PFP is poor, with more than 50% of people reporting persistent pain 5 years after treatment. Clinicians and researchers need to do more to understand how PFP is best managed [4]. The aim of this study was to analyse the validity and reliability of the smartphone goniometer application in patients with patellafemoral pain by comparing it with universal goniometer and digital inclinometer.

Materials-Methods: This research is a descriptive single-blind study. This methodological study was approved by Necmettin Erbakan University Drug and Non-Medical Device Clinical Research Ethics Committee. The study was conducted in accordance with the Declaration of Helsinki. Verbal and written informed consent was obtained from the participants. In addition, verbal information about the study was given. In the calculation of the sample size, using the *Power 3.1.9.2 programme with an effect size of 0.35, a standard error of 0.05 and a power of 95 percent, ANOVA: Repeated measures, within-between interaction in the form of 3 groups and 2 measurements, it was determined that the calculation should be made with 24 people. It was foreseen that there might be losses during the study and 27 patients were included in our study [5].

Participants in our study included individuals with anterior knee pain who came to Seydişehir State Hospital. A total of 27 patients with PFP were evaluated in the physical therapy unit of Seydişehir State Hospital or in the laboratory of Seydişehir Vocational School of Health Services between January and August 2024. Only one extremity of each patient was evaluated

Evaluation Universal goniometer (Saehan SH5110 Steel goniometer set), Inclinometer (Baseline digital Inclinometer), Smartphone (Electrogoniometer) application were used to evaluate knee and hip range of motion by 2 assessors. Patients rested for 5-10 minutes after the first assessor [6].

The flexion range of motion for the knee joint and the flexion/extension, abduction, internal/external rotation range of motion for the hip joint were measured by two examiners using a smartphone application, a digital inclinometer, and a universal goniometer. To assess

interobserver reliability, two measurements made by the first observer were evaluated at 24-48 hour intervals. To assess intraobserver reliability, measurements from both observers were compared. Statistical analysis was performed using intraclass correlation coefficient (ICC) and Pearson correlation analyses.

Results: A single extremity was evaluated for 27 participants between the ages of 19 and 55. The results of intra- and interrater reliability analyses are shown. For the universal goniometer, good inter-observer and intra-observer reliability was observed for hip extension, abduction and external rotation, and excellent intra-observer and inter-observer reliability was observed for other measurements. In our analysis of the concurrent validity of the ROM measurements of knee flexion, hip flexion, extension, abduction, external rotation and internal rotation made with the three methods used in our study, the strongest correlations were determined between the smartphone and the inclinometer (r=0.952, r=0.993, r=0.842, r=0.807, r=0.970, r=0.963, respectively).

Discussion-Conclusion: Clinical measurements should be accurate, reliable, reproducible and sensitive to change the results, easily applicable and accessible [20]. In our study, we aimed to determine the concurrent validity, inter- and intra-observer reliability by comparing the smartphone application with inclinometer and universal goniometer to measure active knee and hip joint ROMs of PSS patients. Smartphone application, inclinometer and universal goniometer showed valid and reliable results for measuring lower extremity ROM in PSS patients. It was determined that the smartphone application had excellent intraobserver (ICC>0.80, SEM>1.60) and interobserver reliability (ICC>0.81, SEM>2.18) for all directions in knee flexion, hip flexion, abduction, internal and external rotation. Furthermore, the smartphone application was shown to have excellent correlation with the inclinometer and universal goniometer. These results suggest that the smartphone application may be a useful application for measuring knee and hip ROM in the clinical setting.

Keywords: smartphone application, digital inclinometer, range of motion, validity and reliability, lower extremity joint

REFERENCE

[1] Kamel, A. M., Ghuiba, K., Abd Allah, D. S., Fayaz, N. A., & Abdelkader, N. A. (2024). Effect of adding short foot exercise to hip and knee focused exercises in treatment of patients with patellofemoral pain syndrome: a randomized controlled trial. *Journal of Orthopaedic Surgery and Research*, *19*(1), 207.

[2] Xie P., István B, Liang, M. The relationship between patellofemoral pain syndrome and hip biomechanics: a systematic review with meta-analysis. In *Healthcare*. 2022;11(1): 99

[3] Willy RW, Hoglund LT, Barton CJ, Bolgla LA, Scalzitti DA, Logerstedt DS, Lynch AD, Snyder-Mackler L, McDonough CM. Patellofemoral Pain. J Orthop Sports Phys Ther. 2019;49(9):CPG1-95

[4] Davis IS, Tenforde AS, Neal BS, Roper JL, Willy RW. Gait retraining as an intervention for patellofemoral pain. Curr Rev. Musculoskelet Med. 2020;13:103–114

[5] Vaidya, S. M. (2022). Effect of foam rolling of quadriceps, hamstring, and IT band on knee passive range of motion and physical function in patients with patellofemoral pain syndrome-randomized controlled trial. *Archives of Medicine and Health Sciences*, *10*(1), 37-41

[6] Acar, S., Aljumaa, H., Şevik, K., Karatosun, V., & Ünver, B. (2024). The Intrarater and Interrater Reliability and Validity of Universal Goniometer, Digital Inclinometer, and Smartphone Application Measuring Range of Motion in Patients with Total Knee Arthroplasty. *Indian Journal of Orthopaedics*, *58*(6), 732-739.

EVALUATION OF IMPACTED THIRD MOLARS WITH OPEN AND CLOSED APEX USING CBCT AND A DEEP LEARNING APPROACH: A PILOT STUDY

Suay Yağmur, Ünal[1]*, Filiz Mediha[1], Namdar Pekiner[1]

[1] Maxillofacial Radiology, Marmara University Faculty of Dentistry, Istanbul, Türkiye

** Corresponding author: suayyagmurunal@gmail.com*

Introduction-Aim: The assessment of impacted third molars (M3), particularly the differences between those with closed and open apexes, is a critical component in dental and maxillofacial evaluations. CBCT is a widely utilized imaging modality for such assessments, offering detailed three-dimensional images that assist in diagnostic precision [1].

However, manual interpretation of CBCT images is time-consuming and may be subject to variability between observers. Recent advancements in deep learning techniques present opportunities to automate and enhance the accuracy of these interpretations. This study aims to develop and validate a deep-learning model to differentiate between impacted third molars with closed and open apexes using CBCT images [2].

Materials-Methods: In this retrospective study, the data set was composed of CBCT images of 300 patients aged 18 and over, available in the archives of Marmara University Faculty of Dentistry, Department of Oral and Maxillofacial Radiology. It was classified into two categories according to the apex development. Labeling was done with the segmentation method and the data sets were divided into two: training (n = 270) and test (n = 30) data sets. With the nnU-Net architecture, training, and validation data sets were used to estimate and produce optimal artificial intelligence algorithm weight factors. The success of the model was checked with the test data set, and the obtained accuracy, sensitivity, precision, Jaccard index, and Dice score values gave information about the success of the model.

Results: The study developed a highly effective deep learning model for detecting M3 using CBCT. For M3 with closed apexes, the model demonstrated exceptional performance metrics, achieving an impressive 99% accuracy. Additionally, the model exhibited a sensitivity of 90%, precision of 85%, Jaccard index of 0.78, and Dice score of 0.85. These results indicate that the model was highly effective in identifying M3 with open apexes, with strong sensitivity and precision values that underscore its robustness in detecting these specific cases.

However, for M3 with open apexes, the model maintained a high overall accuracy of 99%, but the sensitivity decreased to 30%. Despite the drop in sensitivity, the model still showed precision of 71%, with Jaccard index of 0.30 and Dice score of 0.37. These findings suggest that while the model is highly accurate in general, its performance in detecting unerupted M3 with open apexes may be less consistent, likely due to the complexity and lower distinguishability of these cases in the CBCT images.

Discussion-Conclusion: In the study by Lahoud et al., their own developed convolutional neural network model was used to segment mandibular third molars in 314 CBCT scans, with the mean IoU value for segmentation success reported as 0.87 [3].

Overall, the deep learning model proved to be a valuable tool for detecting M3, particularly ones with closed apexes, in CBCT. The potential application of such models in clinical practice could significantly reduce the diagnostic workload of dental and maxillofacial surgeons, improving decision-making processes and increasing efficiency in treatment planning [2].

No similar study has been conducted on CBCT and it will provide a new perspective to the literature on this subject. Future study in this area will involve expanding the dataset to include a broader range of cases, enhancing the diversity of the training data, and refining the model's architecture. This will enable the model to better handle the variations in tooth morphology and positioning encountered in clinical scenarios. Furthermore, steps will be taken toward real-world clinical integration, including the development of user-friendly interfaces and validation studies to ensure the model's performance in diverse clinical environments. Ultimately, this research represents a significant step toward leveraging artificial intelligence in dental diagnostics, promising to enhance the precision and efficiency of dental and maxillofacial surgery in the future.

Keywords: Deep learning, Artificial Intelligence, CBCT, Mandibular Third molar

REFERENCE

[1] K. Orhan, E. Bilgir, I. S. Bayrakdar, M. Ezhov, M. Gusarev, and E. Shumilov, "Evaluation of artificial intelligence for detecting impacted third molars on cone-beam computed tomography scans," *J. Stomatol. Oral Maxillofac. Surg.*, vol. 122, no. 4, pp. 333–337, 2020. doi: 10.1016/j.jormas.2020.12.006.

[2] G. Keser and F. N. Pekiner, *Current Researches in Health Sciences-I*, Özgür Yayınları eBooks, 2023. doi: 10.58830/ozgur.pub91.

[3] P. Lahoud, M. EzEldeen, T. Beznik, H. Willems, A. Leite, A. Van Gerven, and R. Jacobs, "Artificial intelligence for fast and accurate 3-Dimensional tooth segmentation on cone-beam computed tomography," *J. Endod.*, vol. 47, no. 5, pp. 827–835, 2021. doi: 10.1016/j.joen.2020.12.020.

ADEQUACY OF CHATGPT'S RESPONSES TO RESIDENCY TRAINING DEVELOPMENT EXAM QUESTIONS OF THE TURKISH SOCIETY OF ORTHOPEDICS AND TRAUMATOLOGY

Anıl Özgür[1]*, Batuhan Göçer[2], Ömer Erşen[1], Cemil Yıldız[1]

[1]*Orthopedics and Traumatology, University of Health Sciences, Gülhane Training and Research Hospital, Ankara, Türkiye*

[2] *Orthopedics and Traumatology, Dokuz Eylül University Hospital, İzmir, Türkiye*

* *Corresponding author: dr.anilozgur@gmail.com*

Introduction-Aim: Artificial intelligence chatbots gained popularity due to their ability to analyse substantial scientific data from machine learning techniques and generate human-like responses in medicine. Chat Generative Pre-Trained Transformer (ChatGPT) has become the most popular chatbot after its release in November 2022, and every new version of ChatGPT has increased its ability to respond appropriately.

The Residency Training Development Exam (RTDE) aims to evaluate the theoretical training of assistants, consisting of 200 multiple true-false questions. In this exam, there are three options for the resident to choose: true, false, or don't know. Wrong answers decrease resident's scores to avoid guessing in the exam. Recent studies have shown that multiple false positive questions can provide similar assessment methods to multiple choice questions for evaluating students. Also, it has been shown that exams with multiple true-false questions can better demonstrate the student's thinking system. This study aims to evaluate the accuracy and validity of different versions of ChatGPT (ChatGPT 3.5, ChatGPT 4, and ChatGPT 4o) responses to questions of RTDE 2024.

Materials-Methods: The questions on spinal surgery of RTDE 2024 were asked and responses of ChatGPT versions were recorded in June 2024. We used a new window in the internet browser to avoid the memory bias of artificial intelligence for each question. The authors graded the responses using a rating system like the one introduced by Mika et al. The grading system included four grades:

1- A correct answer with an excellent explanation,
2- Correct answers with inadequate explanation or
   Correct answers with an explanation that ChatGPT is not definitive and requires expert interpretation,
3- Wrong answer due to incorrect understanding of the statement in the question by artificial intelligence or
   Wrong answer with an explanation that ChatGPT is not definitive and requires expert interpretation or
   Wrong answers that include only a part of the exact answer
   Wrong answer due to inability to distinguish between anatomical structures with similar names (i.e., deltoid ligament and deltoid muscle)
4- Wrong answer with wrong explanation.

Since patient data was not used in this study, ethical approval was not required or obtained. We used Microsoft Excel (Microsoft Corp, Redmond, WA) for statistical analysis. Scoring and variables are presented as percentages and continuous variables as mean ± standard deviation (SD) or median (quartiles).

Results: The mean score of ChatGPT-3.5 was 2.14±1.4, while the mean score of ChatGPT-4o was 1.96±1.3. ChatGPT-4 achieved the highest score of 1.44±1.0. The rate of correct answers was 52.5% for ChatGPT-3.5, 74.5% for ChatGPT-4, and 71% for ChatGPT-4o.

If the answers to ChatGPT's hesitant statements were accepted as "don't know", it was evaluated that ChatGPT-3.5 would answer "don't know" to 10 questions, ChatGPT-4 would answer "don't know" to 5 questions, and ChatGPT-4o would answer "don't know" to two questions. Twelve of the ‘don’t know’ responses were true in the answer key, while five were false.

When the wrong answers were evaluated, it was seen that the statements with the answer false were often answered incorrectly. ChatGPT-3.5 evaluated sixty statements that should have been marked as false incorrectly (63%). Thirty-eight wrong answers of ChatGPT-4 were the statements that should be marked as false (74%). The ratio of false statements that ChatGPT-4o incorrectly responded to was 77% (45/58). These results have revealed that ChatGPT versions evaluate false statements as true. The most important reason for this was the inability of artificial intelligence to perceive the nuances in the statements fully. Information containing only part of the true statement was accepted as true by ChatGPT, but the correct answer was false.

Discussion-Conclusion: As the resident's education period increases, the scores are expected to increase. Similarly, ChatGPT can be expected to achieve higher scores with each new version. ChatGPT-4 and 4o had higher scores than ChatGPT-3.5.

Nowadays, when almost everyone can easily access artificial intelligence through their smartphones or mobile devices, artificial intelligence is rapidly entering medicine and health sciences, including education.

Although ChatGPT versions fell short of expectations in true-false questions of RTDE, with the developments of new learning and language models, the exam performance of artificial intelligence will improve.

Keywords: artificial intelligence, ChatGPT, multiple true false, question format

REFERENCE

[1] Ahmed W, Saturno M, Rajjoub R, Duey AH, Zaidat B, Hoang T, Restrepo Mejia M, Gallate ZS, Shrestha N, Tang J, Zapolsky I, Kim JS, Cho SK (2024), “ChatGPT versus NASS clinical guidelines for degenerative spondylolisthesis: a comparative analysis,” *Eur Spine J.*, Mar 15

[2] Brassil C, Couch B (2019). “Multiple-true-false questions reveal more thoroughly the complexity of student thinking than multiple-choice questions: a Bayesian item response model comparison,” *International Journal of STEM Education* 6:16.

[3] Couch B, Hubbard JK, Brassil C (2018). “Multiple–True–False Questions Reveal the Limits of the Multiple–Choice Format for Detecting Students with Incomplete Understandings,” *BioScience* 68:6, 455–463.

[4] Fiedler B, Azua EN, Phillips T, Ahmed AS (2024). “ChatGPT performance on the American Shoulder and Elbow Surgeons maintenance of certification exam,” *J Shoulder Elbow Surg.* 33(9):1888-1893.

[5] Johns WL, Kellish A, Farronato D, Ciccotti MG, Hammoud S (2024). “ChatGPT Can Offer Satisfactory Responses to Common Patient Questions Regarding Elbow Ulnar Collateral Ligament Reconstruction,” *Arthrosc Sports Med Rehabil.* 13;6(2):100893.

[6] Mika AP, Martin JR, Engstrom SM, Polkowski GG, Wilson JM (2023). “Assessing ChatGPT Responses to Common Patient Questions Regarding Total Hip Arthroplasty,” *J Bone Joint Surg Am.* 4;105(19):1519-1526.

[7] Schaap, L., Verkoeijen, P., Schmidt, H. (2014). “Effects of different types of true–false questions on memory awareness and long-term retention,” ****Assessment & Evaluation in Higher Education*, 39***(5), 625–640.

[8] Zhang S, Liau ZQG, Tan KLM, Chua WL (). “Evaluating the accuracy and relevance of ChatGPT responses to frequently asked questions regarding total knee replacement,” *Knee Surg Relat Res.* 2;36(1):15.

PREDICTION OF HEALTH RISKS IN PREGNANCY WITH MACHINE LEARNING MODELS: A LIFE-SAVING DIGITAL SOLUTION AGAINST MATERNAL MORTALITY

Gözde Özsezer[1]*, Gülengül Mermer[2]

[1] *Çanakkale Onsekiz Mart University, Department of Public Health Nursing, Çanakkale, Türkiye*

[2] *Ege University, Department of Public Health Nursing, İzmir, Türkiye*

* *Corresponding author: gozdeozsezer@hotmail.com*

Introduction-Aim: Many pregnancies involve health risks for both mother and baby. Globally, maternal mortality is a serious public health problem, especially in undeveloped and developing countries. Although pregnancy and childbirth are physiological events, the risks faced by women in this age group can cause maternal and neonatal disabilities and, in the most advanced dimension, deaths. In terms of public health, protecting maternal health, predicting the risks, and reducing both mortality and morbidity have an important contribution to maternal, child, and national health. This study aims to use machine learning (ML) models to predict health risks during pregnancy, thereby reducing maternal mortality.

Materials-Methods: The study used 'Maternal Health Risk Data' collected from different hospitals, community clinics, and centers providing maternal health services through an IoT-based risk monitoring system. The dataset was created by Marzia et al. at Daffodil International University in Dhaka, Bangladesh [1]. The dataset contains data collected from pregnant women (n = 1014) on 7 characteristic variables: age, systolic blood pressure, diastolic blood pressure, blood glucose level, body temperature (Fahrenheit), heart rate, and estimated risk intensity during pregnancy.

In this study, ML models were applied for the prediction of health risks during pregnancy. Logistic regression (LR), K-nearest neighbor (KNN), support vector machine (SVM), artificial neural networks (ANN), classification and regression tree (CART), random forest (RF), gradient boosting machine (GBM), XGBoost (XGB), Light GBM, and CatBoost algorithms were applied to the same dataset to determine the most appropriate ML method that gives the highest classification performance result. Prediction rates of ML models were evaluated. In this study, 80% of the data were randomly allocated for training and 20% for testing to analyze the model performances. 10-fold cross validation was applied. The process consists of four stages: data processing, hyperparameter tuning, modeling, and comparative analysis. This forms the high-risk pregnancy prediction model to be used later. The method used Python 3.0 as the main programming language, with libraries such as Numpy, Pandas, and Sci-Kit. Learn on which the algorithms are based. The author's Windows-based personal computer with an Intel i5 7th generation processor and an NVIDIA GeForce 940MX graphics card was used to analyze the ML models using Google Colab as the main IDE.

In this study, different performance comparisons, including accuracy, precision, recall, and F1 score, were calculated separately to determine the best-performing model.

Results: Maternal mortality is a major public health problem that continues to affect women worldwide. To address this issue, it is important to identify women at high risk of maternal mortality and provide them with appropriate care and support. One of the main advantages of using ML models to predict maternal health risks is their ability to analyze large amounts of data and identify patterns and relationships that human analysts may not immediately

notice. This allows for early identification of women at high risk of maternal mortality, early intervention, and improved outcomes.

In this study, pregnant women under 24 years of age had mostly low health risks. From the age of 25 years onwards, it is seen that the risks start to increase in pregnancy, and conversely, the risk decreases in pregnancies over 59 years of age. When systolic blood pressure values are considered, the risk decreases in pregnant women with an upper value lower than 100 mmHg. The higher the systolic blood pressure during pregnancy, the higher the health risk. Diastolic blood pressure is quite similar to systolic blood pressure. Almost all pregnant women with a blood glucose level of 8 have a high health risk. However, the risk level decreases when the blood glucose level is lower than 8. Most pregnant women have a body temperature of 98 F. Pregnant women with a body temperature greater than 100 F are at increased risk. In the dataset, the HeartRate variable has an outlier of 6 bpm, a value that appears unnatural. It is seen that health risks increase with the number of heart rates.

There is a different correlation between each characteristic and high, medium, and low health risk. Most of the correlations had correlation coefficients less than 0.5 and were not very significant. According to the correlation matrix, body temperature is negatively associated with almost all the characteristics in the dataset. There is a low negative correlation between systolic BP, diastolic BP, and heart rate. There is a high positive correlation between systolic BP and diastolic BP characteristics. However, it is important to emphasize the correlation of 0.59 between BS levels and risk level. By looking at these analyses, it can be said that it responds to the fact that high blood glucose levels cause gestational diabetes in the mother. Gestational diabetes that is not well controlled causes the sugar in the baby's blood to rise. The baby becomes overfed and large. This can lead to difficult labor, premature birth, and pre-eclampsia in the mother.

Numerous experiments were conducted to improve classification accuracy. Limitations in the data significantly affected the models. When the accuracy, precision, recall, and F1 score performance metrics of the ML models used for the experiments were compared, LightGBM and CatBoost algorithms provided the highest prediction value with 88% accuracy. CART and RF algorithms with 87% and 86% accuracy, respectively.

Discussion-Conclusion: In the meta-analysis of Sufriyana et al. (2020), it is stated that RF, GBM, and XGB perform well in determining pregnancy risks; LR is also frequently used to develop prediction models, and its prediction rate is low compared to other algorithms [2]. In a prediction study conducted with decision trees, RF, SVC, KNN, and MLP algorithms to predict high-risk pregnancies with data collected from pregnant women living in a village in the Philippines (n = 90), it was reported that RF showed the highest performance [3]. In a study (n = 19331) in Tiajin, China, where gestational diabetes during pregnancy was predicted by machine learning models, it was reported that the XGB algorithm performed better than LR [4]. Marić et al. (2020) compared LR, elastic net, and GBM machine learning models in a retrospective cohort study using data from 16370 deliveries for the early prediction of pre-eclampsia and emphasized that LR had a lower prediction accuracy than other algorithms [5]. Similarly, LR showed low performance in our study. Therefore, considering the results of this study, it can be said that the RF algorithm gives results with high accuracy in predicting pregnancy risks and can be used in the field and clinics.

In conclusion, the use of ML models in the prediction of maternal health risks has the potential to greatly improve maternal health outcomes and reduce maternal mortality rates. While there are some challenges to overcome, the potential benefits make it an area of research worth pursuing. Further research and development in this area has the potential to have a significant impact on the lives of women and their families. In addition, the study could form the basis of a decision support system for nurses and healthcare professionals. It is important to ensure that the use of these models is seamlessly integrated into existing care delivery systems and does not cause disruption.

Keywords: health protection, health risks, machine learning, maternal mortality, pregnancy

REFERENCE

[1] UCI Machine Learning Repository: Maternal Health Risk Data Set Data Set. (n.d.). UCI Machine Learning Repository. Retrieved January 25, 2023, from https://archive.ics.uci.edu/ml/datasets/Maternal+Health+Risk+Data+Set

[2] Sufriyana, H., Husnayain, A., Chen, Y. L., Kuo, C. Y., Singh, O., Yeh, T. Y., ... & Su, E. C. Y. (2020). Comparison of multivariable logistic regression and other machine learning algorithms for prognostic prediction studies in pregnancy care: Systematic review and meta-analysis. JMIR Medical Informatics, 8(11), e16503. https://doi.org/10.2196/16503

[3] Macrohon, J. J. E., Villavicencio, C. N., Inbaraj, X. A., & Jeng, J. H. (2022). A semi-supervised machine learning approach in predicting high-risk pregnancies in the Philippines. Diagnostics, 12(11), 2782. https://doi.org/10.3390/diagnostics12112782

[4] Liu, H., Li, J., Leng, J., Wang, H., Liu, J., Li, W., ... & Yang, X. (2021). Machine learning risk score for prediction of gestational diabetes in early pregnancy in Tianjin, China. Diabetes/Metabolism Research and Reviews, 37(5), e3397. https://doi.org/10.1002/dmrr.3397

[5] Marić, I., Tsur, A., Aghaeepour, N., Montanari, A., Stevenson, D. K., Shaw, G. M., & Winn, V. D. (2020). Early prediction of preeclampsia via machine learning. American Journal of Obstetrics & Gynecology MFM, 2(2), 100100. https://doi.org/10.1016/j.ajogmf.2020.100100

THE ROLE OF GAMIFICATION AND VIRTUAL REALITY IN EARLY DIAGNOSIS OF ALZHEIMER'S DISEASE: A SYSTEMATIC REVIEW

Ceylin Alak[1]*, Emre Olca[2]

[1] Robert College, Istanbul, Türkiye

[2] Software Engineering Department, Maltepe University, Istanbul, Türkiye

** Corresponding author: alacey.26@robcol.k12.tr*

Introduction-Aim: Gamification is an innovative approach that enhances experiences in non-game contexts, such as education and healthcare, by incorporating elements from digital game design—such as scoring, competition, and rewards. By making tasks more engaging and motivational, gamification has proven effective at boosting user involvement and improving outcomes. Unlike traditional games, which focus solely on achieving game-specific objectives, gamification incorporates game-like mechanics into other areas to enhance the experience. It is currently among the top trends in fields like marketing, ranking fourth after artificial intelligence, video content, and influencer marketing, with an adoption rate of 37%.

In healthcare and education, gamification has demonstrated significant impact. In gaming, mobile applications commonly include leveling, rewards, challenges, and leaderboards. In education, platforms like Duolingo and Buddy.ai employ gamification to make language learning more interactive and enjoyable. Likewise, in healthcare, gamification aids patient care and facilitates early diagnosis of conditions such as diabetes, cancer, and Alzheimer's. Specifically, for Alzheimer's, gamification provides a novel approach to tracking cognitive changes over time in a patient-friendly manner.

In recent years, the intersection of gamification and emerging technologies such as AR and VR has propelled the development of engaging applications across various platforms, including web-based, desktop, and mobile games. However, many existing games offer limited player-game interaction, despite its crucial role in capturing user responses effectively. This limited interaction restricts the accuracy with which user reactions can be measured and evaluated. Addressing this gap, there is a growing need for games and virtual environments that can interact with users in more immersive ways to gather data efficiently. AR and VR technologies are particularly well-suited to address this requirement. Virtual Reality (VR) creates specialized, immersive 3D environments, allowing users to engage with alternate realities. Patients can be deeply immersed in VR games featuring gamification elements, where their actions are monitored in real time. VR-compatible tools, such as hand and foot tracking devices, enable the precise collection and evaluation of user movement and interaction data throughout gameplay. Augmented Reality (AR), on the other hand, enhances the real world by blending virtual elements with the physical environment, facilitating interactions with both virtual and real objects. When combined with gamification, AR and VR amplify user engagement by embedding game-like mechanics within these interactive virtual spaces—making them particularly useful in fields like education and healthcare.

This paper focuses on the promising application of gamification and VR in Alzheimer's diagnosis. Gamified VR tools provide innovative methods for tracking cognitive changes over time, contributing to early detection efforts. By incorporating cognitive tasks specifically designed to assess memory, problem-solving, and other critical abilities, this project aims to create a VR game that supports early-stage diagnosis of Alzheimer's disease. Working with neurology specialists, the project integrates these tasks into an engaging gameplay environment. As users interact with the game, data is continuously collected, which is then processed using machine learning algorithms, such as random forests and long short-term memory (LSTM) networks, to detect early signs of cognitive decline associated with Alzheimer’s.

Materials-Methods: Gamification is an innovative technique to improve learning by incorporating digital game design elements, making it more efficient, engaging, and sustainable. Andrzej Marczewski, who is a senior digital experience consultant and labeled as a gamification expert, defines gamification as “the application of gaming metaphors to real-life tasks to influence behavior, improve motivation and enhance engagement” in his book “Gamification: A Simple Introduction”. According to Zicherman and Cunningham, “gamification is the use of game thinking and game mechanics in non-game contexts to engage users in solving problems”. The gamification strategy aims to increase engagement and motivation by integrating reward systems, competition, and interactive challenges, ultimately creating a more dynamic and enjoyable learning experience. Unlike traditional games, which focus solely on task completion, gamification is defined as “a non-game activity, which is established via using game principles”. This distinction emphasizes gamification’s ability to turn monotonous work into entertaining and rewarding experiences.

Gamification has progressed from its original applications, such as incentive systems, to a strong tool in various sectors. Initially, point systems and badges were used to boost client involvement. Gamification 2.0 has gradually moved into industries such as education, healthcare, and business, including more complicated game mechanics such as leaderboards, challenges, storytelling, and personalized experiences. Gamification 2.0 mainly focuses on the growing path and knowledge of the gamer rather than simple reward systems. It consists of techniques that aim for gamers’ growth and development. As Gamification 2.0 developed, Gamification 3.0 was released. Gamification 3.0 uses modern technology like virtual reality and artificial intelligence to deliver immersive experiences that promote behavior change while also improving learning and productivity. Gamification 3.0 also provides more detailed personal statistics and customized feedback, which boosts user engagement and motivation.

Gamification has been applied in various fields, including mobile games, healthcare, and education. In the health industry, it plays an important role in fields such as early Alzheimer’s disease diagnosis. Traditional methods for diagnosing Alzheimer’s disease often involve invasive tests and lengthy diagnostic procedures, which can be uncomfortable and distressing for patients. These problems can be addressed by introducing gamification into Alzheimer's detection, resulting in a more engaging method for early diagnosis. This strategy enhances the patient experience and encourages engagement, potentially leading to earlier diagnoses and improved outcomes.

This project aims to develop an innovative VR game to aid in the early diagnosis of Alzheimer’s disease. By working with neurology specialists, the game includes tasks that assess key cognitive functions like memory and problem-solving. Data from these tasks will be processed using random forests and LSTM networks to detect early signs of Alzheimer’s. The non-invasive nature of the game, coupled with its interactive design, allows for the tracking of cognitive decline over time, offering valuable data to support early diagnosis and intervention. The tasks will include:

- Identifying simple geometric shapes
- Arranging objects neatly
- Performing basic arithmetic calculations

These tasks, designed based on recommendations from neurology and psychology specialists, will help detect early cognitive impairments.

Results: This project proposes the development of a virtual reality (VR) game aimed at facilitating the early diagnosis of Alzheimer's disease by integrating gamification elements. The game is intended for individuals in the early stages of Alzheimer's or those without a formal diagnosis and will include tasks designed in collaboration with neurology specialists. These tasks will assess key cognitive functions such as memory and problem-solving, which are commonly affected by Alzheimer's disease. The data generated from these tasks is expected to be processed using random forests and long short-term memory (LSTM) network algorithms to detect early signs of cognitive decline. By tracking task performance, such as accuracy and response times, the game would offer a non-invasive, engaging method for monitoring cognitive health and supporting early diagnosis.

Discussion-Conclusion: The anticipated impact of this proposed project lies in the potential for VR technology, combined with machine learning algorithms, to transform Alzheimer's diagnosis. Traditional cognitive assessments can often be stressful or mentally taxing for patients, but the interactive nature of the VR game would aim to create a more comfortable and engaging environment, reducing the anxiety typically associated with diagnostic testing. This immersive experience could improve patient participation, leading to more reliable data collection.

The proposed use of random forests and LSTM networks is expected to enhance diagnostic accuracy. Random forests can manage multi-dimensional data effectively, identifying features that are most indicative of cognitive decline, while LSTM networks are suited to track performance changes over time, ideal for capturing gradual deterioration—a hallmark of Alzheimer's disease. Together, these algorithms are anticipated to provide a comprehensive analysis of cognitive health, offering healthcare professionals valuable insights into both current performance and long-term trends.

Compared to existing methods that lack gamification or AI integration, this VR-based approach represents a novel and promising direction for making early detection more accessible and less intrusive. The design of tasks, based on recommendations from neurology and psychology specialists, would increase the relevance of the game in detecting early cognitive impairments. Future work could include expanding the range of cognitive challenges and refining game scenarios, further improving the tool's accuracy and the overall assessment process.

In conclusion, this study proposes the development of a VR game, enhanced by gamification and machine learning algorithms, to advance the early diagnosis and monitoring of Alzheimer's disease. The use of engaging, interactive tasks paired with the robust data analysis capabilities of random forests and LSTM networks is expected to offer a significant step forward in diagnostic methods. With future refinements, this tool has the potential to improve patient outcomes by enabling earlier interventions and more personalized treatment approaches, ultimately enhancing the quality of life for those affected by Alzheimer's.

Keywords: Gamification, Virtual Reality (VR), Alzheimer's Disease.

REFERENCE

[1] Alzheimer's Research UK. (2021, September 29). Sea Hero Quest - Alzheimer's Research UK [Sea Hero Quest - Alzheimer's Research UK]. Alzheimer's Research UK. Retrieved August 21, 2024, from https://www.alzheimersresearchuk.org/research/forresearchers/resources-and-information/sea-heroquest/#:~:text=Although%20the%20free%20consumer%20version,dementia%20research%20and%20other%20fields.

[2] Bozkurt, A. & Genç-Kumtepe, E. (2014). Oyunlaştırma, Oyun Felsefesi ve Eğitim: Gamification. Akademik Bilişim 2014 (s.147- 156). Mersin Üniversitesi, Mersin.

[3] Cloke, H. (2019, August 29). The History of Gamification: From the Beginning to Right Now. Growth Engineering. https://www.growthengineering.co.uk/history-of-gamification/

[4] Cogstate. (2024). Cogstate. Retrieved August 21, 2024, from https://www.cogstate.com/

[5] Dehganzadeh, H., & Dehganzadeh, H. (2020). Investigating effects of digital gamification-based language learning: A systematic review. Journal of English Language Teaching and Learning, 12(25). https://doi.org/10.22034/elt.2020.10676

[6] Ding, H., Wang, B., Hamel, A. P., Melkonyan, M., Ting, Au, R., & Lin, H. (2023). Prediction of progression from mild cognitive impairment to Alzheimer's disease with longitudinal and multimodal data. Frontiers in Dementia, 2. https://doi.org/10.3389/frdem.2023.1271680

[7] Exploring Mobile App Gamification Strategies [Exploring Mobile App Gamification Strategies]. (2024, January 16). MoldStud. Retrieved August 29, 2024, from https://moldstud.com/articles/p-exploringmobile-app-gamification-strategies

[8] Gamification -An Innovative Teaching Method. (2021). https://ec.europa.eu/programmes/erasmus-plus/project-resultcontent/63bd8c6a-86c0-4a75-9196-7ac990770101/Gamification_book__Gamest.pdf

[9] Gamification: A Simple Introduction. (2024). Google Books. https://books.google.com.tr/books?id=IOu9kPjlndYC&printsec=front cover#v=onepage&q&f=false

[10] Kamińska, D., Sapiński, T., Wiak, S., Tikk, T., Haamer, R., Avots, E., Helmi, A., Ozcinar, C., & Anbarjafari, G. (2019). Virtual reality and its applications in education: Survey. Information, 10(10), 318. https://doi.org/10.3390/info10100318

[11] Kutbay, E., & Bozbuğa, N. (n.d.). HEALTH EDUCATION: GAMIFICATION, HEALTH LITERACY, AND THE NEW ERA; METAVERSE SAĞLIK EĞİTİMİ: OYUNLAŞTIRMA, SAĞLIK OKURYAZARLIĞI VE YENİ DÖNEM; ÖTE EVREN. https://doi.org/10.26650/B/ET07.2022.013

[12] Kutbay, E., & Bozbuğa, N. (2021). Yaygın e-Öğrenme tabanlı sağlık eğitimi: Oyunlaştırma ve sağlık okuryazarlığı. Tıp Bilişimi, 743-758. https://doi.org/10.26650/b/et07.2021.003.36

[13] Lumosity: Daily Brain Games. (2018). Lumosity. https://www.lumosity.com/en/

[14] Marache-Francisco, C., & Brangier, E. (2013). Process of Gamification. CENTRIC.

[15] Marougkas, A., Troussas, C., Krouska, A., & Sgouropoulou, C. (2023). How personalized and effective is immersive virtual reality in education? A systematic literature review for the last decade. Multimedia Tools and Applications, 83(6), 18185-18233. https://doi.org/10.1007/s11042-023-15986-7

[16] Preeti GT. (2023). *The Key Difference Between Serious Games and Gamification in eLearning*. Elblearning.com. https://blog.elblearning.com/the-key-difference-between seriousgames-and-gamification inelearning#:~:text=The%20key%20difference%20between%20the,edu cational%20value%20and%20not%20simply

[17] Ravikiran HK, Deepak R, A, D. H., Kumar, P., Sharath S, & H, Y. G. (2024). A robust framework for Alzheimer's disease detection and staging: incorporating multi-feature integration, MRMR feature selection, and Random Forest classification. *Multimedia Tools and Applications*. https://doi.org/10.1007/s11042-024-19875-5

[18] Statista Search Department. (2023, February). *Leading marketing trends according to marketers in Germany, Italy, and the United Kingdom (UK) as of February 2023* [Infographic]. Statista. https://www.statista.com/statistics/1380590/marketing-trendseuropean-countries/

[19] Statista Search Department. (2024). *Leading language learning apps worldwide in January 2024, by downloads* [Infographic]. Statista. https://www.statista.com/statistics/1239522/top-language-learningapps-downloads/

[20] Strickland, E. (2014, April 22). *Start-up Profile: Akili Diagnoses Alzheimer's With a Game*. IEEE Spectrum; IEEE Spectrum. https://spectrum.ieee.org/startup profile-akili-diagnoses-alzheimerswith-a-game

[21] VirtuAAL Research Group. (2019, December 12). *Goals*. VIRTUAAL. Retrieved September 8, 2024, from https://virtu-aal.eu/

[22] What "Gamification" is and what it's not. (2017). *European Journal of Contemporary Education*, *6*(2). https://doi.org/10.13187/ejced.2017.2.221

[23] Zicherman, G. & Cunningham, C. (2011). Gamification by Design: Implementing Game Mechanics in Web and Mobile Apps (1st ed.). Sebastopol, California: O'Reilly Media

HOW READABLE THE CHATGPT AND BARD GENERATED RESPONSES FOR NUTRITION

Volkan Hancı[1†*], Nevin Borzan[2*], İsmail Erdemir[1], Bişar Ergün[3], Ferid Baran Hancı[4†]

[1] *Dokuz Eylül University, Department of Anesthesiology and Critical Care, Faculty of Medicine, İzmir, Türkiye*

[2] *Specialist Dietitian, Aydın, Türkiye*

[3] *Dr. Ismail Fehmi Cumalıoğlu City Hospital, Department of Internal Medicine and Critical Care, Tekirdağ, Türkiye*

[4] *Ostim Technical University, Artificial Intelligence Engineering Department, Faculty of Engineering, Ankara, Türkiye*

[†]*Volkan Hancı and Nevin Borzan are co-first authors*

* *Corresponding author: vhanci@gmail.com*

Introduction-Aim: Our hypothesis in this study is that the readability values of answers which produced by the AI chatbot ChatGPT and Bard according patients' questions related to nutrition are not within the recommended limits. To test this hypothesis, we aimed to evaluate the readability of answers to questions asked ChatGPT and Bard about nutrition using scores.

Materials Methods: For this study, two independent authors (N.B. and V.H) investigated the term "Nutrition" on ChatGPT and Bard on July 6, 2023. Two independent authors (N.B. and V.H) asked ChatGPT and Bard separately "what are most frequently asked top 100 questions about nutrition". After identifying the 100 most frequently asked questions, each question was asked to both ChatGPT and Bard. Two calculators were used to evaluate the readability of the websites. Calculator 1: https://readabilityformulas.com/free-readability-formula-tests.php, and Calculator 2: https://www.online-utility.org/english/readability_test_and_improve.jsp. Texts were evaluated using both calculators. The readability scores were compared and analyzed with the sixth-grade level of readability recommended by the American Medical Association and National Institutes of Health. [1, 2, 3, 4].

The study includes a methodology that does not require the use of human participants, human or animal data. The study includes anonymized data collected on the website, and analyzed from open sources. Ethics committee approval was not required for this study. SPSS Windows 24.0 (SPSS Inc., Chicago, IL, USA) was used for the statistical analysis. A p value of <0.05 was determined as a significant difference.

Results: When the median readability indexes of Chat GPT and Bard were compared with the sixth-grade reading level, a statistically significant difference was observed with all formulas compared to the sixth-grade level (p<0.001). According to all formulas, AI responses had a readability level above sixth grade.

When the readability of the answers given by Chat GPT and Bard was evaluated, it was determined that there was a significant difference between the readability of the answers of both artificial intelligences. Bard's answers were found to be more readable for all formulas of the evaluations. (p<0.001).

Discussion-Conclusion: Readability is an important factor in understanding patient education materials (PEM). Reading materials should be appropriate to the reader's education level. It is recommended to use sentences of 8-10 words to increase the readability of health information. It is reported that simpler words should be used instead of complicated medical

terms. When PEMs on different subjects were examined in previous studies, it was found that their readability levels were above the recommendations, similar to our study [1-4] Investigating the appropriateness of ChatGPT and Bard's answers to medical questions is important because they are used by many patients in a short time to obtain information about their diagnosed diseases [1-4]. To our knowledge, no studies have evaluated the readability of ChatGPT or Bard's nutritional responses.

In the past, the readability and quality of PEM provided by different AI chatbots in different healthcare services have been evaluated in many studies. Most of the studies emphasized that the PEM provided by AI chatbots was of low quality and the readability level was above the recommended Standards [1-8]. According to the results of our current study, when different readability indices were evaluated, the readability of the answers given by AI chatbots about Nutrition was also significantly above the recommended level. In this study, Chat GPT readability classification was found to be Flesh reading ease (23.49-35.29) and Flesch Kincaid class level (13.06-15.28). Considering Google Bard readability, Flesh reading ease (53.41-63.17) was obtained as Flesch Kincaid grade level (7.80-10.19). In terms of all formulas evaluated, Google Bard's responses were found to be more readable. It has been determined that a university degree is required in order to understand the Chat GPT responses. The data of our study are also compatible with previous literature. The readability of both chatbots' responses is above sixth grade. This is above the recommended level. We also compared ChatGPT and Bard's responses and found that Bard's responses were more readable than ChatGPT's responses.

In conclusion; ChatGPT and Bard may offer the possibility of improving health outcomes and patient satisfaction in nutrition by serving as an interactive tool for providing medical information online. However, the current capabilities of ChatGPT and Bard are not sufficient for the readability of responses about nutritione. Efforts should be made to ensure that responses to both are at the appropriate level of readability.

REFERENCE

[1] V. Hancı, B. Ergün, Ş. Gül, Ö. Uzun, İ. Erdemir, FB. Hancı, “Assessment of readability, reliability, and quality of ChatGPT®, BARD®, Gemini®, Copilot®, Perplexity® responses on palliative care,” *Medicine (Baltimore)*, vol. 103(33), pp. e39305, Aug 2024.

[2] V. Hanci, B. Otlu, A.S. Biyikoğlu, “Assessment of the Readability of the Online Patient Education Materials of Intensive and Critical Care Societies,” *Crit Care Med*, vol. 52(2), pp. e47-e57, Feb 2024.

[3] D. Ömür Arça, I. Erdemir, F. Kara, N. Shermatov, M. Odacioğlu, E. İbişoğlu, et al., “Assessing the readability, reliability, and quality of artificial intelligence chatbot responses to the 100 most searched queries about cardiopulmonary resuscitation: An observational study,” *Medicine (Baltimore)*, vol. 103(22), pp. e38352, May 2024.

[4] Ş. Gül, I. Erdemir, V. Hanci, E. Aydoğmuş, Y.S. Erkoç, “How artificial intelligence can provide information about subdural hematoma: Assessment of readability, reliability, and quality of ChatGPT, BARD, and perplexity responses,” Medicine (Baltimore), vol. 103(18), pp. e38009, May 2024.

[5] C.J. McCarthy, S. Berkowitz, V. Ramalingam, M. Ahmed, “Evaluation of an Artificial Intelligence Chatbot for Delivery of Interventional Radiology Patient Education Material: A Comparison with Societal Website Content,” *J Vasc Interv Radiol*, vol. 34(10),pp. 1760-1768.e32, Oct 2023.

[6] B. Momenaei, T. Wakabayashi, A. Shahlaee, A.F. Durrani, S.A. Pandit, K. Wang, et al., “Appropriateness and Readability of ChatGPT-4 generated Responses for Surgical Treatment of Retinal Diseases,” *Ophthalmol Retin*, vol. 7(10), pp. 862-868, Oct 2023.

[7] D. Musheyev, A. Pan, S. Loeb, A.E. Kabarriti, “How Well Do Artificial Intelligence Chatbots Respond to the Top Search Queries About Urological Malignancies?” *Eur Urol,* vol. 85(1), pp. 13-16, Aug 2024.

[8] R. Davis, M. Eppler, O. Ayo-Ajibola, J.C. Loh-Doyle, J. Nabhani, M. Samplaski, et al., “Evaluating the Effectiveness of Artificial Intelligence-powered Large Language Models Application in Disseminating Appropriate and Readable Health Information in Urology,” *J Urol, vol.* 210(4), pp. 688-694, Oct. 2023.

OPTIMIZING TREATMENT OUTCOMES USING MACHINE LEARNING IN PERSONALIZED MEDICATION: CURRENT STATUS AND FUTURE DIRECTIONS

Hunaıda Avvad [1*]

[1] Department of Management Information Systems, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: Hunaida.awwad@bakircay.edu.tr*

Introduction-Aim: The rise of advanced machine learning (ML) technologies has significantly transformed personalized medication, enabling more precise and efficient treatment strategies. By integrating the use of ML, the personalized medicine frameworks can integrate patients' different data like the electronic health records, genomes, and wearables devices to come up with drug response prediction, personalized drug dosing, and treatment plan. This study covers the current stats of how supervised, unsupervised and reinforcement learning algorithms have enabled medical decision making to become more real time and precise especially when it comes to prediction of individual reactions and dosage. However, there are challenges that need to be overcome if ML has to realise its full potential in healthcare and these include; data complexity, Data privacy, and Data ethics. The paper explores also trends and challenges that are possibly to improve the treatment outcomes and integrated outcome of patients such as explainable AI (XAI) and federated learning.

Materials-Methods: This study used a comprehensive literature review to assess the performance of ML applications and challenges to personalized medication. The review looked at various types of ML supervised, unsupervised, and reinforcement techniques among them support vector machine (SVM) which is widely used in classifying health related information as well as techniques such as decision trees (DT). These models were also assessed for their accuracy in forecasting drug responses and categorizing patients into clinically relevant subgroups with reference to their health data: electronic health records (EHRs); genomics data; and patient demographics. Clustering analysis has been demonstrated and categorized in terms of its ability to find patterns of patient datasets and uncover of new subtypes of diseases different treatments plans for the patients [1]. Reinforcement learning paradigms were also considered especially in the analysis of how it can be used to optimize treatment in communicative and interactive approaches based on the patient response and health status thus enhancing optimization of treatment in complex and adaptively structured treatment scenarios [2].

The methodology also examined cutting-edge ML approaches that concern existing gaps in personalized medication including data privacy and interpretability. Examined were Explanatory AI (XAI) approaches to address the capability to improve model interpretability, which would help healthcare professionals trust the outcomes produced by the elaborate models [3]. Another main focus topic was federated learning, which was considered due to emergence of the problems related to patient data privacy and allows to train machine learning models on decentralized data without the need to consolidate it [4]. These are milestones in technology and showed that while we are developing machine learning model for diagnosis, we are also making sure that our models can diagnose the diseases without having to violate the privacy of the patient and also without violating the existing regulations.

Results: ML, as a subdiscipline of AI, has radically enhanced personalized medication management to improve the decision-making process of care providers. With other data sources such as EHRs, genetic information, and real-time data from wearable devices, the ML models can predict a patient's outcomes when given certain treatments. Elimination of the trial-and-error of handling of medications and patients are accorded the opportunity for quicker treatment by their individual health plans. Supervised machine learning, such as decision trees and the support vector method, have been implemented in optimizing drug-response predictions and making recommendations for individual drug dosage. Also, the fact that the application of the ML algorithm is useful for biomarkers and patient classification, identifying those who could benefit from specific therapies. This is especially significant when it comes to diseases such as cancer and autoimmune diseases; nothing is better than customized programs when it comes to medication. Majority of the unsupervised learning methods have been applied in the task of identifying the structure in the patient datasets by partitioning them into various treatment groups. This allows recommendation of treatments that produce the best therapeutic outcomes with the least adverse effects right on time.

The most impactful use of ML has been the computerized decision support systems through which the prediction results in clinicians' treatment plans. Moreover, the reinforcement learning has applied to modify the intervention strategies according to the overall patient improvement, and therefore, the care is being improved continually. This also helps to reduce a cardinal issue affecting the healthcare sector because it provides instant notifications and recommendations to healthcare providers regarding patient care. The importance of using ML as a resource has been established due to the following: the creation of personalized medicine, improving patient satisfaction, improving the rate of success, and its ability to give efficient use of the resources in the health sector.

Discussion-Conclusion: The future of machine learning (ML) in healthcare looks promising, especially with the integration of explainable AI (XAI). This is especially in deep learning where the decision making is quite complex and thus clinicians cannot easily understand the rationale behind them. To this end, XAI seeks to make these models more explainable, to increase the levels of trust between clinicians and engineers as well as clinicians and computers. Future development of XAI is critical for ensuring the application of ML in clinical practices, the need for which concerns more applied aspects of healthcare technology.

ML is set to revolutionize drug discovery and personalized treatment, since AI scans chemical databases and predicts the connection between drug and target faster compared with traditional methods. The integration of AI with genomics will make drugs safer for patients and at the same time will be more effective in the longer term. Also, the effectiveness of the ML's application in telemedicine and home monitoring will further improve in the healthcare delivery to make it less location-based and more patients focused.

Effective ML implementation requires close collaboration among healthcare providers, researchers, technologists, and policymakers. Ethical issues are important especially when it comes to data acquisition, model transparency and fairness. Training healthcare professionals through awareness creation via workshops, interdisciplinary education and research will help in ensuring that those trained employ the ML in delivering better health care services and the promotion of innovation of the heath care tools.

Keywords: machine learning, personalized medicine, personalized treatment

REFERENCE

[1] I. Bica, A. M. Alaa, C. Lambert, and M. van der Schaar, "From real-world patient data to individualized treatment effects using machine learning: current and future methods to address underlying challenges," *Clinical Pharmacology & Therapeutics*, vol. 109, no. 1, pp. 87-100, 2021.

[2] J. Peng, E. C. Jury, P. Dönnes, and C. Ciurtin, "Machine learning techniques for personalised medicine approaches in immune-mediated chronic inflammatory diseases: applications and challenges," *Frontiers in Pharmacology*, vol. 12, p. 720694, 2021.

[3] A. Shinozaki, "Electronic medical records and machine learning in approaches to drug development," in *Artificial Intelligence in Oncology Drug Discovery and Development*. IntechOpen, 2020.

[4] G. S. Ginsburg and K. A. Phillips, "Precision medicine: from science to value," *Health Affairs*, vol. 37, no. 5, pp. 694-701, 2018.

HAND GESTURE RECOGNITION BASED ON DEEP LEARNING TECHNIQUES FOR EMERGENCY REPORT GENERATION

Sherdl Abdalla Hamad[1], M. A. Balafar[1]*, M. Alper Selver[2]

[1] Department of Computer, University of Tabriz, Tabriz, Iran

[2] Department of Electrical and Electronics, University of Dokuz Eylul, İzmir, Türkiye

** Corresponding author: balafarila@tabrizu.ac.ir*

Introduction-Aim: Hand gesture recognition in computer vision is a significant challenge that involves analysing and interpreting human gestures by machines. Existing literature on gesture recognition often works with individual gestures captured in images or videos for classification. This study focuses on identifying human gestures in a continuous stream of data from a live camera in a Realtime, without predefined limits. The challenge is heightened by various lighting conditions, backgrounds, and gesture positions within the same data stream. The study proposes an innovative method for hand gesture recognition to classify gestures from different viewpoints and object sizes. using the fastest and most accurate deep learning framework based on object identification and tracking called THP-YOLOv5. This algorithm is widely known for its ability to identify objects in various backgrounds. To train the model, a Hagrid2023 dataset used which includes a set of accurately annotated hand gestures, considering diverse characteristics for comprehensive gesture recognition. This dataset consists of 18 classes and 5400 images.

Materials-Methods: The main goal of the study is to evaluate the effectiveness of the THP-YOLOv5 algorithm in identifying hand gestures. A real-time system was implemented to capture, analyse, and classify live gesture videos frame by frame. Comparisons with other deep learning models and benchmark datasets validate the effectiveness of this approach, which exhibits superior performance in detecting gestures in challenging conditions. In the discussions on combining image processing algorithms and data augmentation, we will examine algorithms that have the best results for identification and detection and select the best combination of image processing algorithms. Our focus is heavily on evaluating the final accuracy, achieving a high accuracy of over 99.5% on the TPH-YOLOv5 model, and analysing metrics such as mean Average Precision (mAP), precision, recall, and F1 score to comprehensively analyse the model's effectiveness. To visually present the results, we will use implemented charts and figures with Power BI to provide a good understanding of the problem and demonstrate the algorithm's ability to accurately identify the location of hand gestures.

Results: Compared with existing gesture detection and recognition methods, our method show that the model achieved an impressive accuracy exceeding 99.5%, with metrics such as mean Average Precision (mAP), precision, recall, and F1 score providing a thorough evaluation of its effectiveness. Furthermore, this study underscores the potential application of the gesture recognition system in various real-world scenarios, utilizing Raspberry Pi and a camera module to create a compact and efficient solution for gesture detection and monitoring. By programming the gesture recognition algorithm in Python and leveraging OpenCV libraries, we ensured a seamless experience for real-time applications.

Discussion-Conclusion: We proposed an innovative hand gesture recognition method utilizing the advanced THP-YOLOv5 deep learning framework, which enables the fast, accurate, and robust detection and recognition of gestures with complex backgrounds. By training our model on the Hagrid2023 dataset, which encompasses 5400 images across 18 gesture classes, we achieved a robust foundation for comprehensive gesture recognition. The algorithm utilizes both the efficient layer aggregation network module and the CBAM attention mechanism module, which reduces the model's parameters while enhancing the feature extraction ability of the baseline network and the robustness in detecting complex backgrounds. Our research demonstrated the effectiveness of the THP-YOLOv5 algorithm in capturing and analyzing live gesture videos frame by frame, resulting in outstanding performance even under challenging conditions.

Keywords: YOLOv5, computer vision, gestures

INVESTIGATION OF 8-WEEK ARTIFICIAL INTELLIGENCE BASED STRUCTURED EXERCISE PROGRAM IN HEALTHY ADULTS

Dilek Nur Kara[1]*, Elif Dilem Ocak[1] Tezel Yıldırım Şahan[2]

[1] University of Health Sciences, Gulhane Institute of Health Sciences, Ankara, Türkiye

[2] University of Health Sciences, Gulhane, Orthopedic Physiotherapy and Rehabilitation Department, Ankara, Türkiye

** Corresponding author: pt.dileknurkara@gmail.com*

Introduction: Posture is defined as the alignment of body parts relative to each other and depends on gravity, muscle tension and the integrity of bone structures [1,2]. Posture disorders have become widespread in young adults due to modern lifestyle and these disorders are usually caused by prolonged sitting, low physical activity and use of digital devices [3,4].

The most common disorders in the literature include head-forward posture, rounded shoulders, kyphosis and lumbar lordosis [5,6]. Proper posture prevents injuries and improves the functions of organs by ensuring efficient functioning of muscles and joints. On the contrary, poor posture can cause muscle and headaches, fatigue and permanent deformations [5,9]. Studies have shown that a properly planned exercise program is effective in correcting posture disorders, especially in young adults [10,11]. A meta-analysis by Dimitrijević et al. revealed that corrective exercise programs were effective in reducing lumbar lordotic angle in individuals with lumbar lordosis and hyperlordosis [11]. Similarly, another study by Sheikhhoseini et al. showed that therapeutic exercises improved craniovertebral angle and alleviated neck pain in individuals with head-forward posture [12].

The positive effects of exercise programs on posture have been shown, and artificial intelligence-based exercises are thought to play an important role in the future for posture correction [7,8]. Regular exercise is an effective method for posture correction and musculoskeletal health [7,10]. Artificial intelligence-based exercise programs are also becoming increasingly popular and are used to create personalized plans [13,14]. AI-based exercise programs are thought to be effective in correcting posture disorders in young adults [14]. However, studies on AI-based exercise programs are still limited and more research is needed in this field [8].

In this study, the effects of an 8-week artificial intelligence-based exercise program were examined in young adults. This study, which was conducted to evaluate the effects of AI-based programs on correcting posture disorders and general physical health in young adults, will make an important contribution to the scientific literature. AI-based exercises are thought to have great potential in creating customized exercise programs for individuals with posture disorders. This study will provide an important basis for future research.

Methods: In this study, an AI-supported exercise program was implemented on 44 voluntary young adults aged between 18-25 (mean age 20.97±0.35 years). Participants were divided into two groups: those who regularly exercised (n=19) and those who were sedentary (n=25). The exercise program was conducted three times a week for 8 weeks under the supervision of a physiotherapist, using the MuscleWiki AI application. Postural assessments, including craniovertebral angle, shoulder angle, pelvic tilt, and knee angles, were measured using the Posture Screen Mobile application. Pre- and post-treatment values were compared.

Results: The exercise group and the control group were similar in terms of demographic characteristics (p>0.05). After 8 weeks, when comparing post-treatment craniovertebral, shoulder, pelvic tilt, and knee angles between the groups, a significant difference was observed only in knee angles (p>0.05, p_knee=0.007). In the within-group comparisons of pre- and post-treatment values, a significant difference was found solely in knee angles (p<0.001).

Conclusion: This study demonstrated that an AI-structured exercise program has positive effects on healthy adults. It was also found that AI-structured exercise programs could be utilized to promote the continuity of a healthy lifestyle.

Keywords: artificial intelligence, posture, exercise therapy

REFERENCE

[1] G. Yagci and N. Bek, “Genç yetişkinlerde postürün sağlıkla ilgili yaşam kalitesi üzerine etkisinin araştırılması,” *Türk Fizyoterapi ve Rehabilitasyon Dergisi*, vol. 32, pp. xx-xx, 2021.

[2] C. Raine and L. Twomey, “Attributes and qualities of human posture and their relationship to dysfunction or musculoskeletal pain,” *Critical Reviews in Physical and Rehabilitation Medicine*, vol. 6, pp. 409-409, 1994.

[3] F. R. França, T. N. Burke, E. S. Hanada, A. P. Marques, and L. A. Ramos, “Effect of global posture reeducation and of static stretching on pain, range of motion, and quality of life in women with chronic neck pain: a randomized clinical trial,” *Clinics*, vol. 70, pp. 674-679, Oct. 2015.

[4]. M. Salsali, R. Sheikhhoseini, P. Sayyadi, J. A. Hides, M. Dadfar, and H. Piri, “Association between physical activity and body posture: a systematic review and meta-analysis,” *BMC Public Health*, vol. 23, p. 1670, 2023.

[5] D. Sharan, M. Mohandoss, and R. Ranganathan, “Postural health in university students: a cross-sectional study,” *Journal of Back and Musculoskeletal Rehabilitation*, vol. 27, pp. 331-338, 2014.

[6] O. Hershkovich et al., “Associations of body mass index and body height with low back pain in 829,791 adolescents,” *American Journal of Epidemiology*, vol. 178, pp. 603-609, Oct. 2013.

[7] A. V. Bruyneel and N. C. Duclos, “Effects of the use of mobile phone on postural and locomotor tasks: a scoping review,” *Gait & Posture*, vol. 82, pp. 233-241, 2020.

[8] H. K. Chen, F. H. Chen, and S. F. Lin, “An AI-based exercise prescription recommendation system,” *Applied Sciences*, vol. 11, p. 2661, Mar. 2021.

[9] R. Hokmabadi and P. Sepehr, “Assessing the posture and predicting the factors affecting musculoskeletal disorders in computer users by neural networks,” *Journal of Health and Safety at Work*, vol. 11, pp. 700-719, 2021.

[10] N. Miyakoshi et al., “Back extensor strength and lumbar spinal mobility are predictors of quality of life in patients with postmenopausal osteoporosis,” *Osteoporosis International*, vol. 18, pp. 1397-1403, 2007.

[11] V. Dimitrijević, T. Šćepanović, V. Milankov, M. Milankov, and P. Drid, “Effects of corrective exercises on lumbar lordotic angle correction: a systematic review and meta-analysis,” *International Journal of Environmental Research and Public Health*, vol. 19, p. 4906, Aug. 2022.

[12] R. Sheikhhoseini, S. Shahrbanian, P. Sayyadi, and K. O’Sullivan, “Effectiveness of therapeutic exercise on forward head posture: a systematic review and meta-analysis,” *Journal of Manipulative and Physiological Therapeutics*, vol. 41, pp. 530-539, Jun. 2018.

[13] G. Prabhu, N. E. O’Connor, and K. Moran, “Recognition and repetition counting for local muscular endurance exercises in exercise-based rehabilitation: A comparative study using artificial intelligence models,” *Sensors*, vol. 20, p. 4791, Sep. 2020.

[14] S. Jungreitmayr, C. Kranzinger, V. Venek, and S. Ring-Dimitriou, “Effects of an app-based physical exercise program on selected parameters of physical fitness of females in retirement: a randomized controlled trial,” *Frontiers in Physiology*, vol. 13, p. 821773, 2022.

KEY CHALLENGES, DRIVERS, AND SOLUTIONS IN INTEGRATING KNOWLEDGE GRAPHS WITH EXPLAINABLE AI FOR HEALTHCARE

***Fares A. Dael*[1*]**

[1]*Department of Management Information Systems, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: fares.dael@bakircay.edu.tr*

Introduction-Aim: The application of AI in medical diagnosis has been rising due to possible improvements in diagnostic precision and treatment plans [1], [2]. Nonetheless, most AI models are unclear regarding decision-making, which causes clinicians' anxiety and hinders the models' deployment. One of them is the overwhelming volume of medical information which poses an issue of incorporating such knowledge into the black-box like architecture of AI systems. Knowledge Graphs (KGs) provide a clear and organized model for representing medical knowledge, they appear to have the potential to solve the problem [3]. This paper identifies major concerns, potential, and recommendations on how to advance the interoperability of knowledge graphs and Explainable AI (XAI) in healthcare decisions. Challenges covered are data quality and alignment, knowledge graphs scalability in terms of large-scale knowledge graphs, and more importantly trust regarding AI-produced clinical guidance. Drivers are the desire to reveal AI integration in healthcare, the transition to precision medicine, new forms of health data, and more requirements for transparency in technology. Potential solutions include the organization of the healthcare data, a two-approach model employing AI, a system with a graph database that can be horizontally scaled up, a clinician-in-the-loop model, and improvement to the explanation abilities such as graph neural networks (GNNs) where necessary. The integration of KGs with XAI may revolutionize healthcare decision support in the study's findings.

Materials-Methods: This paper relies on conceptual research methodology, examining existing literature on KGs and XAI in healthcare decision support systems. It emphasizes frameworks and methodologies in the construction of KGs and XAI but focuses on health care. The first part provides an overview of limitations for creating and using KGs in healthcare including concerns with data collection from various sources including electronic health records (EHRs), medical ontologies, and clinical decision support systems (CDSS), and the level of complexity of knowledge graphs that are expected to increase with time [3], [4]. This stage also includes case studies of AI models using knowledge graphs in disease diagnosis, treatment recommendations, and patient outcome prediction.

The second part defines the driving forces leading to KGs and XAI integration, including the need for transparency in prediction, development of personalized medicine, availability of standard datasets, and call for explainability due to regulatory compliance requirements [5], [6]. This survey specifically focuses on techniques and approaches of how to build KGs for practical use in clinical decision-supporting systems and their embedding into healthcare applications. Proposed strategies are the use of standard health care information using Fast Healthcare Interoperability Resources (FHIR) and Systematized Nomenclature of Medicine (SNOMED), the usage of rules in combination with machine learning, and the integration of clinician's feedback to update the knowledge graph [7], [8].

Results: There are difficulties arising from the application of KGs with XAI in health care, including difficulty in managing complex medical knowledge and quality data, integration difficulties arising from the use of KGs, and the issue of scalability of knowledge graphs due to their large size. Still, some drivers can be named including the need for AI transparency,

the progress of healthcare data, and the pressure from regulators, which also contribute to using XAI in clinical practice. Data standardization, employment of modern graph database technologies, feedback from clinicians as part of a model's output, and usage of more elaborate methods based on GNNs for better explainability all serve to help integrate AI into the clinical decision-making process.

Discussion-Conclusion: The use of KGs combined with XAI has further potential to deliver effective transparent, efficient healthcare decision support systems. With the advent of KGs and XAI, there is the potential to increase the positive impact on clinical decision-making pertaining to transparency, trust, and patient care. Future research should focus on improving the scalability of these systems and evaluating their real-world impact across diverse clinical environments.

Keywords: Decision support, Explainable AI, Healthcare, Knowledge Graphs, Transparency

REFERENCE

[1] M. Khalifa and M. Albadawy, "AI in diagnostic imaging: Revolutionising accuracy and efficiency," *Comput. Methods Programs Biomed. Update*, vol. 5, p. 100146, Jan. 2024, doi: 10.1016/j.cmpbup.2024.100146.

[2] Z.-H. Chen, L. Lin, C.-F. Wu, C.-F. Li, R.-H. Xu, and Y. Sun, "Artificial intelligence for assisting cancer diagnosis and treatment in the era of precision medicine," *Cancer Commun.*, vol. 41, no. 11, pp. 1100–1115, 2021, doi: 10.1002/cac2.12215.

[3] E. Rajabi and S. Kafaie, "Knowledge Graphs and Explainable AI in Healthcare," *Information*, vol. 13, no. 10, Art. no. 10, Oct. 2022, doi: 10.3390/info13100459.

[4] M.-E. Vidal, E. Niazmand, P. D. Rohde, E. Iglesias, and A. Sakor, "Challenges for Healthcare Data Analytics Over Knowledge Graphs," in *Transactions on Large-Scale Data- and Knowledge-Centered Systems LIV: Special Issue on Data Management - Principles, Technologies, and Applications*, A. Hameurlain, A. M. Tjoa, O. Boucelma, and F. Toumani, Eds., Berlin, Heidelberg: Springer, 2023, pp. 89–118. doi: 10.1007/978-3-662-68014-8_4.

[5] J. H. Bettencourt-Silva *et al.*, "Exploring the Social Drivers of Health During a Pandemic: Leveraging Knowledge Graphs and Population Trends in COVID-19," in *Integrated Citizen Centered Digital Health and Social Care*, IOS Press, 2020, pp. 6–11. doi: 10.3233/SHTI200684.

[6] S. Garg, S. Parikh, and S. Garg, "Navigating Healthcare Insights: A Bird's Eye View of Explainability with Knowledge Graphs," in *2023 IEEE Sixth International Conference on Artificial Intelligence and Knowledge Engineering (AIKE)*, Sep. 2023, pp. 54–61. doi: 10.1109/AIKE59827.2023.00016.

[7] B. Abu-Salih, M. AL-Qurishi, M. Alweshah, M. AL-Smadi, R. Alfayez, and H. Saadeh, "Healthcare knowledge graph construction: A systematic review of the state-of-the-art, open issues, and opportunities," *J. Big Data*, vol. 10, no. 1, p. 81, May 2023, doi: 10.1186/s40537-023-00774-9.

[8] C. Peng, F. Xia, M. Naseriparsa, and F. Osborne, "Knowledge Graphs: Opportunities and Challenges," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 13071–13102, Nov. 2023, doi: 10.1007/s10462-023-10465-9.

A SCALE DEVELOPMENT STUDY: ETHICAL SENSITIVITY TOWARDS ARTIFICIAL INTELLIGENCE AND ROBOT NURSES

Eda Ergin[1*], Gamze Göke Arslan[2], Şebnem Çınar Yücel[3]

[1] *İzmir Bakırçay University Faculty of Health Science, Department of Fundamentals of Nursing, İzmir, Turkiye*

[2] *Department of Nursing, Faculty of Health Sciences, Karamanoğlu Mehmetbey University, Karaman, Turkiye*

[3]*Faculty of Nursing, Ege University, İzmir, Türkiye*

** Corresponding author: eda.ergin@bakircay.edu.tr*

Introduction-Aim: Artificial intelligence and robotic technologies have controversial ethical issues and some disadvantages for the nursing profession. Robots with artificial intelligence can observe patients, and record and transfer patient information. However, without appropriate regulations, procedures, and protocols, these capabilities may pose a threat to the privacy of patients and other individuals who interact with them [1,2], which brings about the issue of patient privacy in addition to ethical problems. The present study was aimed at developing a valid and reliable measurement tool that can assess ethical sensitivities about artificial intelligence and robot nurses in the field of nursing and testing this scale in nursing students.

Materials-Methods: A methodological study was conducted in two phases: (1) development of the scale through a literature review and interviews related to Ethical Sensitivity towards Artificial Intelligence and Robot Nurses; 2) confirming construct validity, criterion-related validity and reliability of the developed scale. The data were collected from 356 nursing students studying at the Nursing Department of a university in Türkiye between November 2022 and December 2022. Descriptive data were presented in numbers, percentages, arithmetic mean, standard deviation, and minimum and maximum values. The significance level was accepted as $p<0.05$. Shapiro-Wilk test was used to test the normality of scale averages. The Content Validity Index was used in the compatibility analysis of expert opinions about the ESSAIRN.Kaiser-Meyer-Olkin (KMO) Sampling Appropriateness Measure and Bartlett's Test of Sphericity were used to determine whether the data were suitable for principal component analysis. Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used to determine whether the items and sub-dimensions could explain the original structure of the scale. Since the draft scale consisted of 19 items, the data was divided into two and analysis was carried out based on the data obtained from 178 students for the EFA and 178 students for the CFA. Covariance Matrix and Principal Axis Factoring estimation methods were used for the EFA. Maximum Likelihood and Covariance Matrix were also used for the CFA. Pearson correlation analysis was used for the item-total score analysis of the scale and sub-dimensions. The Cronbach's Alpha coefficient was used to determine the internal consistency of the scale and sub-dimensions. To analyze the data, the SPSS (Statistical Package for the Social Sciences) 22.0 (SPSS Inc., Chicago, IL, USA) and IBM SPSS Amos 24 analysis programs were used [3].

Results: In the study, 356 nursing students were included. Of them, 102 (28.7%) were first-year students, 89 (25.0%) were second-year students, 74 (20.8%) were third-year students and 91 (25.6%) were fourth-year students. Their mean age was 20.41 (SD 1.45) years. Of the participants, more than half were women (n=240, 67.4%), 32.9% (n=117) had lived in a metropolis for the longest time and slightly more than half (n = 194, 54.5%) chose the nursing department willingly. Content validity analysis of the 19-item ESSAIRN was conducted and

the content validity index for the overall scale was determined as 0.93. For the EFA, the KMO coefficient was calculated and the Bartlett's Test of Sphericity was used [3]. The KMO value of the ESSAIRN was 0.76 and the results of Bartlett's Test of Sphericity were χ2=1174.25, p= 0.000. According to the first exploratory factor analysis of the 19 items, five factors explained 59.85% of the total variance and 46.45% of the common variance, and the eigenvalue of all the five factors was higher than 1. According to Büyüköztürk [4], the difference between the load values should be as high as possible and at least 0.10. After the exploratory factor analysis conducted with the remaining 17 items, the scale items consisted of four sub-factors with eigenvalues greater than 1. According to the results of the Confirmatory Factor Analysis, fit indices were determined as follows: χ^2/SD=1.979, Root Mean Square Error of Approximation =0.074, Comparative Fit Index =0.894, Incremental Fit Index =0.897 and Goodness of Fit Index =0.880. The Cronbach's alpha for the overall ESSAIRN was calculated as 0.830.

Discussion-Conclusion: In the present study, it was concluded that the ESSAIRN, developed to determine ethical sensitivity towards artificial intelligence and robot nurses, was a valid and reliable measurement tool. It is recommended that the ESSAIRN should be used in different sample groups from different cultures and societies in order to assess ethical sensitivity towards artificial intelligence and robot nurses. It is also recommended that, in future studies, the test-retest method should be used to measure ethical sensitivity towards artificial intelligence and robot nurses, and that the parallel form reliability method should be used to test construct validity from different aspects.

Keywords: Artificial Intelligence, Robot Nurse, Validity, Reliability

REFERENCE

[1] Güvercin CH. Artificial Intelligence in Medicine and Ethics. Yapay Zeka ve Tıp Etiği 2020;1:7–13.

[2] Kandemir F, Azizoğlu F, Terzi B. Use of Artificial Intelligence and Robotic Technologies in Nursing. J Intensive Care Nurs 2023;27:118–27.

[3] Field A. Discovering statistics using IBM SPSS statistics. sage; 2013.

[4] Büyüköztürk S, Kılıc Cakmak E, Akgün ÖE, Karadeniz S, Demirel F. Scientific Research Methods in Education. 25. Baskı. Ankara: Pegem Akademi; 2018.

DESIGN OF CARDIAC PACEMAKER CONTROLLER BASED ON REINFORCEMENT LEARNING

Kağan Orbay[1*], Mehmet Sağbaş[1], Murat Demir[1]

[1] Electrical and Electronics Engineering, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: mehmet.sagbas@bakircay.edu.tr*

Introduction-Aim: Heart attacks represent a critical global health issue, leading to millions of fatalities each year due to cardiac arrest and arrhythmias, with the latter playing a significant role in these conditions (Ponikowski et al., 2014; Bajpai, Alam, & Ali, 2017). Medical devices, particularly pacemakers, have proven essential in managing arrhythmias by stabilizing heart rhythms through electrical stimulation (Arunachalam et al., 2016). These devices are typically implanted to continuously monitor and regulate heartbeats, providing a more stable heart rate (Bikki, Dhiraj, & Kumar, 2023). However, advancements in pacemaker efficiency are sought to address heart rate control accurately. Consequently, researchers have explored various control strategies, including PID (Proportional-Integral-Derivative) controllers, to improve pacemaker responsiveness and adaptability to patients' physiological needs. This study aims to evaluate the effectiveness of PID controller parameters derived using both Genetic Algorithms (GA) and Reinforcement Learning (RL) methods, comparing their performances in terms of response times, overshoot values, and settling times. By analysing these parameters, we seek to identify the optimal method for enhancing pacemaker functionality.

Materials-Methods: In this study, the control parameters for a PID controller were derived using Genetic Algorithms (GA) and Reinforcement Learning (RL) techniques. GA was chosen for its wide acceptance and effectiveness in optimizing complex systems, while RL was included due to its adaptability and learning capabilities in dynamic environments. Several mathematical models of cardiac dynamics served as a basis for developing control algorithms, drawing from the model established by Biswas, Das, and Guha in 2006, which uses a closed-loop negative unit feedback system. Both GA and RL methods were applied to tune PID controller parameters aimed at optimizing pacemaker performance. The effectiveness of each approach was then assessed by simulating the step response of the pacemaker control system, focusing on key metrics such as rise time, settling time, and overshoot. These simulations provided a comparative analysis of the control parameters' effectiveness for both methods.

Results: The simulation results demonstrate that both the GA and RL-optimized PID controllers enhance pacemaker performance significantly compared to an uncontrolled system. The RL-optimized PID controller produced a step response with an overshoot of 1.14% and a maximum peak value of 72.83, which is notably lower than the overshoot of 28.82% and peak value of 92.75 achieved by the GA-optimized controller. However, both methods exhibited improvements in control performance, though with differing strengths across measured parameters.

Discussion-Conclusion: This study highlights the effectiveness of using PID controllers optimized by GA and RL techniques in regulating pacemaker heart rhythms. The findings indicate that while RL provides superior control over overshoot and peak response, GA excels in terms of response time and settling time. These differences suggest that each method offers distinct advantages: RL for cases requiring precision control with minimal overshoot, and GA for applications demanding rapid response times. In conclusion, the choice of optimization technique may depend on specific clinical requirements, with a potential for integrating both methods in a hybrid approach to achieve an ideal balance between response time and accuracy in pacemaker performance. Further research into adaptive hybrid models could enhance pacemaker responsiveness and adaptability, optimizing patient outcomes in diverse physiological conditions.

Keywords: Heart rhythm regulation, pacemaker control system, PID controller optimization, reinforcement learning

REFERENCE

Ponikowski, P., Anker, S. D., AlHabib, K. F., Cowie, M. R., Force, T. L., Hu, S., et al. (2014). Heart failure: preventing disease and death worldwide. ESC Heart Failure, 1(1), 4–25.

Bajpai, S., Alam, S., Ali, M.A.: Intelligent Heart Rate Controller using Fractional Order PID Controller Tuned by Genetic Algorithm for Pacemaker. International Journal of Engineering Research & Technology. 6(05), 715-720 (2017).

Arunachalam, S. P., Kapa, S., Mulpuru, S.K., Friedman P.A. ve Tolkacheva, E.G., Intelligent Fractional-Order PID (FOPID) Heart Rate Controller for Cardiac Pacemaker, 2016 IEEE Healthcare Innovation Point-Of-Care Technologies Conference (HI-POCT), Cancun, Meksika, 2016, s. 105-108.

Bikki, P., Dhiraj, Y., & Kumar, R. N. (2023). Implementation of a Dual-Chamber Pacemaker for Low-Power Applications. https://doi.org/10.1109/icecct56650.2023.10179677

Lima, G.S., Savi, M.A., Bessa, W.M.: Intelligent control of cardiac rhythms using artificial neural networks. Nonlinear Dynamics. 111(12), 11543–11557 (2023). https://doi.org/10.1007/s11071-023-08447-1

Momani, S., Batiha I.M., El-Khazali, R.: Design of PIλDδ-Heart Rate Controllers for Cardiac Pacemaker, 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), Ajman, United Arab Emirates, pp. 1-5, (2019). https://doi.org/10.1109/ISSPIT47144.2019.9001785

Nako, J., Psychalinos, C., Elwakil, A.S.: Minimum Active Component Count Design of a PIλDμ Controller and Its Application in a Cardiac Pacemaker System. J. Low Power Electron. Appl. (2023). https://doi.org/10.3390/jlpea13010013

Srivastava, R., Kumar, B.: Design of Anfis based pacemaker controller having improved transient response and its FPGA implementation. Biomedical Signal Processing and Control (2022). https://doi.org/10.1016/j.bspc.2021.103186

Chen, E.Z., Wang, P., Chen, X., Chen, T., Sun, S.: Pyramid Convolutional RNN for MRI Image Reconstruction. in IEEE Transactions on Medical Imaging. 41(8), 2033-2047 (2022). https://doi.org/10.1109/TMI.2022.3153849

Yürdem, B., Aksu, M.F. & Sağbaş, M. Design of Fractional/Integer Order PID Controller Using Single DVCC and Its Cardiac Pacemaker Application. Circuits Syst Signal Process (2024). https://doi.org/10.1007/s00034-024-02810-2

Biswas, S.C., Das, A., Guha, P., Mathematical Model of Cardiovascular System by Transfer function Method, Calcutta Medical Journal, 2010.

DIAGNOSIS OF MALIGNANT MESOTHELIOMA USING MACHINES LEARNING

Murat Demir[1*]

[1]_Electrical and Electronics Engineering, İzmir Bakırçay University, İzmir, Türkiye_

[*]_Corresponding author: murat.demir@bakircay.edu.tr_

Introduction-Aim: Malignant mesothelioma (MM) is an aggressive and rare cancer that affects the lining of the lungs (pleura), abdomen (peritoneum) and, less commonly, the lining around the heart (pericardium) [1-3]. The main cause is long-term exposure to naturally occurring mineral fibres such as asbestos and erionite. These minerals pose a high risk, particularly to people working in industries such as construction, mining and shipbuilding. Erionite, another fibrous mineral, is found in specific regions, notably in Türkiye, where its use in construction materials has led to high rates of mesothelioma among residents [4-6].

Inhalation of asbestos and erionite fibres causes inflammation and cellular damage in the pleural tissue, which can eventually lead to cancer. Because the disease progresses slowly over several decades, malignant mesothelioma often presents long after initial exposure, making it difficult to diagnose in its early stages. Its symptoms, such as coughing, shortness of breath, chest pain, and fatigue, are non-specific and can easily be mistaken for other respiratory conditions. As a result, mesothelioma is often diagnosed at an advanced stage when treatment options are more limited [7].

Materials-Methods: In recent years, artificial intelligence (AI) technologies have become increasingly important in the diagnosis of complex diseases such as malignant mesothelioma, which is difficult to detect in its early stages. AI uses machine learning algorithms trained on large datasets to diagnose diseases and identify patterns based on clinical data [3,8-9]. This study focuses on improving the detection of malignant mesothelioma by analysing data, particularly medical reports, collected from patients at Dicle University, Faculty of Medicine. For this purpose, data from a total of 324 patients with malignant mesothelioma (MM) were obtained and classified by AI to determine whether they had the disease. The dataset includes 34 of the most important features, selected by physicians, for all samples. AI's ability to process and interpret complex datasets quickly and accurately offers significant advantages, particularly in non-invasive diagnostic procedures, and can be used as a decision support system [10].

Results: Support vector machine (SVM), XGBoost and logistic regression models have shown exceptional performance in diagnosing malignant mesothelioma, achieving highest accuracy rate. These models analyse various factors such as patients' age, gender, the type of mineral exposure, and other clinical findings to assess the likelihood of the disease. All models rely on clinical data, not imaging techniques or biopsies, making the diagnostic process less invasive and more efficient.

The use of AI in mesothelioma diagnosis offers another key benefit: reducing the workload on doctors and minimizing human error. Diagnosing mesothelioma is a complex process that involves considering numerous variables and interpreting complicated medical data. However, AI models can quickly process these variables and provide reliable diagnostic support, helping doctors make faster and more informed decisions. As a result, patients can begin treatment earlier, potentially improving their quality of life and increasing their chances of survival. Early detection is especially vital in malignant mesothelioma, where treatment outcomes are closely tied to how early the disease is diagnosed [11].

Discussion-Conclusion: In conclusion, malignant mesothelioma remains a serious public health concern due to widespread past exposure to asbestos and erionite, and new cases continue to emerge. However, advances in AI technology offer new hope for early detection and diagnosis. AI models such as SVM, XGBoost and logistic regression, with their high accuracy rates, are changing the way clinicians approach mesothelioma, enabling faster and more effective diagnoses. Integrating AI into the diagnostic process not only saves time, but also speeds up the treatment process, giving patients a better chance of living longer, healthier lives.

Keywords: AI Models for cancer detection, artificial intelligence in diagnosis, asbestos exposure, malignant mesothelioma.

REFERENCE

[1] Pass, H. I., Vogelzang, N. J., & Carbone, M. (2010). Malignant pleural mesothelioma: biology, diagnosis, and emerging therapies. Annals of Internal Medicine, 129(4), 250-260.

[2] Zervos M.D., Bizekis .C, Pass H.I. Malignant mesothelioma 2008. Curr Opin Pulm Med 2008; 14:303–9.

[3] Er O, Tanrikulu AC, Akbay A, Temurtas F, An approach based on probabilistic neural network for diagnosis of Mesothelioma's disease. Computers and Electrical Engineering, 2012;38, 75-81.

[4] Carbone, M., Baris, Y. I., Bertino, P., Brass, B., & Pass, H. (2011). Erionite exposure in North Dakota and Turkish villages with mesothelioma. Proceedings of the National Academy of Sciences, 108(33), 13618-13623.

[5] Metintas S, Metintas M, Ucgun I, Oner U. Malignant mesothelioma due to environmental exposure to asbestos. Follow-up of a Turkish cohort living in a rural area. Chest 2002; 122:2224–9.

[6] Tanrikulu AC, Senyigit A, Dagli CE, Babayigit C, Abakay A. Environmental malignant pleural mesothelioma in Southeast Türkiye. Saudi Med J 2006;27(10):1605–7.

[7] Kamp, D. W., & Weitzman, S. A. (1999). Asbestos-induced lung injury: mechanisms and cellular targets. Environmental Health Perspectives, 107(Suppl 5), 785-788.

[8] Gill, T. S., Shirazi, M. A., & Zaidi, S. S. H. (2023). Early detection of mesothelioma using machine learning algorithms. Engineering Proceedings, 46(1), 6.

[9] Er O, Yumusak N, Temurtas F. Chest diseases diagnosis using artificial neural networks. Expert Systems Appl 2010;37(12):7648–55.

[10] Faisal, F. E., & Abo-Hamed, Z. (2021). Applications of machine learning for mesothelioma diagnosis. Journal of Biomedical Informatics, 116, 103719.

[11] Bibault, J. E., Giraud, P., & Burgun, A. (2016). AI in medical diagnostics: reducing errors and improving outcomes in cancer detection. Cancer Research, 76(11), 322-330.

CLASSIFICATION OF DIABETIC RETINOPATHY DISEASE WITH DEEP LEARNING METHODS

***Metin Tuncel*[1*]*, Murat Uçar*[1]**

[1] Computer Engineering, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: tuncel155@gmail.com*

Introduction-Aim: Diabetes is defined as a chronic disease that occurs as a result of elevated blood sugar levels (hyperglycaemia), in which the organism cannot make sufficient use of carbohydrates, fats and proteins due to the inability of the pancreas to produce enough insulin hormone or to perform its function. According to the chronic diseases report published by the World Health Organisation, diabetes ranks first in terms of intensity. One of the complications of medicine tip1 diabetes is that it causes diabetic retinopathy. Diabetic retinopathy is defined as an eye condition caused by damage to the blood vessels in the light-sensitive tissue (retina) located at the back of the eye due to diabetes complications. According to the International Diabetes Federation (2021) Diabetes Atlas 10th Edition, diabetes is among the top three diseases that cause blindness. Blindness caused by diabetes is mostly caused by the destruction of small vessels in the retina due to long-term hyperglycaemia. Approximately 25% of diabetic patients worldwide have diabetic retinopathy at any level. There are approximately 2 million diabetic patients in our country and 25% of these patients have diabetic retinopathy. There are 5 classes of diabetic retinopathy. These are non-proliferative diabetic retinopathy (npdr), mild non-proliferative retinopathy, moderate non-proliferative retinopathy, severe non-proliferative retinopathy, proliferative diabetic retinopathy (pdr) from the lowest to the most severe. In this study, using the APTOS2019 dataset, a computer-aided diagnosis system is created to help doctors make early diagnosis with convolution-based deep learning models. Two- and five-class classification was performed using state of the art models VGG16, InceptionResNetV2, ResNet152V2, EfficientNetB0, MobileNetV2, which are frequently preferred in the classification of medical images in the literature. Since the amount of data in the five-class classification in diabetic retinopathy disease images was not equal, the data were equalised by using data augmentation techniques using the albumentations library in the training dataset. Accuracy, precision, sensitivity and f1-score values of state of the art models in two-class classification were 0.96. Among the models used in five-class classification, VGG16 was the best model due to its accuracy, precision, sensitivity and f1-score 0.78 and precision 0.79 metric values.

Materials-Methods: The publicly available Aptos2019 dataset was used in the study. There are 3662 fundus images in this dataset. For our first scenario, two-class classification, the images in the dataset are divided into the healthiest (npdr) and the other 4 stages of diabetic retinopathy disease and 72% of the images are reserved for training the models, 8% for validation to prevent overlearning problems during training, and 20% for testing the success of the models. In the second scenario, the images constituting the samples from 5 stages of the disease were divided into 5 different classes with the same ratios. In the third scenario, synthetic data were created from the images in the training data set and the data set was balanced by creating an equal number of images for each level.

In recent years, deep learning methods have been frequently used in object recognition, image classification and segmentation in medical and ophthalmological images and very successful results have been obtained. In particular, deep convolutional neural networks (DNNs) have been used for early detection and identification of retinal diseases such as diabetic retinopathy, age-related macular degeneration and glaucoma from retinal images. [1,2,3,4]

In this study, VGG16, InceptionResNetV2, ResNet152V2, EfficientNetB0, MobileNetV2 architectures were selected from the state of the art models and these models were used by retraining with transfer learning technique. Two- and five-class classification was performed for diabetic retinopathy. Since the amount of data in the five-class classification was not equal, the number of data was equalised by data augmentation using the albumentations library in the training dataset. The performances of the models were evaluated using different metrics such as accuracy, precision, sensitivity and f1-score according to the classification report metrics.

Results: In two-class classification, the accuracy, precision, sensitivity and f1-score values of the state of the art models were 0.96. In five-class classification, accuracy, precision, sensitivity and f1-score values were 0.78, precision 0.79, accuracy, precision, sensitivity and f1-score values were 0.78 in VGG16 model, accuracy, precision, sensitivity and f1-score values were 0.78 in MobileNetV2 model, accuracy and sensitivity were 0.78, precision and f1-score were 0. 77, accuracy, precision, sensitivity 0.74, f1-score 0.73 in InceptionResNetV2 model, accuracy and sensitivity 0.73, precision 0.59, f1-score 0.65 in InceptionResNetV2 model without data augmentation, accuracy, sensitivity and f1-score 0.75, precision 0.77 in EfficientNetB0 model.

Discussion-Conclusion: In the test results, the highest successes were obtained in the two-class classification scenario in which the images in the dataset were considered as the healthiest (npdr) and sickest. The results obtained without data augmentation were low in the five-class classification experiments where the level of the disease was also tried to be predicted. The fact that the data set used in the level classification of the disease was an imbalanced data set caused the models to be biased towards classes with more samples during the learning process. In the experiments where the training datasets were equalised with data augmentation techniques, satisfactory achievements were obtained in the test results. It has been demonstrated that the data augmentation method is beneficial in handling diabetic retinopathy as a multi-class disease.

Keywords: Convolutional Neural Networks, Deep Learning, Diabetic Retinopathy, Transfer Learning.

REFERENCE

[1] V. Gulshan ve diğerleri, “Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs,” Jama, vol. 316, no. 22, pp. 2402-2410, 2016.

[2] C. S. Lee, D. M. Baughman ve A. Y. Lee, “Deep learning is effective for classifying normal versus age-related macular degeneration OCT images,” Ophthalmol. Retin., vol. 1, no. 4, pp. 322-327, 2017.

[3] X. Chen, Y. Xu, D. W. K. Wong, T. Y. Wong ve J. Liu, “Glaucoma detection based on deep convolutional neural network,” in 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), 2015, pp. 715-718.

[4] Uçar M. (2021). Detection of Glaucoma Disease with Convolutional Neural Network Architectures. DEÜFMD 23(68), 521-529.

MOTIVATIONS AND BARRIERS TO USING GENERATIVE AI FOR SELF-DIAGNOSIS: A BEHAVIOURAL REASONING THEORY APPROACH

Ouranıa Areta Hızıroğlu[1*], Ali Emre, Aydın[2]

[1] Department of Management Information Systems, İzmir Bakırçay University, İzmir, Türkiye

[2] Department of Business Administration, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: ourania.areta@bakircay.edu.tr*

Introduction-Aim: The rapid growth of Generative AI technologies, such as large language models, has resulted in their applications having an increasing scope within the healthcare sector, including self-diagnosis [1] [2]. This has led to critical questions regarding patient safety, patient healthcare-seeking behaviour, and the possibility of the transformation of traditional healthcare delivery models [3][4]. While these AI systems continue to grow in capabilities, there is a significant knowledge gap about how people perceive and interact with generative AI for self-diagnosis purposes [5] [6]. The misalignment of technological advancement and user acceptance can create the risk of under-utilization of useful AI tools, or over-reliance on possibly inaccurate self-diagnostic systems [7] [8]. Considering this existing knowledge gap, this paper aims to study the factors that shape the adoption of generative AI for self-diagnosis. With the usage of the Behavioural Reasoning Theory (BRT) [9], we examine the motivations and barriers to using such technologies and offer an in- depth understanding of how users make decisions in this context. Within this framework, the study employs a quantitative approach in measuring key constructs, such as perceived usefulness, interactivity, convenience, safety related risks, and concerns for AI accuracy (AI hallucination). Furthermore, it also accounts for the effects of tradition and familiarity with AI on users' attitudes and behavioural intentions [10][5].

Materials-Methods: The data for this cross-sectional study will be collected through an online survey. In line with the scope of the study, individuals who have previously used generative AI tools will be included in the research. Considering the significant use of these tools by younger people, the study will focus on individuals aged 18-25. Additionally, to assess the role of health education in attitudes towards self-diagnosis in generative AI tools, students enrolled in health-related programs will also be included in the study. This study will employ convenience sampling for data collection. To test the hypotheses developed in line with the research model, the questionnaire will include 31 items representing 9 dimensions. To start with, innovation-related values will be measured using three items [11] [12]. The factors that may positively influence the use of generative AI, such as compatibility, will be measured by three items [13]; interaction will be measured using three items [12], and perceived usefulness will be measured by four items [14]. Factors that may negatively influence the use of generative AI, such as security risk, will be measured using three items [13]; AI hallucination will be measured using four items [14], and tradition will be measured using three items [13]. Attitudes toward use will be measured with four items, and behavioural intention with six items [14]. Each item in the questionnaire will be tailored to the research context. Responses will be recorded on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). Additional items will also be included to capture participants' knowledge and usage levels of AI, as well as their demographic characteristics. For data analysis, descriptive statistics, internal consistency levels, and correlation analyses will be conducted using SPSS software. Additionally, structural equation modelling (SEM) will be employed to test the hypotheses, using the AMOS statistical software package. SEM is a technique that allows for the observation of relationships between multiple dependent and independent variables while controlling for measurement errors, thus providing a holistic perspective [15][16]. Accordingly,

the measurement model will first be tested, and if the measurement model fits the data, the structural model will then be tested. Furthermore, composite reliability and convergent validity will be examined. A p-value of < 0.05 will be considered statistically significant.

Results: This study is expected to reveal specific patterns in the adoption of generative AI for self-diagnosis among young adults. Through structural equation modelling, we anticipate identifying the relative strength of various factors influencing adoption intentions. Specifically, we expect to quantify: the impact of innovation-related values on reasons for and against adoption, the relative importance of convenience, interactivity, and perceived usefulness as motivators for adoption, and the significance of security risks, AI hallucination concerns, and tradition as barriers to adoption. The analysis will also determine how these factors collectively influence attitudes and behavioural intentions toward using generative AI for self-diagnosis. Additionally, the comparison between health students and other participants may reveal how health education influences perceptions and intended use of these technologies. The statistical analysis will provide reliability coefficients, validity measures, and path coefficients that quantify these relationships, offering a comprehensive model of adoption behaviour in this context.

Discussion-Conclusion: The findings of this work aim to provide important insights into the decision-making process individuals undertake when considering the use of generative AI for self-diagnosis. The study's focus on young adults and health students offers valuable understanding of how future healthcare consumers and professionals view these technologies. Moreover, it will support the advancement of the literature on AI adoption in healthcare [17], and deliver actionable recommendations for AI developers, healthcare providers, and policymakers. By understanding the factors driving adoption or resistance, stakeholders can better address user concerns, leverage potential benefits, and develop more user-centric AI solutions for health self-management. In the end, this study aims to provide support for the safe and effective incorporation of generative AI in the healthcare sector while maintaining the benefit of pervasive health information access alongside the need to protect patients' safety, and the continued relevance of human clinical expertise. Future research should explore how these adoption patterns evolve over time and vary across different cultural contexts and age groups. These insights will be crucial for the responsible development and implementation of AI-driven healthcare solutions that balance innovation with patient safety and clinical expertise.

Keywords: Generative AI, Self-Diagnosis, Healthcare, Behavioural Reasoning Theory

REFERENCE

[1] T. B. Brown, B. Mann, N. Ryder, M. Subbiah, J. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, et al., "Language models are few-shot learners," arXiv preprint arXiv:2005.14165, 2020.

[2] T. Hirosawa and T. Manabe, "The potential, limitations, and future of diagnostics enhanced by generative artificial intelligence," Diagnosis, vol. 11, no. 1, pp. 145–148, 2024.

[3] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," Nature Medicine, vol. 25, no. 1, pp. 44–56, 2019.

[4] S. Aboueid, R. H. Liu, B. N. Desta, A. Chaurasia, and S. Ebrahim, "The use of artificially intelligent self-diagnosing digital platforms by the general public: Scoping review," JMIR Medical Informatics, vol. 7, no. 2, p. e13445, 2019.

[5] O. Asan, A. E. Bayrak, and A. Choudhury, "Artificial intelligence and human trust in healthcare: Focus on clinicians," Journal of Medical Internet Research, vol. 22, no. 6, p. e15154, 2020.

[6] Y. Shahsavar and A. Mozaffari, "User intentions to use chatgpt for self-diagnosis and health-related purposes: Cross-sectional survey study," JMIR Human Factors, vol. 10, p. e48317, 2023.

[7] M. Baldauf, P. Fr¨ohlich, and M. Przybylski, “Trust me, i’m a doctor - user perceptions of ai-driven apps for mobile health diagnosis,” in Proceedings of the 19th International Conference on Mobile and Ubiquitous Multimedia, pp. 167–178, 2020.

[8] M. Skjuve, A. Følstad, and P. B. Brandtzaeg, “Why do people use chatgpt? exploring user motivations for generative conversational ai,” First Monday, vol. 29, no. 1, 2024.

[9] J. D. Westaby, “Behavioral reasoning theory: Identifying new linkages underlying intentions and behavior,” Organizational Behavior and Human Decision Processes, vol. 98, no. 2, pp. 97–120,2005.

[10] M. C. Claudy, R. Garcia, and A. O’Driscoll, “Consumer resistance to innovation—a behavioral reasoning perspective,” Journal of the Academy of Marketing Science, vol. 43, no. 4, pp. 528–544, 2015.

[11] B. Sivanthanu, “Adoption of internet of things (iot) based wearables for healthcare of older adults – a behavioural reasoning theory (brt) approach,” Journal of Enabling Technologies, vol. 12, no. 4, pp. 169–185, 2018.

[12] A. Anayat, K. Hasan, and M. H. Bhutta, “Behavioral intention to use ai-enabled content generators: Role of personality traits, performance, and process satisfaction,” Technology in Society, vol. 74, p. 102241, 2023.

[13] A. Gupta and N. Arora, “Understanding determinants and barriers of mobile shopping adoption using behavioral reasoning theory,” Journal of Retailing and Consumer Services, vol. 36, pp. 1–7,2017.

[14] J. Christensen, J. M. Hansen, and P. Wilson, “Understanding the role and impact of Generative Artificial Intelligence (AI) hallucination within consumers’ tourism decision-making processes,” Current Issues in Tourism, pp. 1–16, 2024.

[15] R. P. Bagozzi and Y. Yi, “Specification, evaluation, and interpretation of structural equation models,” Journal of the Academy of Marketing Science, vol. 40, no. 1, pp. 8–34, 2012.

[16] D. Iacobucci, “Structural equations modeling: Fit indices, sample size, and advanced topics,” Journal of Consumer Psychology, vol. 19, no. 1, pp. 90–98, 2009.

[17] S. Gao, L. He, Y. Chen, D. Li, and K. Lai, “Public perception of artificial intelligence in medical care: Content analysis of social media,” Journal of Medical Internet Research, vol. 22, no. 7, p. e16649, 2020

IMPROVING SKIN LESION SEGMENTATION MAPS WITH CLUSTERING ANALYSIS AND IMAGE PROCESSING

Feridun Pözüt[1]*, M. Kemal Güllü[2]

[1] Electrical and Electronics Engineering, İzmir Bakırçay University, İzmir, Türkiye

[1] Data Scientist, Amatis, Bayraklı/İzmir, Türkiye

[2] Department of Circuits and Systems, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: feridun.pozut@bakircay.edu.tr*

Introduction-Aim: The soft boundaries frequently observed in lesion segmentation maps are often due to the structure of the reference segmentation maps used during the training phase. The segmentation of lesions plays a crucial role in clinical applications and dermatological diagnoses, as the accurate identification of clear boundaries is essential for disease diagnosis. However, existing methods often fail to delineate boundary details sufficiently, complicating the accurate calculation of the ABCD (Asymmetry, Border irregularity, Color variation, Diameter) characteristics. The ABCD features are an important tool for assessing the malignancy risk of skin lesions, and their accurate computation allows for proper evaluation of patients. Therefore, recovering boundary details in segmentation maps is of great importance both in research and clinical practice.

Materials-Methods: In this study, a transition from the RGB color space to other color spaces like HSV and YCrCb was made to improve lesion boundaries, utilizing color saturation information. Color saturation expresses the intensity of colors in images and contributes to the enhancement of lesion boundaries. A deep learning-based segmentation model was employed to obtain segmentation maps of the lesions. These maps were improved by comparing them with saturation images in local regions and performing binary clustering with blocks of different sizes. The K-Means algorithm, an effective clustering method commonly used in such image processing applications, was utilized to determine the optimal box size based on Intersection over Union (IoU) scores, resulting in the final masks. IoU is a metric used to measure the accuracy of the model and plays a significant role in evaluating the quality of segmentation results. This method has the potential to correct mislabeling and aims to create more accurate segmentation maps by recovering boundary details. During the study, open-source [3] ISIC 2016 (International Skin Imaging Collaboration) lesion images and their corresponding reference segmentation maps were used.

Results: As a result of the improvements made, the boundary details in the obtained segmentation maps have allowed for a more precise detection of lesion boundaries, positively impacting the calculation of ABCD features in the literature. This facilitates the evaluation of skin lesions in clinical practice and enables more effective identification of cases at risk of malignancy. While existing studies in the literature often rely on reference segmentation maps, this study demonstrates that addressing mislabeling can provide significant enhancements in lesion analysis.

Discussion-Conclusion: The developed method allows for the enhancement of the outputs of a deep learning-based hybrid segmentation model while ensuring that the segmentation maps yield more accurate results through a binary clustering approach that analyzes additional features such as color saturation. Studies like those of Messadi et al. (2021) and Majumder & Ullah (2019) indicate that these methods can play a significant role in improving metric results in mask-based ABCD feature extraction applications. The findings of this study reveal that clarifying the boundaries of segmentation maps not only enhances model performance but also has the potential to correct labeling errors. In conclusion, this study presents an innovative approach for recovering boundary details in lesion segmentation maps and contributes to the existing literature in this field. Such improvements are expected to enable a more accurate analysis of skin lesions, thereby providing significant benefits in clinical applications.

Keywords: Lesion segmentation, K-Means, Color saturation, HSV color space, YCrCb color space

REFERENCE

[1] Messadi, M., Cherifi, H., & Bessaid, A. (2021). Segmentation and ABCD rule extraction for skin tumors classification. arXiv preprint arXiv:2106.04372.

[2] Majumder, S., & Ullah, M. A. (2019). A computational approach to pertinent feature extraction for diagnosis of melanoma skin lesion. Pattern Recognition and Image Analysis, 29, 503-514.

[3] Gutman, David; Codella, Noel C. F.; Celebi, Emre; Helba, Brian; Marchetti, Michael; Mishra, Nabin; Halpern, Allan. "Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)". eprint arXiv:1605.01397. 2016.

CLASSIFICATION OF SKIN LESIONS WITH DEEP HYBRID MODELS AND IMAGE-BASED ANALYSIS

Feridun Pözüt[1]*, M. Kemal Güllü[2]

[1] Electrical and Electronics Engineering, İzmir Bakırçay University, İzmir, Türkiye

[1] Data Scientist, Amatis, Bayraklı/İzmir, Türkiye

[2] Department of Circuits and Systems, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: feridun.pozut@bakircay.edu.tr*

Introduction-Aim: In this study, two popular pre-trained deep learning architectures, [1] VGG16 and [2] ResNet50, were utilized. Separate inputs were provided to the VGG16 and ResNet50 models, and the data corresponding to these inputs were processed. Global Average Pooling was applied to the output layers of both models, and the resulting features were reduced to 128 dimensions. These 128-dimensional features obtained from the two models were concatenated and used for the classification process. The final layer aimed to perform binary classification using the sigmoid activation function to classify outputs as 0 and 1. This approach allows for the classification of skin lesions as malignant and benign.

During the model training, a labeled skin lesion dataset was used, and commonly used optimization algorithms were preferred for model optimization. The model achieved a classification accuracy of 89% for distinguishing between benign and malignant cases.

Materials-Methods: In this study, two popular pre-trained deep learning architectures, [1] VGG16 and [2] ResNet50, were utilized. Separate inputs were provided to the VGG16 and ResNet50 models, and the data corresponding to these inputs were processed. Global Average Pooling was applied to the output layers of both models, and the resulting features were reduced to 128 dimensions. These 128-dimensional features obtained from the two models were concatenated and used for the classification process. The final layer aimed to perform binary classification using the sigmoid activation function to classify outputs as 0 and 1. This approach allows for the classification of skin lesions as malignant and benign.

During the model training, a labeled skin lesion dataset was used, and commonly used optimization algorithms were preferred for model optimization. Throughout the training, the data was validated against a validation set, and various metrics (accuracy, precision, recall, F1-score) were employed to monitor the model's performance. Additionally, the [3] ISIC (International Skin Imaging Collaboration) open-source datasets from 2018 and 2016 were utilized to enhance the model's performance.

Results: During the training and validation phases, a continuous improvement in the model's performance metrics was observed. As a result, it was determined that the developed model achieved a classification accuracy of 89% on the validation set. According to the confusion matrix analysis, the misclassification of benign and malignant classes was evenly distributed. This indicates that the model was able to classify both classes with similar precision.

Additionally, the feature maps produced by the model were used to calculate the Euclidean and Cosine distances for each lesion. The feature map was designed to have 256 dimensions, and the five closest lesions to each lesion were identified based on this map. Through the analysis of the Euclidean and Cosine distances among these lesions, the lesions that were closest in class to each lesion were determined. Thus, the model not only contributed to classification performance but also aided in understanding the similarities and differences among skin lesions.

Discussion-Conclusion: The developed model successfully classified skin lesions with high accuracy rates by combining the VGG16 and ResNet50 architectures. An accuracy rate of 89% is competitive with many commonly used methods in this field. The model's ability to classify both benign and malignant lesions in a balanced manner can provide health professionals with an important decision support system for accurate diagnosis.

Moreover, analyses based on Euclidean and Cosine distances derived from the feature maps produced by the model enabled the identification of visual and structural similarities among skin lesions. This can assist dermatologists in making more accurate diagnoses by comparing similar skin lesions.

In conclusion, this study demonstrated the applicability of deep learning-based models in the classification of skin lesions and offered potential areas for improvement for future research. Further training of the model with larger and more diverse datasets could enhance classification performance even more. Additionally, integrating other feature extraction methods could improve the accuracy of lesion classification.

Keywords: Skin lesions, deep learning, classification, VGG16, ResNet50, Euclidean distance, Cosine distance.

REFERENCE

[1] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770-778. https://doi.org/10.1109/CVPR.2016.90

[2] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations (ICLR). https://arxiv.org/abs/1409.1556

[3] Gutman, David; Codella, Noel C. F.; Celebi, Emre; Helba, Brian; Marchetti, Michael; Mishra, Nabin; Halpern, Allan. "Skin Lesion Analysis toward Melanoma Detection: A Challenge at the International Symposium on Biomedical Imaging (ISBI) 2016, hosted by the International Skin Imaging Collaboration (ISIC)". eprint arXiv:1605.01397. 2016.

MACHINE LEARNING MODELS USED IN THE PREDICTION OF CHILDHOOD VACCINATION RATES: LOOKING TO THE FUTURE IN NURSING WITH A SYSTEMATIC REVIEW

Gözde Özsezer[1*], Gülengül Mermer[2], Başak Süslü[3]

[1] *Çanakkale Onsekiz Mart University Faculty of Health Sciences, Department of Public Health Nursing, Çanakkale, Türkiye*

[2] *Ege University Faculty of Nursing, Department of Public Health Nursing, İzmir, Türkiye*

[3] *Çanakkale Onsekiz Mart University Faculty of Health Sciences, Department of Nursing, Çanakkale, Türkiye*

* *Corresponding author: basakssl712@gmail.com*

Introduction-Aim: Childhood immunizations have a vital role in protecting public health and preventing the spread of infectious diseases. Childhood immunization protects children against diseases such as measles, diphtheria, whooping cough, and polio, which can lead to fatal and serious health problems, and enables individuals to step into a healthy life. Vaccines not only provide individual protection but also contribute to the control of infectious diseases in the community by supporting herd immunity. Societies with high vaccination rates prevent the spread of these diseases and reduce the risk of epidemics.

This study aims to conduct a systematic review to comprehensively examine ML models used in the prediction of childhood vaccination rates. The main objective of the study is to reveal the contexts in which different ML algorithms are most effective in predicting childhood immunization rates, to analyze the accuracy of these models in detail, and to contribute to the existing body of knowledge in this field. This analysis aims to determine which models are more appropriate and effective under certain populations, geographical regions, or socioeconomic conditions by comparing the success of various models.

Materials-Methods: This systematic review was conducted in accordance with the 'Preferred Reporting Items for Systematic Reviews (PRISMA)'. The inclusion and exclusion criteria of the studies included in this study were determined according to the PICOS method. Inclusion criteria for this study: (1) Population: Children, (2) Intervention: Studies involving vaccine intervention, (3) Comparison: Studies using machine learning methods, (4) Outcomes: Prediction by machine learning methods, (5) Study design: Studies published in English between 2014 and 2024 that included machine learning methods of original artificial intelligence were included. In this systematic review, articles published in English between 2014 and 2024 were included. In the study, a literature search was performed in the 'Web of Science, Google Scholar, Pubmed, and Scopus' databases between 10-20 October 2024 using different combinations of the keywords 'child', 'children', 'vaccine', 'vaccination', 'immunization', 'machine learning', 'rate' and 'prediction'. It was aimed to reach all studies related to the subject in the search of databases. Reference lists of included studies and previous systematic reviews were checked for additional searches. Three investigators independently performed the selection of studies.

Initially, duplicate studies were excluded, and studies were selected if they fulfilled the search criteria when screened by title, abstract, and full text, respectively. The titles and abstracts of all relevant publications retrieved by electronic search were independently reviewed by the researchers. As a result of the search, 19165 studies (Google Scholar: 18600, Pubmed: 549, Scopus: 2, WOS: 14) were reached. The studies were firstly analyzed according to their titles, and 16625 studies that were not related to the research topic were excluded. Abstracts and full texts of the remaining 2540 studies were screened for inclusion and exclusion criteria. A total of 226 studies including reviews, letters to the editor, meta-analyses, and conference proceedings were excluded. A total of 416 duplicated articles were identified and removed using the Mendeley Reference Manager program. A total of 6 studies were found to meet the criteria for systematic review. The methodological quality of the articles included in this systematic review was evaluated by both researchers. Articles using artificial intelligence techniques in the research were accepted as diagnostic test accuracy studies. The Joanna Briggs Institute (JBI) Critical Appraisal Checklist for Diagnostic Test Accuracy Studies was used to assess the quality of the included studies. A score of 0% to 50% was considered low quality, 50% to 70% was considered medium quality, and any text article with a score of 70% and above was considered high quality. RoBvis 2 tool was used for risk of bias in the study. The decision was expressed as 'Low' or 'High' or 'Some concerns'. The study did not require ethical approval as the research articles included in the sample were obtained from openly accessible electronic databases and search engines. All stages of the study were conducted in accordance with the principles of the Declaration of Helsinki.

Results: Of the 19165 studies initially identified, 6 were included. Six articles were critically appraised. The methodological quality of the articles was high, and all scored 70% or more. The systematic review included five studies using RF algorithm [1-5], four studies using SVM algorithm [1,4-6], DT algorithm in four studies [2,4-6], NB algorithm in four studies [3-6], LR algorithm in three studies [3-5], MLP algorithm in two studies [3,6], XGB algorithm in two studies [2,4], recursive partitioning and C-forest algorithm in one study [1], GNB, BNB and Lightgbm algorithm in one study [2], PART, J48, LogitBoost and AdaBoost algorithm in one study [3], KNN and ANN algorithm in one study [4], LASSO regression in one study [5].

Chandir et al. (2018) reported that the RF model provided 94.9% sensitivity and 54.9% specificity, while the recursive partitioning algorithm achieved the highest AUC value (0.791, 95% CI 0.784-0.798) [1]. Hasan et al. (2021) reported that the optimized LightGBM model performed best with 84.60% sensitivity and 80.0% AUC, and the performance improved with the combination of XGBoost and LightGBM [2]. Demshah et al. (2023) found that the PART algorithm gave the best results with 95.53% accuracy, followed by J48, MLP, and random forest models [3]. Tadese et al. (2024) stated that the XGBoost model stands out with 79.01% accuracy, 89.88% recall, 81.10% F1 score, 73.89% sensitivity, and 86% AUC [4]. Wang et al. (2024) emphasized that the RF model performed best on the training set, while logistic regression and Naive Bayes models stood out on the validation set [5]. Qazi et al.

(2021) reported that the MLP model correctly predicted the probability of children defaulting in the immunization series with 98.5% accuracy and 0.994 AUC [6].

Discussion-Conclusion: These results show that various ML algorithms are effective in predicting childhood vaccination. However, the performance of each model varies depending on the dataset and methodology. While some algorithms, such as recursive partitioning and LightGBM, stand out, combinations and ensembles play a critical role in increasing prediction accuracy. This systematic review examines the effectiveness of ML methods in childhood vaccination prediction and provides important findings for nursing practice. The results show that algorithms such as RF, SVM, and XGBoost are particularly effective in predicting vaccination with high accuracy and sensitivity. In addition, it was determined that ensemble and combination approaches have the potential to increase vaccination prediction performance.

Keywords: child, machine learning, nursing, rate, vaccination

REFERENCE

[1] S. Chandir *et al.*, “Using predictive analytics to identify children at high risk of defaulting from a routine immunization program: Feasibility study,” *JMIR Public Health and Surveillance*, vol. 4, no. 3, p. e9681, 2018, doi: 10.2196/publichealth.9681.

[2] M. K. Hasan *et al.*, “Associating measles vaccine uptake classification and its underlying factors using an ensemble of machine learning models,” *IEEE Access*, vol. 9, pp. 119613-119628, 2021, doi: 10.1109/ACCESS.2021.3108551.

[3] A. W. Demsash, A. A. Chereka, A. D. Walle, S. Y. Kassie, F. Bekele, and T. Bekana, “Machine learning algorithms’ application to predict childhood vaccination among children aged 12–23 months in Ethiopia: Evidence 2016 Ethiopian Demographic and Health Survey dataset,” *PLOS ONE*, vol. 18, no. 10, p. e0288867, 2023, doi: 10.1371/journal.pone.0288867.

[4] Z. B. Tadese, A. M. Nigatu, T. Z. Yehuala, and Y. Sebastian, “Prediction of incomplete immunization among under-five children in East Africa from recent demographic and health surveys: A machine learning approach,” *Scientific Reports*, vol. 14, no. 1, p. 11529, 2024, doi: 10.1038/s41598-024-62641-8.

[5] Q. Wang *et al.*, “A prediction model for identifying seasonal influenza vaccination uptake among children in Wuxi, China: Prospective observational study,” *JMIR Public Health and Surveillance*, vol. 10, no. 1, p. e56064, 2024, doi: 10.2196/56064.

[6] S. Qazi, M. Usman, and A. Mahmood, “A data-driven framework for introducing predictive analytics into expanded program on immunization in Pakistan,” *Wiener Klinische Wochenschrift*, vol. 133, pp. 695-702, 2021, doi: 10.1007/s00508-020-01737-3.

INNOVATIVE REMOTE NECK PAIN RELIEF USING MOBITHERA: LEVERAGING AI-BASED FACE MESH ON MOBILE DEVICES

Eminullah Yaşar[1*], Hamdi Yalın Yalıç[1], Baran Datlar[1], Alaettin Uçan[1], Ali Yaşar Yiğit[1], Adem Ali Yılmaz[1]

[1] Tiga Information Technologies, Health Information Science and Systems, Ankara, Türkiye

** Corresponding author: eminullah.yasar@tigahealth.com*

Introduction-Aim: In today's technological world, digital health solutions play a critical role in improving individuals' quality of life. Neck pain, in particular, has become one of the most common health issues in daily life, requiring innovative and personalized solutions for effective management [1].

If neck discomfort is not addressed in a timely manner, it can evolve into chronic conditions over time, significantly impacting individuals' quality of life without them realizing it. The increasing use of mobile devices has contributed to the rise of neck-related problems. As it has become nearly impossible to separate people from their smartphones, encouraging them to spend a small portion of their screen time on their health could be highly beneficial. Therefore, a mobile, AI-powered exercise application, MobiThera, has been designed to help users reduce neck pain.

The aim of this paper is to introduce MobiThera, an AI-assisted exercise application for mobile devices, leveraging Google FaceMesh technology to offer remote, personalized neck pain relief [2]. The paper will explore the application's potential in promoting user health by providing neck exercises through a familiar platform-mobile devices-that users already engage with frequently.

Materials-Methods: MediaPipe Face Mesh model was used to detect and track users' neck movements. MediaPipe is an open-source machine learning library developed by Google that provides a variety of pre-trained models for real-time performance on mobile and web platforms.

The Face Mesh model in MediaPipe detects 468 facial points, each represented as 2D (x, y) or 3D (x, y, z) coordinates. These points correspond to key facial features such as eyes, lips, eyebrows, and facial features. The following specific points were selected to detect neck movements: 1, 33, 263, 61, 291, and 199. These points are evenly distributed across the face, allowing for accurate motion analysis of the head and neck regions.

For real-time camera feeds on mobile devices, OpenCV for Unity was integrated to process rotation and translation vectors that facilitate head position estimation [3].

The following steps were used to estimate neck movements (yaw, pitch, and roll) using MobiThera:

1. Face Landmark Detection: The landmarks of the face were detected using the MediaPipe Face Mesh model. This model provided a set of 468 key points representing facial features, and the landmarks (indexes 1, 33, 263, 61, 291, and 199) were used for head motion detection.

2. Rotation and Translation Vector Calculation: The rotation vector and translation vector were calculated using the solvePnP function of OpenCVForUnity, which estimates the

pose of the object in 3D space [4]. These vectors represent the motion and position of the face relative to the camera.

3. Rotation Matrix Calculation: The rotation and translation vectors were passed to the OpenCVForUnity Rodrigues function to create a rotation matrix. This matrix represents the orientation of the face in 3D space and is important for calculating the angular movements of the neck [5].
4. Euler Angle Decomposition: The rotation matrix was decomposed to obtain Euler angles representing the yaw (horizontal rotation), pitch (vertical tilt), and roll (tilt along the axis of the face) angles of the face [6]. These angles were used to track the user's neck movements.
5. Normalization and Thresholding: After calculating the x, y, and z rotation values, the matrix was normalized to ensure consistent scaling [7]. Thresholds were then determined for each value to identify neck movements. The z coordinate of the first landmark (index 1) was taken into account to estimate the user's distance from the camera, which played a role in setting the thresholds.
6. Exercise Development: Based on the detected movements, specific exercises were created by determining the appropriate repetitions, sets, and angle variations for each neck movement. These exercises are designed to relieve neck pain by encouraging appropriate head rotation, tilt, and posture correction.

By integrating these techniques, MobiThera provided users with real-time feedback on their neck movements, ensuring that exercises were performed correctly and effectively.

Results: The MobiThera application successfully utilized Google's MediaPipe Face Mesh model to detect and analyse users' neck movements in real-time through mobile or tablet cameras. The system accurately identified six key facial landmarks (indices 1, 33, 263, 61, 291, and 199) and processed them to estimate head pose angles-yaw, pitch, and roll-enabling precise detection of neck rotation, tilt, and posture changes.

Neck Movement Detection: By employing the OpenCV for Unity library's solvePnP and Rodrigues functions, MobiThera generated rotation and translation matrices from the facial landmarks. These matrices were then decomposed into Euler angles, providing real-time measurements of neck movements. The normalized x, y, and z rotation values, along with pre-defined thresholds, allowed the system to accurately categorize neck movements into specific rotational patterns, such as left-right turns (yaw), up-down tilts (pitch), and side-to-side tilts (roll).

Accuracy of Detection: Initial tests of the application with different users demonstrated that the system maintained high accuracy in detecting suitable neck movements. The z-coordinate of the first landmark (index 1) effectively adjusted for variations in the user's distance from the camera, ensuring consistent performance. The distance-based adjustments allowed MobiThera to maintain reliable detection even when the user moved closer to or farther from the camera, improving the flexibility and usability of the application. The exercises were focused on improving neck mobility, strength, and pain relief, with real-time feedback provided to ensure proper movement execution.

Discussion-Conclusion: The development of MobiThera has demonstrated the feasibility of using facial landmark detection for real-time neck movement tracking on mobile devices. By leveraging the Google MediaPipe Face Mesh model, we were able to accurately capture and analyse neck movements without needing to account for individual physical limitations or head sizes. This universality is one of the key strengths of the approach, as it ensures that the same method can be applied to a wide range of users, regardless of their physical characteristics.

The detection of neck movements (yaw, pitch, and roll) was achieved with high accuracy. The adjustment of exercise parameters—such as movement duration, repetitions, sequencing, and complexity—ensured that exercises were appropriate for improving mobility and reducing neck pain. The mobile platform made these exercises easily accessible, offering users the convenience of performing neck exercises anywhere and at any time.

In conclusion, MobiThera successfully implemented an AI-powered solution for detecting and analysing neck movements on mobile devices using facial landmarks. The system proved capable of capturing neck motions with high accuracy, providing users with exercise routines aimed at improving neck mobility and pain. The ease of access and use on mobile devices encourages individuals to incorporate these exercises into their daily routine, promoting better neck health through technology.

Despite the overall success, there remains room for improvement, particularly in enhancing the system's robustness under challenging environmental conditions. Further research and testing will be focused on refining the model's sensitivity to lighting and partial occlusions, ensuring the application can function accurately in a wider range of real-world scenarios. By addressing these limitations, MobiThera can offer even more reliable and effective neck pain management, making AI-driven healthcare solutions increasingly accessible and practical for everyday use [8].

Keywords: Neck Pain Relief, Neck Mobility, Neck Strengthening, FaceMesh

REFERENCE

[1] Cohen, S. P., & Hooten, W. M. (2017). Advances in the diagnosis and management of neck pain. Bmj, 358.

[2] Google Research. (2022, Nov.). MediaPipe: Real-time face, hands, and body tracking. https://mediapipe.dev (accessed: Oct 25, 2024)

[3] S. Saha, "Head Pose Estimation using OpenCV and Dlib," *LearnOpenCV.* https://learnopencv.com/head-pose-estimation-using-opencv-and-dlib/ (accessed: Oct 25, 2024)

[4] Zingoni, A., Diani, M., & Corsini, G. (2019). Tutorial: Dealing with rotation matrices and translation vectors in image-based applications: A tutorial. IEEE Aerospace and Electronic Systems Magazine, 34(2), 38-53.

[5] Evans, P. R. (2001). Rotations and rotation matrices. Acta Crystallographica Section D: Biological Crystallography, 57(10), 1355-1359.

[6] J. Crassidis, "The Euler Angle Parameterization," Berkeley University. https://rotations.berkeley.edu/the-euler-angle-parameterization/ (accessed Oct 25, 2024).

[7] Ali, P. J. M., Faraj, R. H., Koya, E., Ali, P. J. M., & Faraj, R. H. (2014). Data normalization and standardization: a technical report. Mach Learn Tech Rep, 1(1), 1-6.

[8] Kasula, B. Y. (2024). Advancements in AI-driven Healthcare: A Comprehensive Review of Diagnostics, Treatment, and Patient Care Integration. International Journal of Machine Learning for Sustainable Development, 6(1), 1-5.

UTILIZATION OF THE WELCH PEDIOGRAM METHOD IN THE PRELIMINARY DIAGNOSIS OF SLEEP APNEA

Orhan ER[1], Kemal GÜLLÜ[2], Bülent ÇİFTÇİ[3], Feridun PÖZÜT[4], Yalım ERDEM[4]

[1] Department of Computer Engineering İzmir Bakırçay University, İzmir, Türkiye

[2] Department of Electrical and Electronics Engineering İzmir Bakırçay University, İzmir, Türkiye

[3] Department of Chest Diseases, Faculty of Medicine, Yuksek Ihtisas University, Ankara, Türkiye

[4] Amatis Information Technology Research Center, İzmir, Türkiye

** Corresponding author: y.erdem@amatis.nl*

Introduction-Aim: The aim of this study is to establish a preliminary diagnosis of suspected sleep apnea in patients by analyzing 24-hour ECG data obtained from a Holter device through a Welch periodogram using an expert system. To achieve this objective, Heart Rate Variability (HRV) data are derived from patients' RR interval data to determine a threshold value indicative of sleep apnea suspicion.

Materials-Methods: The Welch method is a spectral analysis technique used to estimate the power spectral density (PSD) of a signal. The signal is divided into specific time segments (data in high and low frequencies), these segments overlap, and each segment undergoes a windowing process. The resulting spectra are averaged to obtain the power spectrum. This method enables a clearer view of high-frequency components and potential features related to sleep. In this study, which has no similar precedent in the literature, data from 600 patients were used to provide a preliminary diagnosis of sleep apnea based on information within the High Frequency (HF) range of the Welch periodogram, derived from signals recorded during patients' labeled sleep periods.

Results: The power spectral density (PSD) of the signal obtained from the RR intervals derived from the ECG signal was calculated using the Welch periodogram in the frequency domain for 600 patients. It was found that there is a linear relationship between the increase in the intensity values in the high-frequency region and the occurrence of sleep apnea. It was observed that the signal is affected by respiratory patterns and particularly by abnormalities during sleep. The graph illustrates how the power is distributed across these frequency bands, revealing sleep-related conditions such as sleep apnea.

Discussion-Conclusion: This study has demonstrated that high-frequency components are significant in the preliminary diagnosis of sleep apnea using Welch periodogram applied to 24-hour ECG data. It was found that an increase in power density within the high-frequency range shows a linear relationship with the occurrence of sleep apnea. The results suggest that this method could be an effective tool for detecting respiratory abnormalities such as sleep apnea.

Keywords: expert systems, welch pediogram, sleep apnea, heart rate variability.

KIDNEY REGION PREDICTION USING THE IMAGE-TO-IMAGE REGRESSION-BASED NETWORK

Coşku Öksüz[1*], Artun Narter[1], M. Kemal Güllü[1], Bünyamin Ece[2], Mustafa Koyun[2], İsmail Taşkent[2]

[1] *Electrical-Electronics Engineering, University of Bakırçay, İzmir, Türkiye*

[2] *Department of Radiology, Kastamonu Research and Training Hospital, Kastamonu, Türkiye*

* *Corresponding author: cosku.oksuz@bakircay.edu.tr*

Introduction-Aim: Kidney stones (nephrolithiasis) are rigid mineral structures that form due to the accumulation of crystallized substances in the urinary tract and can pass into the ureter, bladder, or urethra over time [1], [2], [3]. Early diagnosis is of vital importance in terms of preserving renal function, reducing the risk of stone recurrence, and reducing costs by avoiding possible future invasive procedures. Artificial intelligence in kidney stone detection can increase accuracy by enabling physicians to make quick and effective decisions [4], [5]. This initial work aims to develop a deep learning-based image regression model, the first stage of our following framework for isolating kidneys from remaining image regions for more effective stone categorization.

Materials-Methods: The pre-trained DeepLabV3+ [6] model, previously developed for the image segmentation task, has been reconfigured to perform image-to-image regression in the study. Since many high-level features have already been learned, the pre-trained ImageNet models are considered for the encoder network of our proposed regression model. In this context, two separate network configurations are obtained where MobileNetv2 [7] and EfficientNetB0 [8] are utilized, respectively, to evaluate the regression performance of our framework. The weights of the encoder portion of the framework were frozen entirely, and weight updates were only made for the remaining layers.

Results: The abdominal CT kidney image data set, including 753 images from 88 kidney stone patients and 1571 images from 36 healthy cases, was acquired from the Kastamonu Research and Training Hospital. The data set was divided into training, validation, and test sets using 80:10:10 proportions. The two network configurations were trained on the same training set for 150 epochs at a relatively small learning rate, i.e., 1e-4. Then, commonly used regression metrics such as MAE, MSE, and RMSE measured the pixel estimation performance. In addition, the segmentation metric, i.e., IoU, was used to evaluate the overlap between the ground truth masks and the thresholded output of the model. After 150 epochs, it was observed that both network configurations had ensured good convergence in which the validation loss followed the training loss with a small gap. The scores attained in the test set were MSE: 0.0004, MAE: 0.0018, RMSE: 0.0187, and IoU: 0.3452 for MobileNetv2-based network configuration. For the EfficientNetB0-based configuration, the scores achieved on the same test set were MSE: 0.0003, MAE: 0.0018, RMSE: 0.0181, IoU: 0.3529.

Discussion-Conclusion: Image-to-image regression is one of the challenging tasks, especially when the kidney region is considered, as it covers a relatively small percentage of the medical image. Experimental results achieved by two separate network configurations are promising yet need further investigation. In future work on our project, we are planning for the encoder portion of the regression network to be replaced by pre-trained other architectures and re-training each network configuration by the expanded data set.

Keywords: Kidney, regression, deep learning, feature extraction.

Acknowledgement: The outputs in this study were produced within the project scope supported by TUBITAK ARDEB with the number 123E442.

REFERENCE

[1] S. R. Khan *et al.*, 'Kidney stones', *Nat. Rev. Dis. Primer*, vol. 2, no. 1, pp. 1–23, Feb. 2016, doi: 10.1038/nrdp.2016.8.

[2] A. T. Devi, R. Nagaraj, A. Prasad, D. B. Lakkappa, F. Zameer, and N. P. M. Nagalingaswamy, 'Nephrolithiasis: Insights into Biomimics, Pathogenesis, and Pharmacology', *Clin. Complement. Med. Pharmacol.*, vol. 3, no. 2, p. 100077, Jun. 2023, doi: 10.1016/j.ccmp.2022.100077.

[3] Y. Liu, M. Li, L. Qiang, X. Sun, S. Liu, and T. J. Lu, 'Critical size of kidney stone through ureter: A mechanical analysis', *J. Mech. Behav. Biomed. Mater.*, vol. 135, p. 105432, Nov. 2022, doi: 10.1016/j.jmbbm.2022.105432.

[4] S. Isha and S. Z. Shah, 'Use of Artificial Intelligence for Analyzing Kidney Stone Composition: Are We There Yet?', *Mayo Clin. Proc. Digit. Health*, vol. 1, no. 3, pp. 352–356, Sep. 2023, doi: 10.1016/j.mcpdig.2023.06.007.

[5] G. Cil and K. Dogan, 'The efficacy of artificial intelligence in urology: a detailed analysis of kidney stone-related queries', *World J. Urol.*, vol. 42, no. 1, p. 158, Mar. 2024, doi: 10.1007/s00345-024-04847-z.

[6] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, 'Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation', *ArXiv180202611 Cs*, Aug. 2018, Accessed: Jan. 25, 2021. [Online]. Available: http://arxiv.org/abs/1802.02611

[7] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, 'MobileNetV2: Inverted Residuals and Linear Bottlenecks', *ArXiv180104381 Cs*, Mar. 2019, Accessed: Nov. 18, 2020. [Online]. Available: http://arxiv.org/abs/1801.04381

[8] M. Tan and Q. V. Le, 'EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks', *ArXiv190511946 Cs Stat*, Sep. 2020, Accessed: Nov. 17, 2020. [Online]. Available: http://arxiv.org/abs/1905.11946

THE BEST OF BOTH WORLDS: NAMING PERFORMANCES OF NEUROTYPICAL INDIVIDUALS THROUGH AI-GENERATED IMAGES

Şevket Özdemir[1*], Eda İyigün-Uzunöz[1], Şükriye Kayhan-Aktürk[1], Aylin Müge Tunçer[1]

[1] Department of Speech and Language Therapy, Muğla Sıtkı Koçman University, Muğla, Türkiye

* Corresponding author: sevketozdemir@mu.edu.tr

Introduction-Aim: The ever-growing technology of Artificial Intelligence (AI) holds important implications for the field of Speech and Language Therapy (SLT), similar to many fields of study in Health Sciences. One of these implications is the use of images as a therapeutic tool both in child and adult populations, which has particular repercussions on the management of acquired language disorders. A recent narrative review by Özdemir et al. [1] highlight the potential use of AI-generated image tools by embodying image generation process and previous image research, suggesting that image generation tools need to be promoted among clinicians to foster the use of AI-generated images in practice. Based on this motivation, the current study aims to examine the naming performances of neurotypical individuals (including naming accuracy and latency) through images generated by AI, specifically Bing Image Creator (BIC) that uses the DALL-E 3 system according to a number of variables including demographic (age and education), word classes (noun and action), and psycholinguistic parameters (imageability, familiarity, and subjective age of acquisition (AoA)).

Materials-Methods: Ethical approval was granted by Muğla Sıtkı Koçman University (MSKU; Protocol no. 240101). The study employed quantitative methodology (descriptive, comparative, and correlational). Prior to conducting the study, a number of procedures were conducted. (a) First, the images of 81 nouns and 52 actions were formed as a result of entering the relevant prompts for these stimuli in BIC. The prompts were entered into BIC. For each prompt, BIC produced four images respectively, which in turn were reviewed by four specialists; three of whom were from the research team and the remaining one was a speech and language therapist working at a special education and rehabilitation center. The team agreed on the appropriate depiction of 71 nouns and 48 actions. (b) 50 third or fourth-year undergraduate students of the Department of SLT at MSKU rated these nouns and actions according to the parameters of naming agreement, image agreement, and visual complexity, which was in accordance with the methodology of Reymond et al. [2]. Following the rating procedures, the list of stimuli was updated, including 67 nouns and 46 actions for final administration.

The final list was administered to 152 neurotypical individuals. Purposive sampling was used in the recruitment. The inclusion criteria for the participants were (a) being between the ages of 18 to 60, (b) speaking Turkish as the mother tongue, (c) normal range of vision as well as no history of neuropsychological disorders that might affect their participation, as reported by the participants. Those above 60 were not included due to the reasons that a cognitive screening test was required to eliminate any signs of neurological disorders and the available cognitive screening tests in Turkish required paper-based administration (i.e., the Cognitive Screen section of the Turkish version of Comprehensive Aphasia Test, [3]). The invitation for the study was disseminated to the faculty and administrative staff members of the MSKU. The participants did not receive any compensation for their participation. The data were collected in September 2024.

The administration was conducted via Qualtrics™ with appropriate necessary security arrangements to prevent multiple submissions. During the administration, (a) the participants provided information related to their demographics which was considered for subsequent data analyses; (b) the written responses to the images were recorded for assessing naming accuracy; (c) the timing question feature of Qualtrics™ was used to evaluate naming latency in seconds [4]. The order of nouns and actions was randomized for each participant as the randomizer feature of Qualtrics™ was activated [5].

IBM SPSS v26 was used in data analysis. A number of procedures were carried out: First, the correct and incorrect responses given to each noun and action were scored as 1 and 0, for each participant respectively. The scores were summed for nouns and actions separately, which referred to naming accuracy. Also, these scores were transformed into ratios through the following formula, to ensure comparison due to differing number of items in assessing naming abilities of nouns and actions: Naming accuracy score / total number of items (67 for nouns and 46 for actions, respectively). Furthermore, the frequency of naming accuracy for each item was computed (i.e., the number of participants correctly naming each item). The mean value of naming latency (in seconds) for each item was also identified. The psycholinguistic values of the items including imageability, familiarity and subjective AoA values were extracted from the study of Selvi-Balo et al. [6].

The data did not show normal distribution, therefore non-parametric tests were used in data analysis. The scores including accuracy and latency were examined according to two age (18-44 and 45-60) and three education groups (those receiving associate degree and below; those with undergraduate degrees; and finally, those with graduate degrees including master's and doctoral education). The noun and action naming performances were compared in terms of accuracy and latency, as well. The correlation between naming accuracy and naming latency, imageability, familiarity, and subjective AoA values of the items was investigated.

Results: The results were derived from 152 neurotypical individuals: (a) Regarding age groups, the participants in the younger group (n= 93) scored higher in noun naming than the other group (n= 59) (p= 0.046). The same was observed in action naming, which was in favor of the younger group (p= 0.011). As regards latency, younger adults named nouns and actions faster than the older ones ($p<0.001$ for nouns and actions). (b) The noun naming scores of those in the graduate group scored significantly higher than those in the remaining two groups (those that received undergraduate education and below; p= 0.006 and 0.041, respectively). The action naming scores of those in the graduate group scored significantly higher than those with associate degrees and below (p= 0.004). Regarding latency, those receiving the lowest education spent more time in naming nouns and actions compared to the remaining education groups ($p<0.01$). (c) The mean ratio of noun naming was significantly higher than action naming, which showed that participants were more successful in naming nouns compared to actions ($p<0.001$). Also, the nouns were named faster than the actions, according to the mean values of latency in seconds ($p<0.001$). (d) The naming accuracy showed a negative and statistically significant correlation with naming latency (r_s= -0.688; $p<0.001$) and subjective AoA (r_s= -0.246; p= 0.009), as well as positive and statistically significant correlation with imageability (r_s= 0.520; $p<0.001$) and familiarity (r_s= 0.358; $p<0.001$).

Discussion-Conclusion: This study follows that of Pierce [7] who investigated the feasibility of the AI-generated images produced by DALL-E 2 system. Pierce [7] reports that the images of items that referred to nouns were generated with highest efficiency and accuracy compared to verbs and sentences. Current research moves beyond the scope of this study in that the AI-generated images were used in actual practice and administered to neurotypical individuals to assess their naming skills, which was apparently successful. The findings of this study show that AI can be promising in the field of SLT, especially in the preparation of individualized therapy materials.

Keywords: Image generation, naming abilities, psycholinguistic variables, Speech and Language Therapy, artificial intelligence.

REFERENCE

[1] Ş. Özdemir, Ş. Kayhan-Aktürk, E. İyigün-Uzunöz, and A. M. Tunçer, "Beyond line drawings and photographs: Exploring the potential use of AI-assisted images in acquired language disorders through the narrative review of image research" 2024, *OSF*.

[2] C. Reymond, S. Widmer Beierlein, C. Müller, R. Reutimann, K. P. Kuntner, N. Falcon Garcia, et al., "Naming images in aphasia: effects of graphic representations and photographs on naming performance in persons with and without aphasia", *Aphasiology*, vol. 37, pp. 993-1015, 2023.

[3] İ. Maviş, A. M. Tunçer, S. Selvi-Balo, S. D. Tokaç, and Ş. Özdemir, "The adaptation process of the Comprehensive Aphasia Test into CAT-Turkish: psycholinguistic and clinical considerations", *Aphasiology*, vol. 36, pp. 493-512, 2022.

[4] Qualtrics. (2024). *Timing question* [Online]. Available: https://www.qualtrics.com/support/survey-platform/survey-module/editing-questions/question-types-guide/advanced/timing/

[5] Qualtrics. (2024). *Randomizer* [Online]. Available: https://www.qualtrics.com/support/survey-platform/survey-module/survey-flow/standard-elements/randomizer/

[6] S. Selvi-Balo, İ. Maviş, and A. M. Tunçer, "586 Türkçe sözcügün imgelenebilirlik, tanıdıklık ve öznel edinim yaşı norm değerleri", *Dil, Konuşma ve Yutma Araştırmaları Dergisi*, vol. 3, pp. 301-334, 2020.

[7] J. E. Pierce, "AI-Generated images for Speech Pathology—An exploratory application to aphasia assessment and intervention materials", *American Journal of Speech-Language Pathology*, vol. 33, pp.443-451, 2024.

THE IMPORTANCE OF TECHNOLOGY DEVELOPMENT CENTRES IN DIGITAL TRANSFORMATION IN THE FIELD OF HEALTH

Gözde Tuysuz[1], Melisa Gül Aydemir[1]

[1] University of Health Sciences Technology Development Centre, Ankara, Türkiye

**Corresponding Author: gozde.tuysuz@sbu.edu.tr*

Introduction-Aim: Digital transformation is a process that aims to make healthcare services more efficient, accessible and sustainable by developing innovative solutions in the healthcare sector. With the development of the Internet, a digital transformation is taking place in the health sector as in many sectors. In this process, artificial intelligence applications have started to be used in solving problems in health services and management, diagnosis and treatment of diseases, prediction of diseases and many other issues. One of the most important application areas of information technologies in health is artificial intelligence applications (Akalın & Veranyurt, 2023, p. 3). Technology Development Centres (TEKMERs) in Türkiye play an important role in the development of artificial intelligence (AI) and digital health solutions in this transformation process. TEKMERs accelerate the digitalisation processes of health systems by providing environments that support innovation and collaboration in the field of digital health for entrepreneurs, researchers and academics. Research shows that enterprises that focus on technological innovation attach importance to cooperation with TEKMERs. (Bengisu, 2023, p.5) The aim of this study is to evaluate the contributions of TEKMERs in Türkiye in the field of digital health and to examine their contributions to the integration of digital health solutions into healthcare services. Furthermore, by emphasising the importance of software and informatics sectors in health technologies, the contributions of TEKMERs to the development of digital health solutions and their integration into health systems are discussed.

Materials-Methods: This research addresses the role of Technology Development Centres (TEKMERs) in the digital transformation process in the field of healthcare in Türkiye with a descriptive research method. In the study, the effects of TEKMERs on digital health solutions were examined and analysed using the literature review method. By evaluating the collaborations of TEKMERs with entrepreneurs and companies developing health technology, the effects of these collaborations on the production of digital health solutions were analysed. In addition, the functionality and strategic importance of TEKMERs in accelerating the digital transformation process in the health sector are emphasised. In the study, examples of companies operating in the field of digital health and clustering activities were used as a basis for the analyses.

Results: This research addresses the role of Technology Development Centres (TEKMERs) in the digital transformation process in the field of healthcare in Türkiye with a descriptive research method. In the study, the effects of TEKMERs on digital health solutions were examined and analysed using the literature review method. By evaluating the collaborations of TEKMERs with entrepreneurs and companies developing health technology, the effects of these collaborations on the production of digital health solutions were analysed. In addition, the functionality and strategic importance of TEKMERs in accelerating the digital transformation process in the health sector are emphasised. In the study, examples of companies operating in the field of digital health and clustering activities were used as a basis for the analyses.

Discussion-Conclusion: This study highlights the central role of Technology Development Centres (TDCs) in Türkiye in the digital transformation process in the healthcare sector. As digital transformation makes healthcare services more efficient, sustainable and accessible, TEKMERs are emerging as important building blocks of this transformation. The innovation and collaboration environment offered by TEKMERs in the development and dissemination of digital health solutions enables entrepreneurs and technology producing companies to develop innovative health solutions. In particular, clustering activities increase the innovative capacities of these companies and increase their competitiveness both in the domestic market and in the international arena.

Thanks to the support and consultancy provided by TEKMERs, digital transformation in the field of healthcare is gaining momentum, contributing to the transformation of healthcare systems into more sustainable and flexible structures. In this context, TEKMERs' continued support for innovation-oriented strategies in healthcare technologies will further advance digital transformation in Türkiye's healthcare sector. In conclusion, the strategic role of TEKMERs contributes to the digitalisation of healthcare services, creating opportunities for the implementation of more effective healthcare policies across the country. In the future, developing strategies to increase the functionality of TEKMERs in the field of digital health can help Türkiye become a pioneer in the field of digital transformation in healthcare.

Keywords: Digital transformation, health technologies, technology development centers, innovation, healthcare

REFERENCES

[1] Bengisu, M. (2004). The Contribution of Technology Development Centers and Technoparks to Technological Innovation and Success Factors in Türkiye. İzmir University of Economics Publications.

[2] Akalın, B., & Veranyurt, U. (2021). Artificial Intelligence in Health Services and Management. Acta Infologica, 5(1), 231-240.

ENHANCED DETECTION OF MEDICAL PERSONAL PROTECTIVE EQUIPMENT IN HEALTHCARE USING YOLO11

Öyküm Akar[1*], Hasan Selim[2], Orhan Er[3]

[1] *İzmir Bakırçay University, İzmir, Türkiye*

The Graduate School of Natural and Applied Sciences, Dokuz Eylül University, İzmir, Türkiye

[2] *Department of Industrial Engineering, Dokuz Eylül University, İzmir, Türkiye*

[3] *Department of Computer Engineering İzmir Bakırçay University, İzmir, Türkiye*

** Corresponding author: oykum.akar@bakircay.edu.tr*

Introduction-Aim: Healthcare is a challenging field to work in due to its rapidly changing dynamics and potential risks to healthcare workers. Particularly in recent years, the increase in infectious diseases has necessitated a greater emphasis on personal protection measures for healthcare workers. Therefore, effective and correct use of personal protective equipment (PPE) plays a vital role in ensuring the safety of healthcare workers and patients. Medical PPE includes respiratory, hand-arm, eye/face and body protection.

In recent years, research has been increasing into medical PPE recognition using deep learning techniques. In this study, YOLO11, one of the most advanced object detection algorithms, was used to detect five different PPE essential for healthcare workers, namely face shields, goggles, masks, gloves and coveralls, on the CPPE-5 dataset [1]. The main objective of this research is to improve the accuracy of the object recognition model with an extended medical PPE dataset.

Materials-Methods: This study aims to detect medical PPE by training YOLO11, the latest model of the YOLO algorithm, on the CCPE-5 dataset and the CCPE-5 AUG dataset with data augmentation. The CCPE-5 dataset, consisting of 1029 images, is used to improve the data quality and enable more effective model training. Removing extraneous and duplicate images reduced the dataset to 1000 high-quality images. The images were manually labeled on the Roboflow platform [2] by domain experts to ensure that each piece of medical PPE was correctly labeled. Various data augmentation techniques were applied to the CPPE-5 dataset to ensure that the model performed better in detecting healthcare equipment and generalized better to different scenarios. The result of this augmentation process is the CCPE-5 AUG dataset, which contains 3800 images. The labeled data is randomly divided into 80% training, 10% test, and 10% validation sets. The YOLO11 nano configuration was trained with the original and augmented datasets for 50 epochs in the next step.

Results: The YOLO11 model was trained on the original and augmented datasets for 50 epochs in the nano configuration. As a result of the training sessions, 80.5% Mean Average Precision (mAP50) performance was achieved on the CPPE-5 dataset. Although the model showed generally high accuracy in detecting PPE, it performed poorly on small objects such as masks and goggles. Therefore, the model must be improved in detecting small and complex PPE. However, by applying data augmentation techniques, the model's accuracy was increased to 97.3%, significantly improving mAP50 performance. When the other performance metrics are evaluated on the test dataset, the precision value for the original dataset is 0.890 and the recall is 0.664, while the precision value for the augmented dataset is 0.959 and the recall is 0.949. The training test results on the CPPE-5 AUG dataset reveal that the data augmentation significantly improved the model's ability to detect objects under different conditions.

Discussion-Conclusion: In this study, the performance of the YOLO11 nano configuration used for medical PPE detection was evaluated on both the original CPPE-5 dataset and the CPPE-5 AUG dataset augmented with data augmentation techniques. Data augmentation was one of the most essential factors that improved the model's performance. The model trained on images from different perspectives, brightness, and angles could generalize better to different real-world situations, thus providing higher accuracy in medical PPE detection. The results reveal that by enriching the datasets for future research and applying different data augmentation methods, deep learning algorithms can perform very well in image recognition and analysis in the field of health and safety.

Keywords: Computer vision, Object detection, Deep learning, Occupational health and safety, Medical personal protective equipment, YOLO11

REFERENCE

[1] R. Dagli and A. M. Shaikh, "CPPE-5: Medical personal protective equipment dataset," SN Computer Science, vol. 4, no. 263, pp. 1-5, 2023. doi: 10.1007/s42979-023-01748-7.

[2] J. Nelson, "Getting started with Roboflow," *Roboflow Blog*. Available: https://blog.roboflow.com/getting-started-with-roboflow/. [Accessed: 3-Oct-2024]

MITIGATING ADVERSARIAL ATTACKS ON ECG CLASSIFICATION IN FEDERATED LEARNING VIA ADVERSARIAL TRAINING

Eyüpcan Çelik[1*], Mehmet Kemal Güllü[1]

[1] *Department of Electrical and Electronics Engineering, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: eyupcancelik20@gmail.com*

Introduction-Aim: Federated Learning (FL) [1] has become an important research area in recent years, especially when working with sensitive information such as healthcare data. Since healthcare data contains critical information that needs to be protected, FL offers a great advantage by enabling training on local devices without being collected on a central server. When analyzing healthcare data, such as electrocardiography (ECG), this method enables local processing of data. However, despite its positive impact on privacy, FL can be vulnerable to attacks. Malicious inputs that prevent the model from producing accurate results, known as adversarial attacks, can pose a major threat. Adversarial Training (AT) can be applied as a defense mechanism against these attacks. AT makes the model more robust against such malicious inputs. Federated Adversarial Training (FAT) [2] is the integration of AT into the FL environment. In this work, we propose the use of FAT to provide privacy and security when classifying ECG signals and for robustness against adversarial attacks. For this purpose, a AT is performed in the client by adding the Projected Gradient Descent (PGD) [3] attacked version to the clean ECG data. A Convolutional Neural Network (CNN) architecture was used for local training. The MIT-BIH Arrhythmia Database (MIT-DB) [4] was studied. We also train a federated learning model without using FAT. And we test these structures on the original test data, PGD attacked version and Fast Gradient Sign Method (FGSM) [5] attacked version of the test data and compare the results. The results show that the FL system with FAT achieves much higher performance on adversarial attacks than the FL system without FAT, with some compromise on the performance of the original test data, thus demonstrating the effectiveness of the FL system with FAT against adversarial attacks for the ECG classification task.

Materials-Methods: MIT-DB was used as the data set. The labels in the MIT-DB were modified according to the Association for the Advancement of Medical Instrumentation (AAMI) standard. Convolutional Neural Network (CNN) was used as the deep learning architecture.

In this study, the use of the FAT structure is proposed to enhance robustness against adversarial attacks while classifying ECG signals. In this structure, PGD attacked versions are added to the clean data of the clients and local training is performed in this way. In other words, each of the clients participating in the training performed AT. A FL system without FAT was also trained.

Training of FL models was performed for 10 rounds, and in each round, the clients selected for training performed 5 epochs of local training. There are 10 clients in total. And in each round, 5 randomly selected clients participated in the training.

Results: Both the FL system with FAT and the FL system without FAT were tested with the original test data, the PGD attacked version and the FGSM attacked version and the results were compared. Accuracy, Precision, Recall and F1 Score metrics were used during the tests. Performances achieved against clean data in the 10th round of training with FL system with FAT were as follows: 97.60% Accuracy, 90.55% Precision, 88.79% Recall, and 89.06% F1 Score. Against PGD attacked data: 94.82% Accuracy, 81.72% Precision, 80.13% Recall, and

79.75% F1 Score. Against FGSM attacked data: 95.44% Accuracy, 83.76% Precision, 81.85% Recall, and 81.70% F1 Score. In the FL system without FAT, the performances achieved were as follows: Against clean data 98.44% Accuracy, 93.95% Precision, 92.80% Recall, and 92.94% F1 Score. Against PGD attacked data: 35.14% Accuracy, 30.12% Precision, 21.27% Recall, and 22.61% F1 Score. Against FGSM attacked data: 73.87% Accuracy, 47.05% Precision, 47.29% Recall, and 45.46% F1 Score.

Discussion-Conclusion: In this study, we aim to build a high-performance ECG classifier on adversarial attacked data by using AT in clients in federated learning, i.e. FAT, with very little compromise in performance against the original data. For this purpose, in addition to clean data, PGD attacked version was also used during local training on the clients. The proposed framework was tested against the original test data, PGD attacked version and FGSM attacked version and its performance was compared with that of the federated learning system without FAT. It was observed that the FL system without FAT achieved very high performance in Accuracy, Precision, Recall and F1 Score on the original test data. However, it could not achieve this performance in all four metrics against PGD and FGSM attacked data. In the proposed structure, i.e. the FL system using FAT, high performance is achieved in all four metrics for the original test data, PGD and FGSM attacked data. When both FL systems are compared, the FL system without FAT is more successful on the original test data. On the other hand, the FL system with FAT is more successful on the data subjected to PGD and FGSM attacks. However, the FL system with FAT is also successful on the original test data. The FL system with FAT achieves very high performance against adversarial attacked data with very little performance compromise from the original test data.

With the work done here, while classifying ECG signals, both the desired privacy and security can be achieved with FL and a more robust model can be obtained against adversarial attacks that may occur with FAT. This study can shed light on future studies with different health data.

Keywords: Federated Adversarial Training, Federated Learning, Adversarial Attacks, ECG

REFERENCE

[1] H. B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Y. Arcas, "Communication-Efficient Learning of Deep Networks from Decentralized Data," International Conference on Artificial Intelligence and Statistics, pp. 1273–1282, Apr. 2017, [Online]. Available: http://proceedings.mlr.press/v54/mcmahan17a/mcmahan17a.pdf

[2] G. Zizzo, A. Rawat, M. Sinn, and B. Buesser, "Fat: Federated adversarial training," 2020, arXiv:2012.01791.

[3] A. Madry, A. Makelov, L. Schmidt, D. Tsipras, and A. Vladu, "Towards deep learning models resistant to adversarial attacks.," arXiv (Cornell University), Feb. 2018, [Online]. Available: http://arxiv.org/pdf/1706.06083.pdf

[4] G. B. Moody and R. G. Mark, "The impact of the MIT-BIH Arrhythmia Database," IEEE Engineering in Medicine and Biology Magazine, vol. 20, no. 3, pp. 45–50, Jan. 2001, doi: 10.1109/51.932724.

[5] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," International Conference on Learning Representations, Mar. 2015, [Online]. Available: https://ai.google/research/pubs/pub43405

COMPARATIVE PERFORMANCE OF EXPERT-ENGINEERED AND AUTO MACHINE LEARNING MODELS IN CLASSIFYING VITREOMACULAR INTERFACE DISORDERS IN OPTICAL COHERENCE TOMOGRAPHY IMAGES

Ufuk Beşenk[1], Ceren Durmaz Engin[2], Seher Köksaldı[3], Mustafa Kayabaşı[3], Alper Selver[1]

[1.] Department of Electrical and Electronics Engineering, Faculty of Engineering, Dokuz Eylul University, İzmir, Türkiye

[2] Department of Ophthalmology, Democracy University Buca Seyfi Demirsoy Training and Research Hospital, İzmir, Türkiye

[3] Ophthalmology Clinic, Muş State Hospital, Muş, Türkiye

** Corresponding author: contact.ufuk.besenk@gmail.com*

Introduction-Aim: Vitreomacular interface (VMI) diseases, including epiretinal membrane (ERM), full-thickness macular hole (FTMH), lamellar macular hole (LMH), and vitreomacular traction (VMT), present distinct structural changes that can be precisely visualized through high-resolution optical coherence tomography (OCT). Automated, artificial intelligence (AI) -driven analysis of OCT images holds promise for enhancing diagnostic precision for these pathologies. This study aims to evaluate the performance of AI models, comparing those created with Google Vertex AI AutoML with expert-engineered deep learning models in the classification of OCT images from patients with VMI disorders.

Materials-Methods: This retrospective study included 900 OCT images collected from patients diagnosed with VMI diseases at Buca Seyfi Demirsoy Education and Research Hospital from October 2022 to October 2024. The images represented five distinct categories: ERM, FTMH, LMH, VMT, and normal controls. A total of 200 images were included for each of the ERM, FTMH, LMH, and normal classes, while 100 images were used for the VMT class. To ensure consistent data usage across all models, images were partitioned into training (80%) and testing (20%) sets.

Six classification models were developed for evaluating different diagnostic tasks. These models included four binary classifiers that differentiated each of the four pathological conditions (ERM, FTMH, LMH, and VMT) from the normal control group. Additionally, a specialized binary model was constructed to distinguish between FTMH and LMH, two conditions that often present overlapping morphological features. Finally, a multi-class classification model (Model VI) was created to categorize all five classes (ERM, FTMH, LMH, VMT, and normal) in a single classification task, providing a comprehensive assessment of model performance across all diagnostic categories.

The AutoML model was developed using Google Vertex AI, allowing an ophthalmologist to create machine learning models without coding through an automated pipeline. Vertex AI handled all aspects of image processing, model selection, and hyperparameter optimization. The expert-engineered model, however, was designed using Python and implemented using an EfficientNetV2 architecture. This model underwent custom preprocessing, including resizing and augmentation, to enhance feature extraction capabilities. Data augmentation, such as random flipping and brightness adjustments, was applied. Hyperparameters were optimized manually by the engineer using learning rate schedules, dropout layers, and batch

normalization. Both models were evaluated using standard performance metrics, including accuracy, precision, recall, F1-score, and confusion matrices.

Results: The expert-engineered EfficientNetV2 model outperformed the AutoML model in all models. In binary classifications, the expert model achieved near-perfect results, reaching 100% in accuracy, recall, precision, and F1-scores for distinguishing the normal control group from ERM, FTMH, and VMT. In contrast, the AutoML model achieved accuracy rates ranging from 80–90%, with lower recall for pathological cases, indicating potential diagnostic limitations.

In differentiating between FTMH and LMH, the EfficientNetV2 model achieved 85% accuracy, correctly identifying all FTMH cases with a recall of 1.00 but misclassifying 12 out of 40 LMH cases, resulting in an LMH recall of 70% and an F1-score of 0.82. In comparison, the AutoML model attained 81.2% accuracy, misclassifying 13 out of 40 LMH cases, yielding an LMH recall of 67.5% and an F1-score of 0.80. The expert-engineered model showed a better balance of precision and recall in this model.

In the multi-class classification task involving all five categories, the EfficientNetV2 model achieved an overall accuracy of 86.7%, showing high precision and recall for normal (precision: 0.93, recall: 1.00, F1-score: 0.96) and FTMH (precision: 0.91, recall: 1.00, F1-score: 0.95). The model performed strongly for VMT, with precision and recall of 0.95 and an F1-score of 0.95, and for ERM, with a precision of 0.71, recall of 0.97, and F1-score of 0.82. However, the model struggled to classify LMH accurately, achieving a perfect precision of 1.00 but with a recall of only 45%, resulting in an F1-score of 0.62, as misclassifications tended to confuse LMH with ERM and FTMH. In comparison, the AutoML model reached an overall accuracy of 83.3% and excelled in normal classification (precision: 0.975, recall: 1.00, F1-score: 0.98). However, it underperformed in detecting ERM, FTMH, and LMH, with LMH showing a precision of 0.76, recall of 55%, and F1-score of 0.64, frequently misclassifying LMH as ERM or FTMH, highlighting a limitation in distinguishing between similar retinal pathologies.

Discussion-Conclusion: This study demonstrates the superior diagnostic performance of an expert-engineered EfficientNetV2 model over an AutoML-generated model in OCT image classification for VMI diseases. While AutoML provides efficiency, the expert-designed model showed higher sensitivity and accuracy across all categories, emphasizing the importance of domain expertise in creating effective AI tools for specialized medical applications. Future research should explore hybrid approaches integrating expert guidance with AutoML, potentially enhancing clinical applicability and facilitating reliable AI-assisted diagnostics in ophthalmology.

THE IMPACT OF AI-SUPPORTED MICROPROCESSOR PROSTHESIS ON REHABILITATION: A CASE REPORT

Murat Ali Çınar[1*], Kezban Bayramlar[1]

[1] *Physiotherapy and Rehabilitation, Hasan Kalyoncu University, Gaziantep, Türkiye*

* *Corresponding author: muratali.cinar@hku.edu.tr*

Introduction and Aim: Microprocessor-controlled prostheses represent a significant advancement in prosthetic technology, designed to provide individuals with lower extremity amputations a more natural and functional movement. Traditional prostheses often lack the dynamic adaptability necessary to cater to the varying demands of real-life activities. In contrast, microprocessor-controlled devices leverage artificial intelligence (AI) to analyze the user's movements in real time, optimizing balance and movement control.

The primary aim of this study is to evaluate the functional gains associated with the rehabilitation of a patient with bilateral transfemoral amputations using the Genium X3 microprocessor-controlled prosthesis. This analysis focuses on key performance metrics, including balance, walking distance, energy efficiency, and overall ambulation skills. Additionally, this study investigates how AI-enabled adaptability and movement optimization in response to environmental conditions can enhance the patient's quality of life. By highlighting the contributions of AI-supported prostheses to rehabilitation processes, this study aims to provide insights that can inform future treatment strategies in the field of prosthetic rehabilitation.

Materials and Methods: This study focuses on a 21-year-old patient who has undergone bilateral transfemoral amputations. After the surgery, the patient began a comprehensive rehabilitation program. By the 8th week of this program, the patient was fitted with the Genium X3 AI-supported prosthesis, allowing for a deeper examination of its effects on rehabilitation outcomes.

The rehabilitation regimen included weekly sessions comprising resistance exercises, balance training, and mobilization practices. These sessions were structured to gradually improve the patient's strength, stability, and coordination. Measurement tools employed in the assessment included the 6-Minute Walk Test (6MWT) and the Rate of Perceived Exertion (RPE) scale. These tools provided quantitative data on walking distance and perceived exertion levels, allowing for a comprehensive evaluation of the patient's progress. The Genium X3 prosthesis utilizes advanced AI algorithms that analyze the user's movements in real time, thereby developing personalized movement strategies tailored to the surrounding environment. Equipped with sensors, the prosthesis monitors walking patterns and balance control, dynamically adjusting to meet the individual needs of the user. This adaptability is crucial in promoting optimal function during various activities, from walking on flat surfaces to navigating inclines and uneven terrain. Throughout the rehabilitation process, data were collected weekly on walking distance, balance duration, and overall prosthetic use efficiency. This ongoing assessment allowed for the identification of trends and the quantification of improvements resulting from the introduction of the AI-supported prosthesis.

Results: The findings of this study indicate significant improvements in the patient's functional performance following the integration of the Genium X3 prosthesis. In the initial 6-Minute Walk Test, the patient managed to walk 100 meters. Remarkably, by the end of the 8-week rehabilitation program, this distance increased to 320 meters, demonstrating a marked

enhancement in mobility. Such improvement reflects not only the physical capabilities of the prosthesis but also the patient's increasing confidence and proficiency in using it. Balance training also yielded positive results. The duration for which the patient could maintain balance increased from an initial 10 seconds to 35 seconds by the conclusion of the program. This enhancement underscores the prosthesis's effectiveness in providing stability and support, enabling the patient to engage more fully in daily activities. Additionally, the AI technology in the prosthesis played a crucial role in minimizing energy consumption while maximizing stability. The integration of AI allowed for a seamless adjustment to the user's individual movement patterns, thereby facilitating smoother and more energy-efficient ambulation. Ultimately, it was concluded that the Genium X3 AI-supported prosthesis provided greater functional gains compared to traditional prosthetic options, significantly contributing to the rehabilitation process.

Discussion and Conclusion: This case report illustrates the transformative potential of AI-supported prostheses within rehabilitation settings. The advancements in prosthetic technology, particularly those incorporating AI, can profoundly impact the quality of life for individuals with amputations. The ability of these devices to adapt dynamically to the user's needs and environmental conditions offers a more holistic approach to rehabilitation.

As the field of prosthetics continues to evolve, it is imperative to expand research to include diverse patient populations and varying levels of amputations. Future studies should focus on the integration of AI algorithms into personalized treatment plans, ensuring that all patients can benefit from the advancements in prosthetic technology. By doing so, we can pave the way for more effective rehabilitation strategies that enhance both physical function and overall quality of life.

Keywords: Artificial Intelligence, Rehabilitation, Prosthesis,

REFERENCE

[1] X. Wang et al., "Hybrid Active–Passive Prosthetic Knee: A Gait Kinematics and Muscle Activity Comparison with Mechanical and Microprocessor-Controlled Passive Prostheses," Journal of Bionic Engineering, vol. 20, no. 1, pp. 119-135, 2023.

[2] J. Lee and M. Goldfarb, "The effects of swing assistance in a microprocessor-controlled transfemoral prosthesis on walking at varying speeds and grades," Wearable Technologies, vol. 4, no. e9, 2023.

[3] T. D. Klenow et al., "An enhancement of the Genium™ microprocessor-controlled knee improves safety and different aspects of the perceived prosthetic experience for unilateral and bilateral users," Frontiers in Rehabilitation Sciences, vol. 5, no. 1342370, 2024.

[4] Y. Bader et al., "Development of an Integrated Powered Hip and Microprocessor-Controlled Knee for a Hip–Knee–Ankle–Foot Prosthesis," Bioengineering, vol. 10, no. 5, p. 614, 2023.

[5] A. Kannenberg, B. Zacharias, and E. Pröbsting, "Benefits of microprocessor-controlled prosthetic knees to limited community ambulators: systematic review," Journal of Rehabilitation Research and Development, vol. 51, no. 10, pp. 1469-1476, 2014.

CLASSIFICATION OF RETINAL DISEASES USING A MULTI-MODEL APPROACH

Burak Kaya[1*], Murat Uçar[1]

[1] *Computer Engineering Department, Bakırçay University, İzmir, Türkiye*

* *Corresponding author: bkay5292@gmail.com*

Introduction-Aim: The retina is a sensory neural layer located in the innermost part of the eye wall, composed of photosensitive cells. It consists of blood vessels that nourish its nerve cells and multiple layers, each responsible for a different phase of the vision process. Its primary function is to detect light, convert it into electrical signals, process these signals, and transmit them to the vision-related area of the brain via optic nerves, thus enabling the vision process.

Advancements in medicine, science, and technology continue to bring innovations to both software and hardware solutions used in disease diagnosis and treatment. Various innovations, including artificial intelligence-supported simulations, robotic applications, and personal health systems, play crucial roles in disease prevention, treatment improvement, and enhancement of quality of life. These developments, prevalent in numerous medical fields from cardiology to dentistry and from neurology to orthopedics, are frequently observed in ophthalmology as well. In this context, there is intensive and widespread interest in retinal diseases, which are recognized as one of the leading causes of vision loss [1]. Promising research is being conducted on the potential application of various artificial intelligence algorithms for eye diseases, particularly diabetic retinopathy, age-related macular degeneration, and retinopathy of prematurity [2]. In a study published in 2023 by researchers at Hamad Bin Khalifa University, various methods were applied for the detection, classification, and prediction of retinal detachment using OCT images through Machine Learning and Deep Learning techniques, contributing significantly to the literature in this field [3]. Existing studies have predominantly focused on image-based classification, notably lacking the integration of other influential parameters (such as age, gender, etc.) in disease diagnosis.

Considering the increasing prevalence of artificial intelligence technologies, there is a perceived need for increased research efforts not only for rapid and easy disease detection but also for prevention through early diagnosis. Despite various promising studies [4], [5], challenges in algorithm development, data privacy concerns, and lack of clinical validation have led to limitations in the literature from this perspective, preventing widespread adoption in clinical settings.

This study aims to develop a more meaningful prediction model by incorporating personal data such as age and gender in addition to retinal fundus images, thus including more parameters that could be effective in disease labeling.

Materials-Methods: With the initial conception of artificial intelligence in the 1950s, the subsequent increase in computational power and advancement of technologies such as big data analytics since the early 2000s has accelerated the proliferation and widespread adoption of AI applications across various domains [6]. Medicine stands as one of the most significantly impacted fields. The applications of artificial intelligence, encompassing machine learning, are evident across numerous medical areas, from disease diagnosis to personalized treatment, and from drug discovery to radiology. Various studies conducted in these areas promise a more effective, efficient, and personalized healthcare system [7].

Medical AI applications demonstrate their influence across numerous specialties including cardiology, dentistry, neurology, and others. Along with traditional machine learning algorithms, deep learning's capacity to analyze large and complex data sets more rapidly under appropriate conditions has led to its increasing popularity in healthcare [8]. Notable in recent research is the widespread utilization of CNN architecture, a subset of deep learning algorithms, for analyzing images obtained from medical imaging devices such as computed tomography (CT), magnetic resonance imaging (MRI), ultrasonography (USG), and optical coherence tomography (OCT) [9]-[11].

Convolutional Neural Networks (CNNs), fundamental building blocks of deep learning methods used for various computer vision tasks such as object recognition and classification, effectively process image data to identify and classify complex patterns. In these architectures, the training process is completed according to selected parameters, resulting in the development of a deep learning model based broadly on CNN architecture [12].

In this study, a multi-model approach is proposed which consists of two main parts where the first part is using ConvNet based models to extract features from OCT images and the second part is, after normalizing personal data, using machine learning base models to classify diseases using the Ocular Disease Recognition[13] dataset.

Results: This section presents comparisons of four of the State of the Art ConvNet architectures (DenseNet, InceptionNet, ResNet50 and VGG16) and the multi-model method proposed. After all these comparison results are analyzed together, a multi-model architecture is proposed.

Discussion-Conclusion: This study aims to propose a multi-model approach to classify retinal fundus diseases where input data include more than only image data.

Since classification on different retinal fundus diseases using different types of data can benefit to machine learning methods in aspect of robustness and these machine learning methods get bigger roles in the medicine after each day, the proposed model can have use in applications regarding the field of medicine.

The proposed model in this study can be used in such applications in the field of medicine which provides a quick early classification chance in order to examine the patients.

Keywords: retinal disease, classification, convolutional networks, prediction

REFERENCE

[1] Schmidt-Erfurth, U., Sadeghipour, A., Gerendas, B. S., Waldstein, S. M., & Bogunović, H. (2018). Artificial intelligence in retina. Progress in retinal and eye research, 67, 1-29.

[2] Keskinbora, K., & Güven, F. (2020). Yapay Zeka ve Oftalmoloji. Turk J Ophthalmol, 50(1), 37-43.

[3] Zaky, H., Salem, A., Alzubaidi, M., Shah, H. A., Alam, T., Shah, Z., & Househ, M. (2023). Using AI for Detection, Prediction and Classification of Retinal Detachment. Healthcare Transformation with Informatics and Artificial Intelligence, 636-639.

[4] Ettiyappan, A., & David, J. A. (2023). Improved Lion Optimization and Faster Mask Recurrent CNN Developed for Diabetic Retinal Detachment Prediction. Revue d'Intelligence Artificielle, 37(6).

[5] Bilal, A., Imran, A., Baig, T. I., Liu, X., Long, H., Alzahrani, A., & Shafiq, M. (2024). Improved Support Vector Machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification. Plos one, 19(1), e0295951.

[6] Kaul, V., Enslin, S., & Gross, S. A. (2020). History of artificial intelligence in medicine. Gastrointestinal endoscopy, 92(4), 807-812.

[7] Cipriano, L. E. (2023). Evaluating the impact and potential impact of machine learning on medical decision making. Medical Decision Making, 43(2), 147-149.

[8] Castiglioni, I., Rundo, L., Codari, M., Di Leo, G., Salvatore, C., Interlenghi, M., ... & Sardanelli, F. (2021). AI applications to medical images: From machine learning to deep learning. Physica medica, 83, 9-24.

[9] Alzubaidi, L., Zhang, J., Humaidi, A. J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., ... & Farhan, L. (2021). Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of big Data, 8, 1-74.

[10] Dhillon, A., & Verma, G. K. (2020). Convolutional neural network: a review of models, methodologies and applications to object detection. Progress in Artificial Intelligence, 9(2), 85-112.

[11] Yao, G., Lei, T., & Zhong, J. (2019). A review of convolutional-neural-network-based action recognition. Pattern Recognition Letters, 118, 14-22.

[12] Phung, V. H., & Rhee, E. J. (2018). A deep learning approach for classification of cloud image patches on small datasets. Journal of information and communication convergence engineering, 16(3), 173-178.

[13] “ODIR-2019 - Grand Challenge,” grand-challenge.org. https://odir2019.grand-challenge.org/

ANALYSIS OF ACOUSTIC BIOMARKERS WITH YAMNET: AUTOMATIC DIAGNOSIS OF LUNG DISEASES USING COUGH SOUNDS

Ayşen Özün Türkçetin [1,2*], Turgay Koç[3]

[1] IT Department, Suleyman Demirel University, Isparta, Türkiye

[2] Graduate School of Natural and Applied Sciences Department of Mechanical Engineering, Suleyman Demirel University, Isparta, Türkiye

[3] Department of Electric and Electronic Engineering, Suleyman Demirel University, Isparta, Türkiye

** Corresponding author: aysenturkcetin@sdu.edu.tr*

Introduction-Aim: Respiratory diseases, which constitute a significant global health problem, require effective and rapid diagnostic methods. As one of the important indicators of respiratory function, acoustic and sound-based digital biomarkers have been increasingly used in diagnostic processes in recent years. Thanks to the development of machine learning (ML) algorithms, the analysis of these biomarkers offers promising approaches for the detection and diagnosis of respiratory disorders. In this study, mobile Android devices were used to collect cough sound data in the 16 kHz frequency band and WAV format from hospitalized patients with asthma, COPD, and pneumonia. The YamNet model, a deep learning model with a wide range of sound classification capabilities, is used to analyze cough sounds and helps classify sounds into different groups. Google developed the YamNet neural network architecture, which was trained on the AudioSet dataset. It can therefore recognize and evaluate sounds like coughing with accuracy. In this study, the time-frequency features of cough sounds collected using the YamNet model were analyzed in order to automatically diagnose lung diseases using cough sounds.

Materials-Methods:

Dataset Collection

For this study, we collected cough sound data from patients diagnosed with pneumonia, Chronic Obstructive Pulmonary Disease (COPD), and asthma. The dataset comprises recordings of cough sounds that were captured using a mobile Android phone, ensuring ease of access and practicality for real-world applications. The recordings were taken in a controlled hospital environment, adhering to ethical guidelines and patient consent. The dataset is divided into four distinct subsets, each targeting specific disease combinations to evaluate the model's performance in various scenarios:

- Dataset 1: Cough sounds from patients with asthma and COPD (total: 119 recordings)
- Dataset 2: Cough sounds from patients with asthma and pneumonia (total: 97 recordings)
- Dataset 3: Cough sounds from patients with COPD and pneumonia (total: 114 recordings)
- Dataset 4: Cough sounds from patients with asthma, COPD, and pneumonia (total: 165 recordings)

Each recording was collected in WAV format at a sampling frequency of 16 kHz, allowing for the effective capture of the acoustic features of cough sounds. The collection process involved instructing patients to cough naturally while minimizing background noise to enhance the quality of the recordings.

Data Preprocessing

Prior to analysis, the recorded audio files were subjected to preprocessing steps to ensure consistency and usability for machine learning applications. This included normalization of audio levels, trimming of silent segments, and segmentation of the cough sounds for focused analysis. The resulting datasets were then used for training, validation, and test sets to facilitate model evaluation in different configurations.

Model Architecture

This study employs a deep learning model designed to classify cough sounds associated with lung diseases. The model architecture consists of a series of fully connected layers that progressively reduce the dimensionality of the input feature space while applying non-linear transformations. To enhance the model's ability to classify cough sounds, we incorporated the YamNet model as a feature extractor. YamNet is a pre-trained convolutional neural network specifically designed for audio classification tasks. By using YamNet, we can leverage its ability to extract meaningful audio features from the cough sound recordings. During the preprocessing stage, cough sounds were processed through YamNet, which transforms the raw audio signals into high-level feature representations. These extracted features are then input into the dense layers of our model, where they undergo further processing to classify the cough sounds based on the presence of specific lung diseases.

Performance Metrics

The model's performance was evaluated using a confusion matrix, which provides a detailed breakdown of the classification results by displaying true positive, true negative, false positive, and false negative counts. This allows for the calculation of key metrics such as precision, recall, and accuracy.

Result: The performance of the proposed model was evaluated using four distinct datasets, each designed to assess the model's ability to distinguish between different respiratory conditions based on cough sound analysis. This demonstrates the YamNet model's promising approach for clinical applications in automating the diagnosis of respiratory diseases based on cough sound analysis.

Discussion-Conclusion: The findings of this study underscore the potential of using machine learning algorithms for the automated diagnosis of lung diseases based on cough sound analysis. The model demonstrated its capability to effectively classify cough sounds associated with asthma, COPD, and pneumonia, suggesting that acoustic analysis can serve as a valuable diagnostic tool in clinical settings. The integration of the YamNet model proved beneficial, as it provided robust feature extraction from the audio data. The ability of YamNet to capture relevant sound characteristics facilitated the model's learning process, enhancing its accuracy in distinguishing between different types of lung diseases. However, several factors may have influenced the model's performance. The imbalance in the dataset, with varying sample sizes for each disease category, could have contributed to the discrepancies observed in the confusion matrix. Future work should focus on balancing the dataset to improve the model's ability to generalize across different classes. Moreover, the study highlights the importance of utilizing diverse audio data sources and environmental conditions to enhance model robustness. As cough sounds can vary significantly among individuals and contexts, expanding the dataset to include a wider range of recordings may further improve classification accuracy. In conclusion, this study demonstrates the feasibility of employing cough sound analysis as a non-invasive method for diagnosing lung diseases. Continued advancements in machine learning and audio processing techniques hold promise for developing more effective diagnostic tools in respiratory medicine.

Keywords: lung disease diagnosis, cough sound analysis, YamNet, machine learning, acoustic biomarkers, asthma, COPD, Pneumonia

REFERENCE

[1] Kapetanidis, P., Kalioras, F., Tsakonas, C., Tzamalis, P., Kontogiannis, G., Karamanidou, T., ... & Nikoletseas, S. (2024). Respiratory Diseases Diagnosis Using Audio Analysis and Artificial Intelligence: A Systematic Review. Sensors, 24(4), 1173.

[2] Hershey, S., Chaudhuri, S., Ellis, D. P., Gemmeke, J. F., Jansen, A., Moore, R. C., ... & Wilson, K. (2017). CNN architectures for large-scale audio classification. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 131-135). IEEE.

[3] Sobahi, N., Atila, O., Deniz, E., Sengur, A., & Acharya, U. R. (2022). Explainable COVID-19 detection using fractal dimension and vision transformer with Grad-CAM on cough sounds. biocybernetics and biomedical engineering, 42(3), 1066-1080.

[4] Deshpande, G., & Schuller, B. W. (2020). Audio, speech, language, & signal processing for covid-19: A comprehensive overview. arXiv preprint arXiv:2011.14445.

[5] Türkçetin, A. Ö., Koç, T., & Çilekar, Ş. (2023, July). The Use of ANN in the Sound Detection of Lung Diseases: Example of COPD, Asthma, Pneumonia. In 2023 31st Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE.

THE INVESTIGATION OF KNOWLEDGE, PERCEPTIONS AND ATTITUDES OF SPEECH AND LANGUAGE THERAPY DEPARTMENT STUDENTS REGARDING ARTIFICIAL INTELLIGENCE TOOLS: PRELIMINARY FINDINGS

Şükriye Kayhan-Aktürk[1*], Beyza Nur-Dükar[2]

[1] *Department of Speech and Language Therapy, Muğla Sıtkı Koçman University, Muğla, Türkiye*

[2] *Department of Speech and Language Therapy, Anadolu University, Eskişehir, Türkiye*

** Corresponding author: sukriyeakturk@mu.edu.tr*

Introduction-Aim: Speech and Language Therapy (SLT) is a healthcare field focusing on assessing, diagnosing, and treating communication, voice and swallowing disorders. Artificial intelligence (AI) tools are expected to have significant impacts in the field of SLT, as in various other healthcare fields [1]. The implementation of AI in the field of SLT can enhance both therapeutic applications and educational processes. As a result, individuals' treatment and educational experiences can become more efficient [2]. In this context, identifying the knowledge, perceptions, and attitudes of SLT students toward AI tools is crucial for providing opportunities for students to take an active role in the development, implementation, and use of these technologies. Building on this, the aim of the current study is to explore the knowledge, perceptions, and attitudes of SLT students in Türkiye regarding AI tools. The lack of studies examining the knowledge, perceptions, and attitudes of SLT students in Türkiye toward AI tools serves as the basis for this research.

Materials-Methods: Prior to the initiation of the study, ethical approval was obtained from Anadolu University (Protocol no. 770370). In accordance with the purpose of the study, a descriptive survey model, one of the quantitative research methods, was used. A total of 257 SLT students, 231 women and 26 men, studying at various universities in Türkiye during the 2024-2025 academic year participated in the study. Prior to participation, participants were presented with an Informed Consent Form, which provided information about the content and procedure of the study. The data were collected through the “SLT Students’ Knowledge, Perception, and Attitude Questionnaire on AI Tools”, which was developed by the research team through a review of the literature. Before the questionnaire was developed, the course contents and objectives of SLT programs at universities in Türkiye were examined. Based on these examinations, the researchers created an item pool, and the questions within this pool were categorized. The final version of the questionnaire was evaluated by two assistant professors from SLT departments at different universities in terms of scientific validity, clarity, and language use. Following the feedback from this evaluation, the final version of the questionnaire was created, consisting of five sections and 40 items. The sections of the questionnaire are as follows:

a) Demographic Information
b) Knowledge of AI
c) AI in SLT Education
d) The Impact of AI on SLT Practices
e) Perceptions of AI and Ethical Considerations

The questionnaire primarily consisted of items rated on a three-point scale (yes, no, undecided) or a five-point scale (strongly agree, agree, undecided, disagree, strongly disagree). Completing the questionnaire took approximately 10 minutes. The questionnaire was shared with SLT students through online platforms, social media, and email via Google Forms. The results obtained from the questionnaire were analyzed using descriptive statistical methods.

Results: The analysis of the research findings revealed that around 45% of participants believe they lacked adequate information regarding AI. The majority of participants utilized AI for homework (86%), translation (73%), and exam preparation (58%) during the educational process. Concerning the role of AI in undergraduate education, 66% of participants indicated the absence of a course specifically focusing on AI, whilst 12% noted its inclusion in certain courses. 70% of the participants thought that they would not have the necessary skills after graduation and that AI should be integrated into the undergraduate curriculum, and that AI would enhance their professional prospects. The prevailing perspective was that AI would positively influence SLT applications especially in the stages of material development, homework, evaluation, reporting and it would be beneficial mostly in the fields of speech sound disorders and developmental language disorders. 38% of the participants believed it would not influence career prospects, while 33% remained uncertain. 67% said that AI tools would not substitute SLT intervention, whereas 43% of those who considered these tools replaceable predicted this would occur within 20 years. This circumstance could impose mild (33%), moderate (26%), severe (23%), and extremely high (6%) levels of anxiety.

Discussion-Conclusion: The integration of AI into healthcare services has been progressively increasing throughout the years. The knowledge, perceptions, and attitudes of SLT students concerning AI are crucial for the adoption of this technology, as these students will significantly contribute to its application and development within the field. Their competencies, desires and expectations regarding these technologies will affect the process of using and developing AI. The integration of AI into healthcare services has been accelerating in recent years.

As a result of the research, 50.2% of the participants stated that they did not have knowledge of AI, 60% stated that there was no course related to AI in the curriculum, and that they did not have the necessary equipment related to AI. The research findings are similar to the AI attitudes of medical students [3,4] and dentistry students [5] in different countries. As a result of the research, it was generally found that students do not have sufficient knowledge about AI, but they have positive attitudes towards learning more about the AI tools [6]. Approximately 70% of the participants think that AI should be included in the curriculum and that this information will benefit their careers. It is emphasized that integrating AI topics into the curriculum in the fields of medicine [3,4] and dentistry [5], as well as in the field of SLT, is necessary to better prepare future health professionals. Participants, like dentistry students [5], stated that AI will positively affect SLT practices and facilitate practices. They think that they can benefit from AI in the field of SLT, especially in material development, homework follow-up, and reporting stages. It was observed that the proportion of opinions that AI will affect the possibility of finding a job in the field of SLT was close, but 43% of the participants thought that AI could replace SLT in 20 years, which caused different levels of anxiety in the majority. Similarly, students in professions such as medicine and dentistry believe that AI will significantly reduce job opportunities [3]. This situation is thought to pave the way for professional and ethical discussions for SLT students in parallel with students in other health fields.

Keywords: Artificial intelligence, speech and language therapy, student attitudes.

REFERENCES

[1] B. Bucholz, B. V. Prabhakar, J. K. Ko, “Speech recognition systems for speech disorders: Applications and advancements”, *Journal of Speech Therapy,* vol. 56, pp. 145-158, 2023.

[2] A. D'Agostino, S. M., Fernandez, T. R. Anderson, “Natural language processing and machine learning in speech therapy: A review”, *Artificial Intelligence in Medicine,* vol. 118, pp. 102-113, 2022.

[3] N. Jha, P. R. Shankar, M. A. Al-Betar, R. Mukhia, K. Hada, S. Palaian, “Undergraduate medical students’ and interns’ knowledge and perception of artificial intelligence in medicine”, *Advances in Medical Education and Practice*, vol. 13, pp. 927, 2022.

[4] N. Mehta, V. Harish, K. Bilimoria, F. Morgado, S. Ginsburg, M. Law, S. Das, “Knowledge of and attitudes on artificial intelligence in healthcare: A provincial survey study of medical students”, MedRxiv, vol.2021-01, pp. 1-14, 2021.

[5] E. Yüzbaşıoğlu, "Attitudes and perceptions of dental students towards artificial intelligence," *Journal of Dental Education*, vol. 85, no. 1, pp. 60-68, 2021.

[6] A. J. Buabbas, B. Miskin, A. A. Alnaqi, A. K. Ayed, A. A. Shehab, S. Syed-Abdul, et al., "Investigating students’ perceptions towards artificial intelligence in medical education," *Healthcare*, vol. 11, no. 1298, 2023.

INSTANT HEALTH CARE SYSTEM INTEGRATED TO SMART HOME SYSTEMS

Vadi Su Yılmaz[1*], Yunus Sühan, Ücüretci[1], Barış Mert Akpınar[1], Reşat Özgür Doruk[1*]

[1] *Electrical Electronics Engineering, Atılım University, Ankara, Turkiye*

* *Corresponding author: resat.doruk@atilim.edu.tr*

Introduction-Aim: Today, interest in smart systems is progressing increasingly [1,2]. As a result of increasing interest, intelligent systems have branched out among themselves. Although IoT systems have been discussed in the literature on communication and many issues [3,4], they are crucially important for healthcare smart systems [5]. Although many advances have been made in smart home systems today, people in need of help and elderly people still sustain their lives mostly depending on patient care assistants. However, the latest technological developments can meet the requirements of these people, cost of these systems can not be affordable for every income level. The integration of smart homes and smart health systems offers a comprehensive approach to instant monitoring of the health of a person and prevents accidents or injuries that may occur in the house. Although the instantaneous transmission of health status to a relevant physician is not broadly implemented, it is anticipated that It will become accessible with the adaptation of technology. In this way, it is possible to make early intervention in case of a possible health abnormality [6,7]. In addition, smart home systems can be controlled simultaneously with instant notification health systems from the application [8]. Thus, it can prevent theft, fire, flooding, and home-related accidents, with instant notification of fainting, falling situations, blood pressure, sugar, pulse, blood oxygen, and diet monitoring from urine pH, vital situations can be brought under control with early intervention in case of any abnormality that may occur regarding the health condition. According to studies conducted by the United Nations, by 2050, 70% of the world's population will begin to live in cities [9], and It is anticipated that 22% of the population will consist of elderly individuals due to advancements in medicine [10]. Intelligent systems developed for the elderly and people in need of care enable them to maintain their living spaces and periods in the best possible way, thus the developments in this field provide great convenience to our living comfort.

Materials-Methods: In this system, one of the main purposes of smart home systems is to facilitate the individual's life. Concerning this, to deal with the key and security problem at the entrance, controllable devices such as fingerprint sensors, and face recognition system, and for the health system, the patient's blood sugar values are taken daily, the urine paste is automatically checked and the blood oxygen and pulse values are continuously monitored. Received data is sent by mobile application to the physician or companion. It aims to prevent any problems that may occur in the processing and interpretation of the data by microprocessor-based systems. Also informs that if the values obtained from the application do not satisfy the reference values, the doctor is contacted directly via SMS. In the literature, these systems are given as telemedicine systems and monitoring systems under the name of telemonitoring systems. In this way, a low-budget system has emerged, and the opportunity to be tested many times has been obtained due to its portability. The systems are provided with the opportunity to work in an integrated manner with each other by communicating via microcontrollers. The received data is stored in the personal database, allowing a trustworthy transmission to the mobile application and the doctor's control. In addition, the coded mobile application is constantly updated with data from the database. Since every change at home will be updated in the database, all the data is transferred to the mobile application. The data received with the condition commands created by the microprocessors and the mobile application are examined. According to the received data, the system is activated or deactivated. In this way, the user provides optimum energy efficiency and usage savings. Additionally, microprocessors are deployed at certain points of the house and they are interconnected with each other over the database. The Internet-based database automatically receives data from the microprocessors used and transfers the data to the microprocessors at the point of need. Thus, sensors and related microcomputing interfaces can receive data instantly and interpret and evaluate it. All this order process took place in the market, and tests were carried out continuously. A design has emerged by working simultaneously between health and smart home systems.

Results: These systems aim to prevent any problem that may occur at any time by collecting data instantly and continuously in the house. This way, it is ensured that situations such as natural gas leaks and fires that may develop in the place are prevented before they begin. Ovens, curtains, doors, heating systems, windows, faucets, lights, and electronic devices are automatically and manually controlled systems in the context of our smart home systems. These systems are generally used to control the appliances in the house by controlling the sockets. It is a system that eliminates the necessity of control by imitating infrared signals, which allows each device controlled by remote control to be controlled by the mobile application. A telemedicine system is a two-way mechanism, of data transfer between the doctor and the user, who is under the doctor's control. Daily data transmission occurs between the doctor and the patient. The environment of this flow is provided generally from mobile applications, but in case of emergency, it is transmitted as SMS by the GSM module. Physicians receive data on blood sugar, oxygen level, blood pressure, heart rate, and diet adjustment according to urine pH level, although they also receive data on fainting and falling situations.

Discussion-Conclusion: To summarize, this system, which was established to facilitate the lives of people in need of help supported by smart home systems, has given a new dimension to telemedicine systems as well as the literature, giving the user a structure based on the doctor's control as a result of dietary recommendations and excessively long-lasting disorders, from the urine pH value. Thanks to this system, the stabilization of the pH value of urine prevents a problem that may arise from an eating disorder. In addition, it is a system that provides convenience, such as instant treatments and prescriptions, when necessary, through routine control, blended with telemonitoring and IoT systems, without the need for a constant doctor's examination. At the same time, this application, which includes smart home systems and control mechanisms that will positively affect the user's health, can be developed and turned into various compact structures.

Keywords: IoT, Healthcare Systems, Telemedicine, Smart Home

REFERENCE

[1] Padrós, M. R. C., Pastor, N., Paracolls, J. A., Peña, M. M., Pergolizzi, D., & Vergès, À. S. (2023). A smart system for remote monitoring of patients in palliative care (HumanITcare Platform): Mixed methods study. *JMIR Formative Research, 7*(1), e45654. https://doi.org/10.2196/45654

[2] Singh, L. K., Garg, H., & Khanna, M. (2023). An artificial intelligence-based smart system for early glaucoma recognition using OCT images. In *Research anthology on improving medical imaging techniques for analysis and intervention* (pp. 1424–1454). IGI Global. https://doi.org/10.4018/978-1-6684-7179-2.ch070

[3] Rejeb, A., Rejeb, K., Treiblmaier, H., Appolloni, A., Alghamdi, S., Alhasawi, Y., & Iranmanesh, M. (2023). The Internet of Things (IoT) in healthcare: Taking stock and moving forward. *Internet of Things, 22*, 100721. https://doi.org/10.1016/j.iot.2023.100721

[4] Balasundaram, A., Routray, S., Prabu, A. V., Krishnan, P., Malla, P. P., & Maiti, M. (2023). Internet of Things (IoT)-based smart healthcare system for efficient diagnostics of health parameters of patients in emergency care. *IEEE Internet of Things Journal, 10*(21), 18563-18570. https://doi.org/10.1109/JIOT.2023.3281457

[5] Akeeb, A. A., King, S. M., Olaku, O., & White, J. D. (2023). Communication between cancer patients and physicians about complementary and alternative medicine: A systematic review. *Journal of Integrative and Complementary Medicine, 29*(2), 80–98. https://doi.org/10.1089/jicm.2023.0011

[6] Got, Y., Borotikar, B., Sandillon-Garétier, C., Le Coat, A., Dulou, R., & Garétier, M. (2024). Use of instant messaging applications by general practitioners during overseas deployment: A survey of the French Military Health Service. *Military Medicine, 189*(7–8), e1745–e1752. https://doi.org/10.1093/milmed/usad137

[7] Gatto, M. C., Frisicale, E. M., Palopoli, P., Sapienza, M., Caroppo, E., Patrizi, C., & Damiani, G. (2024). The role of health institutions in training healthcare personnel for the digital transition: The international training program of the Order of Physicians and Dentists of Rome. *International Medical Education, 3*(1), 92–99. https://doi.org/10.3390/ime3010008

[8] Shoukat, M. U., Yan, L., Zhang, J., Cheng, Y., Raza, M. U., & Niaz, A. (2024). Smart home for enhanced healthcare: Exploring human machine interface oriented digital twin model. *Multimedia Tools and Applications, 83*(11), 31297–31315. https://doi.org/10.1007/s11042-023-15892-4

[9] Baudier, P., Ammi, C., & Deboeuf-Rouchon, M. (2020). Smart home: Highly-educated students' acceptance. *Technological Forecasting and Social Change, 153*, 119355. https://doi.org/10.1016/j.techfore.2020.119355.

[10] Hu, Y., Wang, B., Sun, Y., An, J., & Wang, Z. (2020). Graph-based semi-supervised learning for activity labeling in health smart home. *IEEE Access, 8*, 193655–193664. https://doi.org/10.1109/ACCESS.2020.3032370.

EFFECTS OF DIFFERENT MACHINE LEARNING ALGORITHMS ON BREAST CANCER DIAGNOSIS

Beyza Aslan[1*], Emine Uçar [1]

[1] *Management Information Systems, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: beyza.aslan@bakircay.edu.tr*

Introduction-Aim: Accordingly, to the World Health Organization (WHO), 2.3 million women were diagnosed with Breast Cancer (BC) in 2022, and BC was found to be the most often sorts of cancer in women in 157 out of 185 countries, causing 670,000 deaths in the same year [1]. The late discovery of the disease and the complexity of the methods are among the reasons for the decreased survival rate. Therefore, early detection of BC is vital to reduce the risk of growing cancer in other tissue cells and to provide appropriate treatment [2]. BC can be classified as either benign or malign. The first is categorized as cancerous cells, is toxic and malignant, and can spread to other organs. The second doesn't damage other organs and is non-invasive [3]. Machine learning (ML) algorithms have emerged as excellent methods for diagnosing BC by examining multiple data sources and separating among benign and malignant tumors. These algorithms have the potential to improve BC diagnosis efficiency and accuracy [4]. Lastly, ML methods have played a major role in the diagnosis and prognosis of BC by using grading techniques to recognize persons with BC, differentiate benign from malignant tumors, and predict prognosis. Correct classification can help doctors prescribe the most effective treatment regime [5]. In this research, ML algorithms used for BC diagnosis accuracy have been studied on the Wisconsin Breast Cancer Dataset (WBCD). Support Vector Machine (SVM), Decision Tree (DT), Random Forest (RF), K-Nearest Neighbors (KNN), Logistic Regression (LR) algorithms were used in the research and because of the research, it was concluded that all algorithms gave high accuracy rates. To guide decision-makers and healthcare professionals in this field and to contribute to the literature, this research compares various ML algorithms for BC diagnosis and prediction, and the results are discussed in the following sections

Materials-Methods: In this study, the WBCD was used as the data set. The WBCD consists of 569 patient data and 32 variables. Radius: the radius of the cell nucleus, Texture: the roughness of the cell nucleus surface, Perimeter: the perimeter of the cell nucleus, Area: the area of the cell nucleus, Smoothness: the smoothness of the edges of the cell nucleus, Compactness: a measure of how round the cell nucleus is. Concavity: the concavity of the edges of the cell nucleus, Concave Points: Count of concave points on the edges of the cell nucleus, Symmetry: symmetry of the cell nucleus, Fractal dimension: how complex and irregular the cell edges are (higher values are more complex structures). Within this study, lung cancer data were analyzed using python programming language and NumPy and Pandas libraries. ML algorithms such as SVM, DT, RF, KNN and LR were used in the study. Accuracy scores and comprehensive performance data were obtained for each ML algorithm in the dataset. Based on the calculated metrics (Accuracy, Precision, Sensitivity, F-1 Score, Confusion Matrix) was concluded which model is the best classifier for BC detection.

Below is some brief information about the algorithms and performance metrics used in the prediction of BC:

- *Support Vector Machine:* SVMs are powerful supervised ML models that are widely used for classification and regression missions [6] The primary goal of SVMs is to predict and

classify data into different categories [7]. SVMs work by finding the optimal hyperplane that best separates different classes in the feature space [8].

- *Decision Tree:* DTs are a simple and straightforward model used in ML. DTs create final decisions or classes by branching the data set according to different features [9]. DTs are well-known for being accurate, efficient, and interpretable, especially when utilized in tree ensembles [10].
- *Random Forest:* It is a ML algorithm that is strong and adaptable. A collection of several DTs is called RF. By using this technique, every DT can function independently and ultimately reach a decision based on the vote of the majority [11].
- *K-Nearest Neighbors:* KNN is an effective and basic ML algorithm. KNN classifies a data point by looking at the classes of its nearest neighbors [12]. The KNN algorithm determines the nearest neighbors of a test object in a feature space using a given value k, which indicates the number of neighbors to be considered [13].
- *Logistic Regression:* LR is a simple and effective classification technique used in ML. LR is used to estimate the probability that a data point is owned by a particular class [14]. Its main goal is to predict the probability that a data point will be owned by a particular class. This approach is especially used in issues involving binary classification [15].

Performance metrics are implemented in ML to appraise the effectiveness of methods on different tasks and are very important. These metrics provide an overview of how well a method performs based on different criteria such as accuracy, precision, recall, F1 score, and more [16].

Results: In this article, different ML algorithms used for BC prediction were examined, and their effectiveness was compared. The main objective of the study was to determine the optimal algorithm that provides the highest accuracy and efficiency in BC diagnosis. The models used include SVM, DT, RF, KNN, and LR. These models were evaluated with different metrics on test and validation datasets.

The most successful results in the study were shown by the SVM (Linear) model. It attracted attention with 96.49% validation accuracy, 98.25% test accuracy, and a fast-training time of 0.0053 seconds. The average cross-validation accuracy was 94.13% and the standard deviation value was 0.0294, which means consistent and high performance. The SVM (RBF) model also showed similar success, achieving 94.74% validation accuracy and 96.49% test accuracy. The training time was very short at 0.0008 seconds. These results show that the SVM (RBF) model is reliable and consistent. On the other hand, the SVM (Sigmoid) model performed lower than other SVM models with 87.72% validation accuracy and 95.61% test accuracy.

Discussion-Conclusion: When the results were evaluated, it was observed that the SVM (Linear) and KNN models have the highest accuracy and f1 score values. In particular, the SVM (Linear) model showed very successful performance with its high accuracy rate and short training time. The KNN model also had similar accuracy rates with 97.37% test accuracy and drew attention with its short training time. Random Forest and Logistic Regression models also exhibited strong performance, but their training times were slightly longer than other models. Decision Trees and SVM (Sigmoid) models, on the other hand, exhibited lower performance than other models.

When studies conducted with similar data sets in the literature are examined, it is observed that SVM generally performs better than other algorithms, as in this study. In the study of [17], SVM, LR, KNN, DT, NB, and RF algorithms were compared, and it was found that SVM's accuracy rate was higher than other algorithms. Again, [18] compared SVM, RF, and KNN algorithms and found that SVM had the highest accuracy. Similarly, in the study of [19], it was found that SVM performed better than other algorithms.

As a result, it was observed that all models used in this study had high accuracy rates. It is recommended that future studies compare the same algorithms with more up-to-date BC data.

Keywords: breast cancer, diagnosis, machine learning

REFERENCE

[1] World Health Organization. "Breast cancer". Date of access: 03.04.2024. https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[2] D. Bardou, K. Zhang, and S. M. Ahmad, "Classification of Breast Cancer Based on Histology Images Using Convolutional Neural Networks," IEEE Access, vol. 6, pp. 24680–24693, 2018.

[3] Abdulla, S. H., Sagheer, A. M., & Veisi, H. (2021). Breast cancer classification using machine learning techniques: A review. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(14), 1970-1979.

[4] Dou, Y. and Meng, W. (2021). An optimization algorithm for computer-aided diagnosis of breast cancer based on support vector machine. Frontiers in Bioengineering and Biotechnology, 9. https://doi.org/10.3389/fbioe.2021.698390

[5] Yue, W., Wang, Z., Chen, H., Payne, A., & Liu, X. (2018). Machine learning with applications in breast cancer diagnosis and prognosis. Designs, 2(2), 13.

[6] Eitle, V. and Buxmann, P. (2019). Business analytics for sales pipeline management in the software industry: a machine learning perspective. Proceedings of the Annual Hawaii International Conference on System Sciences. https://doi.org/10.24251/hicss.2019.125

[7] Guenther, N. and Schonlau, M. (2016). Support vector machines. The Stata Journal: Promoting Communications on Statistics and Stata, 16(4), 917-937. https://doi.org/10.1177/1536867x1601600407

[8] Zhang, R., Wang, W., Ma, Y., & Men, C. (2009). Least square transduction support vector machine. Neural Processing Letters, 29(2), 133-142. https://doi.org/10.1007/s11063-009-9099-z

[9] Quinlan, J. R. (1986). Induction of decision trees. Machine Learning, 1(1), 81-106.

[10] Geurts, P., Irrthum, A., & Wehenkel, L. (2009). Supervised learning with decision tree-based methods in computational and systems biology. Molecular BioSystems, 5(12), 1593. https://doi.org/10.1039/b907946g

[11] Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

[12] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1), 21-27.

[13] Tajmouati, S., Wahbi, B. E., Bedoui, A., Abarda, A., & Dakkoun, M. (2021). Applying k-nearest neighbors to time series forecasting: two new approaches.. https://doi.org/10.48550/arxiv.2103.14200

[14] Cox, D. R. (1958). The regression analysis of binary sequences. Journal of the Royal Statistical Society: Series B (Methodological), 20(2), 215-232.

[15] Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied Logistic Regression. Wiley.

[16] Shama, A., Hossain, M. B., Adhikary, A., Uddin, K. M. A., & Hossain, M. A. (2022). Prediction of hypothyroidism and hyperthyroidism using machine learning algorithms.. https://doi.org/10.21203/rs.3.rs-1486798/v1

[17] Ara, S., Das, A., & Dey, A. (2021, April). Malignant and benign breast cancer classification using machine learning algorithms. In 2021 International Conference on Artificial Intelligence (ICAI) (pp. 97-101). IEEE.

[18] Yadav, AR (2023, Aralık). Makine Öğrenimi Tekniklerini Kullanarak Meme Kanseri İçin Erken Tahmin Sisteminin Geliştirilmesi. 2023 Uluslararası Yeni Nesil Elektronik Konferansı'nda (NEleX) (s. 1-6). IEEE.

[19] Kumar, A., Saini, R., & Kumar, R. (2024). A Comparative Analysis of Machine Learning Algorithms for Breast Cancer Detection and Identification of Key Predictive Features. Traitement Du Signal, 41(1). https://doi.org/10.18280/ts.410110

THE USE OF ARTIFICIAL INTELLIGENCE IN TRIAGE PROCESSES: CLINICAL AND ECONOMIC IMPACTS

Simge Kamalı[1*], Cansu Yıldırım[2], Şüheda Baran[3]

[1] Department of Healthcare Management, İzmir Bakırçay University, İzmir, Türkiye

[2] Department of Speech and Language Therapy, İzmir Bakırçay University, İzmir, Türkiye

[3] Department of Audiology, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: simge.kamali@bakircay.edu.tr*

Introduction: The administrative and medical processes of healthcare organizations are rapidly transforming using artificial intelligence (AI) systems [1]. The increase in population has raised the need for rapid delivery of healthcare services; in this context, it has been stated that innovative artificial intelligence solutions should be applied in the healthcare system to ensure effectiveness and efficiency without increasing costs [2]. AI has a high potential to improve healthcare, including medical triage optimization [3], [4]. Triage is a medical decision-making process in which patients are prioritized according to the urgency of their condition in order to make the most efficient use of limited healthcare resources [5]. AI-supported systems developed for patient triage, especially in emergency departments, can offer significant advantages by standing out with their ability to make accurate and fast decisions [6]. A meta-analysis by Adebayo et al. found that AI, machine learning and deep learning models exhibited significantly greater predictive accuracy than traditional trauma triage methods for most mortality, hospitalisation and critical care admission outcomes [7]. This study aims to examine the clinical and economic effects of AI in the triage process and to reveal its possible reflections on healthcare services.

Materials-Methods: In this study, which was performed using the literature review method, the economic and clinical contributions of artificial intelligence applications developed for patient triage to healthcare services were examined. In the study, a literature review was conducted using leading databases such as PubMed, Google Scholar, and Web of Science. The keywords 'triage', 'artificial intelligence', 'healthcare management', 'speech-language pathology', and 'audiology' were used as search terms.

Results: When the existing findings in the literature on the clinical and economic effects of artificial intelligence-supported triage systems are examined, it is stated that the AIMS-OD artificial intelligence system accelerates the diagnosis and treatment processes by screening oropharyngeal dysphagia in elderly patients with 84% AUCROC accuracy and improves the quality of care by reducing costs [8]. Another study examines the ability of AI to diagnose stroke at an early stage using audiological assessment results in acute vestibular syndrome (AVS) patients, and the results obtained highlight the high sensitivity and specificity of AI in this area [9]. In a study conducted in Singapore, the use of semi-automated deep learning models in diabetic retinopathy screening resulted in a 20% saving in annual screening costs compared to human assessment [10]. Similarly, AI-based analysis of chest radiography (CXR) for tuberculosis triage has been estimated to be cost-effective, with the potential to increase TB detections by 0.5% to 1.2% while decreasing costs by 19% compared to standard microbiological testing [11]. Additionally, a significant reduction in the length of stay of patients diagnosed with intracranial hemorrhage (ICH) or pulmonary embolism (PE) has also been observed thanks to the integration of artificial intelligence (AI)-assisted triage and prioritization software into radiological workflows [12]. In this context, the findings of the

studies examined show that AI-supported triage systems increase the efficiency of healthcare services by accelerating diagnosis and treatment processes, reducing hospitalization times, reducing costs, and enabling more efficient use of resources.

Discussion-Conclusion: Studies in the literature suggest that the use of AI-supported triage systems helps to increase the efficiency of healthcare services and more effective use of resources by significantly shortening patient hospitalization times [13], [14]. It also contributes to rapid and accurate assessment and prioritization by preventing long waiting lists of clients [15]. AI-supported automatic triage systems to be developed in diseases such as long-term dizziness, hearing loss, voice disorder, and oropharyngeal dysphagia can improve clinical decision-making processes by quickly analyzing patient data and provide significant advantages in terms of both time and cost compared to manual methods [8], [16], [17], [18]. In conclusion, the integration of AI-supported systems can offer significant benefits in terms of both cost savings and improving the quality of patient care in healthcare. To fully utilize the potential of AI in triage, investments should be made in staff training, ethical standards should be maintained, collaboration should be encouraged, and innovation should be embraced [19]. This will improve the efficiency of healthcare services and make patient care more accessible and sustainable.

Keywords: triage, artificial intelligence, healthcare management, speech-language pathology, audiology.

REFERENCE

[1] O. Ali, W. Abdelbaki, A. Shrestha, E. Elbasi, M. A. A. Alryalat, and Y. K. Dwivedi, “A systematic literature review of artificial intelligence in the healthcare sector: Benefits, challenges, methodologies, and functionalities,” *Journal of Innovation & Knowledge*, vol. 8, 1, Mar 2023.

[2] L. S. Y. Lin, M. R. Mahoney, and C. A. Sinsky, “Ten ways artificial intelligence will transform primary care,” *Journal of General Internal Medicine*, vol. 34, pp. 1626-1630, May 2019.

[3] T. Davenport and R. Kalakota, “The potential for artificial intelligence in healthcare,” *Future Healthcare Journal*, vol. 6, 2, pp. 94-98, Jun 2019.

[4] M. D. Christian, “Triage,” *Critical Care Clinics*, vol. 35, 4, pp. 575-589, Oct. 2019.

[5] S. Tyler et al., “Use of artificial intelligence in triage in hospital emergency departments: A scoping review,” *Cureus*, vol. 16, 5, e59906, May 2024.

[6] K. J. W. Tang, C. K. E. Ang, T. Constantinides, V. Rajinikanth, U. R. Acharya, and K. H. Cheong, “Artificial intelligence and machine learning in emergency medicine,” *Biocybernetics and Biomedical Engineering*, vol. 41, pp. 156-172, Dec 2021.

[7] O. Adebayo, Z. A. Bhuiyan, and Z. Ahmed, "Exploring the effectiveness of artificial intelligence, machine learning and deep learning in trauma triage: A systematic review and meta-analysis," *Digital Health*, vol. 9, Oct. 2023.

[8] A. Martin-Martinez et al., “A systematic and universal artificial intelligence screening method for oropharyngeal dysphagia: improving diagnosis through risk management,” *Dysphagia*, vol. 38, 4, pp. 1224-1237, Dec 2023.

[9] A. Korda et al., “Artificial intelligence for early stroke diagnosis in acute vestibular syndrome,” *Frontier in Neurology*, vol. 13, Sep 2022.

[10] Y. Xie et al., “Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: An economic analysis modelling study,” *Lancet Digital Health*, vol. 2, 5, pp. e240-e249, Oct 2020.

[11] N. P. Nsengiyumva et al., “Triage of persons with tuberculosis symptoms using artificial intelligence–based chest radiograph interpretation: A cost-effectiveness analysis,” *Open Forum Infectious Diseases*, vol. 8, 12, ofab567, Dec 2021.

[12] M. Petry, C. Lansky, Y. Chodakiewitz, M. Maya, and B. Pressman, “Decreased hospital length of stay for ICH and PE after adoption of an artificial intelligence-augmented radiological worklist triage system,” *Radiology Research and Practice*, Aug 2022.

[13] D. C. Classen, C. Longhurst, and E. J. Thomas, “Bending the patient safety curve: How much can AI help?” *NPJ Digital Medicine*, vol. 6, 2, Jan 2023.

[14] C. Pickstone, “Triage in speech and language therapy,” *Prioritising Child Health*, pp. 49-54, 2007.

[15] A. Abdel-Hafez et al., “Artificial intelligence in medical referrals triage based on Clinical Prioritization Criteria,” *Frontiers in Digital Health*, vol. 5, 1192975, Oct 2023.

[16] S. Romero-Brufau et al., “Development of an Automated Triage System for Longstanding Dizzy Patients Using Artificial Intelligence,” *OTO Open*, vol. 8, 3, pp. e7006, Sep 2024.

[17] M. U. Rehman, A. Shafique, S. S. Jamal, Y. Gheraibia, and A. B. Usman, “Voice disorder detection using machine learning algorithms: An application in speech and language pathology,” *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108047, Jul 2024.

[18] S. V. Bhagat and D. Kanyal, “Navigating the future: The transformative impact of artificial intelligence on hospital management - A comprehensive review,” *Cureus*, vol. 16, 2, e54518, Feb 2024.

THE ROLE OF MODERN TECHNOLOGY IN DIABETES DISEASE DIAGNOSIS APPLICATION OF MACHINE LEARNING BASED PREDICTIVE ANALYTICS

Ayşenur Karakaya[1], Burak Keskin[1*], Abdulkadir Hızıroğlu[1]

[1] *Department of Management Information Systems, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: burak.keskin@bakircay.edu.tr*

Introduction-Aim: The global prevalence of diabetes is a growing health crisis, impacting over 537 million individuals as of 2021, with numbers projected to rise to 783 million by 2045 [1]. This chronic disease places a substantial financial burden on healthcare systems due to the high costs associated with its treatment and management, which includes managing severe complications such as cardiovascular disease, neuropathy, kidney failure, and retinopathy [2]. These complications not only reduce patients' quality of life but also lead to increased mortality rates. Given the complexity and progression of diabetes, early detection and timely management are critical in preventing these adverse outcomes [3]. This study aims to investigate the role of modern technologies, specifically artificial intelligence (AI) and machine learning (ML), in the early diagnosis, management, and potential prevention of diabetes. By leveraging predictive analytics, we aim to accurately identify individuals at high risk for diabetes, thereby enabling timely intervention and reducing the likelihood of serious complications.

Materials-Methods: To achieve this, a dataset comprising 100,000 anonymized health records, which includes variables such as age, blood glucose levels, BMI, family history, and lifestyle factors, was utilized [4]. Three machine learning algorithms—Random Forest, K-Nearest Neighbours (KNN), and Logistic Regression—were selected for this study based on their strengths in classification and predictive capabilities [5]. The models were trained on 80% of the dataset, with the remaining 20% used for testing. To ensure robust evaluation, the models were assessed using accuracy, sensitivity, specificity, and the area under the receiver operating characteristic curve (AUC), providing insights into each model's predictive reliability. Cross-validation techniques were applied to fine-tune the models and minimize potential overfitting, ensuring that results are generalizable across different datasets.

Results: The analysis revealed that the Random Forest model outperformed other models, achieving the highest accuracy and AUC, suggesting it as the most effective model for diabetes risk prediction. The Random Forest model's AUC was recorded at 0.962, indicating strong discriminatory power in distinguishing between high-risk and low-risk individuals. Logistic Regression demonstrated balanced performance, with an AUC of 0.937 and respectable sensitivity and specificity scores, making it valuable for broader clinical applications. However, KNN showed comparatively lower performance, achieving an AUC of 0.752, and was sensitive to variations in the data, indicating its limited suitability in this context. The findings suggest that AI-enhanced models can significantly improve the predictive accuracy of diabetes diagnosis, offering healthcare providers valuable tools for early intervention.

Discussion-Conclusion: This study highlights the transformative potential of integrating AI and ML in diabetes care, illustrating how predictive analytics can empower healthcare providers to make proactive, data-driven decisions. With early risk detection, patients can benefit from tailored preventive strategies and lifestyle modifications, potentially delaying or

preventing the onset of diabetes-related complications. For future research, it would be advantageous to incorporate a wider range of data types, such as genetic information, social determinants of health, and additional lifestyle factors, to further enhance the model's predictive accuracy and equity. Additionally, exploring ensemble and hybrid models that combine various ML techniques could address the limitations observed in individual algorithms. Implementing these AI-driven predictive models into healthcare systems not only promises better patient outcomes but also has the potential to reduce healthcare costs by lowering hospitalization rates and the need for intensive care.

Keywords: Diabetes, Artificial Intelligence, Machine Learning, Predictive Analytics, Healthcare Systems

REFERENCE

[1] International Diabetes Federation, IDF Diabetes Atlas, 10th ed. Brussels, Belgium: International Diabetes Federation, 2021. [Online]. Available: https://diabetesatlas.org/atlas/tenth-edition/.

[2] American Diabetes Association, "Economic costs of diabetes in the U.S. in 2017," Diabetes Care, vol. 41, no. 5, pp. 917-928, 2018. [Online]. Available: https://diabetesjournals.org/care/article/41/5/917/36518/Economic-Costs-of-Diabetes-in-the-U-S-in-2017.

[3] World Health Organization, Global Report on Diabetes. Geneva, Switzerland: World Health Organization, 2016. [Online]. Available: https://www.who.int/publications/i/item/9789241565257.

[4] Kaggle, "Diabetes prediction dataset," [Online]. Available: https://www.kaggle.com/datasets/iammustafatz/diabetes-prediction-dataset.

[5] M. Badawy, N. Ramadan, and H. A. Hefny, "Healthcare predictive analytics using machine learning and deep learning techniques: a survey," Journal of Electrical Systems and Information Technology, vol. 10, no. 40, 2023. [Online]. Available: https://doi.org/10.1186/s43067-023-00108-y.

OVERVIEW OF VIRTUAL REALITY IN THE TREATMENT OF PSYCHIATRIC DISORDERS

Zeren Keskin[1*], Yeşim Aygül[1], Onur Uğurlu[1]

[1] *Faculty of Engineering and Architecture, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: zerenkeskin01@gmail.com*

Introduction-Aim: In recent years, Virtual Reality (VR) technology has gained attention as an innovative tool for diagnosing psychiatric disorders and developing treatments. By creating immersive environments, VR contributes to assessing psychiatric disorders by enabling patients to engage in controlled and interactive scenarios that mimic real-world experiences. This study aims to provide an overview of VR's effectiveness in treating psychiatric disorders such as autism [1], schizophrenia [2], social anxiety [3], post-traumatic stress disorder (PTSD) [4], and psychosis [5]. Recent studies emphasize VR's potential to enhance traditional therapeutic applications and improve treatment outcomes for individuals facing these complex conditions.

Materials-Methods: This study examines research conducted between 2020 and 2024, highlighting the transformative impact of VR interventions in enhancing social skills and facilitating cognitive rehabilitation. In VR health applications, essential tools such as head-mounted displays and motion-tracking sensors are frequently used to increase interaction and immersion. In the reviewed studies, head-mounted displays, combined with pre-developed software from external companies and solutions created by academic teams, facilitate patient engagement and deliver an immersive experience. For individuals with autism, VR provides a controlled environment where social interactions can be practiced effectively, addressing challenges related to social presence and gaze behavior. In schizophrenia, VR-supported therapies have been adapted to alleviate negative symptoms, thereby increasing patient engagement and promoting active involvement in recovery processes. For social anxiety disorder, VR has been seamlessly integrated with acceptance and commitment therapy, enabling participants to confront their fears within a safe yet stimulating framework. For individuals coping with PTSD, VR exposure therapy offers a gradual and supportive approach to addressing traumatic memories, fostering core emotional processing. In psychosis, automated VR therapy has been used to help reduce agoraphobic avoidance and distress in patients.

Results: The literature suggests that VR interventions provide significant therapeutic benefits across various psychiatric disorders. Simmons et al. (2023) [1] found that VR-based social cognition training considerably improved social interaction skills in children with high-functioning autism, emphasizing VR's promise as a tool for autism treatment. Similarly, Cella et al. (2022) [2] explored the feasibility and acceptability of VR-based social skills training for patients with schizophrenia, reporting noticeable improvements in social functioning and high levels of patient engagement. For social anxiety disorder, Gorinelli et al. (2023) [3] highlighted the effectiveness of immersive VR environments, where patients experienced gradual reductions in anxiety through controlled simulations of real-world social interactions. Hannigan et al. (2023) [4] assessed a novel VR therapy specifically designed for military veterans with PTSD. They utilized a mixed-methods analysis to identify critical factors influencing treatment outcomes, which demonstrates the VR's potential in addressing trauma-related disorders among specialized populations. Lastly, Smith et al. (2023) [5]

reported positive clinical outcomes from a randomized controlled trial examining VR therapy for persistent auditory hallucinations, which is a significant challenge in psychosis treatment.

Discussion-Conclusion: Integrating VR into psychiatric treatment offers exciting possibilities for enhancing therapeutic outcomes across a range of mental health conditions. The reviewed studies indicate that VR creates controlled environments where patients can engage safely and deeply with their challenges. As VR technology advances, its inclusion in standard clinical practices should be considered to benefit a broader range of individuals.

Keywords: Virtual Reality, Mental Health, Psychiatric Disorder.

REFERENCES

[1] T. Simmons, J. Snider, and L. Chukoskie, “The effects of social presence on gaze, movement, arousal and blink rate in autism: A cooperative virtual reality game-based approach,” IEEE Trans. Games, 2024, doi: 10.1109/TG.2024.3410163.

[2] M. Cella, P. Tomlin, D. Robotham, P. Green, H. Griffiths, D. Stahl, and L. Valmaggia, “Virtual reality therapy for the negative symptoms of schizophrenia (V-NeST): a pilot randomised feasibility trial,” Schizophr. Res., vol. 248, pp. 50–57, 2022. doi: 10.1016/j.schres.2022.08.008.

[3] S. Gorinelli, A. Gallego, P. Lappalainen, and R. Lappalainen, “Virtual reality acceptance and commitment therapy intervention for social and public speaking anxiety: A randomized controlled trial,” J. Context. Behav. Sci., vol. 28, pp. 289–299, 2023. doi: 10.1016/j.jcbs.2023.08.003.

[4] B. Hannigan, R. van Deursen, K. Barawi, N. Kitchiner, and J. I. Bisson, "Factors associated with the outcomes of a novel virtual reality therapy for military veterans with PTSD: Theory development using a mixed methods analysis," *PLOS One*, vol. 18, no. 5, p. e0285763, 2023. doi: 10.1371/journal.pone.0285763.

[5] L. C. Smith, L. Mariegaard, D. L. Vernal et al., “The Challenge trial: The effects of a virtual reality-assisted exposure therapy for persistent auditory hallucinations versus supportive counselling in people with psychosis: Study protocol for a randomized clinical trial,” Trials, vol. 23, p. 773, 2022. doi: 10.1186/s13063-022-06683-1.

USING GPT-4O FOR NAMED ENTITY RECOGNITION IN TURKISH ELECTRONIC HEALTH RECORDS

Adem, Altungeyik[1], Şebnem, Bora[1*], Oğuz, Dikenelli[1]*

[1] *Computer Engineering, Ege University, İzmir , Türkiye*

* *Corresponding author: adem.altungeyik@ege.edu.tr*

Introduction-Aim: The importance of extracting information from texts, like Electronic Health Records (EHR), in the healthcare sector has been shown by developments in natural language processing. Named Entity Recognition (NER) is a critical step in this context, transforming unstructured text data into structured formats that allow healthcare providers to efficiently interpret and manage patient information. It categorizes various elements like medical conditions, treatments and diagnostic tests mentioned in clinical records. This approach serves as a fundamental for extracting valuable information from texts. Research and technology in this field have found English EHR datasets beneficial resources for the NER process. The lack of labelled datasets in Turkish presents considerable difficulties [1]. We overcame this issue by applying the Generative Pre-Trained Transformer (GPT) to the i2b2-2010 clinical concept extraction dataset, which was translated into Turkish by medical professionals.

Materials-Methods: We used the GPT-4o model to perform Named Entity Recognition (NER) on the i2b2 Turkish translation dataset, from 2010 research on healthcare topics like procedures in clinics or hospitals. We applied methods brought by developments in prompt engineering to improve the accuracy and efficiency [2]–[3]. Our instructions in In-Context Learning included simple task descriptions, annotation rules and instructions from the guidelines to aid the model's learning process effectively. To assist the model in identifying the relevant items in texts, we employed Chain-of-Thought prompting. With a small sample size, the model has also observed linguistic variations, such as suffixes and sentence patterns.

Results: Our strategic approach to prompt engineering and adapting our prompt for the most helpful contribution to the model significantly improved GPT-4o's entity recognition capabilities. The model achieved a micro-F1 score of 0.73 when used solely in In-Context Learning to identify entities and their respective categories in Turkish medical texts. The micro-F1 score improved to 0.81 when Chain-of-Thought Prompting was used. Among the medical entities the model could identify are multi-word phrases and context-specific terms from Turkish medical literature. The success of using Chain-of-Thought prompting in combination with In-Context Learning highlighted the model's reliability. These results show how successful entity recognition is possible even in low-resource domains like Turkish medicine, where annotated datasets aren't abundant.

Discussion-Conclusion: This research introduces a way of using GPT models to identify named entities in Turkish medical texts. The successful results suggest that the help of well-built prompts and Chain-of-Thought prompting techniques in GPT-4o can effectively solve these problems arising from the limited availability of annotated medical data in the Turkish language. Although the model's performance is not currently better than the specialized biomedical models trained on large datasets, it exhibits great potential for use in languages that lack sufficient resources. Translation and annotation of the dataset by experts guarantee the clinical significance and precision of our results. We aim to upgrade the dataset by including annotations and improving our approaches, while investigating retrieval-based techniques [4] to improve model performance. This research aims to set a fundamental for testing clinical NLP tools in Turkish texts and potentially other low-resource languages. That will help to improve information extraction and analysis in these contexts.

Keywords: GPT, Named Entity Recognition, Turkish Electronic Health Records, Chain of Thought, Prompt Engineering

REFERENCE

[1] R. Yeniterzi, "Exploiting Morphology in Turkish Named Entity Recognition System," in Proceedings of the ACL 2011 Student Session, Portland, OR, USA, Jun. 2011, pp. 105–110.

[2] J. Wei et al., "Chain-of-Thought Prompting Elicits Reasoning in Large Language Models," arXiv preprint arXiv:2201.11903, 2023. [Online]. Available: https://arxiv.org/abs/2201.11903

[3] Y. Hu et al., "Improving Large Language Models for Clinical Named Entity Recognition via Prompt Engineering," arXiv preprint arXiv:2303.16416, 2024. [Online]. Available: https://arxiv.org/abs/2303.16416

[4] M. Li, H. Zhou, H. Yang, and R. Zhang, "RT: A Retrieving and Chain-of-Thought Framework for Few-Shot Medical Named Entity Recognition," Journal of the American Medical Informatics Association, vol. 31, no. 9, pp. 1929–1938, Sep. 2024, doi: 10.1093/jamia/ocae095.

[5] Y. Wang et al., "Clinical information extraction applications: A literature review," Journal of Biomedical Informatics, vol. 77, pp. 34–49, Jan. 2018, doi: 10.1016/j.jbi.2017.11.011.

THE USE OF ARTIFICIAL INTELLIGENCE IN THE DETECTION OF FACIAL PROSODY: PILOT STUDY

Mümüne Merve Parlak[1], Mete Orçun Bayrakdar[1*], Seren Düzenli Öztürk[2], Cansu Yıldırım[2], Merve Sapmaz Atalar[3], Kubilay Muhammed Sünnetçi[4], Ahmet Alkan[5], Görsev Yener[6,7]

[1]*Department of Speech and Language Therapy, Ankara Yıldırım Beyazıt University, Ankara, Türkiye*

[2]*Department of Speech and Language Therapy, İzmir Bakırçay University, İzmir, Türkiye*

[3]*Department of Speech and Language Therapy, University of Health Sciences, Istanbul, Türkiye*

[4]*Department of Electrical and Electronics Engineering, Osmaniye Korkut Ata University, Osmaniye, Türkiye*

[5]*Department of Electrical and Electronics Engineering, Kahramanmaraş Sütçü İmam University, Kahramanmaraş, Türkiye*

[6]*Department of Neurology, Faculty of Medicine, Dokuz Eylül University, İzmir, Türkiye*

[7]*Department of Neurology, Faculty of Medicine, İzmir Biomedicine and Genome Center, İzmir, Türkiye*

** Corresponding author:* ***morcunnn@gmail.com***

Introduction-Aim: Artificial intelligence has been widely used in emotion analysis through facial expressions, and significant progress has been made in this field in recent years [1]. Research focuses on developing more efficient and accurate models for real-time emotion recognition in different environments [2]. However, to the best of our knowledge, the percentage of facial prosody production in different tasks in Turkish society is unknown [3]. In addition, facial prosody production percentages based on gender and in different tasks have not been investigated [4, 5]. In this study, it is aimed to analyze facial prosodic components, compare facial prosody productions based on gender, and comparatively examine the expression levels of emotions in different tasks with the SCRFD [7] and HSEmotion [8] artificial intelligence models, which are designed as a 2-stage structure to detect emotional expressions and changes and prepared using the WIDER FACE [6] dataset.

Materials-Methods: The study included 22 participants (11 males, mean age 22.27 ± 2.68 years; 11 females, mean age 21.09 ± 0.83 years). In this study, 8 emotions were analyzed: neutral, sadness, happiness, anger, disgust, contempt, fear, and surprise. The video recordings of these expressions were analyzed using SCRFD and HSEmotion artificial intelligence tools, and the percentage data were calculated by the artificial intelligence tools according to the 8 emotions. In order to determine the facial prosody production percentages of the participants in different tasks, different tasks were asked of the participants: spontaneous speech, emotional and monotonous reading of the 'Pinocchio' piece, reading the National Anthem as an anthem and poem.

Results: In this study, a total of 110 video recordings obtained from 22 participants were analyzed. Participants expressed the emotion of disgust at the highest rate in all tasks, with a mean of 25.37 ± 25.61%. In male participants, disgust was the most common emotion, with a mean of 29.97 ± 30.96%, while neutral emotion was the most dominant in female participants, with a mean of 26.26 ± 19.10%. When compared between genders, male participants showed significant differences between fear and neutral emotions in all tasks. In female participants, a significant difference was found between fear and neutral emotions in all tasks except spontaneous speech. Males also showed significant differences between fear-neutral and angry-neutral emotions in all tasks ($p < 0.05$), whereas female participants showed no significant differences in all tasks ($p \geq 0.05$).

Discussion-Conclusion: There are studies in the literature in which facial prosody is identified using artificial intelligence; however, there is no study on the Turkish population. In addition, there is no study focusing on facial prosody analysis using different tasks in the literature. This study is the first of its kind in the national literature in these aspects. In the literature, it has been reported that fear emotion is commonly expressed, women exhibit happiness and sadness emotions more prominently, and men are more prominent in anger emotion [9]. However, in this study, the findings obtained in the Turkish population differ from this picture. It was observed that the most frequently expressed emotion among male participants was disgust, while the most frequently expressed emotion among female participants was neutral emotion. This suggests that, in addition to gender, cultural and social factors may play a determining role in emotional expression. Task-based analyses also supported gender differences. While these results emphasize the effect of gender and task types on emotional expression, cultural and individual differences are also important factors to be considered. The study was conducted with a limited number of participants; however, these are preliminary results and have the potential to inspire further research comparing the facial prosodic production of different cultural groups.

Keywords: artificial intelligence, facial prosody, prosody, speech and language therapy

REFERENCE

[1] J. F. Cohn *et al.*, "Detecting depression from facial actions and vocal prosody," *2009 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops,* pp. 1-7, 2009.

[2] N. V. Babu, E. Grace, M. Kanaga, T. Mandl, S. J. Modha, and E. Grace, "Sentiment Analysis in Social Media Data for Depression Detection Using Artificial Intelligence: A Review," *Sn Computer Science,* vol. 3, 2021.

[3] F. Burkhardt, N. Audibert, L. Malatesta, O. Türk, L. M. Arslan, and V. Aubergé, "Emotional prosody - does culture make a difference?," *Speech Prosody 2006,* 2006.

[4] L. Forni-Santos and F. d. L. Osório, "Influence of gender in the recognition of basic facial expressions: A critical literature review," *World journal of psychiatry,* vol. 5 3, pp. 342-51, 2015.

[5] M. Fitzsimons, N. Sheahan, and H. Staunton, "Gender and the integration of acoustic dimensions of prosody: Implications for clinical studies," *Brain and language,* vol. 78, no. 1, pp. 94-108, 2001.

[6] S. Yang, P. Luo, C.-C. Loy, and X. Tang, "Wider face: A face detection benchmark," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 5525-5533.

[7] J. Guo, J. Deng, A. Lattas, and S. Zafeiriou, "Sample and computation redistribution for efficient face detection," *arXiv preprint arXiv:2105.04714,* 2021.

[8] A. Savchenko, "Facial expression recognition with adaptive frame rate based on multiple testing correction," in *International Conference on Machine Learning*, 2023: PMLR, pp. 30119-30129.

[9] Y. Lin, H. Ding, and Y. Zhang, "Gender differences in identifying facial, prosodic, and semantic emotions show category-and channel-specific effects mediated by encoder's gender," *Journal*

MACHINE LEARNING-BASED GENE SIGNATURE DETECTION FOR ACUTE MYELOID LEUKEMIA PREDICTION

Selahattin A. Uysal[1*], Burçin Kaymaz[1]

[1] Medical Biology, Ege University Graduate School of Medicine, İzmir, Türkiye.

** Corresponding author: burcin.tezcanli@ege.edu.tr*

Introduction-Aim: Acute Myeloid Leukemia (AML) is a highly aggressive hematologic malignancy characterized by abnormal growth and differentiation of myeloid cells, leading to impaired hematopoiesis and adverse clinical outcomes [1], [2]. Traditional diagnostic methods often lack precision, highlighting the need for advanced approaches to accurately identify AML cases, monitor progression, and guide therapeutic strategies [3], [4], [5]. The rapidly growing volume of gene expression studies provides valuable big data that can yield critical clinical insights through robust computational methods [6], [7]. Machine learning has become instrumental in leveraging high-throughput transcriptomic data to uncover specific molecular biomarkers and complex mechanisms that contribute to disease pathology [8], [9]. This study aimed to leverage machine learning to identify a minimal yet highly informative gene set that can accurately distinguish AML samples from normal controls. Additionally, the model is designed as a foundational approach, with potential applicability to other diseases. By creating predictive models that deepen our understanding of disease mechanisms, this study aspires to contribute to more precise diagnostics and informed clinical decision-making across a broad range of conditions.

Materials-Methods: Transcriptomic data for this study were obtained from the TARGET-AML [10] project available on the GDC portal (https://portal.gdc.cancer.gov/), focusing on gene expression quantification (FPKM) data. The data was acquired using the TCGAbiolinks package in the R environment [11], [12]. Data preprocessing was performed in Python, where initial filtering excluded low-expression genes to avoid noise. A machine learning pipeline was developed to evaluate RandomForest, XGBoost, and LightGBM models, each using a carefully selected subset of genes. Gene selection was conducted with SelectKBest, optimized with chi-square scoring, to retain only the top 15 genes significantly associated with AML, including mitochondrial genes (e.g., MT-CO2, MT-CYB) and hemoglobin-related genes (e.g., HBB). After balancing with SMOTE and standardizing, the dataset was split into training and test sets. Models were optimized using grid search, with hyperparameters tuned to maximize accuracy. Training included parameters like learning_rate, max_depth, and n_estimators for XGBoost and LightGBM, while RandomForest parameters included max_depth, min_samples_split, and n_estimators. Model accuracy and robustness were evaluated using training and test accuracy metrics, followed by detailed performance assessment across different sample sizes.

Results: All models performed well in identifying AML cases based on the selected gene set, which included LTF, MT-CO2, MT-CYB, MT-ND2, MT-CO1, MT-ND4, MT-ND1, MT-ATP6, MT-CO3, RN7SKP80, MT-RNR2, MT-RNR1, DEFA3, HBB and RN7SL1. These genes were selected using the chi-square (χ^2) test as implemented in the SelectKBest method, which identifies genes with the strongest association to AML by evaluating their statistical significance in relation to the target class labels.

Specifically:

RandomForest: Achieved test accuracy of 91.68%, using hyperparameters `max_depth=None`, `min_samples_split=2`, and `n_estimators=150`.

XGBoost: Achieved test accuracy of 91.49%, optimized with hyperparameters `learning_rate=0.2`, `max_depth=10`, `n_estimators=150`, `reg_alpha=0.1`, and `reg_lambda=1.5`.

LightGBM: Achieved the highest test accuracy of 92.63%, with `learning_rate=0.2`, `n_estimators=100`, and `num_leaves=100`.

Across all models, the training accuracy was consistently high at 100%, indicating robust gene selection. Despite minor overfitting, the models generalized well on the test sets, demonstrating the feasibility of using transcriptomic data to create AML diagnostic tools with high predictive power.

Discussion-Conclusion: This study successfully identified a set of 15 genes that significantly contribute to differentiating AML cases, showing promise for clinical applications. The gene signature includes mitochondrial and hemoglobin-related genes, highlighting metabolic and oxidative stress pathways previously implicated in AML pathogenesis. These findings align with literature linking AML with altered mitochondrial activity, supporting their diagnostic relevance [13], [14], [15]. Among the tested models, LightGBM outperformed the others, making it a strong candidate for clinical implementation. This model's capacity to achieve high accuracy with only 15 genes could reduce costs and complexity in clinical testing, making AML diagnosis more accessible and timely. However, further validation on independent datasets is necessary to ensure model generalizability. Future work could integrate proteomic or epigenetic data, enabling a multi-omics approach for even greater diagnostic accuracy and insights into AML pathophysiology. This study demonstrates the potential of machine learning to improve AML diagnostics by enabling accurate, interpretable models that utilize minimal gene sets, thus supporting more precise clinical and scientific decision-making.

Keywords: AML, gene signature, machine learning, biomarker detection, transcriptomics

REFERENCE

[1] E. Papaemmanuil *et al.*, "Genomic Classification and Prognosis in Acute Myeloid Leukemia," *N. Engl. J. Med.*, vol. 374, no. 23, pp. 2209–2221, Jun. 2016, doi: 10.1056/NEJMoa1516192.

[2] J. L. Carter *et al.*, "Targeting multiple signaling pathways: the new approach to acute myeloid leukemia therapy," *Signal Transduct. Target. Ther.*, vol. 5, no. 1, p. 288, Dec. 2020, doi: 10.1038/s41392-020-00361-x.

[3] H. Han *et al.*, "Expression level and prognostic potential of beta-catenin–interacting protein in acute myeloid leukemia," *Medicine (Baltimore)*, vol. 101, no. 33, p. e30022, Aug. 2022, doi: 10.1097/MD.0000000000030022.

[4] R. Dillon, N. Potter, S. Freeman, and N. Russell, "How we use molecular minimal residual disease (MRD) testing in acute myeloid leukaemia (AML)," *Br. J. Haematol.*, vol. 193, no. 2, pp. 231–244, Apr. 2021, doi: 10.1111/bjh.17185.

[5] S. D. Gupta and Z. Sachs, "Novel single-cell technologies in acute myeloid leukemia research," *Transl. Res.*, vol. 189, pp. 123–135, Nov. 2017, doi: 10.1016/j.trsl.2017.07.007.

[6] A. Liberzon, C. Birger, H. Thorvaldsdóttir, M. Ghandi, J. P. Mesirov, and P. Tamayo, "The Molecular Signatures Database Hallmark Gene Set Collection," *Cell Syst.*, vol. 1, no. 6,

pp. 417–425, Dec. 2015, doi: 10.1016/j.cels.2015.12.004.

[7] J. Zeng, C. Lai, J. Luo, and L. Li, "Functional investigation and two-sample Mendelian randomization study of neuropathic pain hub genes obtained by WGCNA analysis," *Front. Neurosci.*, vol. 17, p. 1134330, Apr. 2023, doi: 10.3389/fnins.2023.1134330.

[8] C.-H. Chang, C.-H. Lin, and H.-Y. Lane, "Machine Learning and Novel Biomarkers for the Diagnosis of Alzheimer's Disease," *Int. J. Mol. Sci.*, vol. 22, no. 5, p. 2761, Mar. 2021, doi: 10.3390/ijms22052761.

[9] E. Glaab, A. Rauschenberger, R. Banzi, C. Gerardi, P. Garcia, and J. Demotes, "Biomarker discovery studies for patient stratification using machine learning analysis of omics data: a scoping review," *BMJ Open*, vol. 11, no. 12, p. e053674, Dec. 2021, doi: 10.1136/bmjopen-2021-053674.

[10] R. L. Grossman *et al.*, "Toward a Shared Vision for Cancer Genomic Data," *N. Engl. J. Med.*, vol. 375, no. 12, pp. 1109–1112, Sep. 2016, doi: 10.1056/NEJMp1607591.

[11] M. Mounir *et al.*, "New functionalities in the TCGAbiolinks package for the study and integration of cancer data from GDC and GTEx," *PLOS Comput. Biol.*, vol. 15, no. 3, p. e1006701, Mar. 2019, doi: 10.1371/journal.pcbi.1006701.

[12] T. C. S. Antonio Colaprico, *TCGAbiolinks*. (2017). Bioconductor. doi: 10.18129/B9.BIOC.TCGABIOLINKS.

[13] S. Verma, S. Shukla, M. Pandey, G. T. MacLennan, and S. Gupta, "Differentially Expressed Genes and Molecular Pathways in an Autochthonous Mouse Prostate Cancer Model," *Front. Genet.*, vol. 10, p. 235, Mar. 2019, doi: 10.3389/fgene.2019.00235.

[14] A. Taherian Fard *et al.*, "Deconstructing heterogeneity of replicative senescence in human mesenchymal stem cells at single cell resolution," *GeroScience*, vol. 46, no. 1, pp. 999–1015, Jun. 2023, doi: 10.1007/s11357-023-00829-y.

[15] A. Stoccoro and F. Coppedè, "Mitochondrial DNA Methylation and Human Diseases," *Int. J. Mol. Sci.*, vol. 22, no. 9, p. 4594, Apr. 2021, doi: 10.3390/ijms22094594.

VALIDITY AND RELIABILITY ANALYSIS OF MON4T CLINICAL SMARTPHONE APPLICATION COMPARED TO TETRAX STATIC POSTURALGRAPHY IN THE EVALUATION OF STATIC POSTURALGRAPHIES IN ORTHOPEDIC KNEE PATHOLOGY

Musa, Çankaya[1*], Ülkü Barış Alperen[2]

[1] Necmettin Erbakan University, Seydişehir Health Services, Vocational School Therapy and Rehabilitation Department Konya,Türkiye

[2] Eskişehir Osmangazi University, Faculty of Medicine Hospital, Department of Physiotherapy and Rehabilitation, Eskişehir, Türkiye

** Corresponding author: fizyoterapistbaris@gmail.com*

Introduction-Aim: The primary forces affecting the knee joint include supporting body weight, assisting lower limb oscillation and absorbing impact shock. The complex interaction of these structures allows the knee to withstand tremendous forces during various normal movements [1]. Disruption of this structure results in knee-related pathologies. Therefore, it is important to evaluate static posturography from normal and diseased knee joints for the aid or rehabilitation of human locomotor function. In the literature, there are no studies on static postural assessment in knee pathologies using a smartphone application. This study was planned to determine the validity and reliability of the Mon4t application, a smartphone that performs static postural assessment in knee pathologies, compared with the Tetrax static posturagraphy used in the clinic [2,3].

Materials-Methods: This study is a descriptive single-visit methodological study. Permission was obtained from Osmangazi University Non-interventional clinical research ethics committee. The study was conducted in accordance with the Declaration of Helsinki. Verbal and written informed consent was obtained from the participants. Verbal information about the study was also given. Patients with orthopedic pathology who came to the physical therapy department of Eskişehir Osmangazi University Medical Faculty Hospital between August-October 2024 and September 2024 were included in the study. Sociodemographic characteristics of the patients, static postural assessment with smartphone application and Tetrax, visual anologic scale (VAS) for pain intensity, Lysholm Knee Scoring Scale (LDSS) for knee function assessment, and Knee Test of Activities of Daily Living (KOS-ADL) for quality of life were used in the evaluation. Static postural assessment was performed by 2 practitioners using a smartphone application (Mon4t) and Tetrax device.To assess inter-observer reliability, two measurements made by the first observer were evaluated at 24-48 hours intervals[4]. To assess intraobserver reliability, the measurements of both observers were compared. Statistical analysis was performed using intraclass correlation coefficient (ICC) and Pearson correlation analysis.

Results: Eighteen participants between the ages of 19 and 65 were included in our study. The results of intra- and interrater reliability analyses were shown (0.91, 0.81, respectively). It was determined that the Mon4t smartphone application was perfectly reliable in neutral stance y-axis, feet together y-axis with eyes open, Tetrax weight distribution index with eyes open and closed intra- and inter-observer reliability. In our analysis of the concurrent validity of neutral stance y-axis,feet together y-axis eyes open, Tetrax weight distribution index eyes open and closed with the two methods used in our study, the strongest correlations were found between the smartphone and the inclinometer (r=0.91, r=0.92, r=0.81, r=0.807, respectively).

Discussion-Conclusion: The Mont4 clinical smartphone application was found to have excellent intra-observer (ICC>0.80, SEM>2.19) and inter-observer reliability (ICC>0.71, SEM>1.16) in neutral stance X,Y,Z axis, feet together X,Y,Z axis, eyes open, excellent intra-observer (ICC>0.75, SEM>1.12) and inter-observer reliability (ICC>0.67, SEM>1.18) in assessing postural stability in knee pathologies. Furthermore, the smartphone app was shown to have excellent correlation with the Mont 4 clinicTetrax. These results suggest that the smartphone application may be a useful application for assessing postural stability in the clinical setting.

Keywords: Posturalgraphie, Stability, Tetrax, Mon4t Clinic, Validity

REFERENCE

[1] Cusin, F. S., Tomaz, A., Ganança, M. M., Oliveira, E. M., Gonçalves, A. B. F., & Caovilla, H. H. (2022). Postural Control in Relapsing-Remitting Multiple Sclerosis. *International Archives of Otorhinolaryngology*, *26*(04), e592-e604.

[2] Park, J. Y., Lee, J. C., & Cheon, M. W. (2014). Study on the Effects of Tetrax®-based Combined Rehabilitation Exercise on Chronic Back Pain Cases. *Transactions on Electrical and Electronic Materials*, *15*(3), 144-148.

[3] Regev, K., Eren, N., Yekutieli, Z., Karlinski, K., Massri, A., Vigiser, I., ... & Karni, A. (2024). Smartphone-based gait assessment for multiple sclerosis. *Multiple sclerosis and related disorders*, *82*, 105394.

[4] Margareta, K., Hakim, M., Kurniawan, M., & Anindhita, T. (2021). Indonesian version of the PainDETECT Questionnaire in the assessment of neuropathic pain: a validity and reliability study. *Jurnal Sinaps*, *4*(2), 1-13.

OVERVIEW OF AUGMENTED REALITY IN SURGICAL OPERATIONS

Furkan Emir Arslan[1*], Yeşim Aygül[1], Onur Ugurlu[1]

[1] *Faculty of Engineering and Architecture, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: furkanemirarslan@hotmail.com*

Introduction-Aim: Augmented Reality (AR) technology facilitates the integration of digital objects, images, sounds, or other data types with the physical world interaction by utilizing various sensors, cameras, displays, and software algorithms. This technology enriches the user's real-world perception while expanding interaction possibilities. AR technology is gaining attention in the medical field, especially in surgical operations [1]. Integrating AR applications in surgical operations may enhances the performance of surgical procedures and teaching techniques. This study summarizes the recent developments in AR applications in surgical operations.

Materials-Methods: This study examines some exciting studies published between 2018 and 2024 regarding the use of AR technology in surgical operations, sourced from the MDPI, Springer Nature Link, PubMed, ScienceDirect, and Google Scholar databases. The selected studies were chosen based on their relevance and impact. In these studies, surgeons were surveyed regarding the effectiveness of the AR systems after surgical operations using the AR systems. The collected data were analyzed to compare the operational processes conducted with and without AR.

Results: The application of AR systems in surgical operations has significantly reduced operation times and improved performance. For example, Rose et al. (2019) developed an AR system that enables accurate localization of pathological and anatomical structures in head and neck surgery. They reported that the system can potentially enhance safety and efficiency in this field [1]. Iqbal et al. (2021) used AR to improve user interfaces in robot-assisted orthopedic surgeries. Their study highlighted how AR can overlay virtual medical data onto the actual surgical field, offering a more immersive experience for surgeons than traditional 2D screens. Testing with ten surgeons performing robot-assisted patellofemoral arthroplasty revealed that AR did not significantly change surgical techniques, with no notable differences in procedure completion times or surface roughness [2]. In a study conducted by H. F. Al Janabi et al. (2020), a surgical operation was performed using the Microsoft HoloLens AR device. Surgeons using AR technology observed that the AR device improved the performance outcomes of the operation. Additionally, the AR device's feature of automatically adjusting their direction based on head movements eliminated the need to touch the glasses during the operation, thus maintaining sterility. However, it was noted that clinics need to be trained on the appropriate and safe use of this technology [3]. In a study by J. Olexa et al. (2024), the Apple Vision Pro AR devices were used before surgical intervention, and they reported that the AR system improved the operation. In that study, a 3D model of the area to be intervened upon was created before the surgical procedure, assisting surgeons in their operation preparations [4]. Manuel de Jesus et al. (2024) reported that integrating AR in spine surgery significantly advances medical technology. AR facilitates immersive, three-dimensional visualizations of anatomical structures, which aids meticulous planning and execution of spine surgeries. Despite its benefits, challenges such as model accuracy, user interface design, and the learning curve for new technology need to be addressed [5].

Discussion-Conclusion: AR technology advances surgical operations and medical education by enhancing precision, reducing operation times, and facilitating sterile, touch-

free interactions. Devices like Microsoft HoloLens and Apple Vision Pro have improved surgical outcomes by enabling 3D modeling for pre-operative planning, allowing surgeons to understand complex anatomical structures better. Furthermore, AR’s immersive, interactive features enrich medical education by bridging theoretical knowledge and practical skills. However, proper device usage and safety protocol training are essential for optimal integration into clinical practices. As AR technology develops, addressing ergonomic concerns and conducting further research will be critical to fully realizing its potential in healthcare.

Keywords: Augmented Reality, Surgical Procedures, Apple Vision Pro, Microsoft HoloLens

REFERENCE

[1] A. S. Rose, H. Kim, H. Fuchs, and J. M. Frahm, “Development of augmented-reality applications in otolaryngology–head and neck surgery,” The Laryngoscope, vol. 129, pp. S1–S11, 2019.

[2] H. Iqbal, F. Tatti, and F. R. y Baena, "Augmented reality in robotic assisted orthopaedic surgery: A pilot study," *Journal of Biomedical Informatics*, vol. 120, p. 103841, 2021.

[3] H. F. Al Janabi, A. Aydin, S. Palaneer, N. Macchione, A. Al-Jabir, M. S. Khan, and K. Ahmed, "Effectiveness of the HoloLens mixed-reality headset in minimally invasive surgery: a simulation-based feasibility study," *Surgical Endoscopy*, vol. 34, pp. 1143-1149, 2020.

[4] J. Olexa, A. Trang, J. Cohen, K. Kim, M. Rakovec, J. Saadon, and J. Cherian, "The apple vision pro as a neurosurgical planning tool: A case report," *Cureus*, vol. 16, no. 2, 2024.

[5] M. De Jesus Encarnacion Ramirez, G. Chmutin, R. Nurmukhametov, G. R. Soto, S. Kannan, G. Piavchenko, and N. Montemurro, “Integrating augmented reality in spine surgery: redefining precision with new technologies,” Brain Sciences, vol. 14, no. 7, p. 645, 2024.

INVESTIGATION OF THE RELATIONSHIP BETWEEN OPENNESS TO INNOVATION AND ARTIFICIAL INTELLIGENCE LITERACY IN PHYSIOTHERAPY AND REHABILITATION UNDERGRADUATE STUDENTS

Melike Taş[1*], Sevgi Sevi Yesilyaprak[1], Gülbin Ergin Gedik[1], Derya Ozer Kaya[2], Damla Karabay[2]

[1] *Department of Physiotherapy and Rehabilitation, İzmir Bakırçay University, İzmir, Türkiye*

[2] *Department of Physiotherapy and Rehabilitation, İzmir Katip Çelebi Universityi, İzmir, Türkiye*

** Corresponding author: melike.tas@bakircay.edu.tr*

Introduction-Aim: The concept of "artificial intelligence", which emerged in the 1950s, is generally defined as providing computer systems with special abilities similar to humans [1]. In today's technology, supervised learning models are widespread, and artificial intelligence research is thought to be in its early stages [2]. Although it is in its early stages, artificial intelligence-supported tools and methods have begun to be used in many areas where technological support is needed, as well as in the field of physiotherapy and rehabilitation [3]. However, these systems can be defined as a concept that threatens humanity's existence and triggers social anxiety due to the possibility of uncontrolled models getting out of control, despite the many innovations and conveniences they provide [4]. It is of great importance to learn the artificial intelligence literacy and innovative attitude of individuals consisting of physiotherapy and rehabilitation students to determine which topics related to artificial intelligence the sociological attitude towards fears and anxieties triggered by the unknown peaks, to evaluate the reality of the anxiety on these issues, and to plan both precautions and relevant trainings.

This study aims to examine the relationship between openness to innovation and artificial intelligence literacy in students receiving undergraduate education in physiotherapy and rehabilitation.

Materials-Methods: Volunteers from undergraduate physiotherapy and rehabilitation department students were included in this study. First, demographic information of the participants; age, gender, and undergraduate period were recorded.

Then, the assessment of AI literacy was assessed through the AI Literacy Scale. The AI Literacy Scale consists of 31 items and three factors (technical understanding, critical evaluation, practical application) and ranges from “Strongly Agree (7)” to “Strongly Disagree (1)”.

Finally, attitudes towards innovation were assessed with the Individual Innovation Scale, which consists of 20 items, 8 negative and 12 positive.

Data analysis was performed using the IBM SPSS 23 statistical package program. In the data set consisting of 174 participants between the ages of 17-43, it was determined that 78.2% of the participants were female and 21.8% were male according to descriptive analysis. In terms of education level, 28.2% of the participants were in the first grade; 28.2% were in the second grade; 31.0 were in the third grade; and 12.6 were in the fourth grade. A normality test was performed between the total score of the artificial intelligence literacy scale and the innovation scale. As a result of the analysis, it was determined that there was no missing data and the skewness and kurtosis values were normal ($-1<x<+1$). Considering the sample size in the data set, the Shapiro-Wilk Normality test result was examined and a significance level of $p<0.01$ was found. Based on this, it was determined that the distribution was outside the normal distribution curve and it was decided to use non-parametric tests.

The Spearman Correlation Test was selected to examine the relationship between artificial intelligence literacy and innovation attitude. No correlation ($r=0.129$) was found in the analysis of total scores. When the correlation analysis was repeated on an item basis, a weak positive correlation ($r=0.276$) was found between awareness of the importance of artificial intelligence and innovation attitude.

Chi-square analysis was performed to examine whether there was a difference in the relationship and distribution between artificial intelligence literacy and innovation attitude according to gender and undergraduate class level. Since the data did not fit the normal distribution curve, the Pearson-Chi Square test was applied. As a result of the analysis, it was determined that there was no significance between artificial intelligence literacy and innovation attitude depending on gender and undergraduate class level [5].

Results: No correlation was found in the analysis of total scores of artificial intelligence literacy and innovation attitude. When the correlation analysis was repeated on an item basis, a weak positive significant correlation was found between awareness of the importance of artificial intelligence and innovation attitude. As a result of the analyses conducted to examine whether the relationship and distributions between artificial intelligence literacy and innovation attitude differed according to gender and undergraduate class level, it was determined that there was no difference between the observed and expected frequencies depending on gender and undergraduate class level among the total scores of artificial intelligence literacy and innovation attitude. However, it was observed that 4th-grade students contributed more to the critical attitude compared to other grade levels in the critical evaluation sub-dimension of the Artificial Intelligence Literacy Scale with a $p=0.01$ value [5].

Discussion-Conclusion: Özel et al. In their study published in 2022, they examined the opinions of dentistry students about artificial intelligence applications. In the study in which 236 dentistry students participated, 73.88% of the students knew the use of artificial intelligence. The results of the study determined that upper-class students had more knowledge about the use of artificial intelligence than lower-class students [6]. In this study, parallel to the study of Özel et al., it was determined that 4th-grade students approached artificial intelligence studies with a more critical attitude compared to other grade levels.

In the analysis results conducted to examine the relationship between artificial intelligence literacy and innovativeness attitude, although no significant relationship was found between the total scores of both scales, when the item-based analysis of the Artificial Intelligence Literacy Scale was repeated, a weak positive correlation was found between the awareness attitude about the importance of artificial intelligence and the innovativeness view. Even if there is a positive relationship between individuals with innovative views and the level of awareness about the importance of artificial intelligence, the low level of relationship may support the existence of prejudice against artificial intelligence systems despite having an innovative view. At the same time, this result can be interpreted as the awareness level of the importance of artificial intelligence may be higher in individuals with innovative attitudes.

The stated results were reached as a result of analyses conducted on a limited sample size, and further studies are needed on larger sample groups.

Keywords: Artificial Intelligence, Physiotherapy and Rehabilitation, Artificial Intelligence Literacy, Innovativeness

REFERENCE

[1] Pirim, A. G. H. (2006). Yapay zeka. *Yaşar Üniversitesi E-Dergisi*, *1*(1), 81-93.

[2] Demirhan, A., Kılıç, Y. A., & İnan, G. (2010). Tıpta yapay zeka uygulamaları.

[3] Misican, D. Ö. (2024). İnsan Kaynakları Profesyonellerinin Perspektifinden Dijitalleşen Çalışma Hayatında Yapay Zekâ İşgücü İçin Hangi Yol Ayrımında?. Journal of Academic Value Studies (JAVStudies), 6(2), 152-175.

[4] Li, J., & Huang, J. S. (2020). Dimensions of artificial intelligence anxiety based on the integrated fear acquisition theory. Technology in Society, 63, 101410.

[5] Statistics, L. (2015). Kaplan-Meier using SPSS statistics. Statistical tutorials and software guides.

[6] Özel, Ş., & Büyükçavuş, M. H. (2022). Diş hekimliği öğrencilerinin diş hekimliğinde yapay zekâ uygulamaları ile ilgili düşüncelerinin incelenmesi.

A NEW ERA IN DIABETES MANAGEMENT: PRODUCTIVE ARTIFICIAL INTELLIGENCE

Meleknur Göktaş[1*], Tuğba Bilgehan[2]

[1] *Internal Medicine Service, Elmadag Dr.Hulusi Alatas State Hospital, Ankara, Türkiye*

[2] *Department of Nursing, Ankara Yıldırım Beyazıt University, Ankara, Türkiye*

* *Corresponding author: meleknur.gkts@gmail.com*

Introduction-Aim: Diabetes Mellitus (DM) is an endocrine and metabolic disorder that disrupts carbohydrate, protein, and lipid metabolism due to insulin resistance, absence or insufficiency of insulin secretion, or defects in insulin action, and is characterized by hyperglycemia. Both acute and chronic complications may develop as a result of DM [1, 2, 3, 4]. Over the past three decades, the incidence of DM has increased dramatically worldwide [5]. Acute complications of DM include hypoglycemia, diabetic ketoacidosis, hyperosmolar hyperglycemic state, and lactic acidosis. Chronic microvascular complications include diabetic nephropathy, retinopathy, and neuropathy, while macrovascular chronic complications encompass cardiovascular disease, peripheral arterial disease, and cerebrovascular disease. Uncontrolled monitoring, treatment, and self-management in DM can lead to complications, increasing morbidity and mortality rates [4, 5, 6]. Because DM affects nearly all bodily systems, it requires ongoing medical care [7, 8, 9].

Self-management behaviors in DM involve practices such as maintaining a healthy diet, engaging in regular physical activity, monitoring blood glucose (BG) levels, taking medications correctly and consistently, and developing healthy coping strategies and problem-solving skills [3, 10, 11]. Studies have shown that these behaviors can prevent or delay complications [6, 8]. The preferences, values, and goals of each diabetic individual are special and unique to ensure self-management. In this case, it is important to create an individualized management plan [12]. All aspects of DM management must also consider the person's personal DM history [13]. Following individualized education, individuals with diabetes are responsible for managing their condition [14, 15]. However, time, finances, or other constraints may prevent individuals with diabetes from regularly meeting with a DM educator [16]. In such instances, technology can support daily DM self-management activities, such as BG monitoring, exercise, healthy eating, medication adherence, complication monitoring, and problem-solving [17]. Because all self-care and self-management behaviors performed by individuals with diabetes are affected by the current health services, technological devices, mobile applications, and renewed technological developments [18].

Given the complexity of DM process and management, emerging artificial intelligence (AI) technologies have become essential tools, empowering both individuals with diabetes and healthcare professionals [19, 20]. AI enables computers to mimic human intelligence and encompasses both deep learning (DL) and machine learning (ML) [21, 22]. DL, a subset of ML, uses neural networks to train machines to perform tasks autonomously [23]. One of the types of AI that is very popular today and has started to appeal to a wide audience in the field of health is generative artificial intelligence (GenAI) [24, 25]. GenAI is a subset of AI that can process large amounts of data to learn complex language structures and relationships between texts, using DL techniques to create new content [26]. GenAI is defined as machines that have human-like language skills and mimic human intelligence [27]. While GenAI's output improves as users learn to create effective commands, its related tools are user-friendly and do not require extensive user expertise [28]. In recent years, GenAI has been used not only in

DM care and treatment but also by patients in DM management, diagnosis of complications, and risk analysis [29, 30, 31]. GenAI can create an individualized DM self-management plan for patients and facilitate the implementation of self-management behaviors [32, 33, 34].

Numerous studies in the literature have examined the contributions of GenAI models to DM self-management, and reviewing these studies is essential to provide a general framework on this topic. The primary aim of this study is to systematically examine research that utilizes GenAI in DM management. This review aims to analyze the reasons for using GenAI in the field of DM, outline the first evidence in the literature, and shed light on the future to make the use of GenAI safe.

Materials-Methods: This systematic review was conducted in accordance with PRISMA guidelines to evaluate the role of GenAI in DM self-management. A comprehensive literature search was performed across Scopus, Web of Science, PubMed, Google Scholar, Ulakbim, and Türk Medline between February and October 2024, using the keywords "diabetes," "generative artificial intelligence," and "diabetes self-management". Studies published between 2018 and 2024 were screened based on predefined inclusion and exclusion criteria, resulting in 19 eligible studies.

Results: The review analyzed 19 studies on GenAI applications in DM self-management, emphasizing its role in personalized recommendations, glucose monitoring, complication detection, insulin management, and patient education. Among the reviewed studies, GPT-based models were the most frequently utilized (53%), followed by Google Bard, GAN, LSTM, Bayesian networks, and other architectures [35, 36, 37, 38]. The findings suggest that GenAI-powered systems enhance HbA1c reduction, glycemic stability, and patient adherence to self-management strategies [39, 40]. Studies also highlight GenAI's effectiveness in generating educational materials, assisting in insulin titration, and improving clinical decision support [41, 42]. Additionally, some models have demonstrated superior performance in BG prediction compared to conventional methods [35, 36]. However, challenges such as data security risks, misinformation generation, and inconsistent detection of hypoglycemia remain significant concerns [43, 44].

Discussion-Conclusion: The findings of this systematic review indicate that GenAI technologies are becoming increasingly important in DM self-management, offering powerful tools that assist patients in managing their own care processes. In this study highlights some of the key advantages of GenAI in DM management, including personalized dietary and exercise recommendations, educational material generation, early complication detection, emotional support, and remote patient monitoring [39, 40, 42, 45]. These capabilities suggest that GenAI could help bridge gaps in traditional DM care by providing real-time, AI-driven interventions. Additionally, research indicates that GenAI has strong predictive capabilities, particularly in forecasting BG levels and optimizing insulin dosage, often outperforming conventional methods in glucose regulation and glycemic stability, which is critical for preventing hypoglycemia and hyperglycemia [36, 40, 46]. Despite these benefits, significant challenges remain. Data security and patient privacy continue to be major concerns, particularly regarding how sensitive health information is collected, stored, and processed. There is also the risk of misinformation, as AI-generated responses may sometimes be inaccurate, potentially compromising clinical recommendations and patient safety [38, 40, 43, 44]. Furthermore, some studies highlight limitations in GenAI's ability to detect and manage severe hypoglycemia, emphasizing the need for further validation and refinement [40]. Another critical issue is integrating GenAI into clinical workflows. While AI-powered metabolic coaches and chatbot systems have shown promise in improving patient engagement and adherence, their adoption in real-world clinical settings remains limited [38, 40]. Additionally, healthcare professionals must be properly trained to interpret and utilize AI-generated insights effectively [44].

As GenAI technologies advance, they have the potential to become more accessible, cost-effective, and patient-centered, ultimately contributing to better DM management and patient outcomes. However, to fully realize these benefits, further improvements are needed to make GenAI-based systems more reliable, transparent, user-friendly, and ethically sound. Future research is recommended to focus on developing strategies that enhance data security, accuracy, interpretability, and user satisfaction to strengthen the integration of GenAI applications into clinical practice.

Keywords: Generative Artificial Intelligence, Diabetes Self-Management, Diabetes

REFERENCE

[1] C. American Diabetes Association Professional Practice, "2. Diagnosis and Classification of Diabetes: Standards of Care in Diabetes-2024," Diabetes Care, vol. 47, no. Suppl 1, pp. S20-S42, Jan 1 2024, doi: 10.2337/dc24-S002.

[2] N. A. ElSayed et al., "10. Cardiovascular Disease and Risk Management: Standards of Care in Diabetes—2023," Diabetes Care, vol. 46, no. Supplement_1, pp. S158-S190, 2022, doi: 10.2337/dc23-S010.

[3] N. A. ElSayed et al., "11. Chronic Kidney Disease and Risk Management: Standards of Care in Diabetes—2023," Diabetes Care, vol. 46, no. Supplement_1, pp. S1-S291, 2022, doi: 10.2337/dc23-S011.

[4] D. M. S. a. E. Group, "Guidelines for the Diagnosis, Treatment and Monitoring of Diabetes Mellitus and Its Complications-2022," Ankara,Türkiye, Online Jul. 2022, vol. 15. Accessed: Nov. 2, 2024. [Online]. Available: https://file.temd.org.tr/Uploads/publications/guides/documents/diabetes-mellitus_2022.pdf

[5] I. D. Federation, "IDF diabetes atlas, tenth," International Diabetes, 2021.

[6] G. A. N. A. ElSayed, V. R. Aroda, R. R. Bannuru, F. M. Brown, D. Bruemmer, et al., "Standards of Care in Diabetes—2023," Diabetes Care, vol. 46, 1, pp. 1-291, Jan. 2023. [Online]. Available: https://www.portailvasculaire.fr/sites/default/files/docs/2023_ada_diabete_standards_of_care_in_diabetes_diab_care.pdf.

[7] R. O. T. M. O. H. G. D. O. P. Health, "TURKEY DIABETES PROGRAM 2023 - 2027," REPUBLIC OF TURKEY MINISTRY OF HEALTH, Ankara, Türkiye, Online 2023, vol. 3. Accessed: Nov. 5, 2024. [Online]. Available: https://hsgm.saglik.gov.tr/depo/birimler/saglikli-beslenme-ve-hareketli-hayat-db/Dokumanlar/Programlar/Turkiye-Diyabet-Programi.pdf

[8] S. D. S. W. Group, "Guidelines for the Diagnosis, Treatment and Monitoring of Diabetes Mellitus and Its Complications-2024," Ankara, Türkiye, Online Jun. 2024, vol. 16. Accessed: Jun. 26, 2024. [Online]. Available: https://file.temd.org.tr/ Uploads/publications/guides/documents/diabetesmellitus2024.pdf

[9] N. D. C. GROUP, "Diabetes Diagnosis and Treatment Guide," Istanbul, Türkiye, Online Apr. 2019, vol. 9. Accessed: Jan. 15. [Online]. Available: https://www.turkdiab.org/admin/PICS/files/Diyabet_Tani_ve_Tedavi_Rehberi_2019.pdf

[10] E. H. Wagner, B. T. Austin, C. Davis, M. Hindmarsh, J. Schaefer, and A. Bonomi, "Improving chronic illness care: translating evidence into action," Health Aff (Millwood), vol. 20, no. 6, pp. 64-78, Nov-Dec 2001, doi: 10.1377/hlthaff.20.6.64.

[11] E. Sabaté, Adherence to long-term therapies: evidence for action. World Health Organization, 2003.

[12] S. Dineen-Griffin, V. Garcia-Cardenas, K. Williams, and S. I. Benrimoj, "Helping patients help themselves: A systematic review of self-management support strategies in primary health care practice," PLoS One, vol. 14, no. 8, p. e0220116, 2019, doi: 10.1371/journal.pone.0220116.

[13] X. Y. Khor, J. M. Pappachan, and M. S. Jeeyavudeen, "Individualized diabetes care: Lessons from the real-world experience," World J Clin Cases, vol. 11, no. 13, pp. 2890-2902, May 6 2023, doi: 10.12998/wjcc.v11.i13.2890.

[14] K. D. Brunisholz et al., "Diabetes self-management education improves quality of care and clinical outcomes determined by a diabetes bundle measure," J Multidiscip Healthc, vol. 7, pp. 533-42, 2014, doi: 10.2147/JMDH.S69000.

[15] F. Ozpulat, "A Contemporary Role of the Nurse in Health Protection and Promotion: Educator Identity," Maltepe University Journal of Nursing Science and Art, pp. 293-297, 2010.

[16] M. Song et al., "An evaluation of Web-based education as an alternative to group lectures for diabetes self-management," Nurs Health Sci, vol. 11, no. 3, pp. 277-84, Sep 2009, doi: 10.1111/j.1442-2018.2009.00458.x.

[17] C. W. Hunt, "Technology and diabetes self-management: An integrative review," World J Diabetes, vol. 6, no. 2, pp. 225-33, Mar 15 2015, doi: 10.4239/wjd.v6.i2.225.

[18] A. Cahn, A. Akirov, and I. Raz, "Digital health technology and diabetes management," J Diabetes, vol. 10, no. 1, pp. 10-17, Jan 2018, doi: 10.1111/1753-0407.12606.

[19] P. o. t. s. W. o. A. I. f. Diabetes, "1st Workshop on Artificial intelligence for Diabetes (ECAI 2016)," ed, 2016.

[20] I. Contreras and J. Vehi, "Artificial Intelligence for Diabetes Management and Decision Support: Literature Review," (in English), J Med Internet Res, Review vol. 20, no. 5, p. e10775, 2018, doi: 10.2196/10775.

[21] A. K. Goel and J. Davies, "Artificial Intelligence," in The Cambridge Handbook of Intelligence, R. J. Sternberg Ed., 2 ed. Cambridge: Cambridge University Press, 2020, pp. 602-625.

[22] Azure. "Deep learning vs. machine learning in Azure Machine Learning." Microsoft Learn. https://learn.microsoft.com/tr-tr/azure/machine-learning/concept-deep-learning-vs-machine-learning?view=azureml-api-2 (accessed Nov. 12, 2024).

[23] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85-117, 2015/01/01/ 2015, doi: https://doi.org/10.1016/j.neunet.2014.09.003.

[24] B. Mesko, "The ChatGPT (Generative Artificial Intelligence) Revolution Has Made Artificial Intelligence Approachable for Medical Professionals," (in English), J Med Internet Res, Viewpoint vol. 25, p. e48392, 2023, doi: 10.2196/48392.

[25] S. Liu, J. Chen, Y. Feng, Z. Xie, T. Pan, and J. Xie, "Generative artificial intelligence and data augmentation for prognostic and health management: Taxonomy, progress, and prospects," Expert Systems with Applications, vol. 255, p. 124511, 2024/12/01/ 2024, doi: https://doi.org/10.1016/j.eswa.2024.124511.

[26] G. Pavlik. "What is Generative Artificial Intelligence and How Does It Work?" Oracle Türkiye. https://www.oracle.com/tr/artificial-intelligence/generative-ai/what-is-generative-ai/ (accessed Jul. 17, 2024).

[27] A. Bozkurt, "ChatGPT, üretken yapay zeka ve algoritmik paradigma değişikliği," Alanyazın, vol. 4, no. 1, pp. 63-72, 2023.

[28] D. M. a. agency, "ChatGPT reaches 100 million users two months after launch," The Guardian. [Online]. Available: https://www.theguardian.com/technology/2023/feb/02/chatgpt-100-million-users-open-ai-fastest-growing-app

[29] L. Mazgouti and N. El Amrani, Comparative analysis of the performance of four learning models for predicting blood glucose levels in type 1 diabetic patients based on blood glucose measurements. 2023, pp. 1-7.

[30] H. V. Dudukcu, M. Taskiran, and T. Yildirim, "Blood glucose prediction with deep neural networks using weighted decision level fusion," Biocybernetics and Biomedical Engineering, vol. 41, no. 3, pp. 1208-1223, 2021/07/01/ 2021, doi: https://doi.org/10.1016/j.bbe.2021.08.007.

[31] A. Mohebbi, T. B. Aradottir, A. R. Johansen, H. Bengtsson, M. Fraccaro, and M. Morup, "A deep learning approach to adherence detection for type 2 diabetics," Annu Int Conf IEEE Eng Med Biol Soc, vol. 2017, pp. 2896-2899, Jul 2017, doi: 10.1109/EMBC.2017.8037462.

[32] F. M. Megahed, Y.-J. Chen, J. A. Ferris, S. Knoth, and L. A. Jones-Farmer, "How generative AI models such as ChatGPT can be (mis)used in SPC practice, education, and research? An exploratory study," Quality Engineering, vol. 36, no. 2, pp. 287-315, 2024/04/02 2024, doi: 10.1080/08982112.2023.2206479.

[33] M. Kowsher, M. Y. Turaba, T. Sajed, and M. M. M. Rahman, "Prognosis and Treatment Prediction of Type-2 Diabetes Using Deep Neural Network and Machine Learning Classifiers," in 2019 22nd International Conference on Computer and Information Technology (ICCIT), 18-20 Dec. 2019 2019, pp. 1-6, doi: 10.1109/ICCIT48885.2019.9038574.

[34] R. Kanitz, K. Gonzalez, R. Briker, and T. Straatmann, "Augmenting Organizational Change and Strategy Activities: Leveraging Generative Artificial Intelligence," The Journal of Applied Behavioral Science, vol. 59, p. 002188632311689, 04/11 2023, doi: 10.1177/00218863231168974.

[35] H. V. Dudukcu, M. Taskiran, and T. Yildirim, "Blood glucose prediction with deep neural networks using weighted decision level fusion," Biocybernetics and Biomedical Engineering, vol. 41, no. 3, pp. 1208-1223, 2021, doi: 10.1016/j.bbe.2021.08.007.

[36] T. Zhu, X. Yao, K. Li, P. Herrero, and P. Georgiou, "Blood glucose prediction for type 1 diabetes using generative adversarial networks," in CEUR Workshop Proceedings, vol. 2675, pp. 90-94, 2020.

[37] X. Wang, Y. Lin, Y. Xiong, S. Zhang, Y. He, Y. He, J. M. Plasek, L. Zhou, D. W. Bates, and C. Tang, "Using an optimized generative model to infer the progression of complications in type 2 diabetes patients," BMC Medical Informatics and Decision Making, vol. 22, no. 1, p. 174, 2022.

[38] M. Shiraishi, H. Lee, K. Kanayama, Y. Moriwaki, and M. Okazaki, "Appropriateness of artificial intelligence chatbots in diabetic foot ulcer management," The International Journal of Lower Extremity Wounds, 15347346241236811, 2024, doi: 10.1177/15347346241236811.

[39] A. K. Dey, "ChatGPT in diabetes care: An overview of the evolution and potential of generative artificial intelligence model like ChatGPT in augmenting clinical and patient outcomes in the management of diabetes," International Journal of Diabetes and Technology, vol. 2, no. 2, pp. 66-72, 2023, doi: 10.4103/ijdt.ijdt_31_23.

[40] A. Shaikh, S. Baluni, N. Malpani, P. Lodha, and A. Meena, "Decoding type 2 diabetes through point-of-care testing, cloud-based monitoring, and generative augmented retrieval model-driven virtual diabetes education: A comprehensive approach to glycemic control," International Journal of Diabetes and Technology, vol. 3, no. 1, pp. 25-31, 2024, doi: 10.4103/ijdt.ijdt_5_24.

[41] S. Mashatian, D. G. Armstrong, A. Ritter, J. Robbins, S. Aziz, I. Alenabi, M. Huo, T. Anand, and K. Tavakolian, "Building trustworthy generative artificial intelligence for diabetes care and limb preservation: a medical knowledge extraction case," Journal of Diabetes Science and Technology, 19322968241253568, 2024, doi: 10.1177/19322968241253568.

[42] A. Nayak, S. Vakili, K. Nayak, M. Nikolov, M. Chiu, P. Sosseinheimer, S. Talamentes, S. Testa, S. Palanisamy, V. Giri, K. Schulman, and K. Schulman, "Use of voice-based conversational artificial intelligence for basal insulin prescription management among patients with type 2 diabetes: a randomized clinical trial," JAMA Network Open, vol. 6, no. 12, p. e2340232, 2023, doi: 10.1001/jamanetworkopen.2023.40232.

[43] S. Khan, "The Future of Diabetes Care: Navigating with generative language models," Indus Journal of Medical and Health Sciences, vol. 1, no. 1, pp. 109-116, 2023. [Online]. Available: https://induspublishers.com/IJMHS/article/view/58.

[44] G. G. R. Sng, J. Y. M. Tung, D. Y. Z. Lim, and Y. M. Bee, "Potential and pitfalls of ChatGPT and natural-language artificial intelligence models for diabetes education," Diabetes Care, vol. 46, no. 5, pp. 103-105, 2023, doi: 10.2337/dc23-0197.

[45] H. Sun, K. Zhang, W. Lan, Q. Gu, G. Jiang, X. Yang, W. Qin, and D. Han, "An AI dietitian for type 2 diabetes mellitus management based on large language and image recognition models: preclinical concept validation study," Journal of Medical Internet Research, vol. 25, p. e51300, 2023.

[46] L. Kopitar, I. Fister Jr, and G. Stiglic, "Using generative AI to improve the performance and interpretability of rule-based diagnosis of type 2 diabetes mellitus," Information, vol. 15, no. 3, p. 162, 2024, doi: 10.3390/info15030162.

USE OF MACHINE LEARNING FOR PARKINSON'S DISEASE DETECTION

Merve Nur, Aras[1], Ali Mert, Erdoğan[1*], Abdulkadir Hızıroğlu[1]

[1] Department of Management Information Systems, İzmir Bakırçay University, İzmir, Türkiye epartment, University, City, Country

** Corresponding author: alimert.erdogan@bakircay.edu.tr*

Introduction-Aim: Machine learning (ML), a branch of artificial intelligence, has transformed various fields, especially healthcare, by enabling the analysis of complex medical data to recognize patterns and make data-driven decisions [1]. ML's potential in disease diagnosis is particularly notable, as it processes large datasets—such as medical imaging and health records—identifying patterns that may be missed by human clinicians. This approach is invaluable for early detection of conditions like Parkinson's disease (PD), where timely diagnosis can significantly enhance treatment outcomes and patient quality of life. PD, a prevalent neurodegenerative disorder, affects millions globally, presenting both individual and public health challenges due to its debilitating symptoms and economic burden on healthcare systems [2]. Early diagnosis of PD, which is characterized by motor and non-motor symptoms, including speech disturbances, is critical for effective intervention[3]. In this context, our study evaluates the effectiveness of three machine learning algorithms—XGBoost, logistic regression, and deep neural networks (DNN)—in predicting PD based on speech-related symptoms. By utilizing multiple datasets with varied attributes, this study aims to provide a robust analysis of each model's predictive accuracy and generalizability in clinical settings.

Materials-Methods: Three different techniques applied to three datasets, each with unique speech features. Dataset 1 [4] includes 756 samples and 754 features, such as jitter, shimmer, and Mel frequency cepstral coefficients (MFCCs), from both PD patients and healthy individuals. Dataset 2 [5] comprises 1040 instances with 26 voice features, accompanied by Unified Parkinson's Disease Rating Scale (UPDRS) scores. Dataset 3 [6] includes 240 instances, emphasizing dependency among repeated measures. Each dataset underwent preprocessing, including min-max standardization and verification for outliers and missing data. Logistic regression, XGBoost, and DNN were employed on the datasets, with 80% of data used for training and 20% for testing. For DNN, a model with dense and dropout layers was trained with TensorFlow and Keras, using binary crossentropy loss and accuracy as metrics. The models' performance was evaluated using accuracy, precision, recall, F1 score, and ROC AUC, enabling a robust comparative analysis for PD detection.

Results: In the first dataset, XGBoost outperformed other models with an accuracy of 0.9013, achieving high precision (0.8960), recall (0.9824), F1 Score (0.9372), and ROC AUC (0.9344), indicating a balanced and reliable classification. Logistic Regression also demonstrated competitive results, with an accuracy of 0.8684 and ROC AUC of 0.8605, while the DNN model showed strong recall (0.9561) but lower precision and ROC AUC (0.8499). For the second dataset, XGBoost again led with an accuracy of 0.7451 and ROC AUC of 0.8195, while Logistic Regression and DNN scored lower across metrics, highlighting XGBoost's robustness. In the third dataset, XGBoost and Logistic Regression both achieved an accuracy of 0.7916, though Logistic Regression had a slightly higher ROC AUC (0.8975). The DNN model had an accuracy of 0.7500 with strong recall but did not surpass the other models. Cross-validation on the first dataset confirmed XGBoost's reliability, with consistently high

performance across all folds. Overall, XGBoost demonstrated the strongest results across datasets, particularly in the first dataset, validating its effectiveness for PD detection from voice data.

Discussion-Conclusion: This study evaluated Logistic Regression, XGBoost, and DNN for PD detection using voice data from three distinct datasets. XGBoost consistently outperformed the other models, especially in the first dataset. Cross-validation further validated XGBoost's robust performance across all folds. The findings underscore the potential of XGBoost as a reliable tool for PD classification. Future research directions include expanding datasets to capture a broader demographic range, integrating multimodal data sources such as genetic and wearable data, and refining algorithms for improved real-time analysis. Collaborating with clinicians for real-world validation will be essential, alongside addressing ethical considerations to ensure practical, secure application in clinical environments.

Keywords: parkinson's disease detection, logistic regression, xgboost, dnn

REFERENCE

[1] I. Ibrahim and A. Abdulazeez, "The Role of Machine Learning Algorithms for Diagnosing Diseases," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 10–19, Mar. 2021, doi: 10.38094/jastt20179.

[2] J. J. Ferreira *et al.*, "Prevalence of Parkinson's disease: a population-based study in Portugal," *Eur J Neurol*, vol. 24, no. 5, pp. 748–750, May 2017, doi: 10.1111/ene.13273.

[3] I. Leroi, K. McDonald, H. Pantula, and V. Harbishettar, "Cognitive Impairment in Parkinson Disease," *J Geriatr Psychiatry Neurol*, vol. 25, no. 4, pp. 208–214, Dec. 2012, doi: 10.1177/0891988712464823.

[4] C. Sakar, G. Serbes, A. Gündüz, H. Nizam, and B. Sakar, "Parkinson's Disease Classification," 2018, *UC Irvine Machine Learning Repository*. doi: 10.24432/C5MS4X.

[5] O. Kursun, B. Sakar, M. Isenkul, C. Sakar, A. Sertbaş, and F. Gurgen, "Parkinson's Speech with Multiple Types of Sound Recordings," 2014. doi: 10.24432/C5NC8M.

[6] C. Prez, "Parkinson Dataset with replicated acoustic features," 2019, *UC Irvine Machine Learning Repository*.

ARTIFICIAL INTELLIGENCE-SUPPORTED PERSONAL CARE PRODUCT CONTENT ANALYSIS SYSTEM

İrem Karadağ[1] , Senanur İriz[1], Şilan Ekin[1]

[1] *Computer Engineering, İzmir Bakırçay University, İzmir, Turkiye*

Corresponding author: silanekinceng@gmail.com

Introduction-Aim: This project presents an artificial intelligence (AI)-powered content analysis system designed to analyze personal care product ingredients. The primary aim is to offer a tool for users to assess the safety of personal care products by identifying potentially harmful ingredients through an intuitive interface. Users can either scan product barcodes or utilize a built-in search engine to find and evaluate products. By comparing similar products, this system will support informed decision-making for healthier product choices.

Recognizing the vital role of the endocrine system in health, experts—including physicians and toxicologists—prioritize assessing the potential of cosmetic ingredients to disrupt endocrine function. Disruption of the endocrine system can lead to various health issues, underscoring the importance of this analysis [1].Additionally, heavy metals in cosmetics pose significant health risks, as they can accumulate in the body over time, leading to serious conditions such as cancer, reproductive and developmental disorders, and neurological issues [2]. While cosmetic products are used for beauty and personal care, they often contain toxic chemicals that can cause allergic reactions and irritation to the skin [2]. Therefore, this AI-powered system aims to enhance consumer awareness and promote safer choices in personal care products.

Materials-Methods: The AI model at the core of this project utilizes a pre-trained dataset of over 200 personal care products. This dataset includes comprehensive ingredient lists, allowing the AI to accurately recognize and categorize components based on their safety profiles. The model will be implemented using Python, with TensorFlow serving as the primary library for machine learning processes. The application will be accessible to users via a mobile app that communicates with the AI system through an API, ensuring real-time responses

Results: To ensure optimal accuracy in categorizing ingredients, a convolutional neural network (CNN) model will be trained and fine-tuned with TensorFlow. This model has been selected for its robust performance in image and text recognition tasks, which are crucial for analyzing ingredient information from scanned product images. Initial tests target an accuracy rate of over 90% in categorizing harmful and safe ingredients, helping users to quickly and reliably identify potential risks in the products they use [3]

Discussion-Conclusion: The AI-supported personal care product content analysis system aims to transform how users engage with product information by providing accessible, accurate, and timely insights into ingredient safety. By fostering greater transparency, this system contributes to a safer and more informed consumer experience in the personal care industry

Keywords: AI, personal care, product safety, ingredient analysis, TensorFlow

REFERENCE

[1] e. a. Andersen FA, "Cosmetic Ingredient Review Expert Panel: Evaluation of Safety of Cosmetic Ingredients in Light of Endocrine Disruption Potential," *SAGE Publications,* 2017.

[2] A. M. Khan AD, "Cosmetics and Their Associated Adverse Effects: A Review.," *Journal of Applied Pharmaceutical Sciences and Research,* pp. 1-6, 2019.

[3] A. I. S. a. G. H. Krizhevsky, "ImageNet classification with deep convolutional neural networks," *Neural Information Processing Systems,* 2012.

AI-BASED GAIT ANGLE ANALYSIS WITH YOLOv11 METHOD

Seda Şahin[1*], Mehmet Durmaz[2]

[1] Computer Engineering, Çankırı Karatekin University, Çankırı, Türkiye

[2] Computer Engineering, Undergraduate student, Çankırı Karatekin University, Çankırı, Türkiye

** Corresponding author: sedasahin@karatekin.edu.tr*

Introduction-Aim: Gait analysis plays an important role in the detection of biomechanical disorders and rehabilitation processes by examining the movement patterns of individuals. In this study, it was aimed to detect certain keypoints in the human body and analyze the angles between these points by using the deep learning-based YOLOv11 model in order to provide early diagnosis of gait disorders. Angle analysis performed on key points helps identify physical disorders by detecting deviations in walking patterns. This model has been optimized for use in various clinical and rehabilitation applications with high accuracy and fast processing time [1], [2].

Materials-Methods: Within the scope of this study, the YOLOv11 model was trained to detect certain key points in the human body such as knees, hips and ankles [3]. A trigonometry-based algorithm has been developed to calculate the angles between detected points. Videos of individuals with various walking disorders were used in the training and testing stages. The performance of the model was evaluated with performance measures such as accuracy, processing time, sensitivity and specificity [4].

Results: According to the results obtained, the system using the YOLOv11 model enabled the identified key points to be detected with a high accuracy rate. Analyzes made on the angles between the points were able to detect significant deviations indicating gait disorders. These results show that the model has a successful performance in detecting gait disorders [5].

Discussion-Conclusion: In this study, the method developed using the deep learning-based YOLOv11 model in detecting gait disorders achieved successful results. These findings show that the system can be used effectively for gait analysis in clinical and sports fields. In the future, it is planned to increase the performance of the model by training it with more data sets and to customize it for different types of gait disorders.

Keywords: Angle analysis, Artificial Intelligence, Gait analysis, Gait disorders, YOLOv11

REFERENCE

[1] Z. Cao, T. Simon, S. Wei, and Y. Sheikh, "Realtime multi-person 2D pose estimation using part affinity fields," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Honolulu, HI, USA, 2017, pp. 7291-7299.

[2] H. Wang, C. Zhang, J. Wang, and Z. He, "Human gait recognition based on body skeleton keypoints and temporal information," IEEE Access, vol. 8, pp. 138102-138114, 2020.

[3] K. Sun, B. Xiao, D. Liu, and J. Wang, "Deep high-resolution representation learning for human pose estimation," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Long Beach, CA, USA, 2019, pp. 5693-5703.

[4] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, "2D human pose estimation: New benchmark and state of the art analysis," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Columbus, OH, USA, 2014, pp. 3686-3693.

[5] A. Toshev and C. Szegedy, "DeepPose: Human pose estimation via deep neural networks," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Columbus, OH, USA, 2014, pp. 1653-1660.

MACHINE LEARNING-SUPPORTED WEB APPROACH FOR CHILD FOOD ALLERGY DETECTION

Seda Şahin[1*], Mert Gürpınar[1]

[1] Computer Engineering, Çankırı Karatekin University, Çankırı, Türkiye

** Corresponding author: sedasahin@karatekin.edu.tr*

Introduction-Aim: Food allergies are reported to impose a significant and growing health burden among children w"orldwide (Prescott et al., 2013). In this study, a web application was developed to predict allergy types in children using Machine Learning techniques based on food allergy data. The application aims to forecast potential allergy types by evaluating parameters such as age, weight, genetic predisposition and food consumption habits collected from children. To enhance the accuracy of the Machine Learning model, the training data was cleaned through data preprocessing steps and model optimization was performed. As a result, it serves as a useful tool for the early detection of potential allergies. This study aims to assist children in developing healthier eating habits and protecting them from risks through early allergy diagnosis.

Materials-Methods: In this study, the Python programming language was used to develop a prediction model based on food allergy data using Machine Learning. The data consists of various parameters collected from children such as age, weight, food consumption habits and genetic predisposition. The data was processed and cleaned using the pandas and NumPy libraries. Data cleaning and preprocessing stages are crucial for improving the accuracy of Machine Learning models (Kuhn & Johnson, 2013). The Support Vector Machine (SVM) algorithm was chosen to solve the classification problem because SVM is a suitable method for accurately separating different classes and performs well with small datasets. The model training was carried out by splitting the data into 80% for training and 20% for testing. The scikit-learn library was used to train the model and its performance was measured by accuracy rate. During the training process, hyperparameter optimization was applied to achieve the best classification results.

Results: The Machine Learning model is developed by using the Support Vector Machine (SVM) algorithm which is achieved a high accuracy rate in evaluations conducted on the test data. During training, the model accurately classified the allergy types present in the dataset and it was able to make successful predictions on new data. The results indicate that the model has the ability to accurately predict potential food allergies based on data such as age, weight and food consumption information collected from children.

Discussion-Conclusion: In this study, a Machine Learning model was developed using the Support Vector Machine (SVM) algorithm to predict potential food allergies in children. The success of the SVM algorithm, particularly in classification tasks and its low error rate which supports the model's potential as a reliable predictive tool. This model can serve as a supportive tool for the early diagnosis of food allergies and can contribute to the development of healthy eating habits in children. In conclusion, it is recommended to train the model with larger datasets and to utilize different Machine Learning algorithms to enhance its accuracy

(Breiman, 2001). Additionally, the application of various Machine Learning algorithms is important to reach broader audiences.

Keywords: Child Health, Food Allergy, Machine Learning, Support Vector Machine (SVM)

REFERENCE

[1] Prescott, S., Pawankar, R., Allen, K., & Campbell, D. E. (2013). A global survey of changing patterns of food allergy burden in children. World Allergy Organization Journal, 6(1), 1-12. doi:10.1186/1939-4551-6-21.

[2] Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. Springer.

[3] Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. doi:10.1023/A:1010933404324.

DATA QUALITY-BASED ADAPTIVE LEARNING RATE: A CASE STUDY ON MEDICAL TEXT CLASSIFICATION

Ali Bayram[1*], Banu Diri[1], Savaş Yıldırım[2]

[1] Computer Engineering, Yıldız Technical University, İstanbul, Türkiye

[2] Computer Engineering, İstanbul Bilgi University, İstanbul, Türkiye

** Corresponding author: malibayram20@gmail.com*

Introduction-Aim: Optimizing the learning rate is crucial for deep learning model performance. Traditional approaches often assume uniform reliability across all training data; however, in real-world applications, data quality and reliability can vary significantly, which is especially relevant in specialized fields such as medical text classification [1, 2]. This study proposes an adaptive learning rate strategy that evaluates data reliability and quality based on the expertise level of data providers, dynamically adjusting the learning rate for each training example. We demonstrate the effectiveness of this approach using a real-world medical text classification dataset containing 167,000 samples, showcasing improvements in convergence rate and model performance through data-quality-driven learning rate adjustments.

Materials-Methods: Traditional deep learning methods use a fixed learning rate for all training samples or gradually decrease the learning rate as training progresses. These methods overlook differences in data quality and reliability, treating all data homogeneously despite diverse sources. Studies have shown that adaptive learning rate techniques, such as those incorporating the Barzilai-Borwein method, can enhance model performance by adjusting for data variability [2, 3].

Proposed Formulation

The adaptive learning rate in this study is computed using a simple, effective formula that factors in the expertise of the data provider:

$$\eta\_i = \eta\_base * w\_title$$

Where:

- η_i: Learning rate for the i-th sample
- η_base: Base learning rate (e.g., 0.001)
- w_title: Weight factor based on the provider's title

Results: The adaptive learning rate approach was tested on a medical text classification dataset, resulting in faster convergence and improved performance [1, 3]. The dynamic learning rate adjustments, informed by provider expertise levels, allowed for more efficient model training.

Discussion-Conclusion: This study introduces a novel adaptive learning rate strategy that integrates data quality indicators into the learning process, enhancing model performance and training efficiency in medical text classification. The approach is particularly beneficial in high-stakes domains with variable data reliability. Future research may focus on optimizing weight factors, developing dynamic weight-update mechanisms, and testing in scenarios with multiple experts and in other domains to enhance the strategy's adaptability.

Keywords: adaptive learning rate, data quality, medical text classification, deep learning, dynamic optimization

REFERENCE

[1] Ghada Ben Abdennour, Karim Gasmi, and Ridha Ejbali. "An Optimal Model for Medical Text Classification Based on Adaptive Genetic Algorithm." Data Science and Engineering, 2024. (SpringerLink)

[2] Jinxiu Liang et al. "Barzilai-Borwein-based Adaptive Learning Rate for Deep Learning." Pattern Recognition Letters, 2019. (Cpb Us W2)

(3) Zhiyong Hao et al. "Adaptive Learning Rate and Momentum for Training Deep Neural Networks." arXiv preprint arXiv:2106.11548, 2021. (arXiv)

AUTOMATED SLEEP ANALYSIS USING DEEP LEARNING METHODS ON VISUAL DATA

Emre Turan[1], Orhan Er[2]

[1] Department of Computer Engineering İzmir Bakırçay University, İzmir, Türkiye

Introduction - Purpose: Sleep is a fundamental requirement for human health, and the biological and physiological changes that occur during sleep provide essential information about an individual's overall health status. Currently, methods such as polysomnography are used to diagnose sleep disorders; however, these procedures are time consuming and require expert interpretation. The aim of this study is to develop a model that can automatically analyze biometric and behavioral changes that occur during sleep using image-based analysis methods to classify sleep disorders. Specifically, it will identify different sleep stages and common sleep disorders such as sleep apnea using deep learning based image processing techniques. This will make sleep analysis quicker and more practical, improving the efficiency of medical processes.

Materials and methods: In this study, visual data from 300 patients with suspected sleep disorders were preprocessed before being applied to deep learning models. These preprocessing steps included noise reduction, contrast enhancement, and highlighting of movements that occur during sleep. The pre-processed images were then analyzed using deep learning models such as the Convolutional Neural Network (CNN). The CNN model is ideal for detecting specific biometric data and behavioral changes during sleep, allowing automatic identification of sleep stages. The dataset used consists of video and image recordings obtained from patients diagnosed with sleep disorders, and each image is labelled according to sleep stages. The datasets used in the training phase were carefully selected to improve the accuracy and generalization of the model.

Results: The experiments showed that deep learning models could classify sleep stages with high accuracy and detect specific sleep disorders such as sleep apnea. The accuracy of the model was further improved using advanced image processing techniques. Analyses of the CNN model showed that it could accurately distinguish sleep stages and capture irregularities that can occur during sleep with high accuracy. The presence of high frequency components played a significant role in the detection of sleep disorders such as sleep apnea. This study demonstrates the effectiveness and potential of image-based assessment methods for sleep analysis.

Discussion-Conclusion: The potential application of deep learning techniques in the medical field is increasing day by day, and this study demonstrates the applicability of these techniques for the detection of sleep disorders. Image-based analysis provides a valuable opportunity for automatic detection of sleep disorders. Image-based deep learning models offer a promising approach with high accuracy for early diagnosis of serious health problems such as sleep apnea. The results of the study show that deep learning methods can be an effective tool for sleep analysis. This technological approach is expected to provide significant support to medical professionals in evaluating patients, ensuring speed and accuracy in the diagnostic process.

Keywords: deep learning, sleep apnea, automated classification, biometric analysis.

PATIENT TRUST AND USER EXPERIENCE IN GENERATIVE AI-BASED HEALTHCARE APPLICATIONS: A LITERATURE-BASED REVIEW ON THE ROLE AND EFFECTS OF HUMAN-MACHINE INTERACTION

Yağmur Duru [1*]

[1] *Department of Management Information Systems, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: 6036016@bakircay.edu.tr*

Introduction-Aim: Today, generative AI technologies, like artificial intelligence technologies, have begun to be seen in every sector, and continue to find an important place in the healthcare sector. Types of artificial intelligence; natural language processing, machine learning, robotic process automation, rule-based expert systems, and various types; offer solutions that facilitate the healthcare system in areas such as early diagnosis, treatment, patient monitoring, electronic health records, personal health management, and drug discovery, and make these processes more creative by producing more creative content with generative artificial intelligence-based applications [1], actively using data by giving human-like answers and making them customizable. We can provide ChatGPT as an example of generative AI technology. Generative AI has capabilities such as producing voice and images [2]. For this reason, it is not the right approach to evaluate artificial intelligence technologies and generative AI technologies together. Because generative AI is not only an information system, but also in broader contexts, beyond providing information, these technologies establish a dynamic interaction with users and have the potential to produce content. This study examines from a theoretical perspective how patient trust and user experience are shaped within the framework of human-machine interaction in generative artificial intelligence-based healthcare applications.

Materials-Methods: This research theoretically examines the approaches and trust levels of patients and healthcare professionals towards generative AI technologies and applications and addresses its role in user acceptance. There is a difference between the trust and acceptance levels of generative AI and artificial intelligence applications. Therefore, it is not a correct approach to classify generative AI applications according to artificial intelligence applications. Considering the characteristics of generative AI, it has some unique features and risks compared to AI applications for patient trust. Considering these, the effects of patients' and healthcare professionals' trust levels on user acceptance are analyzed with models such as the Technology Acceptance Model (TAM) used in literature research. The TAM model defines the basic determinants of user experience by determining the relationship and expectations of users who will use the application with the application and technology [3, 4]. As a result, it affects the reliability of applications and their perceptions of reliability. Studies mostly focus on AI and these analyses are conducted accordingly. Similar perceptions can be mentioned for generative AI, but it is necessary to evaluate generative AI by considering its specific situations. In this context, in the design of generative AI applications in healthcare services, ensuring that users can easily access information and obtain reliable results from the applications are among the important elements that strengthen the trust of patients and healthcare professionals.

Results: Generative AI, although not widespread, offers personalized suggestions for patients' personal needs, while also supporting healthcare professionals in decision-making processes and offers important innovations. Generative AI is used in wearable devices, mobile applications, or e-health services [5]. But one of the features of generative AI is production. It can collect, analyze, and report data and produce new results. ChatGPT-based health chatbots can be given as an example. These applications have the potential to examine patients' health conditions and prevent vital situations [6]. However, in addition to these positive effects, these health applications produced with generative AI must be found reliable by both healthcare professionals and patients to diversify and develop further in the future. Patient trust and user acceptance are of critical importance at this point. One of the most important features that distinguishes the use of generative AI in the field of health from other fields is that it contains vital risks. These technologies can work based on data and produce content, and since they can communicate like humans, they need to receive medical training, and this plays an important role in the trust-building process. The main factors affecting patient trust in the literature include medical error concerns, data privacy, ethical issues, and the accuracy of recommendations provided by AI [7]. Trust, user acceptance, and use of these technologies will have a direct impact on the future and success of generative AI technologies in the healthcare sector [8].

Discussion-Conclusion: As a result of this review, by evaluating the effects of generative artificial intelligence applications on patient and healthcare personnel trust and user experience, it was determined that the most important step of technological advances in the healthcare sector is trust creation. One of the most important differences that distinguishes generative artificial intelligence applications from traditional artificial intelligence applications is that they have a human-like creative and personalized approach. And this is an important need. In this context, for generative AI applications to be used, trust should be increased, they should be evaluated in terms of stability and consistency, and data privacy and ethical principles should be carried out with transparency. This study emphasizes that a trust-focused conceptual framework should be created to improve patient and personnel trust and user experience by highlighting the differences between generative AI-based healthcare applications from traditional AI applications on patient trust.

For future studies, the focus should be on exploring the unique aspects of generative AI applications in the healthcare sector and evaluating them with deeper perspectives on how they improve patient trust, as examining generative AI applications solely as AI information systems may cause us to overlook other factors that affect user engagement and trust. Developing frameworks that specifically address unique features will be important to advance and improve generative AI applications in patient- and user-centered healthcare.

Keywords: Generative AI, Healthcare Applications, Patient Trust, Human-Machine Interaction, User Experience

REFERENCE

[1] T. Davenport and R. Kalakota, "The Potential for Artificial Intelligence in Healthcare," *Future Healthcare Journal*, vol. 6, no. 2, pp. 94–98, Jun. 2019, doi: https://doi.org/10.7861/futurehosp.6-2-94.

[2] Mindy Nunez Duffourc and S. Gerke, "Generative AI in health care and liability risks for physicians and safety concerns for patients," *JAMA*, vol. 330, no. 4, Jul. 2023, doi: https://doi.org/10.1001/jama.2023.9630.

[3] "(PDF) Technology Acceptance Models in Health Informatics: TAM and UTAUT," *ResearchGate*. https://www.researchgate.net/publication/335189610_Technology_Acceptance_Models_in_Health_Informatics_TAM_and_UTAUT

[4] İBozkurt, “Teknoloji Kabul Modeli Çerçevesinde Sağlık Profesyonellerinin Yeni Tedavi Yöntemlerini Kullanma Eğilimlerinin İncelenmesi (Özel Hastane Örneği) Investigation Of The Tendencies Of Health Professionals To Use New Treatment Methods In Terms Of Technology Accepted Model (Special Hospital Example),” *Gevher Nesibe Journal IESDR*, vol. 5, no. 7, pp. 88–100, Aug. 2020, doi: https://doi.org/10.46648/gnj.98.

[5] Gamze YORGANCİOGLU TARCAN, Pınar YALÇIN BALÇIK, and Nihat Barış SEBİK, “Artificial Intelligence in Healthcare in Türkiye and the World,” *Lokman hekim dergisi*, vol. 14, no. 1, pp. 50–60, Jan. 2024, doi: https://doi.org/10.31020/mutftd.1278529.

[6] P. Zhang and Maged, “Generative AI in Medicine and Healthcare: Promises, Opportunities and Challenges,” *Future Internet*, vol. 15, no. 9, pp. 286–286, Aug. 2023, doi: https://doi.org/10.3390/fi15090286.

[7] “Elsevier,” *www.elsevier.com*, 2024. https://www.elsevier.com/resources/hospitals-and-health-system (accessed Nov. 01, 2024).

[8] B. Khan *et al.*, “Drawbacks of Artificial Intelligence and Their Potential Solutions in the Healthcare Sector,” *Biomedical Materials & Devices*, vol. 1, no. 36785697, pp. 1–8, Feb. 2023, doi: https://doi.org/10.1007/s44174-023-00063-2.

THE ROLE OF ARTIFICIAL INTELLIGENCE IN THE DIGITAL TRANSFORMATION OF THE HEALTHCARE INDUSTRY IN TÜRKIYE

***Sevgi Cip*[1*]**

[1] *Department of Management Information Systems, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: 6036016@bakircay.edu.tr*

Introduction-Aim: The digital transformation of the healthcare industry gained rapid momentum due to the needs that emerged during and after the pandemic (2019), along with advancements in artificial intelligence. Developments in science and technology have initiated a swift transformation journey in the healthcare sector, as they have in industries such as manufacturing and services. In healthcare, management and clinical processes have transitioned into digital transformation. Artificial intelligence technology has rapidly adapted to the healthcare sector by providing various services, from diagnosis to personalized treatment plans. AI is reshaping healthcare service processes by reducing both administrative and clinical costs. With AI, processes such as diagnosis, treatment, and therapy are accelerated, while human errors are minimized, aiming to improve the quality of healthcare services. [1] This study examines how AI shapes the digital transformation in healthcare, improves patient outcomes, enhances operational and administrative efficiency, and provides solutions to the challenges faced in healthcare services. Soon, because of the accelerated AI research following the pandemic, the virtualization of healthcare services in administrative and clinical aspects and the widespread use of AI-based applications is expected.

Artificial intelligence has also started to be used in many sectors in Türkiye. The Ministry of Health of Türkiye has implemented AI-based applications, particularly in the healthcare sector. These applications aim to reduce costs while improving the quality of healthcare services.[2]

Materials and Methods: This study systematically examines recent developments in artificial intelligence applications in healthcare, including digital health applications, diagnostic imaging, electronic health record management, and other patient monitoring systems. The research is based on quantitative data from peer-reviewed journals, case studies, and healthcare sector reports. Article research process: the research was conducted in the field of healthcare using keyboard combinations such as ‘’ Artificial intelligence’’, ‘’digital strategy’’, ‘’digital transformation in healthcare’’,’’ computer strategy’ ’and ‘’information technology.’’ The research utilized platforms like Scope and YÖK. Several articles were excluded due to limited access to full content.

Results: Digital transformation involves virtualizing traditional methods and concepts such as autonomous systems and unmanned workflows. Compared to other industries transforming like Industry 4.0, the healthcare sector had lagged. [2]However, factors such as population growth, increasing healthcare needs, the pandemic, and the rise in health literacy have necessitated fundamental changes in healthcare services. Although healthcare has been slower to adopt digital transformation than other sectors, it has quickly adapted. When considering digital technologies in healthcare, applications such as personal health monitoring, e-Nabız,[3] the central hospital appointment system(MHRS), and the concept of virtual hospitals come to mind. In a digitized healthcare environment, unlike the traditional physical hospital setting, health services can be delivered remotely by transferring data

through cloud-based centers, allowing expert healthcare professionals to access this information remotely via the internet. In Türkiye, the Hayat Eve Sığar application, which emerged during the pandemic, is an AI-powered application integrated with the [4] MHRS, e-Nabız, and E-Government platforms.

[5] These digital applications optimize processes outside of hospitals, allowing individuals to access their health data, such as test reports, examination records, and previously issued prescriptions, stored in various data centers nationwide. Through AI, the time spent in healthcare facilities is minimized, unnecessary testing is reduced, costs are lowered, and the doctor-patient relationship is strengthened. Additionally, AI has automated tasks such as data entry and management of electronic health records, alleviating the burden on healthcare professionals.[6] The use of artificial intelligence (AI) in healthcare brings some risks along with its advantages. Integrating AI into decision-making processes for diagnosis and treatment is sometimes perceived as a threat on a global level, and the digitalization of diagnosis and treatment processes may increase job-related concerns. AI applications can lead to reduced physical and mental activity for individuals, potentially increasing the risk of psychological, neurological, or musculoskeletal issues. Negative effects such as technology dependence, social isolation, loss of self-esteem, and reduced physical contact may also occur. Additionally, AI systems may have disadvantages like coding errors, high costs, difficulties in updates, measurement errors, and ethical concerns. AI often uses large amounts of personal or sensitive data; if this data's privacy is not secured, it risks misuse by malicious parties. Lack of diversity or biased data in training sets can cause AI to learn these biases, resulting in unfair decisions. Sometimes, AI may misunderstand interactions with people, which can lead healthcare professionals to overlook their own experiences. When AI models are trained on insufficient or inaccurate data, they can produce incorrect diagnoses or treatment recommendations, potentially putting patient health at risk. For example, if an AI tool assisting a doctor in the diagnostic process makes an error in diagnosis, questions remain unclear regarding who owns the AI system, who oversees it, and who is accountable. Therefore, this study highlights the necessity of establishing international ethical standards on how personal health data will be used, stored, and shared and who will hold responsibility when processing such data.

Discussion-Conclusion: Artificial intelligence has greatly advanced the digital transformation of the healthcare sector by securely recording patient outcomes and streamlining operational workflows.[7] New developments in automation are taking place daily across various subfields. There are numerous applications for both administrative and clinical processes. In clinical workflows, AI speeds up essential tasks such as diagnosis, treatment, and therapy to enhance service quality while minimizing human interaction.

Recent research emphasizes the importance of ongoing education for healthcare professionals to effectively leverage these advanced technologies. [8] This training is crucial for maintaining patient trust and responding quickly and effectively to pandemics or large-scale outbreaks. Some studies also support that there are negative aspects of artificial intelligence in the healthcare field. In a study conducted by Jarahi in 2018, it was noted that although artificial intelligence has made significant advancements, it cannot reach the level of emotional sensitivity, intuition, and creativity that a human can possess.[9]

Another study suggests that artificial intelligence is limited to executing predefined tasks and may be prone to errors or system failures when encountering non-routine situations, potentially leading to challenges in maintaining usage standards.[10]

While artificial intelligence presents certain disadvantages, it is imperative to take swift action to align these applications with societal socio-cultural and ethical values, patient rights, and existing legal frameworks, ensuring they are monitored and implemented appropriately.

In Türkiye and globally, the digital transformation of healthcare enabled by artificial intelligence is reducing costs and facilitating more efficient and rapid outcomes in diagnostic and treatment processes. In conclusion, the integration of AI and similar technologies into healthcare is unavoidable.

Keywords: Digital transformation, Artificial intelligence, Digitalization in healthcare.

REFERENCE:

[1] D. D. E. Y. Altuntaş, *Sağlık Hizmetleri Uygulamalarında Dijital Dönüşüm*. Eğitim Yayınevi, 2019.

[2] "Kurumların dijital dönüşüm süreçlerinin incelenmesi: Bir sağlık kurumu için öneri = Examining the digital transformation of institutions: Proposal for a health instution". Erişim: 31 Ekim 2024. [Çevrimiçi]. Erişim adress: https://platform.almanhal.com/Details/Thesis/2000234032?lang=tr

[3] "Yardım | e-Nabız." Erişim: 31 Ekim 2024. [Çevrimiçi]. Erişim address: https://enabiz.gov.tr/Yardim/Index

[4] "HAKKIMIZDA". Erişim: 31 Ekim 2024. [Çevrimiçi]. Erişim address: https://www.mhrs.gov.tr/hakkimizda.html

[5] A. Holzinger, E. Weippl, A. M. Tjoa, ve P. Kieseberg, "Digital Transformation for Sustainable Development Goals (SDGs) - A Security, Safety and Privacy Perspective on AI", içinde *Machine Learning and Knowledge Extraction*, c. 12844, A. Holzinger, P. Kieseberg, A. M. Tjoa, ve E. Weippl, Ed., içinde Lecture Notes in Computer Science, vol. 12844. , Cham: Springer International Publishing, 2021, ss. 1-20. doi: 10.1007/978-3-030-84060-0_1.

[6] "Dijital Sağlıkta Yapay Zekâ: Albert Health'in Rolü", Türkiye Yapay Zeka İnisiyatifi. Erişim: 31 Ekim 2024. [Çevrimiçi]. Erişim adresss: https://turkiye.ai/dijital-saglikta-yapay-zeka-albert-healthin-rolu/

[7] S. A. Alowais *vd.*, "Revolutionizing healthcare: the role of artificial intelligence in clinical practice," *BMC Med. Educ.*, c. 23, sy 1, s. 689, Eyl. 2023, doi: 10.1186/s12909-023-04698-z.

[8] C. Da Silva *vd.*, "Rethinking the Continuous Education and Training of Healthcare Professionals in the Context of Digital Technologies," içinde *Handbook of Research on Instructional Technologies in Health Education and Allied Disciplines*, 2023, ss. 105-129. doi: 10.4018/978-1-6684-7164-7.ch005.

[9] M. H. Jarrahi, "Artificial intelligence and the future of work: Human-AI symbiosis in organizational decision making," *Bus. Horiz.*, c. 61, sy 4, ss. 577-586, 2018.

[10] S. Bhosale, V. Pujari, ve M. Multani, *Advantages And Disadvantages Of Artificial intelligence*. 2020.

ANALYSIS OF HEALTH DATA FROM MOBILE DEVICES WITH DEEP LEARNING

Beste Tokpınar[1*], Nihan Özbaltan[1]

[1] Department of Computer Engineering İzmir Bakırçay University, İzmir, Türkiye

* Corresponding author: 6053004@bakircay.edu.tr

Introduction-Aim: Scales provide important information about the health status of individuals by providing basic body composition measurements such as weight, body fat percentage, muscle mass and hydration level. This study aims to analyze the data collected from Bluetooth-enabled scale devices using deep learning techniques through a mobile application developed to track personal body mass index. The research will be carried out on a small dataset obtained from 10 individuals, and through the processing of this data, its usability for predicting health risks such as obesity, dehydration, muscle loss, diabetes and cardiovascular problems, monitoring long-term health indicators and providing personalized health recommendations will be evaluated.

Materials-Methods: In the study, deep learning models such as LSTM and CNN and anomaly detection techniques will be used. Real-time data collected via Bluetooth connection will be processed to analyze changes in the body composition of each individual. In the study, time series data will be made suitable for model training by applying data preprocessing and normalization procedures. At this stage, measurements such as weight, body fat, muscle mass and water content collected daily from 10 individuals will serve as the basis for a simple health risk assessment**.**

Results: The results of this preliminary study are planned to be targeted as follows: To evaluate the potential of deep learning models in identifying health risks with data obtained from Bluetooth-enabled scale devices. Using models such as LSTM and CNN, it will be examined that changes in body composition can indicate risks such as obesity, diabetes and cardiovascular problems and that future measurements can be predicted. Analyses on daily data collected from 10 individuals will reveal the model performance and its effectiveness in anomaly detection, aiming to contribute to the monitoring of individuals' health status and the development of practical solutions through mobile applications. In this way, it will be possible for individuals to regularly monitor their health status and take preventive measures against potential health problems.

Discussion-Conclusion: This preliminary study was conducted to understand the potential benefits of analyzing biometric data from Bluetooth-based scale devices with machine learning models for personal health management. Analyses on a 10-person dataset were conducted to assess the capacity of deep learning models to detect specific clues to health risks. It is envisaged that such analyses will help individuals to continuously monitor their health status and take preventive health measures. The use of such deep learning models, especially in personalized assistants and mobile health applications, can enable individuals to make more informed and effective health management by providing user-specific health recommendations. In the future, it is recommended to focus on improving model accuracy with larger data sets, integrating new health metrics and expanding the capabilities of personalized health assistants. These approaches may contribute to making mobile health applications more effective and user-centred.

Keywords: Bluetooth scale device, deep learning, LSTM, CNN, health risk analysis

CLASSIFICATION OF GOUT WITH LOGISTIC REGRESSION AND K-NEAREST NEIGHBOUR ALGORITHMS

Bayram Köse[1], Bahar Demirtürk[2*], Şükran Konca[2], Fatma Demet Arslan[3], Eslem Serra Çolak[1]

[1] Electric-Electronic Engineering Department, İzmir Bakırçay University, İzmir, Türkiye

[2] Fundamental Sciences Department, İzmir Bakırçay University, İzmir, Türkiye

[3] Medical School, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: bahar.demirturk@bakircay.edu.tr*

Introduction-Aim: Gout is a disease that has become a public health problem as it is a disease that accompanies chronic diseases such as chronic kidney disease, kidney stones, hypertension, obesity, diabetes and cardiovascular disease, which constitute a significant burden for global health systems and is increasing every year.

Gout, a painful form of arthritis, is a disease that occurs in periods of exacerbation of symptoms such as severe pain and swelling when there is excess uric acid in the human body, usually with the formation of sharp crystals in the joints. This brings along problems such as a decrease in the quality of life of the individual and loss of labour force in daily life. Therefore, early diagnosis and effective management are critical to slow the progression of the disease and reduce complications. Successful early diagnosis methodologies will contribute to the protection of the country's health systems from high costs by protecting healthy physical structures at the individual level on the one hand, and on the other hand, by minimising the rate of diseases that are very expensive to treat and their permanent effects on individual health structures.

Materials-Methods: In this study, the performance of Logistic Regression (LR) and K-Nearest Neighbour (K-NN) algorithms for gout classification is comparatively investigated. An anonymous data set was used in the study and analyses were performed on 2614 patient records in the data set. The characteristics used in the study are age, gender and uric acid level. The performance metrics of both algorithms are analysed in detail and their effectiveness in clinical applications is compared. The findings of the study reveal that both algorithms can be used reliably for clinical decision support. This study shows that artificial intelligence technologies can be used as an effective tool in the diagnosis of gout, one of the most common forms of inflammatory arthritis that causes pain in joints and soft tissues and provides guidance for future studies. It is presented to draw attention to the potential of artificial intelligence technologies to expand their applications in the field of health and to emphasise the role of artificial intelligence in the management of common health problems such as diabetes. There are two basic elements that shape the reliability of the results of a scientific study. One of these elements is the reliability of the data used in the research and its potential to represent the generality. The other is the validity of the data. Based on these two elements, an anonymised data set consisting of 2614 data was used in this study. In fact, when the studies in the literature are examined, it is seen that there are data sets with a wider patient profile. However, this study derives its originality from the use of real data.

Results: The datasets consisted of various medical predictor variables such as uric acid level, age and gender. These data were obtained from 2614 potential patients in total.

In the analyses performed on the dataset within the framework of the stated purpose, both models were trained using the specified features. The performance of both models was evaluated with metrics such as accuracy, sensitivity, specificity and F1 score.

Discussion-Conclusion: The results of the study indicate that Logistic Regression and K-NN algorithms can be effective tools for the classification of gout. These algorithms can be used in real-time health monitoring systems and clinical settings to contribute to the early diagnosis of the disease. This contribution of artificial intelligence and machine learning techniques to the early diagnosis of gout, as well as various chronic diseases, requires both the restructuring of health systems based on artificial intelligence and the construction of new social structures and relationship networks on the basis of patient-patient relatives and healthcare professionals to manage treatment processes based on artificial intelligence and machine learning. Considering this situation, it is necessary to design processes for the construction of new health eco-systems by taking into account the artificial intelligence-machine learning and social dimension trilogy.

Keywords: Gout, Machine Learning, Logistic Regression, K-Nearest Neighbour Algorithm, Clinical Decision Support Systems

REFERENCE

[1] Cao, S., & Hu, Y. (2024). Interpretable machine learning framework to predict gout associated with dietary fiber and triglyceride-glucose index. *Nutrition & Metabolism*, *21*(1), 25.

[2] Brikman, S., Serfaty, L., Abuhasira, R., Schlesinger, N., Bieber, A., & Rappoport, N. (2024). A machine learning-based prediction model for gout in hyperuricemics: a nationwide cohort study. *Rheumatology*, keae273.

[3] Kumar, M. S., Hudson, B., Priya, V., Kumar, V. A., Tharun, C., & Saran, R. S. (2022, November). Determination of Gout Disease using Machine Learning. In *2022 1st International Conference on Computational Science and Technology (ICCST)* (pp. 11-15). IEEE.

[4] Ichikawa, D., Saito, T., Ujita, W., & Oyama, H. (2016). How can machine-learning methods assist in virtual screening for hyperuricemia? A healthcare machine-learning approach. *Journal of biomedical informatics*, *64*, 20-24.

[5] Wang, M., Li, R., Qi, H., Pang, L., Cui, L., Liu, Z., Li, C. (2023). Metabolomics and machine learning identify metabolic differences and potential biomarkers for frequent versus infrequent gout flares. *Arthritis & Rheumatology*.

[6] Borghi C, Agabiti-Rosei E, Johnson RJ, Kielstein JT, Lurbe E, Mancia G, et al. Hyperuricaemia and gout in cardiovascular, metabolic and kidney disease. Eur J Intern Med. 2020;80:1–11.

[7] Zhang Y, Chen S, Yuan M, Xu Y, Xu H. Gout and Diet: a Comprehensive Review of mechanisms and Management. Nutrients. 2022;14(17):3525.

BONE FRACTURE DETECTION IN MEDICAL IMAGES USING YOLO

Büşra Erdoğan[1], Emre Ölmez[1*]

[1] Biomedical Engineering, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: emre.olmez@bakircay.edu.tr*

Introduction-Aim: Bone fractures are a common occurrence in a variety of situations, including accidents, sporting events, and in cases of osteoporosis [1]. The diagnosis of these fractures is challenging due to the necessity of radiological observation of subtle details within the bone structure. In standard imaging techniques, small fractures, especially in complex anatomical structures, may be overlooked or misinterpreted, which could result in misdiagnoses that could adversely affect patient recovery [2]. The integration of artificial intelligence and machine learning techniques with medical imaging offers the potential for faster and more accurate fracture detection [3]. In particular, deep learning-based models aim to establish a new standard for bone fracture detection.

Materials-Methods: The objective of this study is to detect bone fractures using one of the deep learning algorithms, namely You Only Look Once (YOLO). YOLO is an object detection solution that is renowned for its rapid processing speed and compact model size. The YOLO model processes the image directly in order to determine the coordinates of the bounding box and the category of the object in question [4]. The dataset utilized in this study was procured from the Universe Roboflow platform, which provides support for object detection and classification models. The dataset comprises images of fractures in the elbow, finger, forearm, humerus, shoulder, and wrist [5]. A total of 1,728 images underwent data augmentation, resulting in 4,148. Of these, 3,631 were employed for training, 348 for validation, and 169 for testing. YOLOv9c was selected for training [6].

Results: Table 1 illustrates the precision and recall values obtained with YOLOv9c for each fracture class in the test set. Precision reflects the accuracy of the model's fracture detections, indicating its capacity to maintain a low false-positive rate. In contrast, recall indicates the model's efficacy in identifying existing fractures, reflecting the extent to which the model is able to accurately capture fracture cases.

Tablo 1. YOLOv9c ile Kemik Kırığı Tespit Sonuçları: Sınıf Bazında Precision ve Recall Değerleri

	Precision	Recall
Elbow Fracture	0,50	0,03
Fingers Fracture	0,70	0,15
Forearm Fracture	0,92	0,28
Humerus Fracture	0,95	0,54
Shoulder Fracture	1,00	0,05
Wrist Fracture	0,66	0,07

Discussion-Conclusion: This study offers a comprehensive assessment of the efficacy of the YOLOv9c model in identifying bone fractures. The results demonstrate high precision values, while recall remains relatively low. A high level of precision indicates that the model exhibits a low rate of false positives, suggesting that the majority of detected fractures are accurate. Nevertheless, the low recall rate indicates that the model is unable to detect a considerable number of actual fracture cases.

In general, the YOLOv9c model demonstrates high precision for specific fracture types; however, its detection rate is constrained. The high precision values indicate that the model has the capacity for accurate fracture detection with a low incidence of false positives. However, low recall values indicate instances of missed fracture cases, thereby suggesting potential avenues for improvement. Further research could address these limitations by retraining the model on a more diverse and extensive dataset or by incorporating optimizations to improve recognition of various fracture types. Such optimizations could enhance the model's potential as a reliable tool in clinical applications.

Keywords: Artificial intelligence, Medical Image Processing, Yolov9, Bone Fracture Detection

REFERENCE

[1] C.-Y. Hsu, S.-Y. Huang, and W.-J. Cheng, “Evaluation of the effects of a Chinese herb in nonoperative bone fractures in Taiwan: A Retrospective multi-central database Cohort Study,” *J. Herb. Med.*, p. 100956, Oct. 2024, doi: 10.1016/J.HERMED.2024.100956.

[2] P. M. Ricci *et al.*, “Cone-beam computed tomography compared to X-ray in diagnosis of extremities bone fractures: A study of 198 cases,” *Eur. J. Radiol. Open*, vol. 6, pp. 119–121, Jan. 2019, doi: 10.1016/J.EJRO.2019.01.009.

[3] Q. Yu *et al.*, “Multi-task learning for calcaneus fracture diagnosis of X-ray images,” *Biomed. Signal Process. Control*, vol. 99, p. 106843, Jan. 2025, doi: 10.1016/J.BSPC.2024.106843.

[4] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, “A Review of Yolo Algorithm Developments,” *Procedia Comput. Sci.*, vol. 199, pp. 1066–1073, Jan. 2022, doi: 10.1016/J.PROCS.2022.01.135.

[5] Veda, “bone fracture detection Computer Vision Project,” universe.roboflow.

[6] Ultralytics, “YOLOv9: A Leap Forward in Object Detection Technology,” Ultralytics.

VOICE EMOTION RECOGNITION USING DEEP LEARNING

Seçkin YILDIZ[1*], Emre ÖLMEZ[1]

[1] *Biomedical Engineering, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: 6022013@bakircay.edu.tr*

Introduction-Aim: Emotion recognition is of great importance in artificial intelligence applications, especially in the fields of human-computer interaction and healthcare [1]. In this study, emotion recognition was performed on audio data using deep learning techniques. The accuracy of the emotion classification model was analyzed using the RAVDESS (Ryerson Audio-Visual Database of Emotional Speech and Song) dataset [2]. This study aims to highlight the potential of artificial intelligence applications in the field of emotion analysis.

Materials-Methods: The RAVDESS dataset, which comprises audio recordings expressing a range of emotional states (e.g., anger, happiness, sadness, surprise), was employed in this study. The model employed for emotion classification exhibits a hybrid architectural configuration, integrating Conv1D and LSTM (Long Short-Term Memory) layers. During the training process, dropout layers were employed to prevent overfitting. Initially, in the data preprocessing phase, the audio recordings were divided into one-second segments, and 40 Mel-frequency cepstral coefficients (MFCCs) were extracted from each segment. The dataset was then split into training and validation sets at an 80%-20% ratio, and training was conducted.

Results: The developed model demonstrated a high degree of accuracy, achieving a 90% accuracy rate on the training set and a 66% accuracy rate on the test set. Figure 1 depicts the confusion matrix obtained on the test set, while Table 1 presents the class-based precision, recall values, and F1 scores.

Discussion-Conclusion: The findings of this study indicate that deep learning models can be effectively employed as a means of voice emotion classification. To further enhance the results, it is anticipated that future studies will employ a variety of methods. In this context, the objective is to enhance the model's adaptability to more complex and diverse emotional states by increasing data diversity. The application of data augmentation techniques has the potential to enhance the model's capacity for generalization, thereby improving its performance across a range of scenarios. Additionally, extending the existing model architecture by incorporating deeper layers will enhance the model's capacity to discern more subtle emotional expressions. Moreover, the incorporation of a more extensive data set is anticipated to have a beneficial impact on the model's performance. It is anticipated that these improvements will significantly enhance the model's classification accuracy and reliability.

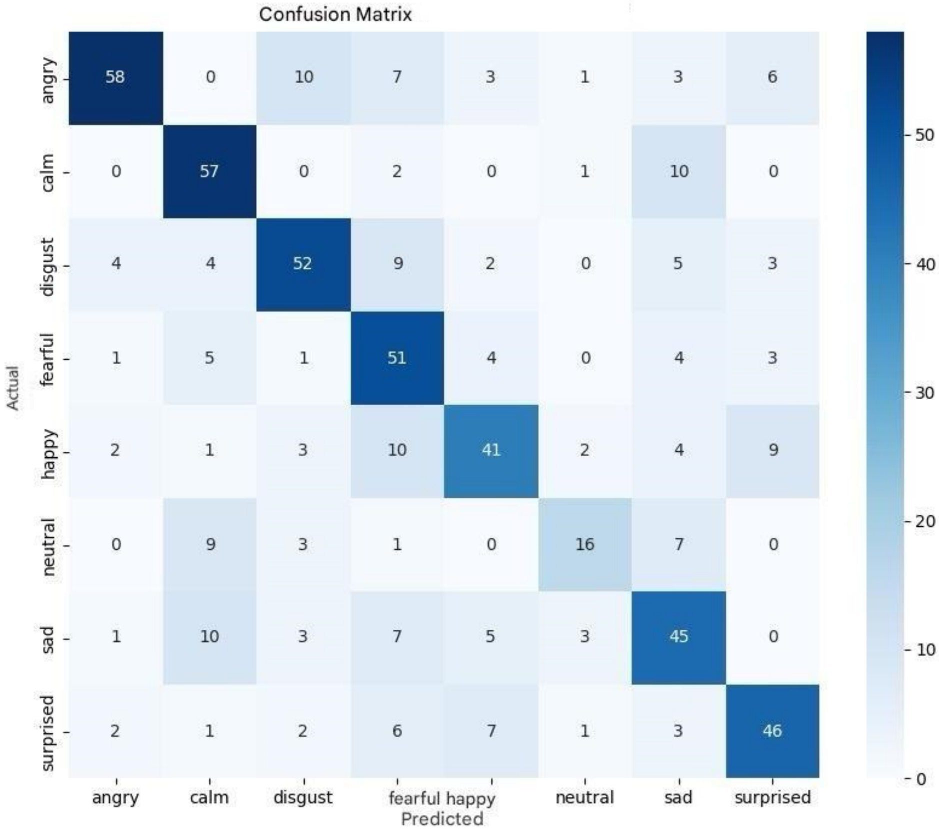

Figure 1: Confusion Matrix of the Test Set

Table 1: Class-Based Evaluation Metrics

Emotional	Precision	Recall	F-1 Score
angry	0.85	0.66	0.74
calm	0.66	0.81	0.73
disgust	0.70	0.66	0.68
fearful	0.55	0.74	0.63
happy	0.66	0.57	0.61
neutral	0.67	0.44	0.53
sad	0.56	0.61	0.58
surprised	0.69	0.68	0.68

Keywords: Emotion Recognition, Deep Learning, Audio Analysis, Artificial Intelligence in Healthcare

REFERENCE

[1] D. Ayata, Y. Yaslan, and M. E. Kamasak, "Emotion Aware Artificial Intelligence for Cognitive Systems: A Case Study Based on Speech Emotion Recognition," Procedia Computer Science, vol. 120, pp. 376–383, 2019, doi: 10.1016/j.procs.2017.11.257.

[2] S. R. Livingstone and F. A. Russo, "The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English," PLOS ONE, vol. 13, no. 5, p. e0196391, 2018, doi: 10.1371/journal.pone.0196391.

[3] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.

EXPERIMENTAL ANALYSIS OF SYNTHETIC DATA GENERATION TECHNIQUES FOR PRIVACY-PRESERVING AI IN MEDICAL IMAGING

Süleyman Yolcu[1*], Elif Özcan[1], Havvanur Dervişoğlu[1], Ruşen Halepmollası[1], Mehmet Haklıdır[1]

[1] TÜBİTAK Informatics and Information Security Research Center, Kocaeli, Türkiye

** Corresponding author: suleyman.yolcu@tubitak.gov.tr*

Introduction-Aim: Medical imaging is an important tool that enables the effective use of machine learning and AI-based models in the early diagnosis and detection of various diseases. However, sharing medical images associated with patients' personal information may compromise privacy. Moreover, using medical images or any patient data without explicit consent for research is unethical. Privacy issues and ethical concerns result in restricted access to medical data. Therefore, there is a critical demand for innovative solutions that respect patient privacy and adhere to ethical guidelines while enabling advancements in AI for healthcare. Meanwhile, synthetic data offers a solution by generating realistic representations to train machine learning models in scenarios where real data is restricted or cannot be used due to privacy or ethical concerns. Also, synthetic data generation can improve model performance and enables more robust and consistent results, especially when large datasets are not available. The approach not only increases data diversity but also addresses cost and privacy issues. In this study, we aim to evaluate three synthetic data generation approaches- Generative Adversarial Networks (GAN), Variational Autoencoders (VAE), and Denoising Diffusion Probabilistic Models (DDPM)- by analyzing their classification performance in diseases (pneumonia, tuberculosis, COVID-19) and the image quality of the synthetic images they produce.

Materials-Methods: The dataset consists of chest X-ray images classified two classes (normal and pneumonia). We applied to images pre-processing steps to ensure data consistency and quality. We employed the synthetic data generation approach when real medical data is limited or unavailable due to data constraints and privacy concerns. It serves as a robust alternative for training machine learning models and improving performance. Particularly when dealing with sensitive data like medical images, synthetic data generation provides a strong alternative to overcome both ethical and legal issues.

In this study, we applied GAN, VAE, and DDPM to generate synthetic images that statistically resemble the original data. These models learned existing data distributions and generated synthetic images that statistically resemble the original data. We conducted a two-step analysis to evaluate the models' performance. First, we tested each model trained on synthetic data with real data, using classification metrics such as accuracy, recall, precision, and F1 score. In the second step, we measured the realism of the generated images using image quality metrics, namely Fréchet Inception Distance (FID) and Kernel Inception Distance (KID). Thus, we evaluated the models in terms of both classification and image generation performance.

Results: We evaluated and compared the performance of classification models in terms of accuracy, precision, recall, and F1-score. Additionally, we compared synthetic data produced by the GAN, VAE, and DDPM with real data to calculate FID and KID scores. According to our results, GAN achieved the best results in terms of accuracy, precision and F1-score with rates 85.10%, 89.23% and 88.21%, respectively. Furthermore, it provides good recall result with

87.27%. In terms of image quality, GAN's FID score was 115.82 and its KID score was 0.13. On the other hand, we observed that VAE exhibited lower performance with 79.65% accuracy, %73.08 precision and %81.78 F1-score. Although it achieved the highest recall rate at 92.83%, it may not be an ideal indicator of overall performance. For image quality, VAE's FID score was 276.66, and KID score was 0.354. DDPM performed close to GAN in classification (86.22% accuracy, 85.19% recall, %94.36 precision and %89.54 F1-score) but showed lower image quality (FID:503.38, KID:0.69). When we tested the classification model with real data, it achieved 81.73% accuracy, 77.60% recall, 99.49% precision, and 87.19% F1- score.

Discussion-Conclusion: Medical imaging plays a critical role in enabling AI and ML models to assist in the early diagnosis and detection of various diseases. However, privacy and ethical issues on patient data often limit access to real medical images. In this study, we demonstrated that synthetic data generation can effectively overcome these limitations by evaluating GAN, VAE, and DDPM models. According to our results, GAN provides the most realistic and effective outcomes for synthetic data generation, particularly when medical images are limited. Our findings indicate the potential of synthetic data—particularly from GANs—as a reliable alternative for training robust AI models when access to real medical data is restricted. Overall, our study highlights synthetic data as a valuable tool to address data constraints in medical imaging and support advancements in AI for healthcare.

Keywords: Healthcare, data privacy, synthetic data

References

[1] Mahaulpatha, P., Abeywardane, T., & George, T. (2024). *DDPM-based X-ray Image Synthesizer*. arXiv. https://arxiv.org/abs/2401.01539

[2] Mohammed, H., & Ali, K. (2022). Generating high-resolution chest X-ray images using CGAN. *Basrah Researches Sciences, 48*(2), 88-101.

[3] Koetzier, L. R., Wu, J., Mastrodicasa, D., Lutz, A., & Chung, M. (2024). Generating synthetic data for medical imaging. *Radiology*. https://pubs.rsna.org/doi/abs/10.1148/radiol.232471

[4] Gajjar, P., Garg, M., Desai, S., & Chhinkaniwala, H. (2024). An empirical analysis of diffusion, autoencoders, and adversarial deep learning models for predicting dementia using high-fidelity MRI. *IEEE Xplore*. https://ieeexplore.ieee.org/document/10400467

[5] Rais, K., Amroune, M., & Haouam, M. Y. (2024). Medical image generation techniques for data augmentation: Disc-VAE versus GAN. *IEEE Xplore*. https://ieeexplore.ieee.org/document/10541221

AI-ASSISTED NECK EXERCISE ASSISTANT

Filiz Meryem Sertpoyraz [1], Berkay Konuk[1], Ayşenur Yağmur Çifci[1*], Büşra ERDOĞAN[2]

[1] Medicine, Physical Medicine and Rehabilitation, İzmir Bakırçay University, İzmir, Türkiye

[2] Biomedical Engineering, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: ygmrciftci@gmail.com*

Introduction-Aim: Chronic neck pain is a prevalent health issue, attributed to factors such as modern lifestyle habits, prolonged computer use, and poor posture. It ranks as one of the most common musculoskeletal disorders after lower back pain, with an incidence exceeding 30% annually. Neck pain adversely affects individuals' quality of life and work performance [1]. During periods of restricted access, such as pandemics, obtaining treatment for neck pain becomes even more challenging [2]. This study aims to develop a personalized exercise assistant that monitors patients to ensure they correctly perform the exercises prescribed by their physicians.

Materials-Methods: In this study, a video database comprising various neck exercises recommended by physiotherapists was created. An AI model, trained using this video dataset, analyzes whether the patient performs the exercise correctly. The system, which uses a camera to monitor the patient’s movements, provides real-time feedback through an interactive interface that encourages correct exercise performance.

Results: Ten neck exercise videos were recorded under the supervision of physiotherapists and analyzed. Movements in these videos were examined using pose detection, and a model to detect patient movements was developed with the MediaPipe library. When the exercise video is played, a camera simultaneously records the patient's movements, which are then evaluated by the model for correctness and reported to the physician. Patients are prompted to repeat the exercises until they perform them flawlessly.

Discussion-Conclusion: The AI-assisted exercise assistant not only increases accessibility to neck pain management but also supports healthcare professionals by reducing their workload and enabling remote monitoring. Future research aims to expand the system's scope to include other musculoskeletal disorders and to evaluate its impact on patient outcomes.

Keywords: AI-Assisted Physiotherapy, Neck Pain Management, Personalized Exercise Assistant, Real-Time Motion Analysis, Health Technologies

REFERENCE

[1] A. M. Wu *et al.*, “Global, regional, and national burden of neck pain, 1990–2020, and projections to 2050: a systematic analysis of the Global Burden of Disease Study 2021,” *Lancet Rheumatol.*, vol. 6, no. 3, pp. e142–e155, Mar. 2024, doi: 10.1016/S2665-9913(23)00321-1.

[2] Y. Lin *et al.*, “Effects of online exercise intervention on physical and mental conditions in young adults with chronic neck pain,” *iScience*, vol. 26, no. 12, p. 108543, Dec. 2023, doi: 10.1016/J.ISCI.2023.108543.

PROGNOSIS PREDICTION WITH MACHINE LEARNING MODELS IN COMMUNITY-ACQUIRED PNEUMONIA

Dilek Orbatu[1], Zeynep İzem Peker Bulğan[1], Sümeyye Akkaya[2], Banu İşbilen Başok[3]

[1] Department of Pediatrics, University of Health Sciences (UHS) İzmir Faculty of Medicine, Dr. Behçet Uz Child Disease and Pediatric Surgery Training and Research Hospital, İzmir, Türkiye

[2] Department of Medical Biochemistry, UHS İzmir Faculty of Medicine, Bozyaka Training and Research Hospital, İzmir, Türkiye

3 Department of Medical Biochemistry, UHS İzmir Faculty of Medicine, Dr. Behçet Uz Child Disease and Pediatric Surgery Training and Research Hospital, İzmir, Türkiye

** Corresponding author: banu.basok@sbu.edu.tr*

Introduction: Community-acquired pneumonia (CAP) is a type of pneumonia acquired from daily life, occurring in individuals who have not been hospitalized in the 7-14 days before symptom onset or developed within the first 48 hours of hospitalization [1]. CAP is one of the most significant causes of morbidity and mortality in childhood worldwide. According to World Health Organization data, 14% of deaths among children under five are due to CAP [2]. Additionally, it significantly impacts the national economy as it is one of the leading causes of hospitalizations

In Türkiye, the guidelines of the Turkish Thoracic Society, which are commonly used, recommend hospitalization for infants under three months [1]. Different guidelines exclude age as a direct indication for hospitalization, recommending admission based on age-specific requirements [3]. Although various guidelines provide recommendations on which children should be treated as inpatients, the parameters for prognosis prediction in these cases are limited.This study aims to develop high-accuracy prognosis prediction machine learning (ML) models for hospitalized CAP patients and identify the variables that influence these models' decisions through explainability analyses.

Materials-Methods: Data on age, gender, complete blood count, C-reactive protein (CRP), procalcitonin, erythrocyte sedimentation rate (ESR), imaging requests (X-ray, CT, and thoracic ultrasonography), and records on outpatient/inpatient treatment of all patients diagnosed with CAP at Dr. Behçet Uz Child Disease and Pediatric Surgery Training and Research Hospital, UHS İzmir Faculty of Medicine, from September 2022 to September 2024 were retrospectively collected from the Hospital Information Management System. The Systemic Inflammatory Index (SII) was calculated using the platelet count x neutrophil/lymphocyte formula [4]. Data analysis was performed using Python's Pandas, Numpy, and Scikit-learn libraries within the Jupyter Notebook interface. ML models such as logistic regression(LR), linear discriminant analysis(LDA), K-nearest neighbor(KNN), decision tree(CART), random forest(RFC), Gaussian naive bayes(NB), support vector machine(SVM), and light GBM(LGB) were developed. Model training and validation were carried out using k-fold cross-validation. The performance metrics accuracy, precision, recall, and F1 scores of the classifiers were calculated. Receiver operating characteristic (ROC) analysis was conducted to evaluate the models' discrimination power. Explainability analyses were performed using SHapley Additive exPlanations (SHAP) and "feature importance" methods.

Results: A total of 14,838 patients were identified with CAP. Due to over 90% data incompleteness in procalcitonin and ESR, these variables were excluded from the dataset. After cleaning incomplete/erroneous data, models were developed on a dataset comprising

9,421 patients. Performance metrics of the developed ML models on training and test sets are presented in Table 1.

Table 1. Machine Learning models performance metrics.

Model	Accuracy		Precision		Recall		F1 Score	
	Train	Test	Train	Test	Train	Test	Train	Test
LR	0.8462	0.851	0.8271	0.8294	0.8462	0.851	0.8173	0.8239
LDA	0.8471	0.8498	0.8275	0.8276	0.8471	0.8498	0.8223	0.8267
KNN	0.87	0.8034	0.8597	0.7631	0.87	0.8034	0.8532	0.7772
CART	1.0	0.7939	1.0	0.8014	1.0	0.7939	1.0	0.7975
NB	0.818	0.8231	0.8027	0.8109	0.818	0.8231	0.8087	0.8161
SVM	0.9998	0.8321	0.9998	0.8603	0.9998	0.8321	0.9998	0.7579
RFC	0.9854	0.8502	0.9856	0.828	0.9854	0.8502	0.9852	0.8258
LGB	0.9685	0.8654	0.9692	0.8523	0.9685	0.8654	0.9675	0.8547

The CART model had 100% metrics in the training set, but its performance in the test set was found to be at the 79-80% level, which was evaluated as overfitting. Although the SVM model achieved very high accuracy in the training set, its accuracy in the test set was found to be 83% with a low F1 score in the test set that exhibited imbalanced performance. The RFC model demonstrated balanced and high performance, with 98.5% accuracy in the training set and 85% in the test set, indicating lower overfitting. The LGB model provided balanced and high performance, with 96.85% accuracy in the training set and 86.5% in the test set. This model particularly achieved good accuracy and F1 scores in the test set compared to other models, with more balanced and higher F1 scores in both the training and test sets. Other models (LR, LDA, KNN, NB) generally provided balanced performance; however, they were not as successful as RFC and LGB in terms of accuracy and other metrics. Since the LGB model demonstrated the most balanced and high performance in the training and test sets, ROC and explainability analyses were performed on this model.

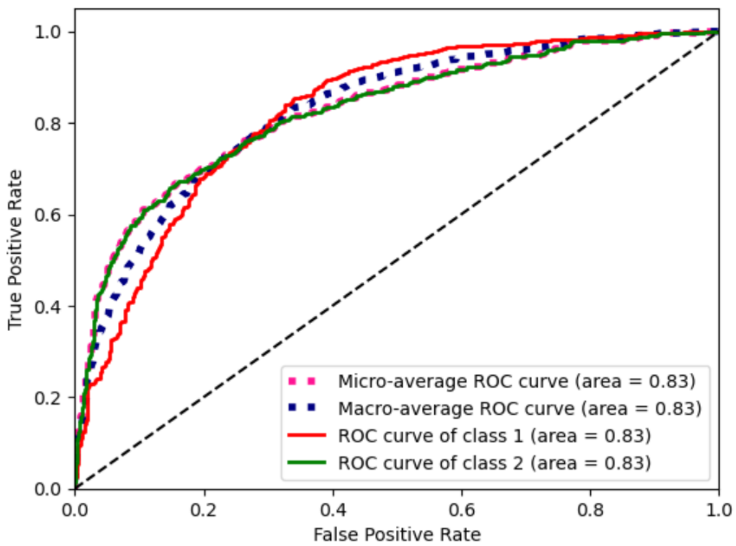

Figure 1 ROC analysis of LGB ML model.

The Area Under Curve (AUC) value of the LGB model was found to be 0.830, indicating that this model has a fairly good classification performance. The close proximity of the micro and macro average values was interpreted as balanced performance across all classes. This model demonstrated good discrimination power between true positive and false positive rates, showing consistent performance in both classes. In the feature importance analysis performed on the LGB model, features related to blood tests such as mean corpuscular volume (MCV), red cell distribution width (RDW-CV), CRP, and platelets were found to have high importance levels, suggesting that these features are more useful for predictions (Figure 2). In the explainability analyses conducted using the SHAP method, it was observed that features such as X-ray, age, and RDW-CV are particularly important for the model (Figure 3).

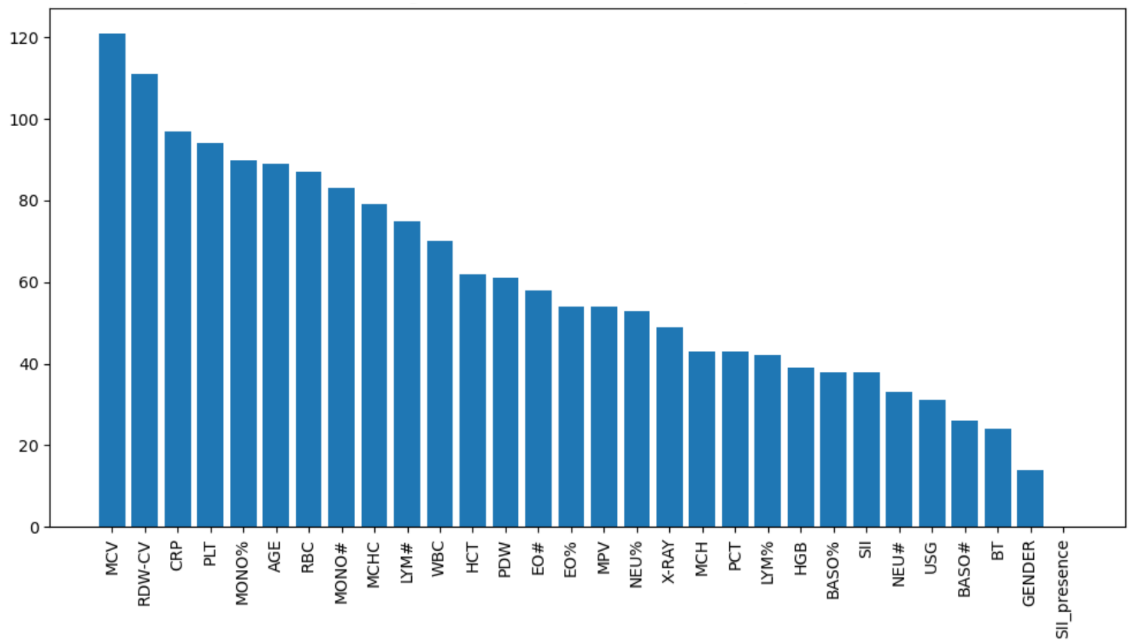

Figure 2. Feature importance analysis for LGB ML model.

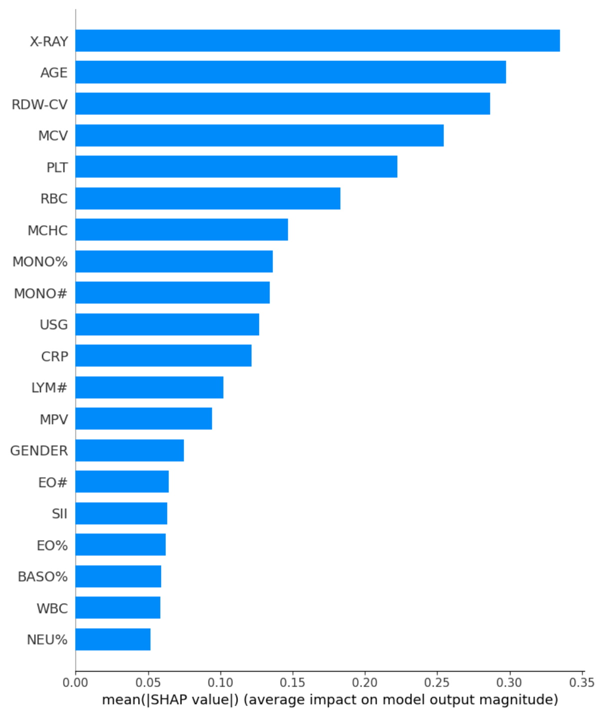

Figure 3. SHAP analysis for LGB ML model.

Conclusion: In our study, the LGB model demonstrated the most balanced and high performance across both the training and test sets. Given its high accuracy and F1 score in the test set, this model was evaluated as the best-performing model for prognosis prediction in the CAP dataset. While the SHAP chart provides a more detailed analysis, the LightGBM feature importance chart highlights the features that the model uses more frequently or considers more informative. To better understand the model's output and detail the influence of features, it may be more beneficial to consider the SHAP chart. Accordingly, in the LGB model's CAP prognostic predictions, the standout parameters were determined as the requirement for X-ray examination, younger age, and higher RDW-CV.

Keywords: Community-acquired pneumonia, machine learning, prognosis, explainability.

REFERENCES

[1] E. Kocabaş, G. Tanır, G. Cinel. Türk Toraks Derneği Çocuklarda Toplumda Gelişen Pnömoni Tanı, Tedavi ve Uzlaşı Raporu 2023. [Online]. Available: https://toraks.org.tr/site/sf/books/2023/12/771b2ce640645948509601bdb319116622032cba0e00afd66955e38c35275bc3.pdf.

[2] World Health Organization. Pneumonia. Who.int. Published November 11, 2021. Accessed August 5, 2024. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/pneumonia.

[3] JS Bradley, CL Byington, SS Shah, et al, The Management of Community-Acquired Pneumonia in Infants and Children Older Than 3 Months of Age: Clinical Practice Guidelines by the Pediatric Infectious Diseases Society and the Infectious Diseases Society of America, *Clinical Infectious Diseases*., vol. 53, no. 7, pp. e25-e76, 2011, doi:https://doi.org/10.1093/cid/cir531.

[4] M. Xie, K. Yuan, X. Zhu, et al, Systemic Immune-Inflammation Index and Long-Term Mortality in Patients with Stroke-Associated Pneumonia, *J Inflamm Res*., vol. 16, pp. 1581-1593, 2023, doi: 10.2147/JIR.S399371.

SMART STETHOSCOPE: A DEEP LEARNING APPROACH FOR CLASSIFYING LUNG DISEASES FROM AUSCULTATION SOUNDS

Gülşah Gökhan Gökçek[1*] ,Mehmet Ali Bayram[2] , Mustafa Ege Şeker[3], Konuralp İlim[4], Yilmaz Onat Koyluoglu[5]

[1] Graduate School of Natural Sciences and Engineering, Yeditepe University, İstanbul, Türkiye

[2] Graduate School of Science and Engineering, Yıldız Technical University, Istanbul, Türkiye

[3] School of Medicine and Public Health, University of Wisconsin, Madison, WI, USA

[4] Graduate School of Sciences and Engineering, Koç University, Istanbul, Türkiye

[5] Seyrantepe Hamidiye Etfal Hospital, Istanbul, Türkiye

** Corresponding author: gulsah.gokcek@yeditepe.edu.tr*

Introduction-Aim: Ambiguous identification and interpretation of sounds in auscultation is a generic issue that should not be neglected as it can lead to inaccurate diagnosis and mistreatment [1]. With the pervasiveness of AI/ML/DL algorithms in our lives, it is revealed that this inaccurate diagnosis and mistreatment problem can be handled with technology. For this purpose, many hospitals and medicine schools benefit from smart stethoscope technology. We are going to develop a plug-and-play sound module that can be plugged into a regular stethoscope, and it will help practitioners diagnose accurately regardless of their experience. On this path, we first create a model to classify the lung disease from respiratory sound.

Materials-Methods: The ICBHI 2017 Challenge Respiratory Sound Database is used. The dataset included 103 participants (62 M, 41 F) of all age groups, out of which 35 participants had no respiratory abnormalities (normal), while 68 had pulmonary diseases including COPD(Chronic Obstructive Pulmonary Disease), Healthy, Pneumonia, URTI(Lower Respiratory Tract Infection), Bronchiolitis, Bronchiectasis, LRTI(Lower Respiratory Tract Infection, Asthma). The whole dataset consists of 6892 sound files (the asthma label and related data were removed due to low sample size). The data is divided into 80% for training and 20% for validation. Preprocessing steps were applied to each recording to extract relevant features: Mel Frequency Cepstral Coefficients (MFCC), chroma features, and mel spectrogram (mSpec) representations. These feature matrices were extracted for each audio sample, capturing both spectral and temporal characteristics essential for respiratory disease classification. For this clasification a model is created that consisting of a multi-path CNN and multi-layer bi-directional LSTM (M-BDLSTM) architecture was used to predict lung disease from respiratory sounds. The model processes MFCC, chroma, and mel spectrogram features through separate CNN and bi-directional LSTM layers to extract spatial and temporal patterns. Each path learns the input sequence in both forward and backward directions, and their concatenated outputs are used to classify lung diseases.

Results: The proposed model achieved an overall accuracy of 91.08%, demonstrating its potential in accurately identifying respiratory diseases. Performance metrics were notably high for COPD, with F1-scores of 87.4%. Classes with fewer samples showed moderate performance, highlighting the challenge of class imbalance in medical datasets.

Feature-specific evaluation revealed that the combined use of MFCC, chroma, and mel spectrogram representations contributed significantly to model performance. MFCCs captured crucial frequency-related features, while chroma and mel spectrograms enhanced

temporal feature representation. The multi-path CNN and M-BDLSTM architecture effectively extracted spatial and temporal patterns from these features, and the bidirectional layers improved sequential understanding of respiratory sounds, leading to more accurate predictions across classes.

The confusion matrix analysis further highlighted the model's and the dataset's strengths and weaknesses. COPD showed high true positive rates, indicating clear differentiation from other classes. However, there was some misclassification between diseases with overlapping symptoms, such as LRTI and URTI, which points to potential areas for improvement in future iteration.

Discussion-Conclusion: This study presents a promising approach to enhancing the accuracy of lung disease diagnosis through a novel plug-and-play sound analysis module integrated with a conventional stethoscope. By leveraging advanced deep learning architectures, specifically a multi-path CNN and multi-layer bidirectional LSTM[3] model, our method effectively classifies respiratory diseases from auscultation sounds. The model's high accuracy, particularly for prevalent diseases such as COPD demonstrates the potential of AI in supporting medical diagnostics, where the ambiguity in sound interpretation has historically led to diagnostic inconsistencies and mistreatments.

The results highlight the importance of feature selection and preprocessing, as MFCC, chroma, and mel spectrogram features each provided unique contributions to model accuracy. MFCCs were instrumental in capturing essential frequency information, while chroma and mel spectrograms aided in the temporal differentiation of respiratory sounds, helping to distinguish subtle acoustic patterns associated with specific lung conditions. The multi-path structure and the bidirectional nature of the LSTM layers also contributed to the model's effectiveness by enhancing its ability to understand sequential patterns in respiratory sounds, which are crucial for accurately identifying diseases that manifest as distinctive auscultatory signs like wheezes and crackles.

One of the main challenges encountered in this study was the class imbalance in the dataset, with diseases like Bronchiectasis and Bronchiolitis being underrepresented. This imbalance likely contributed to the moderate performance for these less prevalent diseases, as indicated by lower F1-scores. The imbalance reflects a broader challenge in medical AI applications, where limited availability of certain pathological data can affect model generalizability. Future work could address this limitation by exploring data augmentation techniques or synthetic data generation to balance class distributions, potentially improving the model's performance for less common diseases.

Another challenge lies in the occasional misclassification between diseases with overlapping symptoms, such as LRTI and URTI. This overlap suggests that certain diseases may have similar acoustic patterns, which could confuse even the most advanced algorithms. Further refinement of the feature extraction process, or the incorporation of additional context such as patient history or environmental factors, could help mitigate this issue in future iterations of the model.

In real-world applications, this plug-and-play sound module could prove particularly valuable in settings with limited access to specialized respiratory diagnostic equipment or experienced practitioners. It has the potential to serve as a decision-support tool, providing clinicians with an additional layer of confidence in diagnosing complex respiratory conditions. The consistent performance on unseen validation data further underscores its robustness, suggesting that the model may generalize well across diverse patient populations and clinical environments.

In the literature, it can be seen that higher accuracies are achieved with similar models but broader datasets [3], indicating that the model could benefit from further improvements with more diverse and extensive data. The model shows promise, particularly when considering the total accuracy and F1 score for the COPD class, which further supports its potential for clinical application.

In conclusion, this study demonstrates that integrating AI with conventional stethoscope technology can significantly enhance diagnostic accuracy in respiratory disease classification, paving the way for more reliable, accessible, and user-friendly diagnostic tools. Future research should focus on addressing dataset limitations, improving feature specificity, and refining model interpretability to enhance both diagnostic performance and clinical adoption. By continuing to bridge the gap between AI technology and healthcare, such advancements could play a crucial role in improving patient outcomes in respiratory medicine.

Keywords: Respiratory Sound Classification, Lung Disease Diagnosis, Plug-and-play stetescope module, Multilayer Bi-directional LSTM, CNN

REFERENCE

[1] Hafke-Dys H, Bręborowicz A, Kleka P, Kociński J, Biniakowski A. The accuracy of lung auscultation in the practice of physicians and medical students. PLoS One. 2019 Aug 12;14(8): e0220606.

[2] F. U. M. Ullah, A. Ullah, I. U. Haq, S. Rho and S. W. Baik, "Short-Term Prediction of Residential Power Energy Consumption via CNN and Multi-Layer Bi-Directional LSTM Networks," in *IEEE Access*, vol. 8, pp. 123369-123380, 2020, doi: 10.1109/ACCESS.2019.2963045.

[3] Fraiwan M, Fraiwan L, Alkhodari M, Hassanin O. Recognition of pulmonary diseases from lung sounds using convolutional neural networks and long short-term memory. Journal of Ambient Intelligence and Humanized Computing. 2022 Oct 1:1-3.

MAXIMUM PACKAGE PLACING AND DELIVERY ROTATION WITH META HEURISTIC ALGORITHMS IN MEDICAL CARGO TRANSPORT

Yılmaz Can METE[1], Bayram KÖSE[1], Bahar DEMİRTÜRK[2], Şükran KONCA[2], Cihat ŞEKER[1]*

[1] *Electric-Electronic Engineering Department,İzmir Bakırçay University, İzmir, Türkiye*

[2] *Fundamental Sciences Department, İzmir Bakırçay University, İzmir, Türkiye*

* *Corresponding author: bahar.demirturk@bakircay.edu.tr*

Introduction-Aim: Medical material transport is a complex process that requires the safe and efficient delivery of sensitive and high-risk materials to their destination. Medical material transfer involves a wide variety of items, from biological samples to sensitive devices, and serious damage can occur as a result of misplacement or transport in inappropriate conditions. The safe transport of such materials directly affects not only financial losses, but also patient safety and continuity of healthcare services. Where manual transport methods are insufficient, optimised placement and routing algorithms supported by intelligent systems should be used. In this study, it is aimed to optimise both the placement scheme and the routing process by using metaheuristic algorithms to ensure the safest and most efficient transport of medical supplies.

Materials-Methods: Genetic Algorithm (GA) and Ant Colony Optimisation (ACO) metaheuristic algorithms were evaluated for the placement of medical supplies and determination of the shortest route. These algorithms aim to provide the most appropriate placement scheme and route planning according to the sensitivity of the materials and the environmental effects they may be exposed to during transport. These algorithms, supported by robotic system integration, ensure the safe placement of materials without the need for human intervention. The metaheuristic algorithms selected to optimise the transport capacity and route of the vehicles are evaluated under different scenarios and the effectiveness of each algorithm is compared with performance analyses.

Results: This study aims to reduce transport costs and minimise time losses by maximising the capacity of cargo vehicles. The performances of different metaheuristic algorithms are compared and analysed which algorithm provides the best results. The algorithms are evaluated according to criteria such as maximum package placement capacity, processing time and minimum distance routing. The results obtained have enabled the selection of the most suitable algorithm in order to increase the safety of the transported materials and to ensure cost-effectiveness. In this process, the performance of the algorithms in various transport scenarios has been analysed in detail and it has been revealed which algorithm provides the most reliable results.

Discussion-Conclusion: The results of the study contribute to maximising safety during the transport of sensitive medical supplies in the health sector. The algorithms developed for this purpose are planned to be integrated with robotic systems to fully automate the placement and delivery of materials. The placement of the materials to be transported through robotic systems without human intervention will allow safe and stable positioning of the loads. This integration allows minimising the risks that may arise from human error, improving operational efficiency by increasing the safety of materials transported in healthcare services. In future studies, it is aimed to contribute to the sustainability of critical transport processes in the healthcare sector by increasing the accuracy and efficiency of these processes with more

advanced algorithms. In conclusion, the advantages provided by placement and routing algorithms in medical cargo transport directly positively affect the quality of healthcare services and offer new opportunities for the safety of transported materials.

Keywords: Medical equipment transport, Metaheuristic algorithms, Optimisation, routing, Robotic system integration

REFERENCE

[1] Acar, E. B., Karabey, C., & Köse, B. (2023). İnsansız Hava Aracı İle Paket Dağıtımında Gezgin Satıcı Probleminin Genetik ve Parçacık Sürü Optimizasyon Algoritmaları ile Çözümü. Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi, 10(20), 168-181. https://doi.org/10.54365/adyumbd.1249391

[2] Karadayı, M. A., Gökmen, Y. G., Kasap, L. G., & Tozan, H. (2019). Sağlıkta güncel simülasyon yaklaşımları: Bir derleme çalışması. International Journal of Advances in Engineering and Pure Sciences, 31(1), 1-16.

[3] Gezici, H., & Livatyalı, H. (2022). İki boyutlu kutu paketleme probleminin çözümü için hibrit çiçek tozlaşma algoritması yaklaşımı. Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 37(3), 1523-1534.

PREDICTING AGITATION IN DEMENTIA PATIENTS USING AI AND IOT FOR ENHANCED REAL-TIME MONITORING AND CARE

Melik Mert Dolan[1*], Burak Ülver[1], Kerem Basım[1], Ruşen Halepmollası[1], Mehmet Haklıdır[1]

[1]TÜBİTAK Informatics and Information Security Research Center, Kocaeli, Türkiye

*Corresponding author: mert.dolan@tubitak.gov.tr

Introduction-Aim: Dementia is a syndrome characterized by a significant decline in cognitive functions such as memory, thinking, and reasoning, which interferes with daily living activities. It is not a single disease but rather a cluster of symptoms associated with various underlying disorders, such as vascular dementia, Alzheimer's disease, frontotemporal dementia, or Lewy body dementia. To the best of our knowledge, there is no definitive treatment for progressive dementias, such as Alzheimer's disease or frontotemporal dementia, and only a few pharmacological interventions are available that can slow the progression of the disease. In the absence of a definitive treatment, a rehabilitation approach focused on preserving existing abilities for as long as possible is essential. In this context, we aim to predict dementia progression and contribute rehabilitation efforts by utilizing Internet of Things (IoT) and Artificial Intelligence (AI) models and identifying data that can guide timely interventions.

Materials-Methods: We used TIHM dataset, which collected from 56 individuals over the age of 50, all of whom had been diagnosed with various forms of dementia. TIHM is a publicly available dataset and includes data collected with PIR sensors, sleep tracking mat and patient routine measurements for remote healthcare monitoring. To prepare the dataset for model training, we applied several preprocessing steps aimed at optimizing the data for maximum predictive performance. First, we combined the separate datasets and linked them together. We performed feature extraction to create new attributes that better capture the underlying patterns in the data. Next, we applied feature selection techniques to reduce data dimensionality, retaining only the most important attributes while eliminating less informative ones. To predict the agitation, we employed LightGBM and Catboost models. Also, we built weighted ensemble voting approach. The proposed system represents an innovation that could provide vital assistance to both patients suffering from dementia and healthcare professionals involved in their rehabilitation.

Results: In this study, we developed a system in which data collected from sensors is transmitted to the cloud via an IoT platform, processed by our effective models, and used to offer real-time feedback. When predicting agitation, we employed weighted ensemble voting approach combining Random Forest, KNeighbors, Support Vector Machine and XGBoost classifiers and also LightGBM and Catboost models. We evaluated and compared the performance of classification models in terms of accuracy, precision, recall, and F1-score. According to our results, the ensemble voting model achieved the best performance, with recall rates of 91.88% for agitation instances and 76.92% for non-agitation instances, and an overall accuracy of 91.34%. On the other hand, for the LightGBM, recall for agitation instances was 89.47% and 73.74% for non-agitation instances and accuracy of 74.58%. Moreover, CatBoost showed a recall of 59.25% for agitation instances, 94.25% for non-agitation instances, and an overall accuracy of 91.62%.

Discussion-Conclusion: In this study, we explored the potential of leveraging advanced data processing techniques and machine learning models to improve dementia care and rehabilitation efforts. The preprocessing steps applied to the dataset, including

standardization, size reduction, and feature extraction. By optimizing the data and selecting the most relevant features, we were able to increase the predictive accuracy of our models, particularly in forecasting agitation labels in dementia patients. These predictions can significantly contribute to real-time monitoring and early intervention, offering valuable support for healthcare professionals working with dementia patients. Our approach underscores the importance of integrating emerging technologies such as IoT and AI in dementia care. By using IoT platforms to continuously collect sensor data and AI models to analyze this data, we can provide healthcare providers with timely, actionable insights. The real-time feedback system we propose could greatly improve patient management by enabling faster responses to changes in a patient's condition.

Keywords: Dementia, Internet of Things (IoT), Artificial Intelligence (AI), Healthcare, Real-time Feedback

REFERENCE

[1] Dementia, U. K. (2021). What is dementia? [Online]. Retrieved from https://keystonemedical.com.sg/dementia/

[2] "The Role of Behavior Analysis in the Rehabilitation of Persons with Dementia." Behavior Therapy 42 (2011): 9–21.

[3] Palermo, F., Y. Chen, A. Capstick, N. Fletcher-Loyd, C. Walsh, S. Kouchaki, J. True, O. Balazikova, E. Soreq, G. Scott, H. Rostill, R. Nilforooshan, and P. Barnaghi. 2023. "TIHM: An Open Dataset for Remote Healthcare Monitoring in Dementia." Nature Scientific Data 10 (1): Article 53. https://doi.org/10.1038/s41597-023-01374-9.

FEDERATED LEARNING FOR DIABETES PREDICTION IN HEALTHCARE

Rabia Arkan Yurtoglu[1*], Havvanur Dervişoğlu[1], Mustafa Aktaş[1], Mehmet Daşçılar[1], Ruşen Hallepmollası[1], İpek Baz[1], Mehmet Haklıdır[1]

[1]*TÜBİTAK Informatics and Information Security Research Center, Kocaeli, Türkiye*

Corresponding author: rabia.arkan@tubitak.gov.tr

Introduction-Aim: According to the NIDDK [1], diabetes, commonly referred to as blood sugar, is a global disease initiated by elevated blood glucose levels and eventually leads to health complications like cardiovascular, neurological, and kidney damage [2, 3]. As of 2024, an estimated 643 million people worldwide suffer from diabetes, and projections suggest this number could rise to 783 million by 2045 [4, 5]. Early detection and effective management of diabetes are crucial to preventing these complications. Recently, machine learning (ML) has played a transformative role within the detection of prediabetes, facilitating the creation of innovative models. Recent research also has demonstrated that ML algorithms can identify diabetes with high success rates [4, 6]. Research in ML allows medical practitioners to anticipate illness risks and take preventative action. However, achieving this potential necessitates access to more comprehensive and sensitive medical data. Currently, however, access to medical data remains restricted due to security and privacy concerns. This constraint impedes the advancement and enhancement of ML models in healthcare, which necessitate extensive, varied datasets for optimal performance. To get around these restrictions, researchers are increasingly turning to federated learning (FL), which offers a privacy-preserving approach to data utilization [7]. Through FL, more inclusive models can be developed using data from various hospitals or patients while keeping critical information confidential. This article presents a comparison of the predictive performance of ML models trained with FL approach against the centralized learning approach for diabetes prediction.

Materials-Methods: This research presents a case study that uses both centralized and federated learning frameworks to build ML models for diabetes diagnosis. Initially, we conducted several exploratory data analysis and preprocessing processes on the diabetes dataset [8] such as handling missing values, identifying outliers, correlation analysis and feature selection. Then, in order to identify the most effective ML algorithms based on the average F1 score, we trained a variety of algorithms using sampling, standardization, and k-fold cross-validation techniques. These algorithms included XGBoost, RandomForest, DecisionTree, SVC, LogisticRegression, LightGBM, AdaBoost, KNN, CatBoost, GradientBoosting, and NaiveBayes. We used GridSearch to determine which of the XGBoost, LightGBM, and CatBoost algorithms had the best metrics and optimal parameters based on their average F1 scores. We evaluated each algorithm's performance and the optimal set of parameters based on the GridSearch results, concluding that the XGBoost approach produced the best classification metrics. Then, in order to compare with the best centralized outcomes in the study, we applied the FL approach, which permits model training in cooperation with other stakeholders without requiring data sharing within the scope of data security. In this context, the data was divided to represent the data of five different hospitals. Every hospital was regarded as a federated client, and the FL training was finished in ten rounds by averaging the model weights (FedAvg) of the clients. The results obtained with FL were evaluated in terms of security, confidentiality and fidelity by comparing them with the results of centralized training, which requires data to be shared and collected in a single

center. Furthermore, the study's results were acquired with 10-fold cross validation in both federated and centralized training, and this was accomplished by employing distinct random seeds in FL.

Results: The results of this study provide a comprehensive analysis of the performance of ML models for diabetes prediction, comparing centralized and FL approaches. Initially, we trained and evaluated several models with default parameters using 10-fold cross validation. The average F1 scores for best models were: CatBoost (0.914), XGBoost (0.911) and LightGBM (0.910). To further improve these high-performing models, we used GridSearch; XGBoost emerged as the best model with an F1 score of 0.92. Then, we compared the performance of XGBoost under the centralized ML approach with its performance in the FL approach. For centralized ML, the XGBoost model, after applying Random Under Sampling and MaxAbsScaling, achieved an F1-score of 0.92 and an AUC value of 0.95. In the FL setting, we used FedAvg as the aggregation method, and the average results over 10 runs yielded an F1-score of 0.88 with a standard deviation of 0.013, and an AUC score of 0.98 with a standard deviation of 0.003.

Discussion-Conclusion: The discussion of the findings highlights both the strengths and challenges of applying centralized and FL approaches to diabetes prediction. Evaluating XGBoost under FL revealed a drop in its F1-score from 0.92 in the centralized setup to 0.89. This decrease can be attributed to the distributed and heterogeneous nature of data in FL, as well as communication limitations and synchronization issues between clients and the server. Interestingly, when comparing the AUC metric, the score increased from 0.95 in centralized learning to 0.98 in FL, indicating that the model's overall discriminative ability improved with exposure to diverse, heterogeneous data distributions. FL's capability to train on diverse data residing on different clients likely contributed to this improvement in the AUC score, suggesting a better ability to distinguish positive and negative classes. However, this same diversity may also have negatively impacted sensitivity and specificity at a given threshold, which is reflected in the decrease in the F1-score. Despite the slightly lower performance of FL in some metrics, it offered significant privacy and security advantages, enabling the use of data from a broader set of users while maintaining confidentiality.

Keywords: *Machine learning, Federated learning, Healthcare, Diabetes*

REFERENCE

[1] Diabetes - NIDDK. (n.d.). National Institute of Diabetes and Digestive and Kidney Diseases. Available: https://www.niddk.nih.gov/health-information/diabetes

[2] Md. A. Uddin et al., “Machine Learning Based Diabetes Detection Model for False Negative Reduction,” Deleted Journal, vol. 2, no. 1, pp. 427–443, Jun. 2023, Available: https://doi.org/10.1007/s44174-023-00104-w

[3] O. T. Kee et al., “Cardiovascular complications in a diabetes prediction model using machine learning: a systematic review,” Cardiovascular Diabetology, vol. 22, no. 1, Jan. 2023, Available: 10.1186/s12933-023-01741-7. Available: https://doi.org/10.1186/s12933-023-01741-7

[4] S. K. S. Modak and V. K. Jha, “Diabetes prediction model using machine learning techniques,” Multimedia Tools and Applications, vol. 83, no. 13, pp. 38523–38549, Oct. 2023, doi: 10.1007/s11042-023-16745-4. Available: https://doi.org/10.1007/s11042-023-16745-4

[5] K. Oliullah, M. H. Rasel, Md. M. Islam, Md. R. Islam, Md. A. H. Wadud, and Md. Whaiduzzaman, “A stacked ensemble machine learning approach for the prediction of diabetes,” Journal of Diabetes & Metabolic Disorders, vol. 23, no. 1, pp. 603–617, Nov. 2023, Available: https://doi.org/10.1007/s40200-023-01321-2

[6] A. Rahman et al., “Federated learning-based AI approaches in smart healthcare: concepts, taxonomies, challenges and open issues,” Cluster Computing, vol. 26, no. 4, pp. 2271–2311, Aug. 2022, Available: https://doi.org/10.1007/s10586-022-03658-4

[7] S. M. Williamso & V. Prybutok, “Balancing privacy and progress: a review of privacy challenges, systemic oversight, and patient perceptions in AI-driven healthcare,” Applied Sciences, 14(2), 675, 2024. Available: https://doi.org/10.3390/app14020675

[8] C-Y Chou, D-Y Hsu, C-H Chou, “Predicting the Onset of Diabetes with Machine Learning Methods”. Journal of Personalized Medicine. 2023; 13(3):406. Available: https://doi.org/10.3390/jpm1303040

AI-DRIVEN DETECTION OF PARKINSON'S DISEASE THROUGH VOCAL FEATURE OPTIMIZATION

Mahmut Lutfullah Özbilen[1,2*], Feyza Özen[1], Ayça Şirin Kındap[1], Hakan Gündüz[3], Ruşen Halepmollası[1], Mehmet Haklıdır[1]

[1] *TÜBİTAK Informatics and Information Security Research Center, Kocaeli, Türkiye*

[2] *Computer Engineering, Kocaeli University, Kocaeli, Türkiye*

[3] *Software Engineering, Kocaeli University, Kocaeli, Türkiye*

* *Corresponding author: mahmut.ozbilen@tubitak.gov.tr,*

Introduction-Aim: Parkinson's Disease is a progressive neurological disorder with motor and non-motor symptoms, including vocal impairments like reduced pitch, breathiness, and changes in speech rhythm. Machine learning techniques offer promising approaches for early detection of these vocal symptoms, enabling timely intervention and potentially improving patient outcomes. In this study, feature selection methods were compared on a high-dimensional dataset containing vocal features. The curse of dimensionality—a phenomenon in which a high volume of feature space can lead to sparse data representations and decreased model performance—makes feature selection crucial. By reducing dimensionality, feature selection improves model accuracy and enhances computational efficiency by retaining only the most informative features. The performance of different feature selection methods and classification algorithms was assessed on a high-dimensional Parkinson's disease dataset.

Materials-Methods: In this study, the dataset containing the features of speech recordings collected from 188 patients diagnosed with Parkinson's Disease and 64 healthy control group participants aged between 33 and 87 from Istanbul University Cerrahpaşa Faculty of Medicine, Department of Neurology was used. The speech samples were taken from the participants in 3 repetitions of continuous phonation of the sound /a/ recorded at a sampling rate of 44.1 kHz. The dataset consists of 755 features including vocal fold measurements, Mel Frequency Cepstral Coefficients wavelet transform features and time-frequency domain features. During the division of the dataset into training and test subsets, it was ensured that data belonging to the same participant was present in either the training or the test set. In this way, data leakage was prevented and an unbiased model evaluation was provided. In addition, the patient and healthy ratio was kept the same in the training and test datasets and the class distribution was preserved. This process was done before feature selection to prevent bias and prevent artificial increases in model performance. Features were scaled with a standard scaler. Feature selection was performed in two different ways to reduce the curse of dimensionality and obtain the most useful features for the model. These are filter-based methods and wrapper-based Genetic Algorithm methods. Filter selection methods included solutions based on statistical relationships with the classification target variable. The Spearman's Rho correlation method was employed in the study. With these methods, 650 features most related to the target variable were selected. While performing feature selection with the Genetic Algorithm method, experiments were made with K-Nearest Neighbors estimators, resulting in the selection of 89 features. Support Vector Machines, K-Nearest Neighbors, Random Forest Classifier, AdaBoost Classifier, Decision Tree Classifier, Multi-Layer Perceptron, XGBoost Classifier models and ensemble models were used as classification algorithms. As an ensemble model, Voting Classifier, which combines the probability outputs of Support Vector Machines, Random Forest Classifier and XGBoost

model results with soft voting to create the final prediction, and Stacking Classifier, which uses Random Forest Classifier as the final decision maker, were used.

Results: The combination of feature selection using filter selector with Spearman's Rho correlation method as the estimator and followed by a voting classifier resulted in an accuracy score of 0.881 and F1 score of 0.873, which are the top results achieved for both metrics with 650 features. Voting Classifier with no feature selection resulted in an accuracy score of 0.877 and F1 score of 0.869. Additionally, the performance of Metaheuristic-based Feature Selection with a K-Nearest Neighbors estimator was notably lower.

Discussion-Conclusion: Overall, when considering both accuracy and F1 scores, the feature selection method using the Filter Selector with Relief algorithm and the classification method using a Voting Classifier demonstrates the best performance. When compared to the case when a voting classifier is applied to the dataset without any feature selection, it improves the accuracy score by 0.4%. Additionally, this approach outperforms the highest accuracy score, which is 0.869, achieved in the literature with a Convolutional Neural Network using leave-one-person-out cross-validation. The increase in the accuracy score is 1.2%.

Keywords: Parkinson's Disease, Feature Selection, Ensemble Learning, Healthcare

REFERENCE

[1] J. Jankovic, "Parkinson's disease: clinical features and diagnosis," *J. Neurol. Neurosurg. Psychiatry*, vol. 79, no. 4, pp. 368–376, 2008, doi: 10.1136/jnnp.2007.131045.

[2] C. Sakar, G. Serbes, A. Gunduz, H. Nizam, and B. Sakar. "Parkinson's Disease Classification," UCI Machine Learning Repository, 2018. [Online]. Available: https://doi.org/10.24432/C5MS4X.

[3] H. Gunduz, "Deep Learning-Based Parkinson's Disease Classification Using Vocal Feature Sets," *IEEE Access*, vol. 7, pp. 115540-115551, 2019, doi: 10.1109/ACCESS.2019.2936564.

BENCHMARKING MACHINE LEARNING AND DEEP LEARNING MODELS FOR DEEPFAKE DETECTION IN LUNG CT SCANS

Merve Zeybel[1*], Havvanur Dervişoğlu[1], Elif Özcan[1], Nida Kumbasar[1], Ruşen Hallepmollası[1], İpek Baz[1], Mehmet Haklıdır[1]

[1]TÜBİTAK Informatics and Information Security Research Center, Kocaeli, Türkiye

**Corresponding author kaya.merve@tubitak.gov.tr*

Introduction-Aim: Computer-generated deepfake videos and images are increasingly applied across various domains including healthcare, economic and political. In healthcare, these techniques are used to alter medical images such as CT and MR scans and raised significant concerns regarding the authenticity and reliability of diagnostic images. Deepfakes of CT and MR images are produced for various purposes, including generating synthetic datasets to train machine learning models, enhancing educational materials, and preserving patient privacy through anonymization. However, they can also be misused to manipulate diagnostic outcomes or commit medical fraud, highlighting the need for effective detection methods. The aim of this study is to reliably detect the manipulations in our CT lung images using machine learning and deep learning methods.

Materials-Methods: To perform our analysis, we used an open source dataset containing both tampered and untampered DICOM CT images of patients, specifically the CT-GAN dataset, which includes synthetic lung images with and without tumor manipulations. We also applied the ±5 range approach. In other words, we selected each labeled slice along with its adjacent slices within a range of [–5, +5] to maximize the number of tampered images in our dataset. To establish a benchmark for binary classification, we employed and compared fourteen machine learning and deep learning algorithms, namely CNN, ConvNeXtTiny, DenseNet201, DenseNet121, DenseNet169, EfficientNetV2S, Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Random Forest, Decision Tree, Gradient Boosting, XGBoost and LightGBM. Due to dataset imbalance, we applied augmentation and undersampling techniques to achieve balanced class representation.

Results: We compared various machine learning and deep learning model results in terms of accuracy, precision, recall, and F1-score. XGBoost outperformed all other machine learning and deep learning models by achieving the highest accuracy (0.95), precision (0.96), recall (0.95), and F1-Score (0.95) rates. LightGBM, Random Forest and Gradient Boosting also showed good results, both achieving accuracy and F1 score of 0.95, 0.94 and 0.94, respectively. Support Vector Machine, Logistic Regression and EfficientNetV2S yielded consistent results across metrics in all categories. Other models, such as ConvNeXtTiny, DenseNet variants, K-Nearest Neighbors, and Decision Tree, achieved slightly lower performance. Our results demonstrated promising performance, indicating that utilized machine learning and deep learning methods can effectively differentiate between original and manipulated images and provide a reliable benchmark for future studies.

Discussion-Conclusion: In this study, we successfully demonstrated the capability of machine learning and deep learning algorithms in detecting manipulations in DICOM CT images. By utilizing the open CT-GAN dataset, we established a reliable benchmark for binary classification between tampered and untampered images. Consequently, this study has demonstrated that fake images generated by deepfake techniques can be effectively detected.

Keywords: Deepfake, Deep Learning, Machine Learning, Healthcare, CT-GAN

A COMPARATIVE ANALYSIS OF OSTEOARTHRITIS CASE EVALUATIONS: INSIGHTS FROM PHYSICAL THERAPISTS AND ARTIFICIAL INTELLIGENCE

Aleyna Kilimci[1], Sena Özdemir[1], Yasmin Günkut[1], Ayşe Kayalı Vatansever[1*]

[1] Department of Physiotherapy and Rehabilitation, İzmir Bakırçay University, İzmir, Türkiye

** Corresponding author: ayse.vatansever@bakircay.edu.tr*

Introduction-Aim: Applications of artificial intelligence (AI) in diagnosing and planning treatment for osteoarthritis are becoming increasingly common. However, there is limited literature on the role of physical therapists in evaluating osteoarthritis cases and developing therapeutic programs. Therefore, this study was designed as a pilot to analyze the perspectives of physical therapists in comparison to AI in the evaluation and treatment planning for osteoarthritis cases.

Materials-Methods: Applications of artificial intelligence (AI) in diagnosing and planning treatment for osteoarthritis are becoming increasingly common. However, there is limited literature on the role of physical therapists in evaluating osteoarthritis cases and developing therapeutic programs. Therefore, this study was designed as a pilot to analyze the perspectives of physical therapists in comparison to AI in the evaluation and treatment planning for osteoarthritis cases.

Results: Since this study is qualitative, statistical data analysis was not used. The purpose of this study was to share the opinions provided. It was found that the responses of one of the six volunteer physical therapists were generated by asking the AI, leading to their exclusion from the study. The remaining five therapists provided their opinions on all three cases. The AI's responses were found to be more systematic and detailed in comparison to those of the physical therapists in terms of evaluation and therapy.

Discussion-Conclusion: As the use of artificial intelligence (AI) in diagnosing and assessing osteoarthritis increases, advancements in physiotherapy and rehabilitation are expected. This study compares the perspectives of physical therapists with AI's views on osteoarthritis cases. The findings show that physical therapists, with at least three years of experience, share common viewpoints on frequently encountered cases, but also have differing perspectives. AI, on the other hand, provided the most comprehensive interpretation, focusing on all cases from the same perspective. Physical therapists, however, approached the cases more personally, incorporating clinical experience to anticipate future outcomes. Given its unique design, this study is expected to guide future research. Based on this pilot study, it is recommended that larger-scale studies with broader data inputs be conducted.

Keywords: Artificial Intelligence, Osteoarthritis, Physiotherapy, Rehabilitation

REFERENCE

[1] Lee, L.S., Chan, P.K., Wen, C. *et al.* Artificial intelligence in diagnosis of knee osteoarthritis and prediction of arthroplasty outcomes: a review. *Arthroplasty* 4, 16 (2022). https://doi.org/10.1186/s42836-022-00118-7

[2] Touahema S, Zaimi I, Zrira N, Ngote MN. How Can Artificial Intelligence Identify Knee Osteoarthritis from Radiographic Images with Satisfactory Accuracy?: A Literature Review for 2018–2024. *Applied Sciences*. 2024; 14(14):6333. https://doi.org/10.3390/app14146333

[3] Binvignat M, Pedoia V, Butte AJ*, et al* Use of machine learning in osteoarthritis research: a systematic literature review *RMD Open* 2022; 8:e001998. doi: 10.1136/rmdopen-2021-001998

ARTIFICIAL INTELLIGENCE AND FUNCTIONAL BIOCHEMICAL FACIAL ANALYSIS

Hayriye Alp [1*], Hayriye Dilek, Akdoğan[2], Serife, Aydın[1]

[1] GETAT Center, Necmettin Erbakan University, Konya, Türkiye,

[2] Medical Education, Izmir Bakırcay University, Izmir, Türkiye

** Corresponding author: hayriyealp@erbakan.edu.tr*

Introduction-Aim: Shüssler minerals are functional salts that have been used by Dr. Shüssler in complementary therapies. Biochemical functional salts are effective in extracellular matrix and humoral pathologies directly in the cell and its surroundings. It is aimed to treat diseases by giving functional substances that are diluted well and distributed to body fluids. Biochemically functional salts are aimed to replace the physiologically deficient ones rather than creating a different stimulus. It is expected that minerals that are sufficiently distributed to the extracellular matrix will be taken from here as needed for their intracellular functions (1,2,3).

Materials-Methods: In order to use Shüssler minerals effectively in treatment, it is necessary to know the functional substances and the physiological areas they serve well. For an effective treatment, performing facial analysis, taking a good anamnesis, understanding pathophysiology, and understanding iris diagnosis can make the doctor's job easier.Deposits may occur in diseases. Complete depletion of mineral substance stores can lead to 'senility fatigue' and even death.

Results: Considering the results of the facial analysis. The human organism tends to store minerals for use in necessary situations due to its internal information. The less the organism stores, the more it restricts itself and dysfunction may occur. (Callus, oily skin, chapped lips, pigment spots, wrinkles, skin granules, air hunger, pale skin) When the storage capacity begins to fill, body functions begin to improve (1,2,3,4). Psychological conditions can also affect mineral metabolism; exposure to long-term stress, pressure for success at home and school can increase mineral deficiency to the extreme. Today, with the extension of years of life, the 70–90-year-olds constitute a significant part of the population; For healthy aging, it would be appropriate to use Shüssler minerals. Mineral deficiencies can be detected with many methods; facial analysis, formation of personal symptoms, symptoms of developing disorders, or kinesiological appropriate mineral detection methods can be investigated. The body gives symptoms in order to reach the missing substance in some way, but it is necessary to understand the language of this organism.

Although Dr. Shüssler also mentioned that deficient minerals can be detected with facial analysis, Kurt Hickethier and Thomas Feichtinger developed the facial analysis we know today. Thus, facial analysis offers us the opportunity to detect and treat deficient minerals and the pathophysiological processes behind them. It may be necessary to support facial analysis with pathophysiology, eye and tongue diagnoses (1). Mineral deficiencies according to facial analysis; The most important indicator in facial analysis is the presence of wrinkles in the inner angle of the eye; wrinkles can be square or fan-shaped, sometimes on the upper eyelid. Another indicator is brown-black dark coloring on the upper eyelids. It is often mentioned that there is a dark ring around the eye that disappears with this mineral supplement. Callus formation on hands and feet, cyanosis lip formation after excessive physical activity and effort are findings supporting nr 1 deficiency. The finding in facial analysis is a waxy appearance on the face. The waxy appearance can be in the ear, cartilage, in front of the ear, and under the nose. The extension of the waxy area to the forehead is informative about the depth of the

deficiency. Tension in the neck and back regions are indicators of nr 2 deficiency. The indication in facial analysis is the presence of bluish black coloration between the eyes and nose. If the coloration extends to the root of the nose, it is an indication that the deficiency is very deep. In facial analysis, sunken temples are a sign of depletion of nr 5 stores. In advanced levels, a gray color settles on the face. It indicates that the person needs to take time to rest for regeneration. If the person has started to get angry easily, he needs nr 6. In facial analysis, there is an earthy color that shows itself under the lower lip. Pigment spots and air hunger have increased. In facial analysis, an increase in the color called magnesium red, which resembles gypsy pink, is evident on the cheeks. In facial analysis, there is a gelatinous shine on the eyelids. The presence of large pores on the skin also indicates a deficiency of nr 8. In children, there are puffy cheeks. With the intake of nr 9, the acid-base balance in the body begins to be re-regulated, when the sufficient level is reached, the acid load accumulated in the body begins to be eliminated naturally, and the body begins to break down sugar-carbohydrates better. The sour taste on the tongue when taking shüssler salt is a finding supporting the deficiency of nr 9. The symptom in the facial analysis is redness on the chin, the presence of a smogy line. The presence of a greenish color on the chin in the facial analysis is an indicator of nr 10 deficiency; the intensity of the color gives information about the degree of deficiency. Under-eye bags are a warning sign, and if hand and foot swellings also begin to occur, it leads to a more definitive diagnosis. In nr 11 mineral deficiency, the person becomes extremely sensitive to light and sound, since connective tissue weakness occurs, easy intra-tissue bleeding develops, and there are bruises. In facial analysis, the presence of accordion-style wrinkles in front of the ears is a warning. Some male patients may have a mirror-like appearance on the top of the forehead due to silica deficiency. In facial analysis, nr 12 deficiency manifests itself as a chalk, plaster, lime-white complexion (1,2,5,6).

Discussion-Conclusion: If mineral deficiency symptoms are uploaded to an artificial intelligence system, the diagnosis can be easily made in the patient who shows his face. Artificial intelligence can help us calculate biochemical deficiency. We think that replacing the missing minerals will improve the quality of life.

Keywords: shüssler, mineral, fasial analysis, artificial intelligence.

REFERENCE

[1] Bütad Bülten Ecz. Fatma HENDEN

[2] Şaroğlu, (2020) Schüssler Tuzları Temel Bilgiler, Onbir Yayınları, İstanbul

[3] Barrera, H., (2010) Scheussler's Biochemistry: An Exploration into the Benefits of CellSaltRemedies, P19,C:/Users/pc/Desktop/Dr.%20Sch%C3%BCssler%20salts/Twenty-Pages-in-the-Life-of-Scheussler-Cell-Salts-Final.pdf

[4] Attard Everaldo.(2011) Homeopathy and Allied Therapies: A Review. Journal OfEuromed Pharmacy, P36-39

[5] Çalış, B. ve Özçelik M. (Eds.) (2021) Schüssler Tuzları Yüz Tanısı Rehberi.Celsus Kitabevi. Türkiye)

[6] Bilgiç, S. (2015). Dr. Schüssler Biyokimyasal Mineralleri. Türkiye.

THE ROLE OF ARTIFICIAL INTELLIGENCE, DIGITAL TRANSFORMATION, AND SUSTAINABILITY IN THE HEALTHCARE INDUSTRY: A SYSTEMATIC REVIEW USING THE PRISMA FRAMEWORK

Arif Söyler[1*], Tarık Semiz[1], Süleyman Dündar[1]

[1] Health Management Department, İzmir Bakırçay University, Izmir, Türkiye

** Corresponding author: arifsoyler@gmail.com.tr*

Introduction-Aim: The integration of artificial intelligence (AI), digital transformation, and sustainability has become an essential focus in the healthcare industry. AI's ability to process and analyze vast amounts of data has enabled significant advancements in diagnostics, predictive analytics, and personalized medicine [1]. Digital transformation, which includes technologies such as electronic health records (EHRs), telemedicine, and the Internet of Things (IoT), has redefined patient care by improving both accessibility and efficiency [2]. However, the real challenge lies in aligning these technological innovations with sustainability goals to create a healthcare system that is not only advanced but also environmentally and economically sustainable [3]. This systematic review aims to address this challenge by using the PRISMA framework [4] to analyze how AI and digital transformation contribute to sustainable practices in healthcare.

Materials-Methods: This systematic review was conducted using the Web of Science database on October 15, 2024. The initial search included the keywords "Artificial Intelligence," "Industry," "Digital Transformation," and "Health," which yielded 60 articles. To refine the results, "sustainability" was added as an additional keyword, narrowing the selection to 9 relevant articles. The inclusion criteria focused on studies published in English between 2018 and 2024 that examined the use of AI and digital transformation in sustainable healthcare practices. Following PRISMA guidelines, titles and abstracts were screened, and these 9 articles were selected for full-text review.

Results: After screening titles and abstracts, 9 articles met the criteria for full-text review. Content analysis of these studies identified recurring themes such as the implementation of AI in predictive healthcare models, integration of EHRs for resource efficiency, and the role of IoT in monitoring environmental impacts within healthcare facilities. The findings highlight that while there are significant advancements in AI and digital transformation, challenges such as data privacy and the need for scalable solutions remain.

Discussion-Conclusion: The analysis demonstrates that the synergy of AI and digital transformation is pivotal for promoting sustainability in healthcare. However, achieving this requires overcoming barriers like regulatory constraints and ensuring equitable access to technological resources [5].

Keywords: artificial intelligence, industry, digital transformation, health, sustainability

REFERENCE

[1] Topol, E. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. Basic Books, 2019.

[2] Wang, J., et al. "Impact of Digital Transformation on Healthcare Services." *Journal of Healthcare Management*, 2021.

[3] Davenport, T., & Kalakota, R. "The potential for artificial intelligence in healthcare." *Future Healthcare Journal*, 2019.

[4] Page, M. J., et al. "PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews." *BMJ*, 2021.

[5] Smith, R., et al. "Challenges in Implementing AI Solutions in Healthcare." *Journal of Technology in Healthcare*, 2022.

THE EFFECTS OF THE INTERNET OF THINGS AND WEARABLE TECHNOLOGY ON ELDERLY HEALTH

Mehmet Ali Şener[1*], Aslı Kılavuz[2]

[1] *Institute of Health Sciences, Internal Medicine, Elderly Health, Ege University İzmir, Türkiye*

[2] *Division of Geriatrics, Department of Internal Medicine, School of Medicine, Ege University, ;zmir, Türkiye*

* *Corresponding author: mehmetalisener627@gmail.com*

Introduction-Aim: With the widespread adoption of the internet in the 21st century, demand for mobile and digital applications has increased. Technological advancements, especially through Health 4.0 and Society 5.0, have become more pronounced in the healthcare field, with notable developments in areas such as artificial intelligence applications, wearable technologies, and robotic surgery. The transformation initiated by Health 4.0 emphasizes using innovative technologies to improve the efficiency of healthcare services and enhance patient treatment processes by making healthcare more accessible to patients [1]. Health 4.0 incorporates various innovative technologies, including artificial intelligence, diagnostic decision support systems, personalized treatment methods, the IoT, health monitoring devices, telemedicine applications, robotic surgery, rehabilitation and assistive devices, health clouds, virtual reality, and augmented reality. However, the limitations of widespread health applications and emotionally intelligent devices have necessitated the integration of smart sensors into the next phase, Health 5.0. Sensor technologies have significantly enhanced healthcare services since their inception and are widely used to capture health data and convert them into observable electrical signals [2]. Health metrics such as heart rate, blood glucose level, stress rate, oxygen saturation, temperature, weight, and blood pressure are commonly captured through sensory smart devices and transmitted as electrical signals for further processing [3]. Sensors have been successfully integrated into smartphones and smart wearable devices with the necessary capabilities for capturing and processing health data remotely. The foundation of this technology is the IoT, enabling any conceivable object to access the internet and communicate with other devices, effectively making every object "smart" [4]. When health data is collected, transmitted, and stored by IoT sensors, IoT enables data analytics and intelligent healthcare, which can enhance risk factor identification, disease diagnoses, treatment, and remote monitoring, empowering individuals to manage their health independently. IoT used in healthcare aims to improve technologies utilized in health services and reduce costs across all aspects of healthcare [5]. Today, IoT is also used to monitor, manage, and detect various aspects of the human body, supported by wearable devices. These wearable devices are particularly significant for disadvantaged individuals. For those in society aged 65 and over, who are often referred to as the elderly, restrictions during the COVID-19 pandemic introduced them to technology more rapidly. During this time, the use of remote patient monitoring devices and the continuation of home healthcare services became increasingly common, encouraging individuals aged 65 and over to adopt wearable technologies and accelerating the spread of these IoT-based technologies within this age group. To facilitate access to these technologies for individuals aged 65 and over, and to make the technology more widespread, it is essential to develop devices suited to their use and ensure they are available at affordable prices. Additionally, data obtained through wearable devices can be stored in cloud-based systems for extended periods, enabling the creation of extensive patient records. This study aims to examine technological developments in the healthcare field, emphasizing the significance of the IoT. Furthermore, it aims to explore

the applications of IoT-based wearable technologies in healthcare, specifically their impact on the health of individuals aged 65 and over.

Materials-Methods: The literature review will be conducted in the PubMed, Scopus, Web of Science, and Google Scholar databases between 2017 and 2024 using keywords such as "Internet of Things (IoT)," "Wearable Technology" and "Geriatric Health." Studies addressing the impact of IoT and wearable technologies on elderly health and containing relevant data will be included, while studies with only abstract information will be excluded.

Results: With the advent of Health 4.0, advanced technologies began entering the healthcare field, and this progression has accelerated with Society 5.0. Developments, particularly in IoT-based technologies, artificial intelligence, and robotic surgery, have made healthcare services more efficient and personalized. Research shows that the use of IoT-based technologies is increasing daily, extending beyond traditional healthcare and empowering individuals aged 65 and older to better manage their health. For elderly individuals, IoT-based wearable technologies provide essential tools for real-time monitoring and managing vital metrics like heart rate, blood glucose, and blood pressure. Studies conducted before and after the COVID-19 pandemic demonstrate that the use of wearable technologies has increased among individuals aged 65 and over, supporting independent living and enhancing quality of life. To support this growth, it is crucial to develop user-friendly, affordable devices that meet the specific needs of older adults. Additionally, storing health data in cloud-based systems provides valuable insights to healthcare providers and contributes to creating long-term, comprehensive patient records. The proactive and inclusive health model initiated by Health 4.0, along with the advancement of Society 5.0 and IoT-based wearable technologies and the removal of accessibility barriers, will help create a more equitable, effective, and sustainable healthcare system that benefits all individuals, regardless of age.

Discussion-Conclusion: The technological advancements in healthcare, beginning with Health 4.0 and progressing to Society 5.0, have gained prominence due to the ever-increasing need for health monitoring. The integration of IoT-based wearable technologies into daily life and their appropriate use mark a new era in healthcare. The growing demand for healthcare personnel, coupled with challenges such as natural disasters and the COVID-19 pandemic, has underscored that wearable technologies can play a crucial role in health monitoring, emergency response, and disaster medicine. IoT and wearable technology also emerge as significant innovations with substantial potential to enhance the health and quality of life for elderly individuals. These technologies not only make the daily lives of older adults safer and more independent but also facilitate easier access to healthcare services. Studies show that wearable devices effectively manage critical health risks in elderly populations, such as cardiovascular health, sleep quality, physical activity levels, and fall prevention. Access to these technologies, the ability to use devices, and perceptions of technology among elderly individuals are critical factors for the adoption of these innovations. Additionally, concerns about privacy and data security could influence the willingness of elderly individuals to embrace these systems. Therefore, future research should focus on developing more user-friendly, reliable, and accessible solutions that consider the needs and expectations of older adults. IoT and wearable technologies offer significant potential in healthcare and contribute to the improvement of overall health by supporting healthy living habits. With the continued progression of technological innovations in healthcare, the positive impact of IoT and wearable technologies on individual and public health is expected to strengthen further.

Keywords: Artificial Intelligence, Internet of Things, Wearable Technologies, Health

REFERENCE

[1] Akalin B, Veranyurt Ü. Sağlık 4.0 ve Sağlıkta Yapay Zekâ. Vol. 4, Sağlık Profesyonelleri Araştırma Dergisi / Journal of Health Professionals Research. 2022.

[2] Li, J., & Carayon, P. (2021). Health Care 4.0: A vision for smart and connected health care. IISE Transactions on Healthcare Systems Engineering, 11(3), 171-180.

[3] Javaid, M., Haleem, A., Singh, R. P., Rab, S., & Suman, R. (2021). Significance of sensors for industry 4.0: Roles, capabilities, and applications. Sensors International, 2, 100110.

[4] Aşkan, L. (2018). Internet of things (Nesneleri interneti) nedir?, (Accessed:23.10.2024) http://www.teknolo.com/internet-things-nesnelerin-interneti-nedir/(

[5] Aceto, G., Persico, V., & Pescapé, A. (2020). Industry 4.0 and health: Internet of things, big data, and cloud computing for healthcare 4.0. Journal of Industrial Information Integration, 18, 100129.

A BIBLIOMETRIC ANALYSIS ON THE EFFECTS OF ARTIFICIAL INTELLIGENCE ON DIGITAL TRANSFORMATION IN THE HEALTHCARE ECOSYSTEM

Ali Alsaç*

[1] R&D, Experteam, Istanbul, Türkiye

** Corresponding author: ali.alsac@experteam.com.tr*

Introduction-Aim: The aim of this study is to conduct a holistic review that reveals the research trends of studies on the digital transformation of the healthcare ecosystem. For this purpose, publications focusing on the digitalisation of subsystems in the healthcare ecosystem were analysed using bibliometric analysis method.

Within the scope of this study, the Web of Science (WoS) database was scanned and the relevant studies in the period 2010-2024 were included in the bibliometric analysis. As a result of the related searches, 201 academic studies written in English were obtained. The studies were analysed in various categories such as distribution by years, number and types of studies, most prolific author, most recurring words, countries and institutions with the highest number of academic publications, and the conceptual, intellectual and social frameworks of the publications were identified and expressed. According to the findings of the analysis, the most prominent concepts are innovation, health services, digital health, e-health and COVID-19. This study also includes the effects of the COVID-19 pandemic process and reveals the current literature. In this context, this study is an important resource as it presents current issues in the field of health economics and digital health.

Materials-Methods: In this study, research articles published between 2010-2024 on digital transformation and artificial intelligence applications in the health ecosystem were evaluated by bibliometric analysis method. Web of Science database was used to obtain the research articles that constitute the data of the study [1].

Results: It was observed that most of the studies were multi-authored. This shows that the studies were written with the combination of different perspectives. Another finding is that case-case-clinical studies are relatively few. In addition, the number of studies on 'robotic surgery', which is an application area, is high. However, it is seen that the studies do not focus on the application area sufficiently. In this context, it is important for researchers to contribute to the increase of knowledge in the field of application in addition to theoretical knowledge. Thus, the practical application effects and results of theoretical knowledge will be revealed.

Discussion-Conclusion: In the conclusions, the findings of the research, discussion (supported by relevant literature) or expected results should be presented in paragraph(s) without indentation. The results of the analysis showing that technological developments positively affect the performance of artificial intelligence applications show that artificial intelligence approaches in health will continue to be a popular research area in the coming years. The results of the detailed bibliometric analysis presented in this study will guide research in this field and motivate researchers and institutions/organisations to collaborate.

Keywords: Artificial Intelligence, Bibliometric Analysis, Digitalization, Healthcare Ecosystem

REFERENCE

[1] S. A. S. AlRyalat, L. W. Malkawi, ve S. M. Momani, "Comparing Bibliometric Analysis Using PubMed, Scopus, and Web of Science Databases", J. Vis. Exp. JoVE, sy 152, s. e58494, Eki. 2019, doi: 10.3791/58494.

CONTRIBUTION OF ARTIFICIAL INTELLIGENCE TO DIAGNOSIS WITH MAGNETIC RESONANCE IMAGING IN LUMBAR DISC HERNIATIONS AND APPLICATION IN COMPLEX REALITY

Emin Uysal[1*], Fatema Shaaban[1], Hayrunnisa Demirbaş[2], Çetin Utlu[3], Sude Uyanık[1]

[1]*Pi Health Technologies Limited Company, Ankara, Türkiye*

[2]*Kırıkkale University, Department of Computer Engineering, Kırıkkale, Türkiye*

[3]*University of Health Sciences, Gülhane Faculty of Medicine, Ankara, Türkiye*

**Corresponding author: eminuysal06@gmail.com*

Introduction – Purpose: Lumbar disc herniation (LDH) is characterized by a change in the normal position of the discs due to overloading of the vertebral discs and usually occurs during degeneration [1,2]. Magnetic resonance imaging (MRI) is the most commonly used method in the diagnosis of disc herniation because MRI provides detailed images of the discs and surrounding tissues, providing high accuracy in the diagnostic process. In particular, MRI analyses supported by deep learning and artificial intelligence methods provide radiologists with significant convenience in the diagnosis process [3].

Today, rapid advances in medical technology offer significant opportunities for integrating new technologies such as artificial intelligence and mixed reality (MR) into diagnostic processes. Alsmirat et al. state that AI-assisted diagnostic systems can automatically detect abnormalities in spine images and are used to reduce the workload of radiologists and increase diagnostic accuracy [4]. By training on large data sets prepared with lumbar MRI, artificial intelligence will enable automatic and rapid detection of diseases such as LDH, minimizing human errors and increasing diagnostic accuracy.

The aim of this project is to develop an innovative diagnostic system using artificial intelligence (AI) and complex reality (MRI) technologies for the diagnosis of lumbar disc herniations. It is aimed to increase the automatic detection and diagnostic accuracy of disc herniations with deep learning algorithms on MRI. With this system, the workload of radiologists will be lightened, time will be saved and possible errors will be minimized.

The application, developed using Unity and C#, will allow experts to analyze the MRI in detail and examine it from different angles, accelerating the diagnosis process and supporting more accurate results. With the integration of artificial intelligence and complex reality, the project will provide the healthcare industry with technological solutions that make diagnostic processes more efficient.

Material – Methods: In this project, artificial intelligence and deep learning methods were applied for the diagnosis of lumbar disc herniations. Deep learning techniques have made significant progress, especially in segmentation and registration applications. Studies have been conducted on the application of these architectures in MRI detection, registration, segmentation and classification [5,6]. In the first stage, the discs and spinal bones on the MRI were labeled by radiologists with the Labeling application. Deep learning based algorithms such as Mask R-CNN, YOLO and Faster R-CNN were used for automatic detection and classification of structures in the lumbar spine. The model development process was carried out on the Google Colab platform, and a complex reality (MR) application was developed using Unity and C# for visual analysis of the results and three-dimensional visualization.

The Mask R-CNN algorithm was used to precisely delimit disc herniations on a pixel-by-pixel basis. During the training process, model performance was optimized by allocating 80% of 4000(open source) lumbar MRIs as training set and 20% as test set.

YOLO algorithm was preferred for fast disc detection and classification. Preliminary analysis with YOLO detected the differences between diseased and healthy discs. In addition, Faster R-CNN was used for the high accuracy operations to accurately detect the spine structures and classify the herniation status of the discs. The dataset includes MRI images labeled with diseased and healthy discs, which were manually labeled with Labelme software.

Finally, the project outputs were presented in a three-dimensional environment in a complex reality (MR) (Figure 1) application developed with Unity and C#. This application aims to speed up the diagnostic process and increase accuracy by allowing radiologists to examine the MRI in more detail.

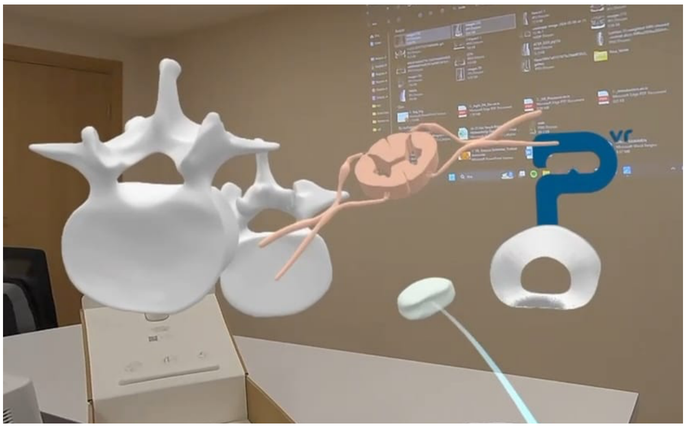

Figure 1

Results: The Mask R-CNN algorithm provides precise delineation of disc and spine structures in MRI, enabling the detection of herniation regions at the pixel level. The YOLO algorithm will significantly accelerate the diagnostic processes of radiologists by obtaining effective results in the identification of healthy and diseased discs with its fast analysis capability. The Faster R-CNN model has made significant contributions to the analysis of spinal structures and classification of herniations with high accuracy rates. In addition, the developed complex reality application will enable three-dimensional visualization of MRI and help experts to examine the images in more detail. With the developed software, it takes an average of 10-15 seconds to identify vertebrae, healthy and diseased discs in the input MRI (Figure 2). These findings clearly demonstrate the significant advantages in the diagnosis of lumbar disc herniation.

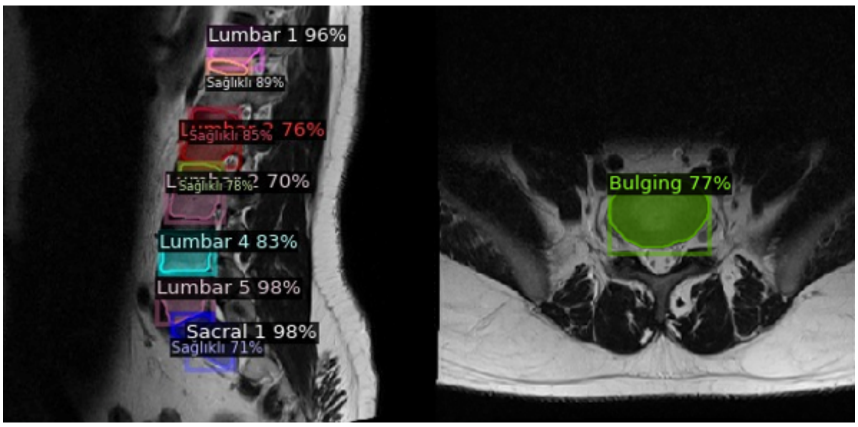

Figure 2

Discussion – Conclusion: This project successfully realized the integration of artificial intelligence and complex reality technologies for the automatic diagnosis of lumbar disc herniations. Deep learning algorithms, specifically Mask R-CNN and YOLO, were used for accurate detection of disc herniations on MRI. The results show a significant increase in diagnostic accuracy and aim to alleviate the workload of radiologists.

The developed complex reality application aims to provide faster and more accurate diagnoses thanks to the three-dimensional visualization of MRI. This project emphasizes the importance of the integration of artificial intelligence and complex reality technologies in the field of healthcare and demonstrates their potential to offer innovative solutions in modern medicine. Expanding these methods with different data sets in future studies aims to improve the quality of healthcare services.

Keywords: Artificial Intelligence, Complex Reality, Lumbar Vertebra, Magnetic Resonance Imaging (MRI), Convolutional Neural Network (CNN)

REFERENCES

[1] Sarı S, Aydoğan M. An important cause of low back pain: Lumbar disc herniation. TOTBID Journal. 2015; 14:298-304.

[2] Deyo RA, Mirza SK. Clinical practice. Herniated lumbar intervertebral disc. N Engl J Med. 2016; 374(18):1763-72.

[3] G. D'Andrea et al, “Intradural lumbar disc herniations: role of MRI in preoperative diagnosis and review of the literature,” Neurosurgical Review, vol. 27, pp. 75-80, 2004.

[4] M. Alsmirat, N. Al-Mnayyis, M. Al-Ayyoub, and A. M. Asma'a, “Deep learning-based computer-aided diagnosis of disc herniation from MRI axial scans,” IEEE Access, vol. 10, pp. 32315-32323, 2022.

[5] J.-G. Lee, S. Jun, Y.-W. Cho, H. Lee, G. B. Kim, J. B. Seo, and N. Kim, “Deep Learning in Medical Imaging: Overview,” Korean Journal of Radiology, vol. 18, no. 4, pp. 570-584, Jul./Aug. 2017.

[6] [J. Liu et al, “Applications of deep learning to MRI images: A survey,” in Big Data Mining and Analytics, vol. 1, no. 1, pp. 1-18, March 2018.

BONE FRACTURE DETECTION IN MEDICAL IMAGES USING YOLO

Büşra Erdoğan[1], Emre Ölmez[1*]

[1] Biomedical Engineering, Izmir Bakırcay University, İzmir, Türkiye

** Corresponding author: emre.olmez@bakircay.edu.tr*

Introduction-Aim: Bone fractures are a common occurrence in a variety of situations, including accidents, sporting events, and in cases of osteoporosis [1]. The diagnosis of these fractures is challenging due to the necessity of radiological observation of subtle details within the bone structure. In standard imaging techniques, small fractures, especially in complex anatomical structures, may be overlooked or misinterpreted, which could result in misdiagnoses that could adversely affect patient recovery [2]. The integration of artificial intelligence and machine learning techniques with medical imaging offers the potential for faster and more accurate fracture detection [3]. In particular, deep learning-based models aim to establish a new standard for bone fracture detection.

Materials-Methods: The objective of this study is to detect bone fractures using one of the deep learning algorithms, namely You Only Look Once (YOLO). YOLO is an object detection solution that is renowned for its rapid processing speed and compact model size. The YOLO model processes the image directly in order to determine the coordinates of the bounding box and the category of the object in question [4]. The dataset utilized in this study was procured from the Universe Roboflow platform, which provides support for object detection and classification models. The dataset comprises images of fractures in the elbow, finger, forearm, humerus, shoulder, and wrist [5]. A total of 1,728 images underwent data augmentation, resulting in 4,148. Of these, 3,631 were employed for training, 348 for validation, and 169 for testing. YOLOv9c was selected for training [6].

Results: Table 1 illustrates the precision and recall values obtained with YOLOv9c for each fracture class in the test set. Precision reflects the accuracy of the model's fracture detections, indicating its capacity to maintain a low false-positive rate. In contrast, recall indicates the model's efficacy in identifying existing fractures, reflecting the extent to which the model is able to accurately capture fracture cases.

Table 1. Bone Fracture Detection Results with YOLOv9c: Class-Based Precision and Recall Values

	Precision	**Recall**
Elbow Fracture	0,50	0,03
Fingers Fracture	0,70	0,15
Forearm Fracture	0,92	0,28
Humerus Fracture	0,95	0,54
Shoulder Fracture	1,00	0,05
Wrist Fracture	0,66	0,07

Discussion-Conclusion: This study offers a comprehensive assessment of the efficacy of the YOLOv9c model in identifying bone fractures. The results demonstrate high precision values, while recall remains relatively low. A high level of precision indicates that the model exhibits a low rate of false positives, suggesting that the majority of detected fractures are accurate. Nevertheless, the low recall rate indicates that the model is unable to detect a considerable number of actual fracture cases.

In general, the YOLOv9c model demonstrates high precision for specific fracture types; however, its detection rate is constrained. The high precision values indicate that the model has the capacity for accurate fracture detection with a low incidence of false positives. However, low recall values indicate instances of missed fracture cases, thereby suggesting potential avenues for improvement. Further research could address these limitations by retraining the model on a more diverse and extensive dataset or by incorporating optimizations to improve recognition of various fracture types. Such optimizations could enhance the model's potential as a reliable tool in clinical applications.

Keywords: Artificial intelligence, Medical Image Processing, Yolov9, Bone Fracture Detection

REFERENCE

[1] C.-Y. Hsu, S.-Y. Huang, and W.-J. Cheng, "Evaluation of the effects of a Chinese herb in nonoperative bone fractures in Taiwan: A Retrospective multi-central database Cohort Study," *J. Herb. Med.*, p. 100956, Oct. 2024, doi: 10.1016/J.HERMED.2024.100956.

[2] P. M. Ricci *et al.*, "Cone-beam computed tomography compared to X-ray in diagnosis of extremities bone fractures: A study of 198 cases," *Eur. J. Radiol. Open*, vol. 6, pp. 119–121, Jan. 2019, doi: 10.1016/J.EJRO.2019.01.009.

[3] Q. Yu *et al.*, "Multi-task learning for calcaneus fracture diagnosis of X-ray images," *Biomed. Signal Process. Control*, vol. 99, p. 106843, Jan. 2025, doi: 10.1016/J.BSPC.2024.106843.

[4] P. Jiang, D. Ergu, F. Liu, Y. Cai, and B. Ma, "A Review of Yolo Algorithm Developments," *Procedia Comput. Sci.*, vol. 199, pp. 1066–1073, Jan. 2022, doi: 10.1016/J.PROCS.2022.01.135.

[5] Veda, "bone fracture detection Computer Vision Project," universe.roboflow.

[6] Ultralytics, "YOLOv9: A Leap Forward in Object Detection Technology," Ultralytics.

AS BOTH INTERVENTION AND QUALITATIVE DATA COLLECTION TOOLS IN IMPROVING MENTAL HEALTH: SOCIAL ROBOTS

Kader Mert [1*]

[1] Department of Nursing, İzmir Bakırçay University, İzmir, Türkiye

**Corresponding author: kader.mert@bakircay.edu.tr*

Introduction-Aim: Talking is the most basic need for socialization. People are sometimes negatively affected by the negative feedback, judgments and facial expressions of the people they talk to. However, people need to be actively listened to and understood by the people they communicate with. Today, social robots are used in research to increase the well-being of disadvantaged groups. The aim of this study is to examine the use of social robots in efforts to improve mental health in society.

Material-Methods: In this study, which was conducted using the literature review method, the use of the developed social robots to improve the mental health of individuals were examined. In the study, a literature review was conducted using leading databases such as PubMed, Google Scholar, and Web of Science. The keywords '*Social robotics*', '*social assistive robots*', and 'mental health' were used as search terms.

Results:The most frequently used social robots in research are Paro, Betty, NAO, Haptic Creature, CRECA [1]. Paro has touch sensors that can detect people and the environment, light, hearing, temperature and posture sensors. This robot is used to provide comfort, companionship and stress reduction for bedridden patients. Betty has touch, motion and visual sensors. Thanks to the touchpad, camera and microphone, this robot has a calming, educational and motivational function that can mimic for motivation, fun and companionship. NAO has sensors such as visual recognition, voice, speech and touch. This robot is used for motivation and companionship in children with mental disorders. Haptic Creature promotes emotional engagement through touch, using it to reduce anxiety and stress similar to animal-assisted therapy. CREA is connected to a computer and microphone to use natural language processing. It has a head nodding feature. This robot is an educational and motivational robot that can mimic verbal and non-verbal interactions between counselors and clients. The functions of social robots can be enhanced to include features such as extracting meaning from sentences, giving positive feedback, summarizing what is said, describing mood, making suggestions such as music, movies, documentaries. In this case, it helps the person to better understand themselves and their emotions, and by recording the emotional states of the individual; it can enable the detection of conditions such as depression tendency, lonelly, social isolation [2] social anxiety [3]. Social robot interventions generally show positive effects on patients with mental health disorders [4,5].

Discussion-Conclusion: The application of valid and reliable tests appropriate for the age group of the person to evaluate the emotions and thoughts of social robots at certain intervals can enable the mental health of people to be monitored. Social robots with these features make it possible to improve the mental health of the elderly, disabled people, people with psychological health problems, people who have been hospitalized for a long time, and especially to evaluate the daily activities of patients who are closed to communication and to examine their emotional states from social robot records. In particular, the examination of social robot records can be used as a data collection tool in research aimed at improving mental health in disadvantaged groups with communication problems.

***Key words**: Social robotics, social assistive robots, mental health,*

REFERENCE

[1] Scoglio AA, Reilly ED, Gorman JA, Drebing CE. (2019) Use of Social Robots in Mental Health and Well-Being Research: Systematic Review. J Med Internet Res. Jul 24;21(7):e13322. doi: 10.2196/13322.

[2] Lederman, Z., & Jecker, N. S. (2023). Social Robots to Fend Off Loneliness?. *Kennedy Institute of Ethics journal*, *33*(3), 249–276. https://doi.org/10.1353

[3] Rasouli, S., Gupta, G., Nilsen, E., & Dautenhahn, K. (2022). Potential Applications of Social Robots in Robot-Assisted Interventions for Social Anxiety. *International journal of social robotics*, *14*(5), 1–32. https://doi.org/10.1007

[4] Guemghar I, Pires de Oliveira Padilha P,Abdel- Baki A, Jutras-Aswad D, Paquette J, Pomey MP (2022) Social robot interventions in Mental Health Care and their outcomes, Bariers,and Facilitators: Scoping Review JMIR Ment Health, 2022;9(4):36094.

[5] Rasouli, S., Gupta, G., Nilsen, E., & Dautenhahn, K. (2022). Potential Applications of Social Robots in Robot-Assisted Interventions for Social Anxiety. *International journal of social robotics*, *14*(5), 1–32. https://doi.org/10.1007