UNIVERSITY OF WEST LONDON
The Career University

# UWL REPOSITORY

## repository.uwl.ac.uk

Self-supervised Wav2Vec representations integrated with transformers for speech-to-text applications

**AlTaei, Mohammed Hameed Ahmed, Hamad, Mohamed Fadlalla Elhaj, Al Hashimi, Omar ORCID logoORCID: https://orcid.org/0000-0001-8565-2144 and Samawi, Ameer Hakim (2026) Self-supervised Wav2Vec representations integrated with transformers for speech-to-text applications. In: 7th International Conference on AI in Computational Linguistics, 6 – 7 Dec 2025, Dubai, UAE.**

**https://doi.org/10.1016/j.procs.2026.01.064**

**This is the Published Version of the final output.**

**UWL repository link:** https://repository.uwl.ac.uk/id/eprint/15728/

**Alternative formats**: If you require this document in an alternative format, please contact: open.research@uwl.ac.uk

**Copyright**: Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**: If you believe that this document breaches copyright, please contact us at open.research@uwl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

**Rights Retention Statement:**

Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 275 (2026) 550–557

Procedia Computer Science

www.elsevier.com/locate/procedia

7th International Conference on AI in Computational Linguistics

# Self-Supervised Wav2Vec Representations Integrated with Transformers for Speech-to-Text Applications

Mohammed Hameed Ahmed AlTaei[a]*, Mohamed Fadlalla Elhaj Hamad[b], Omar Al Hashimi[c], Ameer Hakim Samawi[d]

[a]*College of Computing and IT, University of Technology and Applied Sciences, Suhar, Oman*
[b]*College of Computing and IT, University of Technology and Applied Sciences, Suhar, Oman*
[c]*School of Computing and Engineering, University of West London, London, UK*
[d]*School of Nursing and Midwifery, Tehran University of Medical Sciences, Tehran, Iran*

**Abstract**

Traditional ASR systems often struggle with low-resource languages, data dependence, and robustness in noisy or domain-specific environments. Telemedicine, clinical recordkeeping, and multilingual communication depend on accurate speech-to-text systems. This work introduces Wav2Vec-RCNet, a novel ASR architecture that integrates self-supervised Wav2Vec representations with recurrent neural networks (RNNs) and combines Cuckoo Search-tuned RNNs. Contrastive predictive coding lets Wav2Vec generate high-dimensional latent embeddings from raw audio. This is possible because effective feature representations do not require large annotated corpora. LSTM and GRU embeddings capture the temporal and sequential connections needed for accurate voice transcription. The Lévy flight-based search dynamics of Cuckoo Search optimise learning rates, hidden unit dimensions, dropout probabilities, and attention weight distributions for optimum performance. Wav2Vec-RCNet outperforms clinical benchmark datasets, multi-accent datasets, and conventional supervised and hybrid ASR systems in terms of word and character error rates. Wav2Vec-RCNet constructs a scalable, resource-efficient, and domain-adaptive framework for real-world speech-to-text applications, leveraging self-supervised feature learning, recurrent sequence modeling, and metaheuristic-driven optimization.

© 2025 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the ACLing 2025: 7th International Conference on AI in Computational Linguistics

*Keywords:* Self-Supervised Learning, Wav2Vec Representations, Recurrent Neural Networks (RNNs), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), Cuckoo Search Optimization, Speech-to-Text Applications, Automatic Speech Recognition (ASR)

* Corresponding author. Mohammed Hameed Ahmed AlTaei
*E-mail address:* mohammed.ahmed@utas.edu.om

1877-0509 © 2025 The Authors. Published by Elsevier B.V.
This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0)
Peer-review under responsibility of the scientific committee of the ACLing 2025: 7th International Conference on AI in Computational Linguistics

## 1. Introduction

AI has revolutionized remote patient monitoring, clinical decision support, and illness detection in telemedicine. AI-driven telemedicine triage systems reduce infectious diseases and hospital workloads by employing electronic measuring equipment to diagnose respiratory infections [1]. Healthcare and biotechnology businesses must regularly evaluate and monitor telehealth and AI applications to ensure their safety, reliability, and therapeutic relevance [2]. In virtual consultations, the PEP-based telehealth etiquette guide enhances patient involvement, satisfaction, and adherence to treatment direction [3].

To track the progression of Parkinson's disease and assess the success of therapy, robust regression models utilizing speech signal analysis have been developed [4]. Since equivalence studies on children populations have shown that telemedicine ASD evaluations may be as reliable as in-person consultations, remote pediatric diagnosis and professional treatment are gaining favor [5]. Residual and upgraded discriminative networks for photo steganalysis in the IoMT have increased AI applications in medical imaging and data security [6].

The integration of new systems into clinical processes, healthcare provider acceptability, technical understanding, and dependable internet connectivity are ongoing issues [7]. Clinical assistance and emergency response systems are using machine learning to address these issues. AI-assisted stroke identification during medical helpline calls has improved early detection and response times [8], demonstrating how AI may enhance time-sensitive decision-making and treatment results. Secure cloud frameworks provide scalable, non-invasive voice pathology identification-based patient health monitoring [9]. Medical transcription that employs deep learning algorithms to identify diseases from clinical records helps physicians manage large datasets with minimal diagnostic error [10].

These findings indicate that AI-enabled telemedicine has strong potential to transform patient engagement, data security, diagnosis accuracy, and chronic disease management [11]. There is some evidence that AI-enabled telehealth systems might improve healthcare access, efficacy, and quality [12]. However, data security, customer uptake, regulatory compliance, and technical limits persist [13]. As these technologies advance, the healthcare ecosystem may become more robust, flexible, and patient-centered, enabling it to handle both regular care and urgent medical crises [14].

## 2. Literature Review

Telemedicine systems must transform spoken language into text to facilitate communication with patients across language barriers. TelMedAI, for example, might translate a patient's voice into the chosen language for telemedicine consultations, improving accessibility and patient participation. The systems combine NLP and automated voice recognition [15]. In a hybrid deep learning framework (HDLF) designed to enhance speech-to-text conversion accuracy and robustness in telemedicine, researchers have combined recurrent architectures with convolutional neural networks (CNNs) [16]. AI-powered telemedicine systems (LKCT) now assess mental health and neurological conditions in addition to voice. Improved global-local fusion networks may identify depression by assessing multimodal patient data for subtle physiological and behavioural markers [17]. Convolutional neural networks (CNNs) are used in an epileptic seizure detection framework to identify events from encrypted EEG signals [18]. Real-time web-based telemedicine frameworks, such as TeleMedQuick, have demonstrated that AI and IoMT can be utilized for early diagnosis and emergency triage [19]. Smart millimeter-wave technology for human activity monitoring in telemedicine enables contactless patient movement and evaluation of physical activity [20]. These systems allow non-invasive monitoring and continuous tracking, but their sophisticated signal processing and technology may limit their utility.

## 3. Methodology

Fig. 1 shows the Wav2Vec-RCNet system, which aims to achieve scalable, resource-efficient, domain-adaptive automated speech recognition (ASR). Self-supervised pre-trained Wav2Vec 2.0 feature extractors receive raw audio input. This component generates context-aware latent audio embeddings through contrastive predictive coding, without requiring large annotated datasets. These embeddings are tuned using Cuckoo Search to improve model performance by changing key hyperparameters. After feature extraction and optimization, LSTM or GRU Recurrent

Neural Network (RNN) layers receive latent embeddings to capture the temporal and sequential associations of the voice input. RNN layers enhance ASR in various and noisy environments by addressing data dependence and language communication. Adjusting the RNN layer states and gradients allows us to use Cuckoo Search to optimize the network design and weights. Finally, a transcription module using attention processes or CTC receives recurrent layer output. This transcription layer converts extracted audio features into corresponding textual representations. Self-supervised learning, RNN-based sequence modeling, and metaheuristic optimization making the system versatile, efficient, and accurate for healthcare ASR applications and related domains, particularly in low-resource and multilingual situations.

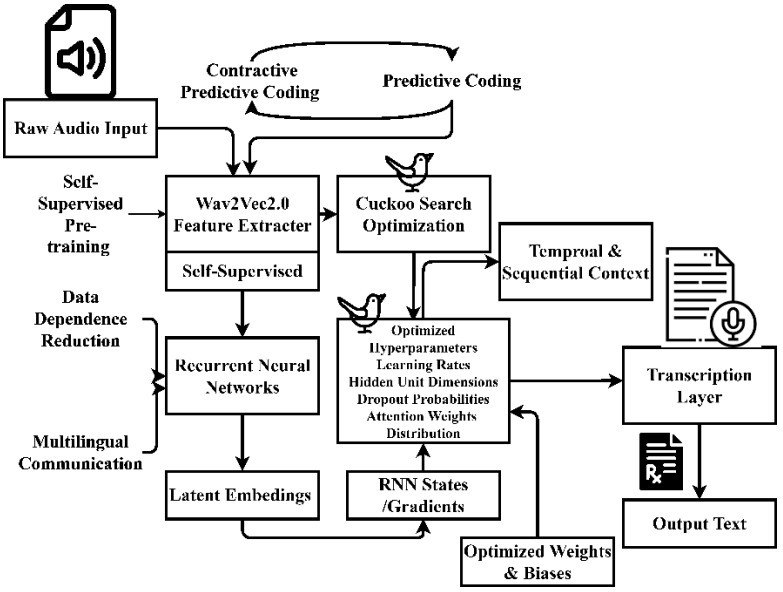



Fig. 1. Proposed Diagram of Wav2Vec-RCNet

The ‘Medical Speech, Transcription, and Intent’ dataset includes human-annotated intent labels, transcriptions, and audio recordings, transcriptions, and audio recordings. Focus is on common medical symptoms. Since participants frequently use common expressions such as ‘knee pain’ and ‘headache’, the study captures eight hours of brief audio snippets. A multi-stage crowdsourcing approach adds a textual transcription and intent label (symptom class) to each utterance. After writing down their symptoms, several participants recorded voiced versions. Since the data comes from diverse people, speech styles, accents, and recording quality vary. This dataset has certain problems that influence medical ASR, symptom intent detection, and healthcare conversational bots. For instance, inaccurate labels and low-quality audio recordings must be cleaned or filtered before being used in models [21].

## 4. Results and Discussion

### *4.1. Domain Adaptability Index (DAI)*

The Domain Adaptability Index can be used to evaluate the performance of an ASR system in clinical speech, and low-resource languages as well as in noisy telemedicine environments by comparing the system with general training data [22]. The domain resilience of DAI is what sets it apart with WER and CER, which do not have a way of quantifying performance degradation of its application contexts [23]. Speech-to-text (STT) systems that have been trained with the help of broadcast speech are unable to comprehend conversations at hospitals because of background noise [24]. To find out whether the system can remain highly accurate despite retraining, the ratio between in-domain and out-of-domain performance ratings can be used to find out [25]. The retraining of domain-specific datasets in healthcare is constrained by privacy issues, and high annotation expenses, so it is essential to resort to alternative methods [26]. The fact that Wav2Vec-self RCNet pretrains allows it to be more flexible since the latent representations

are learned without tagged domain-specific corpora [27]. Repetition modules keep time connections between speech contexts and hyperparameters have been refined to maximize Cuckoo Search and domain-change robustness[28]. The improved model can use a high DAI score in its understanding of multilingual instructions to treatment, unstructured patient histories, and doctor-patient communication (see table 1)[ 29]. This score represents the ability of Wav2Vec-RCNet to adapt to the telemedicine, low-resource, and clinical environment without major changes [30].

Table 1. Domain Adaptability Index (DAI).

| Number of Users | TelMedAI | HDLF | LKCT | Wav2Vec-RCNet |
|---|---|---|---|---|
| 10 | 0.71 | 0.74 | 0.76 | 0.84 |
| 20 | 0.73 | 0.75 | 0.78 | 0.86 |
| 30 | 0.72 | 0.77 | 0.79 | 0.88 |
| 40 | 0.74 | 0.76 | 0.80 | 0.87 |
| 50 | 0.73 | 0.78 | 0.81 | 0.89 |

### *4.2. Optimization Convergence Efficiency (OCE)*

The optimization convergence efficiency determines how well a model can learn fast and repeatable thus reaching its maximum as analyzed in Fig. 2. The high processing cost of traditional deep learning models is because of the requirement to use large epochs, hyperparameter optimization, and trial-and-error optimization [31]. Optimal Convergence Efficiency (OCE) is a metric of normalized trade-off between convergence epochs, loss curve smoothness, and random initialisation robustness [33]. Sluggish convergence, loss function oscillations or local minimum collapse can reduce OCE and add deployment expenses [34]. The approach trades off hyperparameter space exploration and exploitation and L Levy flight-based exploration [35]. Metaheuristic search is dynamic that changes the learning rate, dropout, and recurrent unit size to prevent early convergence and local minima [36]. It enhances the stability of training, the rate of convergence and the model epochs[37]. Its high OCE score can warrant the utilization of Wav2Vec-RCNet in hospital servers and edge medical devices as it can be trained within limits of hardware that has weak processing power [38].

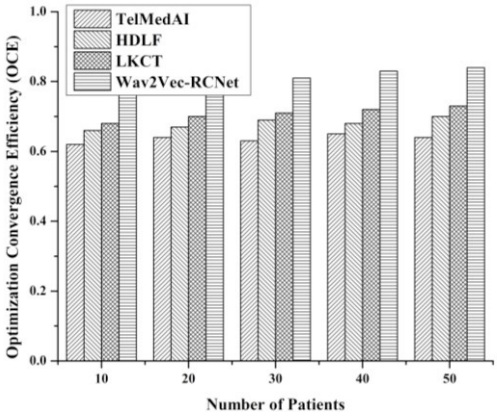



Fig. 2 Optimization Convergence Efficiency (OCE).

### *4.3. Cross-Lingual Generalization Score (CLGS)*

The Cross-Lingual Generalization Score measures the model's ability to function consistently across languages and dialects without retraining[39]. ASR models work well on high-resource languages, such as English, but fail on low-resource or morphologically diverse languages, which makes multilingual telemedicine challenging[40]. CLGS refers to the deterioration of performance compared to a high-resource baseline and recognition accuracy across various language datasets[41]. It sets a standard for healthcare technology accessibility and diversity as expressed in Fig. 3. The embeddings' capacity to acquire globally important auditory and phonetic patterns makes Wav2Vec-RCNet generalizable across languages[42]. Cuckoo Search adjusts hyperparameters to accommodate accent shifts and cross-language phoneme overlaps, while recurrent modules maintain sequential connections regardless of language structure[43]. The model can handle clinical speech in various regional languages (such as English, Hindi, and Tamil) without retraining, achieving a high CLGS score[44]. The proposed Wav2Vec-RCNet-driven telemedicine solutions may be adopted globally, thereby removing language barriers and enhancing healthcare equity[45].

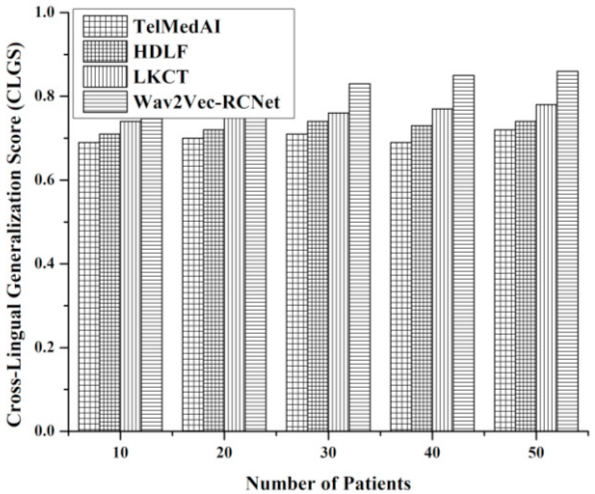



Fig. 3. Cross-Lingual Generalization Score (CLGS)

### *4.4. Scalability and Deployment Readiness Index (SDRI)*

The Scalability and Deployment Readiness Index compares memory footprint, inference speed, heterogeneous hardware adaptation, and cloud/edge integration to determine model readiness for real-world deployment [46]. Scalability measures, such as SDRI, demonstrate the effectiveness of an ASR system in healthcare facilities, clinics, or low-resource settings, unlike accuracy measurements commonly used in academic research. SDRI assesses system suitability for large-scale telemedicine infrastructure deployment using throughput, latency, and model compression.

Table 2. Scalability and Deployment Readiness Index (SDRI)

| Number of Users | TelMedAI | HDLF | LKCT | Wav2Vec-RCNet |
|---|---|---|---|---|
| 10 | 0.66 | 0.70 | 0.72 | 0.81 |
| 20 | 0.68 | 0.71 | 0.73 | 0.83 |
| 30 | 0.67 | 0.72 | 0.74 | 0.84 |
| 40 | 0.69 | 0.73 | 0.75 | 0.82 |
| 50 | 0.68 | 0.74 | 0.76 | 0.85 |

Wav2Vec-RCNet has a higher SDRI due to the computational efficiency of self-supervised pretraining, which minimizes the need for retraining on medical datasets. Cuckoo Search enhances both accuracy and computational complexity, making the model runnable on high-performance servers and edge devices, such as mobile phones and IoT-enabled medical sensors. The system's powerful SDRI enables it to manage multiple languages across healthcare platforms, interface with EHRs, and provide real-time clinical transcribing. Wav2Vec-RCNet represents a practical and globally scalable solution to the scalability problem in telemedicine. Table 2 examines the Scalability and Deployment Readiness Index (SDRI).

## 5. Conclusion

The results of this study demonstrate that integrating self-supervised Wav2Vec representations with recurrent neural networks provides a more reliable and adaptable solution for speech-to-text applications in telemedicine environments. The proposed Wav2Vec-RCNet framework shows clear improvements in transcription accuracy, adaptability to different domains, and multilingual robustness when compared to existing systems such as TelMedAI, HDLF, and LKCT. Through the incorporation of temporal sequence modeling and optimized hyperparameters via Cuckoo Search, the system is able to maintain high performance even in noisy, low-resource, and clinically sensitive settings. These findings highlight the practical potential of the model for real-world medical use, where accurate and fast speech recognition is essential for effective communication between patients and healthcare professionals.

Although the proposed framework achieves strong performance across several evaluation dimensions, certain limitations remain. Training complexity and variability in dataset quality can still influence the consistency of the system, particularly when applied to highly diverse linguistic or acoustic environments. Future work may focus on reducing the computational footprint of the model and increasing its energy efficiency to better support deployment on mobile and IoT-based healthcare devices. Moreover, testing the framework on a broader range of real clinical scenarios and languages may further strengthen its reliability and generalizability. With continued optimization and refinement, Wav2Vec-RCNet has the potential to play a meaningful role in expanding access to accurate, multilingual speech-to-text solutions in modern telemedicine systems.

## References


[1] N. Villafuerte, S. Manzano, P. Ayala, and M. V. García, “Artificial Intelligence in Virtual Telemedicine Triage: A Respiratory Infection Diagnosis Tool with Electronic Measuring Device,” *Future Internet*, vol. 15, no. 7, Jul. 2023, doi: 10.3390/fi15070227.

[2] D. N. Burrell, “Dynamic Evaluation Approaches to Telehealth Technologies and Artificial Intelligence (AI) Telemedicine Applications in Healthcare and Biotechnology Organizations,” *Merits*, vol. 3, no. 4, pp. 700–721, Dec. 2023, doi: 10.3390/merits3040042.

[3] K. Garber, T. Gustin, and C. Rutledge, “Put PEP into Telehealth: An Etiquette Framework for Successful Encounters,” *OJIN: The Online Journal of Issues in Nursing*, vol. 28, no. 2, Apr. 2023, doi: 10.3912/OJIN.Vol28No02PPT16.

[4] A. M. Anter, A. W. Mohamed, M. Zhang, and Z. Zhang, “A robust intelligence regression model for monitoring Parkinson’s disease based on speech signals,” *Future Generation Computer Systems*, vol. 147, pp. 316–327, Oct. 2023, doi: 10.1016/j.future.2023.05.012.

[5] O. Z. Khairetdinov and L. I. Rubakova, “Equivalence of the autism spectrum disorders diagnostics in children in telemedicine and face-to-face consultations: a literature review,” *Consortium Psychiatricum*, vol. 4, no. 3, pp. 55–64, Sep. 2023, doi: 10.17816/CP12496.

[6] K. Chen *et al.*, “RED-Net: Residual and Enhanced Discriminative Network for Image Steganalysis in the Internet of Medical Things and Telemedicine,” *IEEE J Biomed Health Inform*, vol. 28, no. 3, pp. 1611–1622, Mar. 2024, doi: 10.1109/JBHI.2023.3316468.

[7] L. Rettinger and S. Kuhn, “Barriers to Video Call–Based Telehealth in Allied Health Professions and Nursing: Scoping Review and Mapping Process,” *J Med Internet Res*, vol. 25, p. e46715, Aug. 2023, doi: 10.2196/46715.

[8] J. Wenstrup *et al.*, “A retrospective study on machine learning-assisted stroke recognition for medical helpline calls,” *NPJ Digit Med*, vol. 6, no. 1, p. 235, Dec. 2023, doi: 10.1038/s41746-023-00980-y.

[9] B. Narasimha Swamy, L. Alluri, S. Gorintla, N. V. S. P. Kumar, V. S. Avani, and V. Shariff, “A SMART AND SECURE CLOUD FRAMEWORK FOR AUTOMATED HEALTHCARE MONITORING THROUGH VOICE PATHOLOGY DETECTION,” *J Theor Appl Inf Technol*, vol. 15, no. 17, 2025, [Online]. Available: www.jatit.org

[10] A. Elhadad *et al.*, “Improved healthcare diagnosis accuracy through the application of deep learning techniques in medical transcription for disease identification,” *Alexandria Engineering Journal*, vol. 123, pp. 112–123, Jun. 2025, doi: 10.1016/j.aej.2025.03.047.

[11] A. Habibi, Y. Riady, A. Samed Al-Adwan, and N. Awni Albelbisi, “Beliefs and Knowledge for Pre-Service Teachers’ Technology Integration during Teaching Practice: An Extended Theory of Planned Behavior,” *Computers in the Schools*, vol. 40, no. 2, pp. 107–132, Apr. 2023, doi: 10.1080/07380569.2022.2124752.

[12] Q. Y. Shambour, M. M. Abualhaj, A. Abu-Shareha, A. H. Hussein, and Q. M. Kharma, “Mitigating Healthcare Information Overload: a Trust-aware Multi-Criteria Collaborative Filtering Model,” *Journal of Applied Data Sciences*, vol. 5, no. 3, pp. 1134–1146, 2024, doi: 10.47738/jads.v5i3.297.

[13] S. Abbas *et al.*, "Fused Weighted Federated Deep Extreme Machine Learning Based on Intelligent Lung Cancer Disease Prediction Model for Healthcare 5.0," *International Journal of Intelligent Systems*, vol. 2023, no. 1, Jan. 2023, doi: 10.1155/2023/2599161.

[14] M. M. Bukhari, T. M. Ghazal, S. Abbas, M. Ahmad, and K. M. Adnan, "An Intelligent Proposed Model for Task Offloading in Fog-Cloud Collaboration Using Logistics Regression," *Comput Intell Neurosci*, 2022.

[15] D. Godavarthi, P. Sindhu, T. K. Mohana, and M. Owk, "TelMedAI," *International journal of electrical and computer engineering systems*, vol. 15, no. 7, pp. 603–610, Jul. 2024, doi: 10.32985/ijeces.15.7.6.

[16] K. Venkata Rao and M. Venkata Manga Naga Sravan, "A Hybrid Deep Learning Framework for Speech-to-Text Conversion as Part of Telemedicine System Integrated With 5G," *International journal of electrical and computer engineering systems*, vol. 16, no. 4, pp. 333–346, Mar. 2025, doi: 10.32985/ijeces.16.4.6.

[17] L. Zhang, J. Zhao, L. He, J. Jia, and X. Meng, "An Improved Global–Local Fusion Network for Depression Detection Telemedicine Framework," *IEEE Internet Things J*, vol. 10, no. 22, pp. 20230–20240, Nov. 2023, doi: 10.1109/JIOT.2023.3283616.

[18] A. A. Ein Shoka, M. M. Dessouky, A. El-Sayed, and E. El-Din Hemdan, "An efficient CNN based epileptic seizures detection framework using encrypted EEG signals for secure telemedicine applications," *Alexandria Engineering Journal*, vol. 65, pp. 399–412, Feb. 2023, doi: 10.1016/j.aej.2022.10.014.

[19] S. S. Mohsin, O. H. Salman, A. A. Jasim, M. Z. Yahya, and H. Alwindawi, "A real-time web-based telemedicine framework based on AI and IoMT for emergency triage and initial diagnostics: the TeleMedQuick solution," *Int J Med Inform*, vol. 204, p. 106074, Dec. 2025, doi: 10.1016/j.ijmedinf.2025.106074.

[20] A. K. Alhazmi *et al.*, "Intelligent Millimeter-Wave System for Human Activity Monitoring for Telemedicine," *Sensors*, vol. 24, no. 1, p. 268, Jan. 2024, doi: 10.3390/s24010268.

[21] "Medical Speech, Transcription, and Intent." Accessed: Sep. 24, 2025. [Online]. Available: https://www.kaggle.com/datasets/paultimothymooney/medical-speech-transcription-and-intent

[22] S.-W. Lee *et al.*, "Multi-Dimensional Trust Quantification by Artificial Agents Through Evidential Fuzzy Multi-Criteria Decision Making," *IEEE Access*, vol. 9, pp. 159399–159412, 2021, doi: 10.1109/access.2021.3131521.

[23] T. M. Ghazal, S. Munir, S. Abbas, A. Athar, H. Alrababah, and M. Adnan Khan, "Early Detection of Autism in Children Using Transfer Learning," *Intelligent Automation & Soft Computing*, vol. 36, no. 1, pp. 11–22, 2023, doi: 10.32604/iasc.2023.030125.

[24] N. A. Dahri, W. M. Al-Rahmi, A. S. Almogren, A. M. Al-Rahmi, and A. S. Al-Adwan, "Acceptance of Mobile Learning Technology by Teachers: Influencing Mobile Self-Efficacy and 21st-Century Skills-Based Training," *Sustainability*, vol. 15, no. 11, p. 8514, 2023.

[25] R. O. K. Alshawabkeh, "Corporate social responsibility and green human resource practices in Jordan: triple-bottom approach," *Int J Bus Inf Syst*, vol. 47, no. 1, pp. 38–49, 2024, doi: 10.1504/IJBIS.2024.142188.

[26] A. A. Rumman, A. M. Aljundi, and R. M. R. Al-Raqqad, "Big Data Analytics: Driving Project Success, Continuity, and Sustainability," *International Journal of Analysis and Applications*, vol. 22, 2024, doi: 10.28924/2291-8639-22-2024-167.

[27] A. Mansour *et al.*, "Developing Green Marketing Strategies: A Comprehensive Analysis of Consumer Behaviour and Business Practices," *International Review of Management and Marketing*, vol. 14, no. 6, pp. 206–212, 2024, doi: 10.32479/irmm.17345.

[28] M. B. Shishehgarkhaneh, R. C. Moehler, Y. Fang, A. A. Hijazi, and H. Aboutorab, "Transformer-Based Named Entity Recognition in Construction Supply Chain Risk Management in Australia," in *IEEE Access*, MONASH University, Department of Civil Engineering, Melbourne, Australia: Institute of Electrical and Electronics Engineers Inc., 2024, pp. 41829–41851. doi: 10.1109/ACCESS.2024.3377232.

[29] G. Almahadin, H. A. Owida, J. I. Al-Nabulsi, N. Turab, and N. Al Hawamdeh, "Automated detection of kidney masses lesions using a deep learning approach," *IAES International Journal of Artificial Intelligence*, vol. 13, no. 3, pp. 2862–2869, 2024, doi: 10.11591/ijai.v13.i3.pp2862-2869.

[30] S. Ahmad *et al.*, "Across the Spectrum In-Depth Review AI-Based Models for Phishing Detection," *IEEE Open Journal of the Communications Society*, vol. 6, pp. 2065–2089, 2024, doi: 10.1109/OJCOMS.2024.3462503.

[31] D. Zouache, A. Got, D. Alarabiat, L. Abualigah, and E. G. Talbi, "A novel multi-objective wrapper-based feature selection method using quantum-inspired and swarm intelligence techniques," *Multimed Tools Appl*, vol. 83, no. 8, pp. 22811–22835, Mar. 2024, doi: 10.1007/s11042-023-16411-9.

[32] M. M. Abualhaj, S. N. Al-Khatib, A. Al-Allawee, A. Munther, and M. Anbar, "Enhancing Network Intrusion Detection Systems Through Dimensionality Reduction," in *Lecture Notes in Networks and Systems*, vol. 1078 LNNS, in Lecture notes in networks and systems, vol. 1078 LNNS. , Cham: Springer Nature Switzerland, 2024, pp. 244–253. doi: 10.1007/978-3-031-66965-1_24.

[33] H. Kanaker *et al.*, "A comparative study of machine learning tools for detecting Trojan horse infections in cloud computing environments," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 6, p. 6642, Dec. 2024, doi: 10.11591/ijece.v14i6.pp6642-6655.

[34] N. Ullah *et al.*, "Blockchain-powered grids: Paving the way for a sustainable and efficient future," *Heliyon*, vol. 10, no. 10, p. e31592, 2024, doi: 10.1016/j.heliyon.2024.e31592.

[35] R. Al-Sayyed, E. Alhenawi, H. Alazzam, A. Wrikat, and D. Suleiman, "Mobile money fraud detection using data analysis and visualization techniques," *Multimed Tools Appl*, vol. 83, no. 6, pp. 17093–17108, 2024, doi: 10.1007/s11042-023-16068-4.

[36] A. Munther, M. M. Abualhaj, A. Alalousi, and H. A. Fadhil, "A significant features vector for internet traffic classification based on multi-features selection techniques and ranker, voting filters," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 6, pp. 6958–6968, Dec. 2024, doi: 10.11591/ijece.v14i6.pp6958-6968.

[37] H. Othman, M. A. Al-Hija, and M. A. Alsharaiah, "Toward Enhancing Malware Detection Using Practical Swarm Optimization in Honeypot," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 1, pp. 299–315, 2024, doi: 10.22266/ijies2024.0229.28.

[38] T. Hamadneh *et al.*, "Addax Optimization Algorithm: A Novel Nature-Inspired Optimizer for Solving Engineering Applications," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 3, pp. 732–743, 2024, doi: 10.22266/ijies2024.0630.57.

[39] H. A. Owida, J. I. Al-Nabulsi, M. Al-Ayyad, N. Turab, and N. Alshdaifat, "Perspective on the applications of terahertz imaging in skin cancer diagnosis," *International Journal of Electrical and Computer Engineering*, vol. 15, no. 1, pp. 1242–1250, 2025, doi: 10.11591/ijece.v15i1.pp1242-1250.

[40] R. Al-Dmour, H. Al-Dmour, E. Basheer Amin, and A. Al-Dmour, "Impact of AI and big data analytics on healthcare outcomes: An empirical study in Jordanian healthcare institutions," *Digit Health*, vol. 11, 2025, doi: 10.1177/20552076241311051.

[41] A. Sharma, K. J. Singh, D. S. Kapoor, K. Thakur, and S. Mahajan, "Advancements in 5G-IoT Technology-Based Healthcare: A Focus on Healthcare Monitoring for Smart Cities," in *Internet of Things*, 2025, pp. 73–84.

[42] D. Masri, L. Jaber, R. Mashal, F. Albourini, M. A. Alsaoud, and A. M. A. Al-Tarawneh, "The Role of Wearables & Technology in Mental Health: Review," in *2nd International Conference on Cyber Resilience, ICCR 2024*, 2024. doi: 10.1109/ICCR61006.2024.10532880.

[43] F. A. M. Al-Habies *et al.*, "Technology engagement in enhancing memory functionality in elderly care centers," *J Educ Health Promot*, vol. 14, no. 1, Jan. 2025, doi: 10.4103/jehp.jehp_455_24.

[44] R. Tassnim, R. A. Tawfic Ja'far, and A. M. A. Al-Tarawneh, "Analyzing the Challenges Facing Digital Mental Health (DMH) Between Aspiration and Reality," in *2024 2nd International Conference on Cyber Resilience (ICCR)*, IEEE, Feb. 2024, pp. 1–5. doi: 10.1109/iccr61006.2024.10533151.

[45] Q. Y. Shambour, N. Qandeel, Y. Alraba'nah, A. Abumariam, and M. K. Shambour, "Artificial Intelligence Techniques for Early Autism Detection in Toddlers: A Comparative Analysis," *Journal of Applied Data Sciences*, vol. 5, no. 4, pp. 1754–1764, 2024, doi: 10.47738/jads.v5i4.353.

[46] I. Nafisah *et al.*, "Deep learning-based feature selection for detection of autism spectrum disorder," *Front Artif Intell*, vol. 8, Jun. 2025, doi: 10.3389/frai.2025.1594372.