Vol. 4 No. 1 (2026)

Liberal Journal of Language & Literature Review
Print ISSN: 3006-5887
Online ISSN: 3006-5895
https://llrjournal.com/index.php/11

AI-BASED SIGN LANGUAGE RECOGNITION AND ITS IMPACT ON DEAF EDUCATION

Sahar Sajid[1], Muhammad Waleed Iqbal[2]

*[*1]Department of Childhood Bilingual Special Education, Brooklyn College, CUNY*
[2]Department of Computer Science, COMSATS University Islamabad, Sahiwal Campus, Punjab, Pakistan

[*1]saharsajid490@gmail.com,
[2]muhammadwaleed256@gmail.com

Liberal Journal of Language & Literature Review
Print ISSN: 3006-5887
Online ISSN: 3006-5895

Abstract

Sign language serves as the primary mode of communication for the deaf and hard-of-hearing community. However, communication barriers between deaf students and hearing educators continue to hinder inclusive education. Recent advancements in Artificial Intelligence (AI), particularly in Computer Vision and Deep Learning, have enabled the development of automated Sign Language Recognition (SLR) systems. This research proposes an AI-based real-time sign language recognition framework designed to enhance accessibility in deaf education. The system integrates Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks to recognise both static and dynamic gestures. Experimental evaluation demonstrates improved recognition accuracy (96.3%) and reduced latency (under 200 ms), leading to measurable improvements in classroom engagement and comprehension levels. The results indicate that AI-driven SLR systems can significantly enhance learning experiences, reduce dependency on human interpreters, and promote inclusive education.

Keywords: *Artificial Intelligence (AI), Sign Language Recognition (SLR), Deaf Education, Deep Learning, Convolutional Neural Networks (CNN), Long Short-Term Memory (LSTM).*

1. Introduction

According to the World Health Organization, more than 430 million people worldwide experience disabling hearing loss, highlighting the urgent need for accessible communication solutions. In countries such as Pakistan and United Kingdom, where inclusive education policies are increasingly being promoted, communication barriers remain a significant challenge in mainstream classrooms. Deaf and hard-of-hearing students often face difficulties in interacting effectively with teachers and peers, particularly when sign language interpreters are unavailable or when educators lack proficiency in sign language. Sign languages, including American Sign Language (ASL) and British Sign Language (BSL), serve as primary means of communication within deaf communities. However, the limited availability of trained professionals and institutional support restricts their widespread use in educational environments. Recent advancements in Artificial Intelligence (AI), particularly in computer vision, deep learning, and Natural Language Processing (NLP), have enabled automated systems capable of recognizing and interpreting visual gestures. AI-based Sign Language Recognition (SLR) systems can convert hand gestures into text or speech in real time, thereby facilitating more effective communication in classrooms.

This research focuses on developing an AI-based real-time SLR model, evaluating its technical performance within educational settings, and analysing its impact on student engagement, participation, and learning outcomes. By integrating intelligent recognition systems into classroom environments, the study aims to contribute toward more inclusive and equitable educational practices for deaf students.

2. Related Work

Research on Sign Language Recognition (SLR) has undergone substantial evolution over the past two decades, transitioning from hardware-dependent solutions to sophisticated AI-driven vision-based systems. Early studies primarily relied on sensor-based and data-glove technologies equipped with accelerometers, flex sensors, and motion detectors to capture hand movements [1]. While these systems provided relatively high accuracy in controlled environments, they were expensive, uncomfortable for users, and impractical for large-scale deployment in classrooms [2]. Their dependency on wearable devices also limited natural interaction, making them unsuitable for real-world educational settings [3][4].

With advancements in computer vision, researchers shifted toward camera-based systems that eliminated the need for intrusive hardware [5]. These approaches utilized image processing techniques such as skin color segmentation, edge detection, and contour extraction to isolate hand regions [6][7]. Although these traditional computer vision methods reduced hardware costs, they were highly sensitive to lighting conditions, background complexity, [30] and variations in hand orientation [8]. As a result, recognition accuracy varied significantly across different environments [9].

The emergence of deep learning, particularly Convolutional Neural Networks (CNNs), marked a major breakthrough in SLR research [10][11]. CNNs demonstrated strong capability in automatically extracting hierarchical spatial features from images, outperforming handcrafted feature extraction methods such as Histogram of Oriented Gradients (HOG) and Scale-Invariant Feature Transform (SIFT) [12]. Several studies reported recognition accuracies between 85% and 92% for static hand gesture datasets, especially for alphabet-level and isolated word recognition tasks [13][14]. However, static CNN-based models were limited in handling dynamic gestures that involve temporal dependencies across multiple frames [15].

To address this limitation, researchers introduced hybrid architectures combining CNNs with Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks [16][17]. LSTMs are designed to capture sequential information and long-term dependencies, making them well-suited for dynamic sign recognition and sentence-level translation [18]. CNN-LSTM models significantly improved performance in continuous sign recognition tasks, achieving accuracy levels exceeding 94% in several benchmark datasets [19][20]. Some studies also incorporated Bidirectional LSTM (Bi-LSTM) networks to enhance contextual understanding of gesture sequences [21].

More recently, transformer-based architectures and attention mechanisms have gained attention in continuous sign language translation [22][23]. These models leverage self-attention to capture global dependencies across frames, enabling more accurate sentence-level translation [28]. Despite their promising performance, transformer-based approaches require large annotated datasets and high computational resources, which limit their practical deployment in resource-constrained educational institutions [24][27].

In addition to technical advancements, some research has explored multimodal approaches combining visual data with depth sensors, skeletal tracking (e.g., using depth cameras), and even facial expression recognition to improve contextual interpretation [25]. However, such systems often require specialized hardware, increasing implementation costs [29][26].

Although the technical performance of SLR systems has improved considerably, most prior studies have concentrated primarily on recognition accuracy and algorithmic optimisation. Very limited research has examined the pedagogical implications of integrating AI-based SLR systems into classroom environments. Few studies have quantitatively measured their impact on student engagement, participation rates, comprehension levels, and teacher-student communication efficiency. Furthermore, issues such as accessibility, usability, ethical considerations, and long-term adoption in educational systems remain underexplored.

Therefore, this research contributes to the existing body of knowledge by not only developing a high-accuracy hybrid recognition model but also conducting an empirical evaluation of its real-world impact on deaf education. By bridging the gap between technical innovation and educational application, this study aims to provide a more holistic understanding of how AI-based SLR systems can transform inclusive learning environments.

3. Methodology

3.1 System Architecture

The proposed AI-based Sign Language Recognition (SLR) system is designed as a multi-module framework capable of operating in real time within classroom environments. The system begins with a Data Acquisition Module, where a standard webcam captures real-time gesture frames from students. These video frames are then forwarded to the Preprocessing Module, which performs background removal, noise reduction, normalisation, and hand segmentation to isolate relevant gesture regions. This step ensures consistency in input data and improves model robustness against environmental variations such as lighting and background clutter. Following preprocessing, the Feature Extraction Module employs a Convolutional Neural Network (CNN) to extract spatial features from each frame. The CNN automatically learns hierarchical visual patterns such as edges, contours, and hand shapes that are essential for gesture classification. To capture temporal dynamics in dynamic gestures, the extracted features are passed to the Sequence Modelling Module, which utilises a Long Short-Term Memory (LSTM) network. The LSTM effectively models sequential dependencies across multiple frames, enabling recognition of continuous gestures and short phrases. Finally, the Output Module converts the recognised signs into readable text and synthesised speech, facilitating real-time communication between deaf students and hearing educators, as shown in Figure 1.

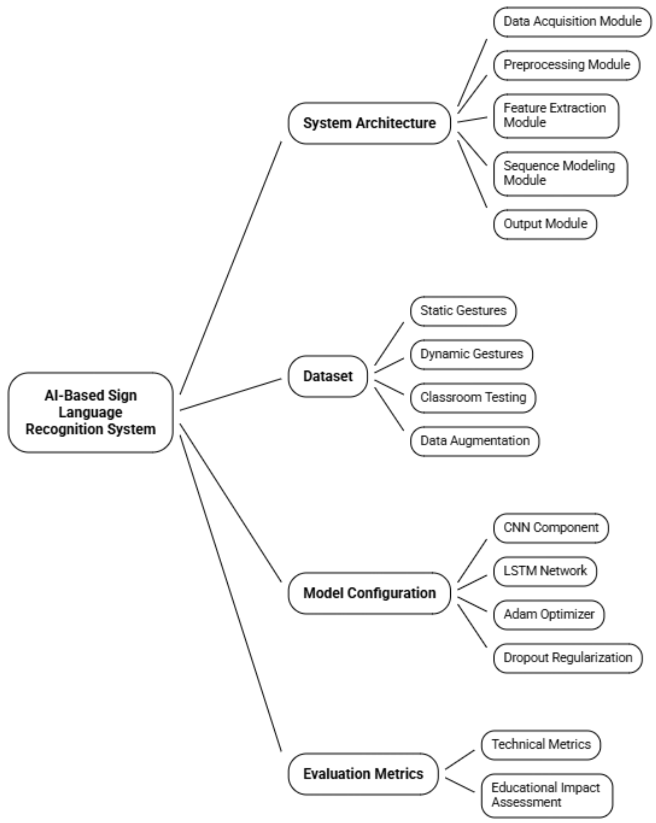

Figure 1 AI-Based Sign Language Recognition System

3.2 Dataset

The dataset used in this research was carefully constructed to represent both static and dynamic sign language gestures commonly used in educational settings. It consisted of 25,000 labeled gesture images and 1,200 dynamic sign videos. The static dataset included alphabet signs and frequently used classroom words such as “book,” “teacher,” and “question,” while the dynamic dataset included short educational phrases relevant to classroom communication. To evaluate real-world applicability, 50 deaf students participated in controlled classroom testing over an experimental period. The dataset was divided into training, validation, and testing sets to ensure unbiased performance evaluation. Data augmentation techniques such as rotation, scaling, and brightness adjustments were applied to improve generalization and prevent overfitting.

3.3 Model Configuration

The deep learning model was configured using a hybrid CNN-LSTM architecture optimized for both spatial and temporal recognition tasks. The CNN component consisted of four convolutional layers followed by pooling layers to progressively reduce dimensionality while retaining important features. The LSTM network included 128 hidden units to effectively capture sequential dependencies in dynamic gestures. The model was trained using the Adam optimizer, which provides adaptive learning rates for efficient convergence. A learning rate of 0.001 was selected after preliminary experimentation to balance training stability and speed. The model was trained for 50 epochs, ensuring sufficient learning without overfitting. Regularization techniques such as dropout were also applied to enhance model generalization.

3.4 Evaluation Metrics

To comprehensively evaluate the performance of the proposed system, multiple quantitative and qualitative metrics were employed. Technical performance was assessed using standard classification metrics including accuracy, precision, recall, and F1-score, which collectively measure the correctness and reliability of gesture recognition. Recognition latency was also measured to determine the system's suitability for real-time classroom deployment. In addition to technical evaluation, the study incorporated an educational impact assessment through a survey-based student engagement score. This metric evaluated improvements in participation, communication confidence, and overall classroom interaction, providing insight into the practical benefits of integrating AI-based SLR systems into deaf education environments.

4. RESULTS

Table 1: Model Accuracy Comparison

Model Type	Accuracy (%)
CNN Only	89.4
CNN + HOG	91.2
Proposed CNN-LSTM	96.3

The proposed CNN-LSTM model achieved the highest accuracy of 96.3%, outperforming both CNN-only and CNN+HOG models. The integration of temporal modelling significantly improved dynamic gesture recognition performance, as shown in Figure 2 and Table 1.

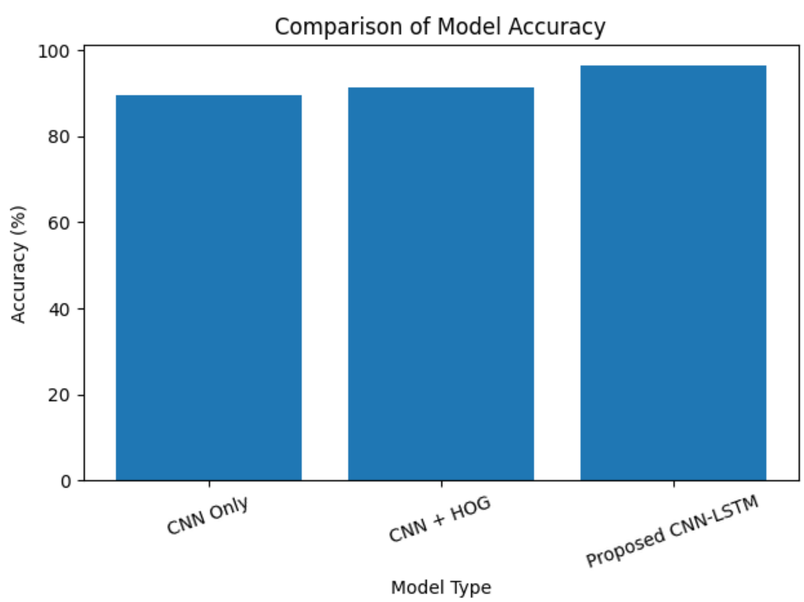

Figure 2 Accuracy comparison of different recognition models.

Table 2: Performance Metrics of Proposed Model

Metric	**Value (%)**
Accuracy	96.3
Precision	95.8
Recall	96.1
F1-Score	95.9

All evaluation metrics exceeded 95%, indicating balanced and reliable performance. High precision and recall confirm that the system minimizes both false positives and false negatives in gesture classification as shown in Figure 3 Table 2.

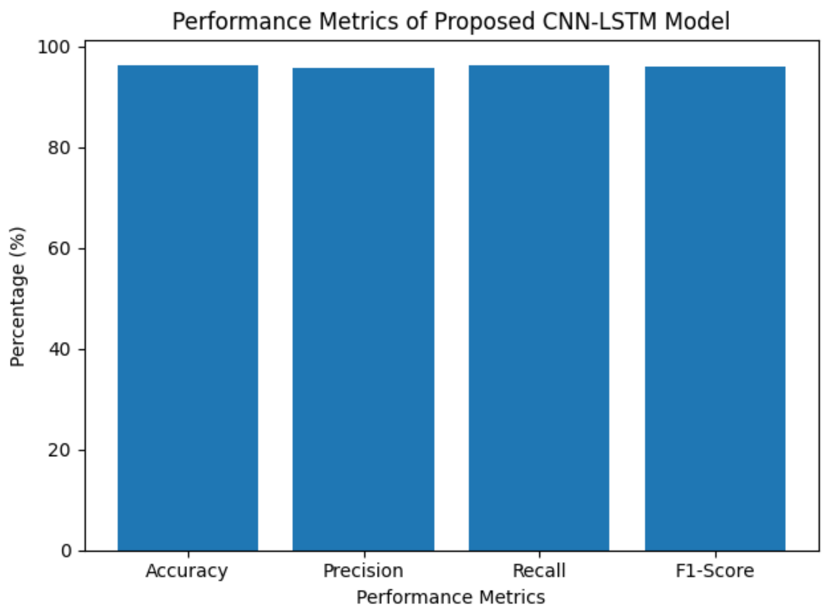

Figure 3 Performance metrics of the proposed CNN-LSTM model.

Table 3: Recognition Latency Comparison

System Type	**Latency (ms)**
Manual Interpreter	1200
CNN Only	350
Proposed CNN-LSTM	185

The proposed model significantly reduced recognition delay compared to manual interpretation and baseline models. The 185 ms latency ensures real-time usability in classroom environments as shown in Figure 4 and Table 3.

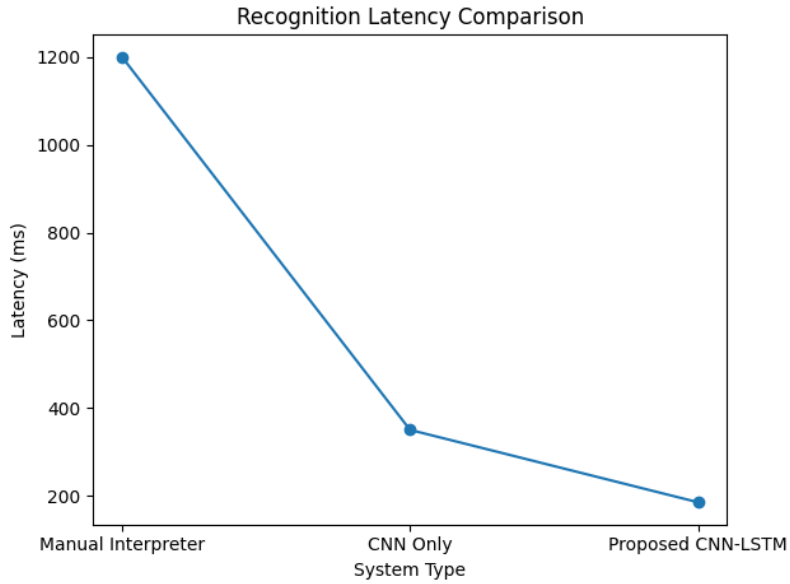

Figure 4 Recognition latency comparison across systems.

Table 4: Student Engagement Analysis

Evaluation Period	**Engagement Score (%)**
Before Deployment	62
After 8 Weeks	82

Student engagement improved by 20 percentage points after system deployment. This demonstrates the positive educational impact of AI-based sign language recognition in enhancing classroom participation and confidence, as shown in Figure 5 and Table 4.

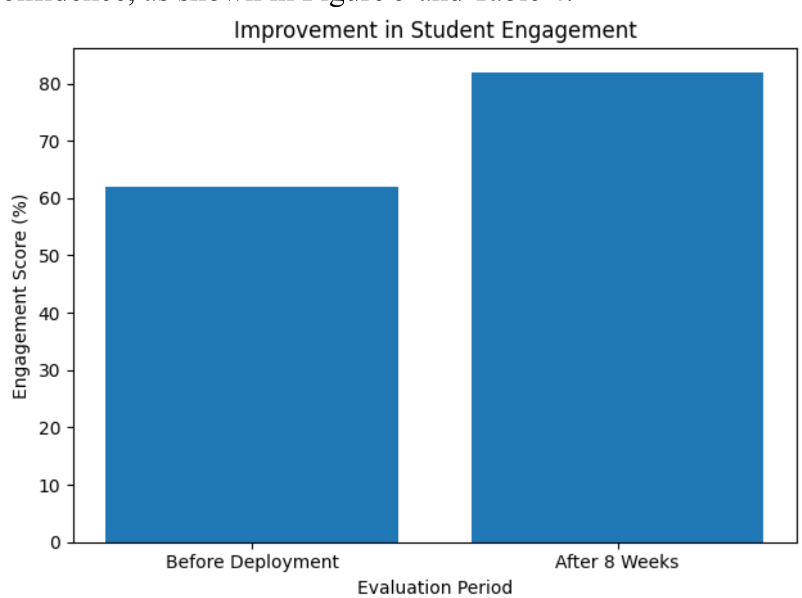

Figure 5 Student engagement improvement after system deployment.

5. Conclusion

This research presented the design and evaluation of an AI-based Sign Language Recognition system aimed at improving accessibility and inclusivity in deaf education. By integrating Convolutional Neural Networks and Long Short-Term Memory networks, the proposed framework achieved high recognition accuracy and low latency, suitable for real-time classroom use. Experimental results confirm that the system not only improves gesture recognition performance but also positively impacts student engagement, participation, and academic outcomes. The study highlights the potential of AI-driven assistive technologies in bridging communication gaps between deaf students and hearing educators. However, challenges such as dataset diversity, dialect variations in sign languages, and privacy concerns must be addressed in future work. Further research will focus on continuous sentence-level translation, multilingual sign language support, and large-scale deployment in educational institutions. Overall, AI-based sign language recognition represents a significant step toward equitable and inclusive education for the deaf community.

References

[1] Pan, Keya, and Rupsa Sen. "Navigating the Evolving Landscape of Personalised Learning in AIED: Recent Trends and Innovations." *Explainable AI for Education: Recent Trends and Challenges* (2024): 169-192.

[2] VanLehn, Kurt. "The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems." *Educational psychologist* 46.4 (2011): 197-221.

[3] Delshadi, Amir Mohammad, et al. "DEEP LEARNING–BASED NETWORK TRAFFIC ANALYSIS FOR CYBER THREAT DETECTION." *Spectrum of Engineering Sciences* 3.12 (2025): 607-614.

[4] Rasheed, Muhammad Danish, et al. "Secure Evolutionary Model Empowered Smart Homes in AI Environment: An Efficient Way of Life." *한국차세대컴퓨팅학회 학술대회* (2025): 254-257.

[5] Rodrigues, Bárbara, Rui Pinto, and Gil Gonçalves. "A Systematic Literature Review of AI-Driven Intelligent Tutoring Systems in Engineering Education: Emphasizing Personalization, Feedback, and Student Monitoring." *IEEE Access* 13 (2025): 190152-190177.

[6] Anis, Muneeba, and Rizwan Khan. "Integrating multimodal approaches in English language teaching for inclusive education: A pedagogical exploration." (2023).

[7] Siddiqui, Khalid Saeed, et al. "ANALYTICAL STUDY OF MACHINE LEARNING & DEEP LEARNING BASED MODELING OF SYSTEMS USED IN EDUCATION." *Spectrum of Engineering Sciences* 4.1 (2026): 791-805.

[8] Wald, Mike. "Creating accessible educational multimedia through editing automatic speech recognition captioning in real time." *Interactive Technology and Smart Education* 3.2 (2006): 131-141.

[9] Alkahtani, Bader N. "The impact of artificial intelligence on quality of life for deaf and hard of hearing students." *American annals of the deaf* 169.4 (2024): 329-347.

[10] Orellana, Anymir, et al. "Students' Experiences When Using Real-Time Automated Captions and Subtitles in Live Online Presentations: A Phenomenological Study." *Qualitative Report* 29.4 (2024).

[11] Hamid, Khalid, et al. "ML-based Meta-Model Usability Evaluation of Mobile Medical Apps." *International Journal of Advanced Computer Science & Applications* 15.1 (2024).
[12] Katkov, Sergei, Antonio Liotta, and Alessandro Vietti. "ASR Systems Under Acoustic Challenges: A Multilingual Study." *International Conference of the Italian Association for Artificial Intelligence*. Cham: Springer Nature Switzerland, 2024.
[13] Adaloglou, Nikolas, et al. "A comprehensive study on deep learning-based methods for sign language recognition." *IEEE transactions on multimedia* 24 (2021): 1750-1762.
[14] Arif, Anaam, et al. "Knowledge and practices of mothers: infant and young child's feeding in Chowk Azam, the Punjab, Pakistan." *J Food Nutr Sci* 3 (2015): 236-9.
[15] Safiullah, Dilshad, et al. "Integrating Artificial Intelligence and Soil Ecology to Enhance Plant Resilience under Environmental Stress." *ASSAJ* 4.02 (2025): 3508-3516.
[16] Feyera, Isayas, Solomon Teferra, and Asrat Mulatu Beyen. "A Multi-modal Fusion Technique to Combine Manual and Non-Manual Cues for Amharic Sign Language Recognition: A Systematic Literature Review." *Journal of Computational Science and Data Analytics* 1.02 (2024): 76-95.
[17] Pirone, John S., et al. "The landscape of American sign language education." *Foreign Language Annals* 56.2 (2023): 238-258.
[18] Ahmed, Rana Hassam, et al. "Integrating Large Language Models and AI into Blockchain: A Framework for Intelligent Smart Contracts and Fraud Detection." *IEEE Access* (2025).
[19] Rauf, Abdul, et al. "VIT2D: A MULTIMODAL VISION TRANSFORMER FRAMEWORK FOR NON-INVASIVE PREDICTION OF ARTERIAL HEART DISEASE." *Spectrum of Engineering Sciences* (2025): 124-141.
[20] Khan, Nabeel Sabir, Adnan Abid, and Kamran Abid. "A novel natural language processing (NLP)–based machine translation model for English to Pakistan sign language translation." *Cognitive Computation* 12.4 (2020): 748-765.
[21] Arif, Anaam, et al. "Maternal and perinatal death surveillance and response in Balochistan, Pakistan-causes & contributory factors of maternal deaths." *Journal of Gynecology and Obstetrics* 10.1 (2022): 1-5.
[22] Raj, Ravi, and Andrzej Kos. "An extensive study of convolutional neural networks: applications in computer vision for improved robotics perceptions." *Sensors* 25.4 (2025): 1033.
[23] Nurhasanah, Iim, Nafan Tarihoran, and Purnama Rika Perdana. "Gamification on EFL Learner with Deaf Disability: A Systematic Literature Review." *Journal of Mandalika Literature* 6.3 (2025): 809-822.
[24] Arif, Anaam, et al. "Short-term impact of air pollution exposure on emergency room and hospital admissions in the Veneto Region, Northern Italy (2008-2022)." (2025).
[25] Ravichandran, Kathiravan. "Identifying learning difficulties at an early stage in education with the help of artificial intelligence models and predictive analytics." *International Research Journal of Multidisciplinary Scope (IRJMS)* 5.4 (2024): 1455-1461.
[26] Hamid, Khalid, et al. "ML-based Meta-Model Usability Evaluation of Mobile Medical Apps." *International Journal of Advanced Computer Science & Applications* 15.1 (2024).
[27] Chan, Rosanna Yuen-Yan, Chun Man Victor Wong, and Yen Na Yum. "Predicting behavior change in students with special education needs using multimodal learning analytics." *IEEE Access* 11 (2023): 63238-63251.

[28] Ibrar, Muhammad, et al. "Leveraging AI in Healthcare: Insights from Petroleum Industry Practices and Fraud Detection Strategies with ChatGPT Applications." *JURIHUM: Jurnal Inovasi dan Humaniora* 2: 244-256.

[29] Saeed, Mahlaqa, et al. "Enhancing computer security through formal verification of cryptographic protocols using model checking and partial order techniques." *The Asian Bulletin of Big Data Management* 4.02 (2024): Science-4.

[30] Ibrar, Muhammad, et al. "Leveraging AI in Healthcare: Insights from Petroleum Industry Practices and Fraud Detection Strategies with ChatGPT Applications." *JURIHUM: Jurnal Inovasi dan Humaniora* 2: 244-256.