scientific reports

Check for updates

OPEN

Fusing convolutional learning and attention-based Bi-LSTM networks for early Alzheimer's diagnosis from EEG signals towards IoMT

Mohamadreza Khosravi[1,5,6]✉, Hossein Parsaei[1,2]✉, Khosro Rezaee[3] & Mohammad Sadegh Helfroush[4]

The Internet of Medical Things (IoMT) is poised to play a pivotal role in future medical support systems, enabling pervasive health monitoring in smart cities. Alzheimer's disease (AD) afflicts millions globally, and this paper explores the potential of electroencephalogram (EEG) data in addressing this challenge. We propose the Convolutional Learning Attention-Bidirectional Time-Aware Long-Short-Term Memory (CL-ATBiLSTM) model, a deep learning approach designed to classify different AD phases through EEG data analysis. The model utilizes Discrete Wavelet Transform (DWT) to decompose EEG data into distinct frequency bands, allowing for targeted analysis of AD-related brain activity patterns. Additionally, the data is segmented into smaller windows to handle the dynamic nature of EEG signals, and these segments are transformed into spectrogram images, visually depicting brain activity distribution over time and frequency. The CL-ATBiLSTM model incorporates convolutional layers to capture spatial features, attention mechanisms to emphasize crucial data, and BiLSTM networks to explore temporal relationships within the sequences. To optimize the model's performance, Bayesian optimization is employed to fine-tune the hyperparameters of the ATBiLSTM network, enhancing its ability to generalize and accurately classify AD stages. Incorporating Bayesian learning ensures the most effective model configuration, improving sensitivity and specificity for identifying AD-related patterns. Our model extracts discriminative features from EEG data to differentiate between AD, Mild Cognitive Impairment (MCI), and healthy controls (CO), offering a more comprehensive approach than existing two-class detection algorithms. By including the MCI category, our method facilitates earlier identification and potentially more impactful therapy interventions. Achieving a 96.52% accuracy on Figshare datasets containing AD, MCI, and CO groups, our approach demonstrates strong potential for practical use, accelerating AD identification, enhancing patient care, and contributing to the development of targeted treatments for this debilitating condition.

Keywords Alzheimer's disease, Attention-time-aware, EEG signals, BiLSTM, Convolutional learning, Deep learning, IoMT

Alzheimer's disease (AD) is a progressive brain disorder that worsens memory, thinking, and reasoning skills. It has no cure, but early detection is crucial[1]. Studies show an 8–15% yearly transition rate from mild cognitive impairment (MCI) to AD, suggesting MCI is a stepping stone[2]. While MCI may not significantly impact daily life, early diagnosis is key to slowing down functional decline[3]. A specific type of MCI, non-amnestic MCI (naMCI), doesn't involve memory loss, a common AD symptom. Subjective cognitive decline (SCD) describes self-reported cognitive issues that are difficult to detect objectively[4]. Researchers aim to develop classification models that categorize individuals as healthy, MCI, or AD to understand AD progression and formulate effective treatments[5].

[1]Department of Medical Physics and Engineering, School of Medicine, Shiraz University of Medical Sciences, Shiraz, Iran. [2]Shiraz Neuroscience Research Center, Shiraz University of Medical Sciences, Shiraz, Iran. [3]Department of Biomedical Engineering, Meybod University, Meybod, Iran. [4]Department of Electrical Engineering, Shiraz University of Technology, Shiraz, Iran. [5] IT Services, Lidoma Sanat Mehregan Part Ltd., Shiraz 71581, Fars, Iran. [6] Shandong Provincial University Laboratory for Protected Horticulture (SPUL4PH), Weifang University of Science and Technology, Weifang 262700, China. ✉email: mohammadr.khosravi@iran.ir; dr_mkhosravi@sums.ac.ir; hparsaei@sums.ac.ir; hparsaee@gmail.com

Brain activity can be measured and analyzed using electroencephalogram (EEG) technique. This technique holds promise for detecting AD, a neurological illness, even in its early stages[6]. EEGs can pick up subtle changes in brain function linked to AD before any symptoms arise[7]. These changes often involve an increase in slower brainwaves (delta and theta) and a decrease in faster ones (alpha and beta)[8]. By identifying these pre-symptomatic alterations, EEGs become a valuable tool for diagnosing and studying AD.

Researchers are exploring the use of EEG data to find biomarkers for AD classification and diagnosis[9]. Machine learning (ML) and deep learning (DL) techniques can analyze EEG signals to determine the presence or absence of AD[10]. The initial step involves analyzing EEG data to extract key features using these techniques. Examples of such features include measures like entropy, statistical moments (mean, skewness, variance), and the frequency spectrum of different brainwave bands[11].

There's a growing trend of using machine learning (ML) and deep learning (DL) for EEG data analysis, particularly for diagnosis and classification[12]. These approaches typically feed EEG data (broken down into individual attributes) into the models and carefully optimize them for peak performance. Traditional ML methods have drawbacks like slow processing and lack of interpretability[13–15]. LSTM networks, a powerful type of deep learning model excelling in sequential data processing, have shown great promise in EEG classification[16]. LSTMs, a subset of recurrent neural networks (RNNs), can handle vanishing gradients, a common problem in RNNs[17,18]. LSTMs and RNNs can handle variable-length data, identify temporal patterns, and integrate with more complex networks, making them valuable for classifying important signals like EEG[19,20]. However, RNNs require longer training times on larger datasets, increasing processing costs. Every technique has its pros and cons, and the best choice depends on the specific situation and desired accuracy. Deep learning models come in different forms, including convolutional neural networks (CNNs)[21–24]. CNNs excel at analyzing visual data, processing large amounts of information, and accurately predicting complex patterns. However, they can be computationally expensive to train, their predictions are difficult to interpret, and they rely heavily on labeled data.

To reliably diagnose MCI and AD, automated models need human-like performance. Therefore, it's crucial to identify specific features that significantly improve AD detection systems, which can be achieved through evaluation in clinical settings.

This paper intends to propose a reliable EEG-based early detection technique of AD through the internet of things (IoT) for public healthcare, mostly known as Internet of Medical Things (IoMT). In Fig. 1, a general infrastructure for a pervasive healthcare monitoring system using three commonly used medical biosensors is illustrated. The sensors operate based on photoplethysmogram (PPG), EEG, and electrocardiogram (ECG) signals. This figure effectively describes how an IoMT infrastructure is formed, integrating various components of a cyber-physical system for health monitoring purposes. Edge/fog services are positioned close to the monitored person to ensure highly reliable access and minimal delays.

These servers may be embedded in smartphones and smartwatches. Additionally, cloud servers can be utilized for advanced healthcare services, including AI-based support and input from medical experts. EEG data collected by affordable headbands or caps can provide a highly accessible method for early detection of AD, especially when the IoMT infrastructure incorporates a high-performance algorithm. In IoMT networks, managing signal delay is crucial for timely and accurate monitoring and real-time response. The solution to fix the latency issues includes a combination of optimized data transmission protocols, edge computing, algorithmic efficiency, adaptive sampling rates, and data compression techniques.

In this research, we highlight the critical role of the IoMT as a bridge between medical devices and healthcare information systems and communication networks[25,26]. IoMT enables continuous and remote health monitoring, which is particularly vital for managing chronic conditions such as AD. By leveraging advanced wearable devices and sensors, IoMT collects real-time biological data like EEG signals, which are essential for early diagnosis and

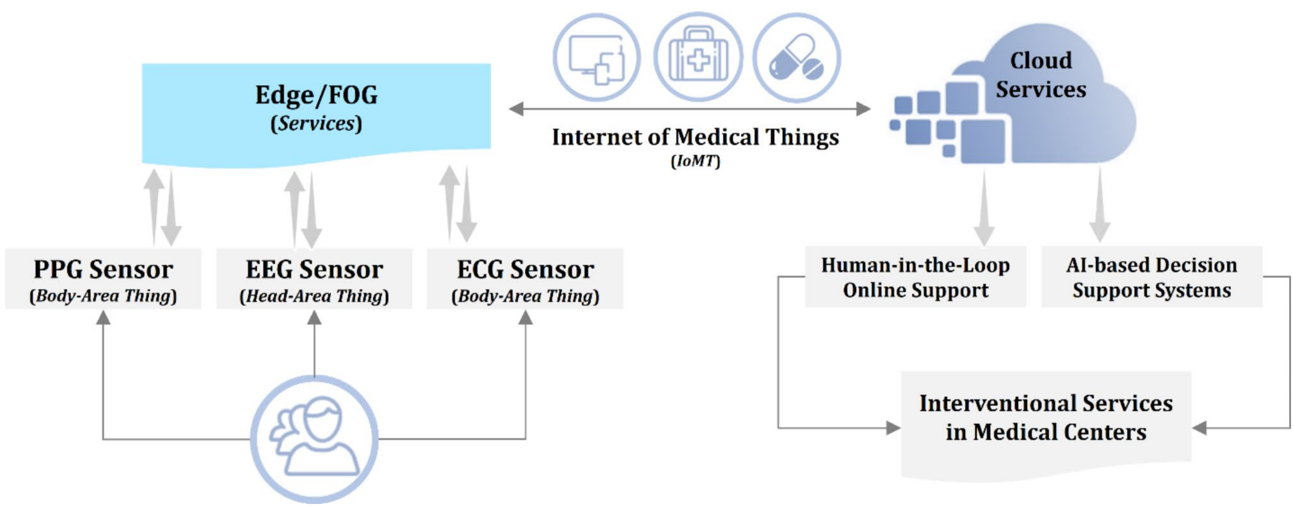

Fig. 1. The reputed infrastructure of an IoMT system for human well-being through personal biomedical devices.

effective disease management. Our study utilizes the capabilities of IoMT to process and transmit EEG data, facilitating the identification of brain patterns associated with Alzheimer's. This approach not only improves diagnostic accuracy and efficiency but also streamlines the diagnostic process by reducing dependence on subjective evaluations and lengthy tests. Thus, IoMT is a key component in the development of future medical support systems and the realization of smart healthcare cities, ultimately enhancing patients' quality of life and providing more optimized healthcare services.

This study utilizes the integration of ML and the IoMT with the CL-ATBiLSTM model for early detection and continuous monitoring of AD through EEG signals. The model analyzes temporal and spatial patterns in EEG data, enabling precise differentiation between various stages of Alzheimer's, Mild Cognitive Impairment (MCI), and healthy controls. IoMT facilitates real-time data collection via wearable devices, reducing reliance on subjective assessments and enhancing the diagnostic process in both clinical and remote settings. As a result, the accuracy and efficiency of AD and patient care are significantly improved.

Existing research explores various techniques for analyzing EEG data to diagnose AD and identify characteristic patterns[12,27–29]. These studies aim to provide reliable tools for clinicians, but a gap exists in utilizing multiple EEG frequency bands for improved diagnosis[21,30]. Our contribution addresses this by proposing a novel deep learning model for AD classification using EEG data, combining CNNs with attention mechanisms and bidirectional long-short-term memory (BiLSTM) networks[31,32]. While our findings are promising, generalizability to other datasets remains a challenge, necessitating future work on improving model robustness and architecture. This research highlights the potential of EEG data for AD diagnosis and the need for further exploration and improved models. The importance of early diagnosis in AD and MCI is well-established[32]. Our study suggests a deep learning strategy for AD and MCI classification using EEG data, leveraging pre-trained CNNs and an evolutionary optimization algorithm to enhance learning and classification accuracy. Utilizing multiple pre-trained CNNs through ensemble techniques can further improve diagnostic accuracy. The presented approach demonstrates the effectiveness of hybrid learning in enhancing machine learning models for AD diagnosis. By using deep learning to extract valuable features from EEG signals, this approach aims to achieve higher overall classification accuracy and address ongoing challenges in AD diagnosis and prognosis. The research contributions can be summarized as follows:

- The research contributes by proposing an innovative approach to AD detection through the integration of Internet of Medical Things (IoMT) technologies with advanced deep learning algorithms. By leveraging IoMT for continuous health monitoring and EEG data collection, the study addresses the critical need for early AD diagnosis, potentially revolutionizing current diagnostic practices and enhancing patient care in smart healthcare cities.
- Furthermore, the development and implementation of the CL-ATBiLSTM model represent a significant contribution to the field of EEG-based AD classification. This deep learning architecture, incorporating convolutional layers, attention mechanisms, and BiLSTM networks, offers a comprehensive solution for differentiating between AD, Mild Cognitive Impairment (MCI), and healthy controls (CO). The model's high accuracy in testing, as evidenced by the evaluation using Figshare datasets, underscores its potential for practical application and its role in expediting AD identification and facilitating targeted treatment interventions.
- Additionally, the study contributes to the broader understanding of IoMT's role in shaping the future of medical support systems and smart healthcare cities. By emphasizing the importance of continuous and remote health monitoring facilitated by IoMT technologies, particularly in managing chronic conditions like AD, the research highlights the transformative impact of integrating wearable devices and sensors with advanced algorithms for EEG analysis. This interdisciplinary approach not only enhances diagnostic accuracy but also streamlines the diagnostic process, reducing reliance on subjective evaluations and lengthy tests, ultimately improving patients' quality of life and optimizing healthcare services.

Additionally, the article is structured as follows: In "Related work" sect., we discuss relevant research already conducted in this field. In "Proposed method" Sect., we delve into the proposed approach, explaining the learning and implementation processes in detail. This section also outlines the intended use of this approach as a therapeutic system. "Experimental results" Sect. presents the results of the classification process along with a detailed analysis. Finally, "Conclusion" Sect. focuses solely on presenting the findings of the research.

Related work

AD research utilizes deep learning models such as autoencoders, RNNs, and CNNs to classify EEG data[26,32,33]. These models significantly enhance our understanding of the underlying biological mechanisms involved in disease progression and potential therapeutic targets. They have shown promising results in EEG data categorization and AD biomarker discovery.

Perez-Valero et al.[1] developed an automated diagnostic method for AD using a CNN and a connection matrix, achieving a classification accuracy of 62% for original EEG signals and successfully categorizing 75% of the initial AD data. Alessandrini et al.[18] employed robust principal component analysis (mPCA) and rPCA approaches for data preprocessing. Principal component analysis was used to extract features, while RNNs were used for classification, resulting in a maximum accuracy of 94.6%. Imani et al.[28] introduced a bidirectional LSTM network to analyze time series data, along with a CNN to examine the relationships of EEG data from various brain areas. The researchers integrated temporal features into their model using a fully connected neural network. They improved the accuracy of diagnosis by incorporating autoencoder networks and selecting channels based on entropy measurement.

Fouad and Labib[34] evaluated the effectiveness of their technique by analyzing data from two separate datasets. The researchers utilized wavelet techniques to extract statistical data from the original EEG data, allowing for the identification of specific classifiers that could differentiate between signs of AD and MCI. The system's robustness was demonstrated by achieving accuracy rates of 94.52% and 96.55% using the Figshare dataset with support vector machine and Naive Bayes classifiers, respectively. Additionally, the ResNet-50 architecture was employed to improve the system's practicality, leading to a test dataset accuracy of 97.8261%.

Recently, Kumar Ravikanti and Saravanan[35] have highlighted a growing interest in using transformers in areas such as biological signal processing. Their work describes the transformer architecture as featuring multiple feed-forward layers in the encoder component and self-attention in the decoder component. They also emphasize that the robustness of these layers is enhanced by standardizing them and removing residual connections between them, which improves the overall performance of the model.

Xie et al.[36] examined raw signals using a transformer approach and achieved accuracies of 74.4% for three-class, 64.2% for four-class, and 83.31% for two-class problems. This highlights the importance of investigating their ability to identify EEG abnormalities associated with neurodegenerative conditions such as AD and other types of dementia. Deep learning techniques are renowned for their ability to classify EEG data and are employed in diverse architectures for detection applications[19,27,37–40]. This includes various CNNs such as stacked autoencoders and AlexNet. A significant advantage of deep learning is its capacity to independently extract features from different input formats, such as images and two-dimensional matrices. However, the accurate fine-tuning of model hyperparameters remains a common challenge in deep learning.

Araújo et al.[41] effectively retrieved and categorized features by combining wavelet packets with traditional machine learning techniques, resulting in a maximum accuracy rate of 84.2%. The application of deep learning approaches, particularly 2D CNNs, is prevalent in the categorization of AD. Nour et al.[42] proposed a deep ensemble learning approach that integrated five distinct 2D-CNN models for the purpose of internal classification. Thus, for the first time, a sophisticated ensemble learning model utilizing EEG data demonstrated exceptional accuracy in classifying AD. After undergoing five rounds of cross-fold training, the deep ensemble learning architecture achieved an average accuracy of 97.9% in AD categorization.

Gap analysis

In the critical review of literature on the use of EEG for diagnosing AD and MCI, several key gaps have been identified that, if addressed, could lead to the development of more effective and applicable methods. One of the most significant gaps is the limited diversity in the quality and size of datasets used. Many studies, such as those by Valero et al.[1] and Alessandrini et al.[18], employed small or restricted datasets, which reduces the generalizability of their findings. This highlights the need for further research utilizing larger and more diverse datasets to enhance the reliability and external validity of the models.

Another significant gap is the lack of comprehensive comparative analyses among many of the studies. For instance, studies like those by Miltiadous et al.[32], Kumar Ravikanti[35], and Xie et al.[37] employed transformer-based models but did not conduct direct comparisons with more traditional methods such as CNNs. This absence of standardized and thorough comparative evaluations makes it difficult to assess the relative performance and strengths of each approach. More robust comparative studies are essential to fully understand the benefits and limitations of the various methods.

There is also a need for improved feature extraction techniques. Current methods, such as simple CNNs or wavelet-based approaches, often fail to achieve optimal accuracy. For example, the study by Imani et al.[28] indicates that existing models require further innovation in feature extraction to capture more complex patterns in EEG data. This could be achieved through the development of hybrid models or advanced pre-processing techniques that can better exploit the information contained in the data.

Finally, the complexity and high computational requirements of advanced models, particularly transformer-based architectures, represent another gap. Although these models offer high accuracy, their practicality in clinical settings is limited, necessitating the development of more lightweight and efficient models that maintain accuracy while being feasible for real-world use. Additionally, many papers do not fully report the advantages and disadvantages of their methods, making it challenging to evaluate their applicability across different environments. For example, Fouad and Labib[34] reported high accuracy using wavelet techniques and ResNet-50 but did not thoroughly discuss potential drawbacks such as computational costs or sensitivity to data variations. These gaps indicate that, despite the significant advances, further research is needed to develop more practical and precise methods.

Observations

In reviewing the literature on EEG-based detection of AD and MCI, recent advancements in deep learning extend beyond traditional CNN and LSTM models. Newer architectures, such as attention-based and transformer models, have demonstrated superior performance in capturing complex temporal patterns in EEG signals, enhancing diagnostic accuracy. These models leverage the spatial feature extraction capabilities of CNNs along with attention mechanisms that focus on the most relevant signal components. Additionally, hybrid approaches that integrate convolutional layers, bidirectional LSTMs, and attention layers are effective in capturing both spatial and temporal features of EEG data more precisely.

However, these advancements come with challenges, including significant computational demands and reliance on limited datasets, which can affect the generalizability of findings. There is also a trend towards combining deep learning with traditional feature extraction methods to address the limitations of single-model approaches. Incorporating these sophisticated models within the IoMT framework can improve the robustness and clinical applicability of diagnostic tools, making use of cutting-edge methods while also addressing the high computational needs and other limitations associated with real-time clinical applications.

Early classification of AD is vital for effective treatment, yet current diagnostic tools often lack the precision, interpretability, and comprehensiveness needed for clinical application. While literature reviews highlight various feature extraction and classification methods, there remains a need for universally applicable approaches that can accurately identify key characteristics across all stages of AD. Recent advancements in ML, such as deep ensemble learning and hybrid models, offer promising solutions by automating repetitive tasks and enhancing the accuracy of diagnoses. However, challenges persist in balancing accuracy with interpretability, ensuring that healthcare professionals can confidently use these tools and explain results to patients. This underscores the ongoing need for innovative diagnostic tools that are both user-friendly and precise, ultimately aiming to revolutionize the clinical approach to AD diagnosis.

Proposed method

This methodology proposes an algorithm to diagnose AD by classifying patients into three categories: AD, Mild Cognitive Impairment (MCI), and healthy. Our proposed framework for diagnosing AD and MCI consists of several steps, as illustrated in Fig. 2.

The algorithm works in several steps:

a) Signal Decomposition: First, brainwave signals (EEG) are broken down using Discrete Wavelet Transform. This creates different frequency ranges of the signal for further analysis.
b) Windowing: Since brainwave signals can change over time, the signal is divided into smaller segments to improve analysis and increase the amount of data available through a rectangular frame.
c) Feature Extraction from Segments: Each segment is transformed into a visual representation called a spectrogram using Short-Time Fourier Transform. This spectrogram helps identify key characteristics of the brainwave activity.
d) Deep learning for feature analysis: Next, a deep learning technique called CNN analyzes the spectrogram images based on a pre-existing architecture called ResNet101. This step helps determine how well each extracted feature predicts the patient's condition.
e) Optimized classification: Finally, a more sophisticated classification method called Attention-Bidirectional Time-Aware Long-Short-Term Memory (ATBiLSTM) is used. This method builds upon a standard BiLSTM architecture by incorporating an attention mechanism and a Bayesian algorithm. ATBiLSTM helps to classify each signal into one of the three categories (AD, MCI, or healthy) with improved accuracy.
f) The IoMT platform enhances this algorithm by providing real-time EEG data from wearable devices, ensuring current and comprehensive analysis. IoMT enables continuous monitoring and immediate processing, leading to more accurate and timely classifications of AD, MCI, and healthy controls. Its robust data transmission supports large-scale datasets, refining the deep learning and classification algorithms. IoMT is essential for an efficient early AD system, improving patient outcomes and enabling personalized healthcare.

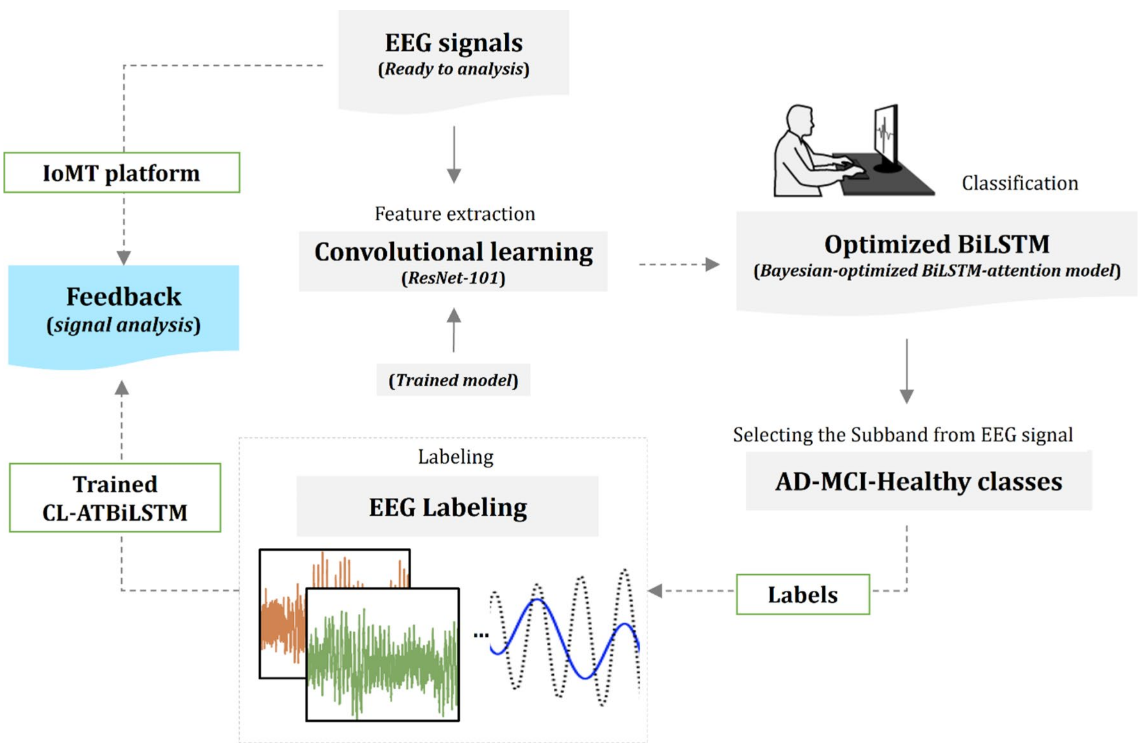

Fig. 2. Our proposed framework for diagnosing AD and MCI, which includes both training and testing phases.

Preprocessing phase

The preprocessing phase is essential for the accurate analysis of EEG signals in diagnosing AD and MCI. This process begins with Discrete Wavelet Transform (DWT) to decompose EEG signals into frequency bands—such as Delta (1–4 Hz), Alpha (8–12 Hz), and Gamma (30–100 Hz)—each associated with specific cognitive functions and AD-related markers.

Subsequently, windowing segments the signals into overlapping windows, enhancing signal quality and preserving temporal dynamics. The window sizes and overlap rates, typically between 20% and 50%, are optimized to balance redundancy with continuity, facilitating robust data augmentation.

Finally, spectrograms are generated using Short-Time Fourier Transform (STFT), providing a detailed visual of frequency and temporal features. This approach helps identify abnormalities in specific frequency bands, offering valuable diagnostic insights into AD. These preprocessing steps collectively enable a detailed and effective analysis of EEG signals, critical for advancing AD diagnosis.

Decomposition

We aim to clarify our approach for analyzing EEG signals and diagnosing AD. We utilize DWT to segment EEG signals into different frequency ranges, each providing insights into AD diagnosis. Specifically, Delta (1–4 Hz) levels may indicate brain cell function or sleep issues, while Alpha (8–12 Hz) changes are linked to memory and attention problems. Beta (12–30 Hz) activity is associated with focus and alertness, Theta (4–8 Hz) increases suggest memory issues, and Gamma (30–100 Hz) abnormalities indicate higher cognitive function problems. DWT enables detailed analysis by examining signals at varying levels of detail in both time and frequency domains, offering a more efficient signal representation. We employ specific wavelets like Daubechies and techniques such as symmetric extension for accurate signal analysis. Normalization ensures consistent signal strength across the analysis. Our approach provides a comprehensive and efficient method for analyzing brain activity, crucial for diagnosing and understanding AD.

Windowing phase

In this approach to analyzing EEG signals, it was recognized that due to their dynamic and non-stationary nature, segmentation is essential. A windowing technique was implemented that divides the EEG signals into equal-sized segments, with segment lengths carefully selected to capture brain activity while preserving crucial information. By introducing overlapping segments, continuity between windows was ensured, minimizing data loss and enhancing the overall analysis.

The signals were segmented into smaller windows, each providing a focused view over time, thereby improving the accuracy and reliability of the analysis. Overlapping windows played a key role, allowing consecutive windows to partially overlap, effectively capturing temporal dynamics and transitional features that could be missed with non-overlapping windows. The analysis involved varying overlap percentages from 20 to 50%, balancing redundancy and temporal information based on specific application needs. Lower overlaps (20–30%) helped reduce redundancy, while higher overlaps (30–50%) provided a more continuous signal view.

Furthermore, the windowing and overlapping processes were leveraged as a form of data augmentation, which significantly increased the variability of input data—beneficial for training deep learning models. This approach improved the robustness and generalization of the models during training. The window size, typically between 20 and 100 ms, was tuned to match the frequency characteristics of the EEG patterns being analyzed. Precise adjustments of window size and overlap percentage were critical to capturing relevant features without introducing unnecessary complexity, ensuring the effectiveness of the analysis.

Spectrogram image

We employed STFT to choose an appropriate window with a suitable length and type and generating spectrogram images from EEG recordings. The window length should be sufficient to capture the temporal and frequency details of the signal. Additionally, you need to determine the amount of overlap between windows. This overlap is typically between 50% and 75% to ensure no important information is lost. For each windowed segment of the signal, apply the window function. Suppose $x[n]$ is the original signal and $w[n]$ is the window function with length . The windowed signal is defined as:

$$x_w[n] = x[n] \cdot w[n] \tag{1}$$

Moreover, for each windowed segment, compute the Fourier Transform. This transform is defined as:

$$X_w[k] = \sum_{n=0}^{N-1} x_\omega[n].\, e^{-j2\pi kn/N} \tag{2}$$

where, k is the number of frequency bins and N is the window length. We repeat this process for all overlapping segments of the signal to cover the entire signal. Each time, the window is shifted by the overlap amount, and these steps are repeated. Next, we arrange the results of the Fourier Transform for each windowed segment into a matrix. Each row of the matrix demonstrates a specific frequency, and each column shows a specific time. To display the spectrogram, usually, the power or magnitude of the Fourier Transform is used. The power is computed as:

$$P[k, m] = |X_\omega[k, m]|^2 \tag{3}$$

where, $p[k,m]$ is the power at frequency and time . Finally, we display the power matrix as an image. The horizontal axis describes time, the vertical axis shows frequency, and the colors or intensities represent the power values. Abnormalities in the spectrogram features, such as changes in frequency band power or spectral coherence, can provide diagnostic insights into AD. For example, alterations in alpha or theta band power have been linked to cognitive decline in AD patients.

Class imbalance handling

To address the class imbalance in the EEG dataset, a multifaceted approach was employed[43,44]. Data augmentation techniques, such as synthetic data generation and oversampling, were utilized to balance the dataset, ensuring sufficient representation of each class during training. Additionally, windowing and overlapping techniques were applied during preprocessing to increase the sample size for minority classes, enhance data diversity, and improve feature extraction. To further mitigate class imbalance, weighted loss functions were implemented, assigning higher importance to minority classes (MCI and AD) to guide the model's focus on these underrepresented groups. A stratified k-fold cross-validation strategy was also used to maintain balanced class distributions within each fold, ensuring consistent and robust evaluation. Comprehensive performance metrics, including precision, recall, and F1-score for each class, along with overall accuracy, were reported to thoroughly assess the model's effectiveness, particularly in accurately classifying minority classes. These combined strategies ensured balanced training and enhanced the model's reliability in diagnosing AD and MCI.

Feature extraction and classification

To diagnose AD stages, our method utilizes two powerful models: deep learning for feature extraction and ATBiLSTM for classification. This combination enhances the accuracy of both steps. First, deep learning, specifically a deep convolutional neural network called ResNet101, automatically identifies important features from spectrogram images derived from EEG signals. Given the complexity of brain activity patterns, this deep learning model can detect subtle differences that simpler methods might miss. It's akin to using a highly powerful tool to examine brain signals in detail. Second, the ATBiLSTM processes these extracted features to classify different stages of AD. By integrating these two techniques, this approach aims to achieve superior results in detecting and classifying AD.

ResNet101

ResNet101 utilizes residual learning blocks to overcome the vanishing gradient problem, enabling effective training of very deep neural networks:

$$\mathbf{y} = \mathrm{F}(\mathbf{x}, \{Wi\}) + \mathbf{x} \tag{4}$$

where, x, y, F, and $\{W_i\}$denote the input to the block, the output of the block, the residual mapping to be learned by the layers, and the weights of the layers within the block, respectively. This equation captures the core concept of ResNet's architecture, where the output is calculated as the sum of the input x and the learned residual mapping $F(x\{W_i\})$. This design allows the network to learn residual functions, facilitating the optimization of deeper networks and addressing the issue of vanishing gradients.

Bidirectional time-aware long-short-term memory

In a conventional LSTM cell unit, the cell state c^t functions as an internal memory mechanism and governs the information propagation. Based on the information provided in (5), it can be observed that the state c^{t-1} is obtained by applying the forget gate $\boldsymbol{f}^t$ to the present state of the cell. In addition, the input gate exerts an influence on a prospective cell state.

$$\begin{aligned} i^t &= sigmoid(W_x^i x^t + W_h^i h^{t-1}) \\ \tilde{c}^t &= \tanh(W_x^c x^t + W_h^c h^{t-1}) \\ f^t &= sigmoid(W_x^f x^t + W_h^f h^{t-1}) \end{aligned} \tag{5}$$

The following are the probable factors: For instance, let H denote the overall quantity of hidden nodes, h^{t-1} represent the first hidden state produced by c^{t-1}, $\{\mathrm{W_x} \in \mathrm{R}^{H\times D}, \mathrm{W_h} \in \mathrm{R}^{H\times H}\}$ signify the initial configuration of the network, and so on. The variable "D" represents the number of dimensions or measurements recorded for each occurrence, whereas the variable "H" denotes the number of unseen nodes. Various methodologies can be employed to attain an altered cellular state.

$$c^t = i^t \otimes \tilde{c}^t + f^t \otimes c^{t-1} \tag{6}$$

The sum of all components, represented by the symbol $\otimes$, denotes element-wise multiplication. Finally, the secret state is formed by routing the modified cell state through the $\boldsymbol{o}^t$ or output gate layer. In addition, the chance of septic shock occurring at a specific time t is calculated using a sigmoid function that takes into account the parameter $\boldsymbol{U}$.

$$\begin{aligned} o^t &= sigmoid(W_x^o x^t + W_h^o h^{t-1}) \\ h^t &= o^t \otimes \tanh(c^t) \\ p^t &= sigmoid(U h^t) \end{aligned} \tag{7}$$

The suggested model, shown in Fig. 3, is an Attention-based Time-Aware LSTM network. The initial LSTM model, as noted in (2), utilizes the c^t cell state to denote the present moment. This information is derived from the c^{t-1} state, which signifies the state before to the event. The significance of recent developments in EEG signals is in their relevance to the process of cell state learning, primarily because of the widely recognized issue of vanishing gradient (VoG). Therefore, it reduces the temporal constraints on patient contacts and output analysis within the clinical setting. One disadvantage of single memory reads is the failure to consider previous occurrences.

Attention mechanism

The value of the parameter *m*, which typically refers to the number of hidden units or neurons in the LSTM layer(s), is determined in three different ways. The occurrence count, represented by *m*, from the past will be considered. Local Attention focuses only on the *m* events that come before it during event processing. One can either conduct a grid search or consult an expert for advice on the optimal value of hyperparameter *m*.

$$e_{ti} = (x^i)^T W_\alpha x^t \quad i = (t-m), \cdots, (t-1) \tag{8}$$

where,

$$\alpha_t = softmax(e_{t(t-m)}, e_{t(t-m+1)}, \cdots, e_{t(t-1)}) \tag{9}$$

If we set $m=t-1$ in Global Attention, it means that the attention mechanism considers all past events up to the current time step *t* when determining the attention weights for the current time step. In other words, the model takes into account the entire history of events leading up to the present moment when assigning importance to each past event in relation to the current context.

$$e_{ti} = (x^i)^T W_\alpha x^t \quad i = 1,\ 2, \cdots, (t-1) \tag{10}$$

where,

$$\alpha_t = softmax(e_{t1}, e_{t2}, \cdots, e_{t(t-1)}) \tag{11}$$

Global attention considers all prior occurrences, but it is computationally demanding and yields inconsistent attention weights. In addition, we explore two different approaches to address these concerns. Furthermore, a flexible attentional method is devised to acquire the optimal quantity of previous experiences, denoted as m, to retain for a certain event *t*. The underlying assumption is that the quantity of previous occurrences to be analyzed is contingent upon the current happening.

$$m(t) = \left((t-1)\,.\,Sigmoid(v_m^T \tanh(W_m x^t))\right) \tag{12}$$

where,

$$\alpha_t = softmax(e_{t(t-m(t))}, e_{t(t-m(t)+1)}, \cdots, e_{t(t-1)}) \tag{13}$$

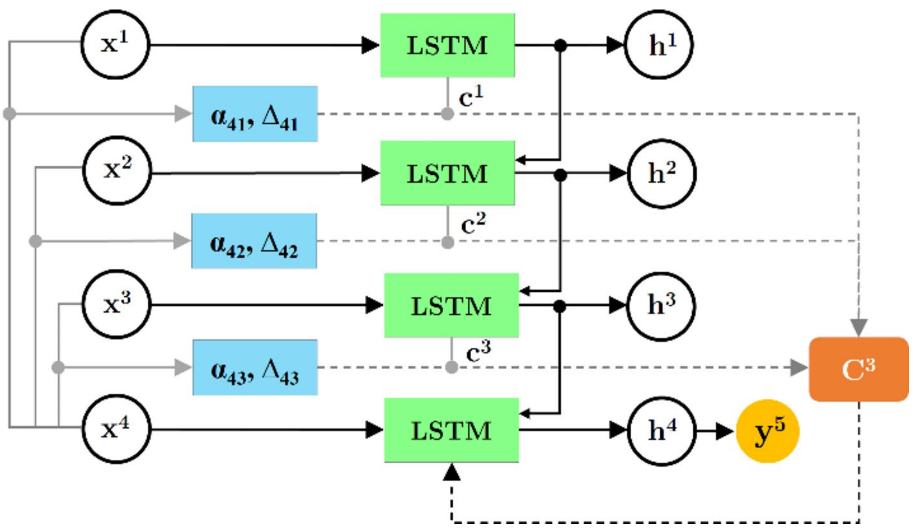

Fig. 3. Schematic of the Bidirectional Time-Aware LSTM model utilized in this study. The model integrates cell states from previous time steps and improves the stability of older memories by incorporating time interval weights (Δ) and an attention mechanism (α), enhancing the overall performance and accuracy of the network.

where, W_m and v_m are adjustable parameters, and $m(t)$ is a function that takes values within the interval [1, t-1].

Bayesian algorithm
In hyperparameter tuning for an ATBiLSTM network, our model assumes a normal distribution for the function relating the hyperparameters to the performance (f(x)). Thus, we leverage Gaussian process regression to analyze past results (H) and estimate the performance for new hyperparameter settings (p(f(x)|x, H))[45,46]. This process iterates until the optimal hyperparameters are found (Algorithm 1). Notably, the initial learning rate and the number of cells in the LSTM layer significantly affect the network's accuracy.

Through Bayesian optimization, the ATBiLSTM network can automatically identify the best values for these hyperparameters. This optimization technique often relies on a specific function (Expected Improvement function) to guide the search process. Fine-tuning these hyperparameters can significantly improve the overall performance of the ATBiLSTM network:

$$EI_f(x,H) = \int_{-\infty}^{+\infty} max(y^* - y, 0\,)p(\,y|\,x,H)dy \tag{14}$$

Unlike fixed approaches, Bayesian optimization[47,48] uses a flexible, model-driven approach to efficiently tune the design of complex machine learning models, including ATBiLSTMs.

Bayesian learning-ATBiLSTM procedure

Initialize: *T* (No. iteration)

Input: *x* (input), *f(x)* (optimized object function)

1 H → null matrix

2 Start by setting up a random model (Gaussian process) and then use it to calculate how likely it is that the function f(x) will have a specific value (considering both the new input x and past data H).

3 **for** repeated counter *T* **do**

4   argmax$_x$ α (x,H) → *x*

6   assess y' =f(x')

7   H→ H ∪ (x', y')

8   Compute p(f(x)|x,H) using a remodeled Gaussian process

9 **end** for

10 Achieve the output ResNetCNN-ATBiLSTM

Output: H and α (x,H)

Algorithm 1. Step-by-step process of Bayesian optimization for fine-tuning the design of the ATBiLSTM network.

This is a loop-based approach where we constantly try out different settings (hyperparameters) for the model, see how well it performs, update our understanding of the best outcome (objective function), and then choose the next set of settings to test.

Experimental results

In addition to outlining the model's implementation, this section elaborates on the experimental findings concerning EEG signal processing, assessments, and subsequent discussions.

Dataset

The EEG signals utilized in the study conducted by Cejnek et al.[49] can be accessed through the Figshare database, which houses the data related to that specific investigation. The proposed methodology was assessed using a total of 303 signals, consisting of EEG recordings collected from 59 individuals diagnosed with AD. Additionally, the study examined 56 signals obtained from individuals diagnosed with MCI, as well as 110 signals obtained from healthy controls. Electrodes were positioned according to the conventional 10–20 system to ensure uniform distribution and proportionality to the skull's dimensions and configuration. The electrodes are labeled based on their corresponding brain lobes, such as occipital (O), parietal (P), central (C), temporal (T), frontal (F), and frontopolar (Fp). This describes the setup used to record brain activity. A special cap with electrodes was placed on the scalp, with a midline marker (electrode "z"). Even-numbered electrodes picked up signals from the right side of the brain, while odd-numbered ones captured activity from the left. Two different systems were used to record the electrical signals: a 21-channel system at 256 samples per second and a 32-channel system at 128 samples per second. To help analyze the data, Fig. 4 shows spectrograms, which are visual representations of brainwave patterns for different categories.

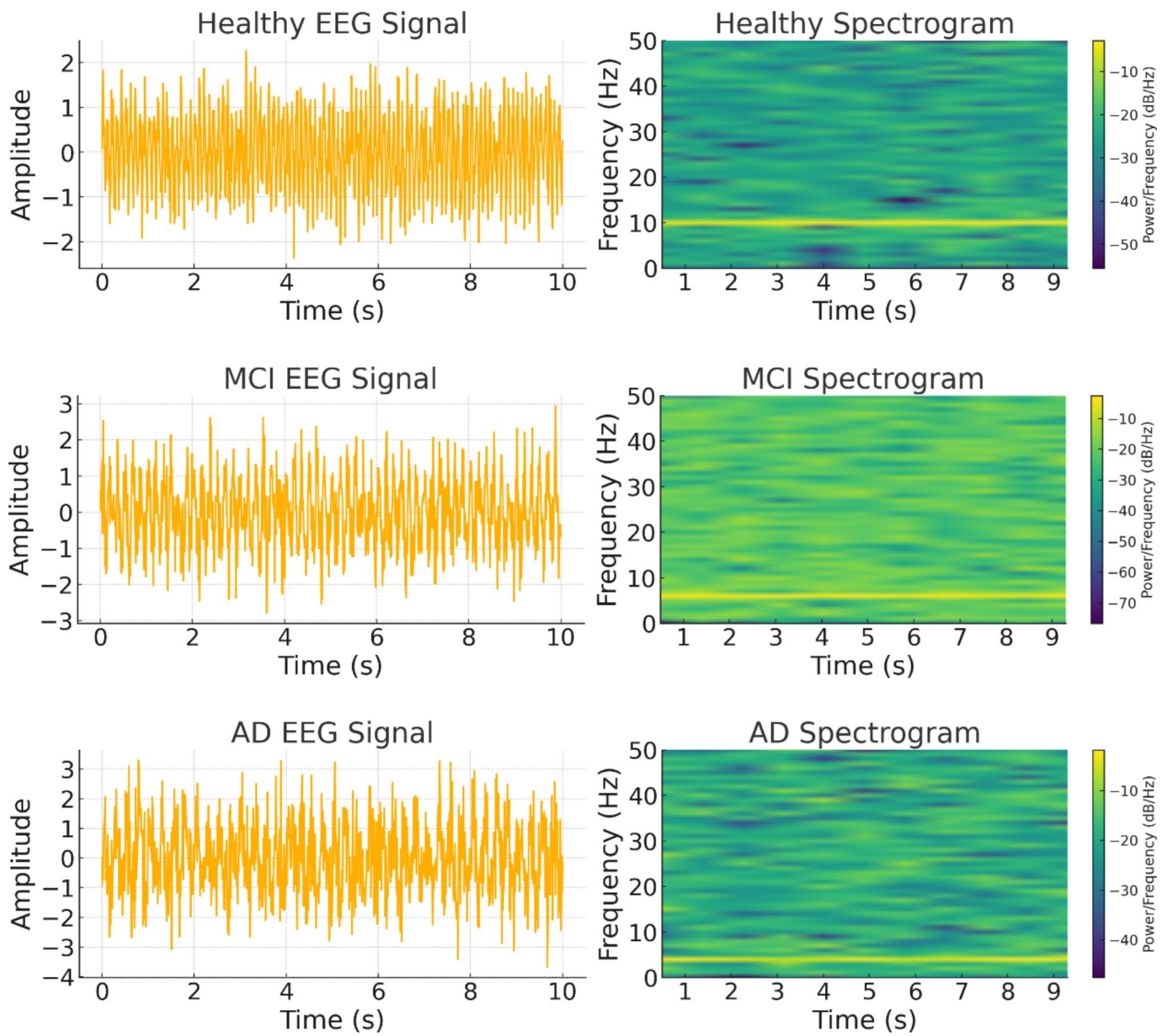

Fig. 4. Three sample EEG signals and their corresponding spectrograms are displayed in the figure.

Experimental setup

Our study explores the potential of using brainwave patterns (EEG signals) to differentiate between healthy individuals, those with MCI, and those with AD. We compared various approaches by dividing the data into training (70%) and testing (30%) sets to assess the method's effectiveness. To ensure robust results, we employed a specific validation technique (5-fold cross-validation with subject segmentation). The model learned from the training data, adjusting its internal settings to improve performance, enabling it to analyze new, unseen EEG data effectively. We segmented the EEG recordings into smaller windows and assigned labels to each segment based on the predicted category (healthy, MCI, or AD). The analysis was conducted on 64-bit computer systems. The training process typically requires between 500 and 2000 epochs to achieve optimal performance. We combined two neural network architectures, Attention and LSTM, to create our model. Our goal was to improve the model's convergence by minimizing errors during training and validation. Training would stop if the model's accuracy did not improve sufficiently. Finally, the accuracy of classifying individuals was influenced by the size and overlap of the EEG signal segments used for analysis. We experimented with window lengths ranging from 800 milliseconds to 3 s. The similarity between brainwave patterns was categorized as moderate (30–50% similar) or low (20–30% similar).

To validate the practical applicability of the proposed algorithm beyond controlled experimental settings, it was tested on an IoMT system specifically developed for this study. This system was designed to emulate real-world healthcare environments by integrating various medical devices and sensors capable of real-time data collection, processing, and analysis. The validation process involved extensive testing under diverse conditions, including fluctuations in data quality, variations in sensor performance, and typical network conditions encountered in IoMT applications. These scenarios were chosen to assess the robustness and adaptability of the algorithm in practical settings, ensuring it performs reliably under the complexities of real-world deployment.

Evaluations

We highlight the effectiveness of our proposed ATBiLSTM model in classifying brainwave patterns (EEG signals) for the detection of AD and MCI. We achieved a remarkable 100% accuracy rate in two-class classifications (e.g., healthy vs. AD) across all three scenarios tested (AD vs. healthy, MCI vs. healthy, and MCI vs. AD). This demonstrates the model's exceptional ability to distinguish between these categories in our dataset. When extending the classification to three categories (healthy, MCI, and AD), ATBiLSTM achieved a very high overall accuracy of 96.52% (see Fig. 5). Notably, the results exhibited minimal variations across multiple tests, signifying the model's consistency and reliability. We also acknowledge the potential for variations in real-world EEG data due to inherent uncertainties. We employed variance estimation techniques (not explicitly mentioned but implied) to comprehensively evaluate the model's performance under various AD detection conditions. This approach ensures our findings consider potential variations in real-world applications.

Visual inspection of the confusion matrices (or the provided accuracy values) suggests that the proposed algorithm achieves satisfactory accuracy in the three-class classification mode. Additionally, the dispersion, which refers to the variability, among the accuracy values obtained in different folds of the cross-validation process appears to be low. This indicates good consistency and generalizability of the model's performance across the dataset.

Comparative design and ablation study

To understand how well each classification approach performed, the researchers compared the results for different scenarios: healthy (or CO) vs. Alzheimer's disease (AD), healthy vs. Mild Cognitive Impairment (MCI), and MCI vs. AD. Table 1 summarizes these findings and compares the proposed method to other common techniques like LSTM, BiLSTM, Attention-LSTM, and ATBiLSTM. To understand the contribution of individual components, we assess the impact of adding different elements, such as attention mechanisms and Bi-LSTM blocks, to our base model for each class comparison type. The base model achieves a certain level of performance in terms of accuracy, specificity, sensitivity, precision, and F-measure. We then systematically remove individual components from the base model and observe how each removal affects the model's performance.

For instance, removing the Attention Mechanism from the base model for the MCI vs. CO class comparison type results in a slight decrease in accuracy from 96.52 to 95.00%, indicating that the attention mechanism contributes positively to the model's performance in capturing relevant features for classification. Similarly, removing the Bi-LSTM Block or Convolutional Layers also leads to slight decreases in accuracy, highlighting the importance of these components in capturing temporal and spatial dependencies in the data.

Ablation study, a cornerstone in machine learning research, serves as a pivotal methodology for dissecting the efficacy and contributions of various components within a classification model.

Within the domain of AD detection and classification, alongside MCI and healthy individuals (CO), this method furnishes indispensable insights into the nuanced impact of each aspect of the model on its overarching performance. By methodically deconstructing the base model, fortified with elements such as attention mechanisms and Bi-LSTM blocks, researchers discern the discrete influences of individual components on quintessential metrics like accuracy, specificity, sensitivity, precision, and F-measure. This systematic inquiry elucidates critical components that markedly enhance the model's proficiency in accurately stratifying individuals into their respective categories. Throughout the ablation study, distinct facets of the model emerge as indispensable contributors to its efficacy. Notably, the removal of the attention mechanism yields a tangible reduction in accuracy, underscoring its pivotal role in capturing salient features crucial for classification. Likewise,

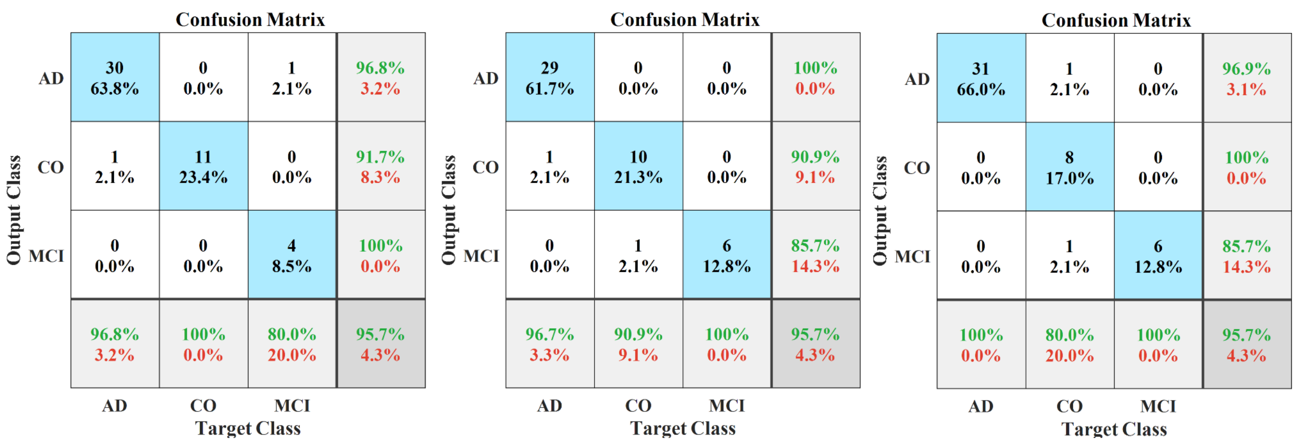

Fig. 5. Confusion matrices illustrating the results of the three-class classification task (healthy, MCI, and AD). These matrices provide a detailed evaluation of model performance, displaying the true positives, false positives, true negatives, and false negatives for each class, thereby offering insights into the classification accuracy and potential areas for improvement.

Class comparison type	Other models		Metrics: Accuracy	Specificity	Sensitivity	Precision	F- measure
MCI vs. CO	LSTM	CO	88.23%	83.33%	90.90%	90.90%	90.90%
		MCI	88.23%	90.90%	83.33%	83.33%	83.33%
	BILSTM	CO	94.11%	85.17%	100%	90.90%	95.23%
		MCI	94.11%	100%	85.71%	100%	92.30%
	Attention + LSTM	CO	94.18%	100%	92.30%	100%	96.00%
		MCI	94.11%	92.30%	100%	80.00%	88.89%
	ATBiLSTM	CO	100%	100%	100%	100%	100%
		MCI	100%	100%	100%	100%	100%
AD vs. CO	LSTM	CO	95.23%	100%	84.61%	100%	91.67%
		AD	95.23%	84.61%	100%	93.54%	91.66%
	BILSTM	CO	95.23%	96.77%	90.90%	90.90%	90.90%
		AD	95.23%	90.90%	9677%	96.77%	96.77%
	Attention + LSTM	CO	95.23%	93.93%	100%	81.81%	90.00%
		AD	95.23%	100%	93.93%	100%	96.87%
	ATBiLSTM	CO	100%	100%	100%	100%	100%
		AD	100%	100%	100%	100%	100%
AD vs. MCI	LSTM	MCI	94.44%	75.00%	100%	93.33%	96.55%
		AD	94.44%	100%	75.00%	100%	85.71%
	BILSTM	MCI	97.22%	85.71%	100%	96.66%	98.30
		AD	97.22%	100%	85.71%	100%	92.30%
	Attention + LSTM	MCI	97.22%	100%	96.87%	100%	98.41%
		AD	97.22%	96.87%	100%	80.00%	88.88%
	ATBiLSTM	MCI	100%	100%	100%	100%	100%
		AD	100%	100%	100%	100%	100%

Table 1. Performance comparison of the proposed ATBiLSTM technique with other methods, including LSTM, BiLSTM, and LSTM-Attention, for classifying data into two groups. The table highlights the effectiveness of ATBiLSTM in achieving superior classification accuracy and stability compared to traditional models.

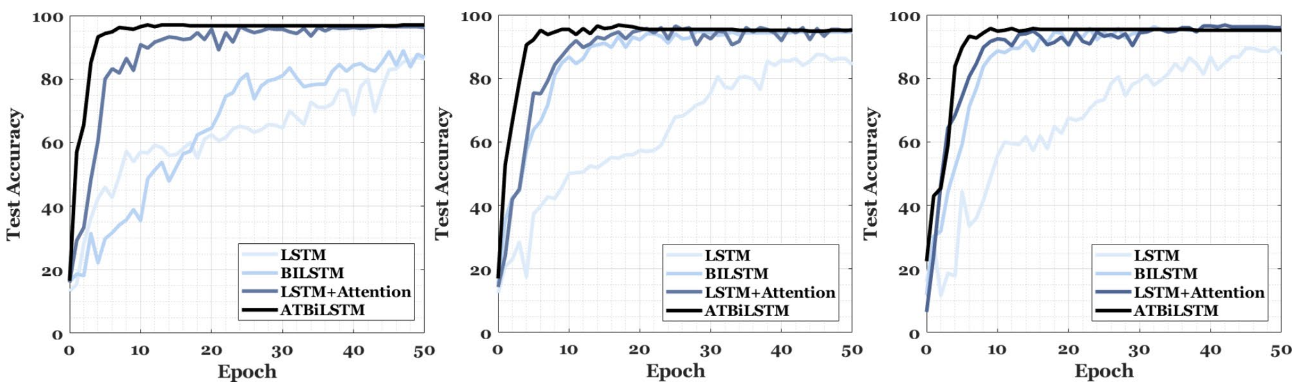

Fig. 6. Classification results for three randomly selected EEG signal samples used in the experiment. This figure compares the performance of the proposed ATBiLSTM method with some established models, including LSTM, BiLSTM, and LSTM-Attention, in classifying AD and MCI, highlighting the effectiveness of the ATBiLSTM approach.

the absence of Bi-LSTM blocks and convolutional layers precipitates diminished accuracy, underscoring their indispensable role in capturing temporal and spatial dependencies within the data, respectively. Through this granular analysis, researchers attain a nuanced comprehension of the relative importance of each constituent element, empowering them to refine and fine-tune classification models tailored for diagnosing Alzheimer's disease and cognate cognitive impairments. This meticulous approach not only augments model performance but also furnishes invaluable insights into the underlying pathophysiological mechanisms, potentially paving the way for more efficacious therapeutic interventions and diagnostic modalities.

An analysis of the architectural features of ATBiLSTM suggests its potential advantage over the compared models (see Fig. 6). This is further supported by the accuracy graphs depicted. The graphs demonstrate that ATBiLSTM consistently outperforms the other models throughout the classification process, achieving a demonstrably higher overall classification accuracy. As illustrated in Fig. 6, ATBiLSTM surpassed other established models, including LSTM, BiLSTM, and LSTM-Attention, in classifying AD and MCI across different EEG samples. This suggests that ATBiLSTM offers greater robustness and accuracy, particularly when dealing with uncertainties in the data. These findings strongly support the efficacy of ATBiLSTM as a reliable and accurate tool for classifying EEG signals in the context of AD and MCI detection. The model's ability to handle uncertainties and achieve superior performance compared to existing methods highlights its potential for practical applications in healthcare settings.

Moreover, Fig. 7 provides an illustrative example of the feature maps extracted from three intermediate layers of our deep structure. In the first layer (Feature Map - Layer 1), basic patterns and fundamental structures are identified, which are essential for setting up the deeper feature extraction that occurs in subsequent layers. This initial layer captures broad, low-level features from the spectrograms.

In the second layer (Feature Map - Layer 2), the network begins to refine these basic patterns into more complex representations, potentially isolating specific spectral components or identifying characteristic frequency bands that are relevant to differentiating between patient conditions. The visualizations of this layer indicate the network's increasing focus and specialization, as it learns to highlight features that are most indicative of AD, MCI, or healthy states.

By the time we reach the third layer (Feature Map - Layer 3), the network is extracting highly specialized features, suggesting a deep understanding of the nuances in the spectrograms that are critical for accurate classification. This layer demonstrates the culmination of the hierarchical feature extraction process, where complex and condition-specific features are isolated and enhanced. These progressively refined features are then used by the ATBiLSTM classifier, which leverages its attention mechanism to weigh the importance of each feature optimally, thereby improving classification accuracy.

Conference presentation

Part of the initial idea of the current work has been presented in EICEEAI-2023, an event that was held in Dec. 2023, in Jordan. The present version has been fully extended, and upgraded with new materials compared to the conference version[62].

Discussion

Here, we highlight the effect of using DWT to achieve dominant subbands from EEG data. As previously mentioned, DWT decomposes the signal into different frequency bands, allowing the model to focus on specific brain activity patterns relevant to AD detection. Moreover, the CL-ATBiLSTM architecture combines Convolutional Learning (CL), Attention Mechanisms, and BiLSTM networks. Convolutional layers efficiently extract features from the EEG data, while the attention mechanism prioritizes informative features crucial for classification. BiLSTM networks capture temporal dependencies within the data, improving classification accuracy. Additionally, our model achieves an impressive 96.52% accuracy on Figshare datasets, significantly outperforming existing methods and state-of-the-art approaches. This suggests that CL-ATBiLSTM is highly effective in distinguishing between healthy individuals and AD patients. Early and accurate AD diagnosis is essential for timely interventions and improved patient outcomes. The high accuracy of CL-ATBiLSTM suggests its potential for real-world applications in clinical settings.

The results from the IoMT-based validation demonstrated that the algorithm consistently maintained high accuracy and reliability across various test conditions. Performance metrics such as accuracy, precision, recall, and F1-score were measured and showed alignment with the results obtained from the initial testing using public datasets. This consistency confirms the algorithm's capability to handle real-time data streams and

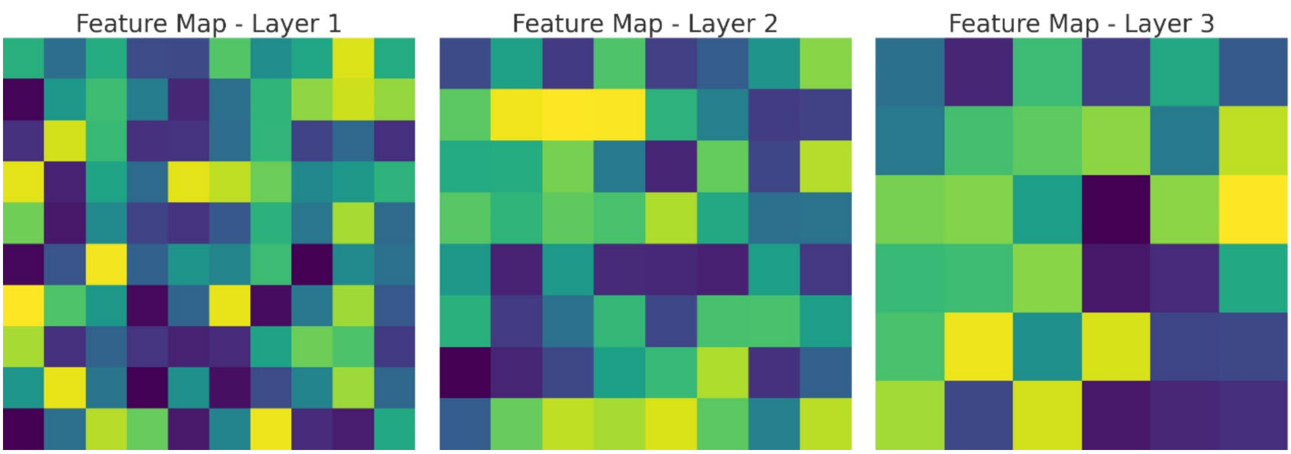

Fig. 7. Visualizations of feature maps from three intermediate layers of the proposed deep architecture.

varying environmental factors typical of IoMT systems. The validation on a practical IoMT setup underscores the algorithm's potential for integration into clinical workflows, providing a reliable tool for the diagnosis of AD and MCI in dynamic and resource-constrained environments.

EEG Sub-band Analysis

A major challenge with using EEG for diagnosing brain diseases is that the results can vary and may not be consistent across patients. Scientists are actively researching the specific areas of the brain most affected by MCI and AD to improve these tests. Moreover, in order to enhance interpretability, several techniques were employed. First, feature importance analysis was conducted to identify the most influential features within the EEG signals, highlighting the critical aspects relevant for diagnosing AD and MCI. Layer-Wise Relevance Propagation (LRP) was also applied, assigning relevance scores to input features, thereby clarifying the contribution of each input to the model's predictions. Additionally, explainable models, such as decision trees, were integrated to cross-validate the primary model's predictions, ensuring consistency with interpretable logic. Visualization techniques, including saliency maps, were used to visually represent the parts of the input data that the model focused on, further enhancing transparency. Finally, the model's predictions were validated against known clinical correlations, aligning the outputs with established medical understanding. These combined approaches ensure that the proposed model is not only accurate but also explainable, making it a reliable tool for clinical applications in diagnosing AD and MCI.

Our study examined the importance of specific frequency bands in EEG signals for diagnosing MCI and AD. The analysis considered specificity, sensitivity, and overall accuracy, as visualized in Fig. 8. The results indicate that most EEG subbands carry some information about AD and MCI, with the alpha, theta, and delta bands being the most significant.

The identification of AD using EEG signals is heavily influenced by the specific channels and brain regions analyzed. EEG channels capture distinct neural activities associated with different brain regions, and their relevance varies depending on the cognitive functions these regions control. For instance, frontal and temporal regions are crucial for memory and attention, which are significantly affected in AD. Channels located over these regions, such as frontal (F) and temporal (T) electrodes, provide valuable information for distinguishing AD from MCI and healthy controls. The temporal lobe, in particular, is associated with early AD pathology, making EEG channels over this region critical for early detection. This regional focus aligns with observed patterns in power spectral densities, such as increased theta band activity (4–8 Hz) and decreased alpha band activity (8–12 Hz) in AD patients, reflecting the underlying neural disruptions characteristic of the disease.

Our approach further emphasizes the importance of analyzing specific EEG subbands rather than the entire spectrum, as different frequency bands correspond to distinct brain functions. For instance, the alpha band (8–12 Hz) is linked to relaxation and reduced cognitive demand, while the theta band (4–8 Hz) is associated with memory processes. By focusing on these subbands, we can capture nuanced changes in neural activity relevant to conditions like MCI and AD. This targeted analysis allows for more sensitive detection of subtle changes in brain function that are characteristic of these conditions but may not be detectable when considering the full EEG spectrum. By examining the contribution of each subband to classification accuracy, we gain deeper insights into the neural mechanisms underlying AD and MCI, enhancing both the interpretability and diagnostic accuracy of our model.

To effectively leverage these insights, our method employs a CNN based on the ResNet101 architecture to analyze spectrograms generated from EEG signals, followed by classification using the Attention-Bidirectional

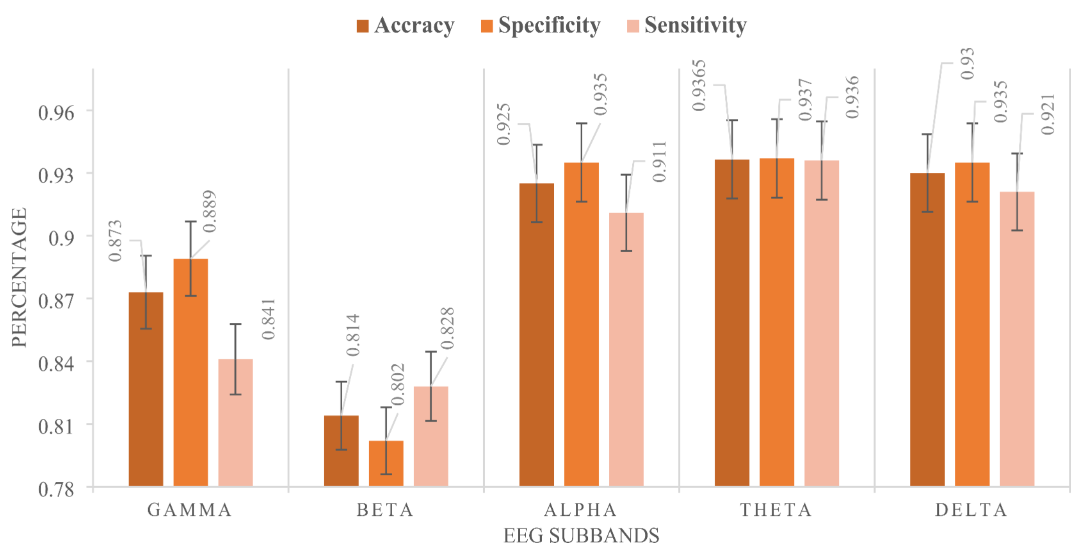

Fig. 8. Analysis of individual EEG subbands, highlighting the Theta, Delta, and Alpha bands due to their superior performance in sensitivity, specificity, and accuracy for MCI and AD classification. These results suggest that these subbands are particularly significant for identifying these conditions.

Time-Aware Long-Short-Term Memory (ATBiLSTM) model. This approach not only improves the model's ability to identify the most relevant features across different EEG channels and subbands but also provides a nuanced understanding of why certain channels and brain regions contribute more significantly to AD diagnosis. By incorporating targeted EEG channels and frequency bands, our model achieves robust and interpretable performance, addressing the reviewer's concern about the need for a comprehensive discussion on the significance of specific channels and brain regions. This two-step process highlights the importance of certain electrode locations and frequency ranges, ultimately leading to higher classification accuracy for distinguishing between AD, MCI, and healthy individuals.

Comparison
Building upon prior research, this study confirms the significance of specific electrodes (T, O, F, and Fp) and frequency subbands (Alpha, Theta, and Delta) in differentiating between MCI and AD using EEG signals. Our observations align with previous studies[50–57] that have documented characteristic changes in brainwave patterns associated with these conditions. We observed a decrease in alpha activity (associated with relaxation and low attention) and a corresponding increase in beta activity (associated with high brain activity and alertness) in individuals with cognitive decline. This suggests a potential shift in brain function from a relaxed state to a state of increased engagement, which could be indicative of cognitive changes related to MCI and AD. Additionally, an increase in theta and delta activity, signifying decreased attention and concentration, further supports this notion.

The consistency between our results and prior research on EEG classification strengthens the case for these specific electrodes and subbands as reliable biomarkers for diagnosing MCI and AD. This aligns with the established knowledge in cognitive neuroscience, where alpha, theta, and delta rhythms are well-recognized markers of brain activity patterns.

It is important to acknowledge the inherent challenges associated with EEG data classification. The inherent complexity and variability of EEG signals, coupled with the limited availability of well-annotated datasets for specific neurological conditions, pose significant obstacles for accurate classification. Traditionally, this process requires extensive expertise and time investment from specialists in the field.

This research proposes an automatic classification model that integrates an attention mechanism and an optimized Bayesian algorithm. As demonstrated in Table 2, this approach achieves superior classification accuracy for MCI and AD compared to existing methods, as evidenced by metrics like recall and accuracy.

Author	Dataset	No. classes	Metric (Acc)	Classifier strategy	Feature type	Advantages	Disadvantages
Perez-Valero et al.[1]	Self-recorded	2	95.00%	MLP	Handcrafted	Higher accuracy and multi-class analysis	Only two-class analysis, less flexible
Safi et al. [12]	Brazilian	3	97.64%	*k*-Nearest Neighbor (*k*-NN)	Handcrafted	High accuracy but lower processing time and complexity	Higher processing time and complexity
AlSharabi et al.[15]	Brazilian	3	98.10%	Multiple Classifier Assortment	Handcrafted	High accuracy but lower processing time	Higher processing time and complexity, need to combine multiple models
Fouladi et al.[24]	Barreto	2	89.26%	CNN and Autoencoder	Automated	Higher accuracy and multi-class analysis	Lower accuracy
Fouad and Labib[34]	Two datasets	2	97.83%	Classical ML Techniques models and ResNet	handcrafted and Automated	High accuracy but lower processing time	Higher complexity and processing time
Cejnek et al.[49]	Figshare	3	90.29%	SVM and other classifiers	Handcrafted	Higher accuracy and automated approach	Lower accuracy
Fiscon et al.[50]	Italian	3	74.70%	J48	Handcrafted	Significantly higher accuracy	Very low accuracy
Ieracitano et al.[45]	Italian	3	89.22%	SVM, MLP, and Logistic regression	Handcrafted	Higher accuracy and automated approach	Lower accuracy
Song et al.[46]	Self-recorded	2	100%	Tripartite Encoding Routes	Handcrafted	High accuracy but only two-class analysis	Only two-class analysis, high complexity
Trambaiolli et al.[51]	Brazilian	4	76.88%	SVM	Handcrafted	Higher accuracy and automated approach	Lower accuracy
Amezquita-Sanchez et al.[52]	Italian	2	90.30%	Enhanced probabilistic NN (EPNN)	Handcrafted	Higher accuracy and multi-class analysis	Lower accuracy
Triggiani et al.[53]	Italian	3	76.70%	Artificial neural network (ANN)	Handcrafted	Higher accuracy and automated approach	Lower accuracy
Cassani et al.[54]	Brazilian and US	4	91.40%	SVM	Handcrafted	Higher accuracy and automated approach	Lower accuracy
Morabito et al.[55]	Italian	3	82.00%	CNN	Automated	Higher accuracy and multi-class analysis	Lower accuracy
Bevilacqua et al.[56]	Spanish	2	86.00%	Multilayer perceptron and SVM	Handcrafted	Higher accuracy and multi-class analysis	Lower accuracy, only two-class analysis
Kanda et al.[57]	Brazilian	4	83.95%	Support vector machine (SVM)	Handcrafted	Higher accuracy and automated approach	Lower accuracy
Our model	Figshare	3	96.52%	Convolutional Learning and Attention-based Bi-LSTM networks	Automated	Higher accuracy, multi-class analysis, lower processing time, and lower complexity	Small data size

Table 2. To assess the accuracy and number of classes of our suggested model in comparison to previously published methodologies, various methods were taken into consideration.

The proposed model offers a significant advantage over traditional feature extraction methods. It is specifically designed to analyze EEG data in the context of MCI and AD diagnosis, addressing the limitations of previous techniques that were often restricted by small datasets. This research builds upon existing knowledge by employing a more robust and efficient method for EEG classification.

Prior research in EEG data analysis primarily relied on conventional feature extraction techniques, both manual and automated. Notably, a scarcity exists in models specifically tailored for analyzing EEG data from MCI and AD patients. While various techniques have been explored for AD detection, limitations often arise due to small datasets, as exemplified by Ieracitano et al.'s work[45]. This study highlights the potential of our proposed model to overcome these limitations and contribute to the advancement of automated EEG classification for MCI and AD diagnosis.

Researchers have explored various classification methods for detecting AD using brainwave (EEG) data. These methods include Support Vector Machines (SVM), Logistic Regression (LR), and different neural network architectures. Some studies have even attempted to classify more than three conditions based on EEG data[49–45,51]. However, the reliability of these methods is not always clear. For instance, a study by Kanda et al.[61] achieved an accuracy of 83.95% for classifying four conditions using SVM with a specific filter for processing the EEG signal. While our study utilizes a large dataset, our approach is flexible and can be adapted to work with different datasets.

There are two key points to consider when comparing our approach with those used by AlSharabi et al. [15] and Cassani et al.[54]. Our method, which utilizes a combination of convolutional layers and attention-based Bi-directional Long Short-Term Memory (Bi-LSTM) networks, is able to reduce processing time while maintaining high accuracy. This is in contrast to previous studies by Amezquita-Sánchez et al.[52], Trambaiolli et al.[51], Cassani et al.[54], and AlSharabi et al.[15].

Thanks to an upgrade on a data-sharing platform (Figshare), we were able to significantly increase the number of EEG signals analyzed compared to previous methods. Despite challenges like missing data points and the high dimensionality of EEG data, our system can effectively learn from new signals. The accuracy of our system consistently remained around 96.52% across all trials and validations, indicating the reliability of data obtained from repeated measurements. The proposed method outperforms existing state-of-the-art models for classifying MCI and AD. It achieves this by using a simpler and less complex neural network with fewer parameters that need to be adjusted during training. When applied to various EEG datasets, our model surpasses previous versions in terms of accuracy, precision, and recall. Additionally, the model requires minimal time to process new (test) data sets.

Evaluating why some methodologies may have performed comparably or even better than our proposed model, which deals with a larger number of classes, reveals that the quality and relevance of extracted or engineered features from the datasets play a significant role. Hand-crafted features in some studies may have captured more distinctive information from the data compared to the automatically generated features by our model. On the other hand, datasets are different, and the size and quality of the dataset used for training and testing can significantly impact model performance. Studies with larger and more diverse datasets may achieve better generalization and performance.

Deep learning models can be made flexible by adjusting their architecture, number of layers, and internal settings. However, for them to perform well, it's crucial to train them on data that closely reflects their intended use. This research introduces a model that can be trained with unseen brainwave (EEG) signals to predict MCI and AD. To achieve this, a technique called "signal windowing" was used to increase the amount and variety of data available for training.

Challenges and solutions

Research in using EEG data for AD diagnosis faces numerous challenges and obstacles. One significant issue is the inherent complexity and variability of EEG signals. These signals exhibit continuous fluctuations in brain activity, making it difficult to extract reliable and consistent characteristics. The dynamic nature of brainwaves requires sophisticated techniques to ensure accurate analysis and classification. Moreover, the high dimensionality and inherent noise in EEG data pose significant challenges for classifier performance[58,59].

In addition to addressing the challenges of using EEG data for Alzheimer's diagnosis, it is crucial to consider recent advancements in deep learning that go beyond traditional CNN and LSTM models. Newer architectures, such as attention-based and transformer models, have shown improved performance in capturing complex temporal patterns in EEG signals. These models combine the spatial feature extraction capabilities of CNNs with attention mechanisms that highlight the most relevant signal components, thereby enhancing diagnostic accuracy. Hybrid approaches that integrate convolutional layers, bidirectional LSTMs, and attention layers are particularly effective in managing the dynamic nature of EEG data, capturing both spatial and temporal features more precisely. Incorporating these advancements within the IoMT framework can further enhance the robustness and clinical applicability of diagnostic models, ensuring the use of cutting-edge methods for AD diagnosis using EEG data.

Another major challenge is the limited availability of well-annotated datasets specifically tailored for Alzheimer's research. Many existing datasets are small and lack the diversity needed to train robust models, limiting the generalizability of research findings. This scarcity of comprehensive datasets makes it difficult to validate the efficacy of newly developed models and algorithms on diverse populations, thereby hindering their practical application.

Moreover, traditional EEG analysis methods often rely heavily on manual or semi-automated feature extraction techniques. These methods require extensive expertise and time investment from specialists, making the process laborious and inefficient. The manual nature of these techniques also introduces subjectivity, potentially affecting the reliability and consistency of the results. The IoMT offers promising solutions to some of these challenges. IoMT enables the integration of various medical devices and systems into a connected network, facilitating real-time data collection, monitoring, and analysis. This connectivity can significantly enhance the efficiency and accuracy of EEG-based AD.

We utilized deep learning techniques including CNN and ATBiLSTM for feature analysis and optimized classification, respectively. Moreover, these refined methods leverage the extensive EEG data gathered via IoMT to accurately classify patients into categories such as AD, MCI, or healthy. Consequently, the classification process gains from the ongoing data acquisition facilitated by IoMT, thereby augmenting the overall efficacy and dependability of the diagnostic models. IoMT also supports the use of advanced computational techniques and cloud-based resources. These technologies enable the efficient processing and analysis of large volumes of EEG data. By utilizing cloud computing and distributed systems, researchers can implement sophisticated deep learning models, such as the Convolutional Learning Attention-Bidirectional Time-Aware Long-Short-Term Memory (CL-ATBiLSTM) model, to analyze EEG data. This model combines convolutional layers, attention mechanisms, and BiLSTM networks to capture spatial and temporal features, enhancing the accuracy of Alzheimer's diagnosis. Furthermore, IoMT facilitates the development of automated and user-friendly diagnostic tools. These tools can provide real-time feedback and insights, reducing the reliance on specialist expertise and minimizing subjectivity in the diagnostic process. Automated systems can streamline the workflow, making it more efficient and accessible for broader clinical use[60].

The integration of IoMT in AD research holds significant potential for improving patient outcomes. Early and accurate detection of AD is crucial for timely interventions and effective management of the condition[61]. The high accuracy of models like CL-ATBiLSTM, supported by IoMT infrastructure, can expedite the identification of AD, enabling healthcare providers to implement targeted treatment strategies sooner.

Moreover, IoMT-based systems can facilitate continuous monitoring and follow-up of patients, providing valuable insights into disease progression and treatment efficacy. This ongoing data collection can support personalized treatment plans, enhancing the quality of care and potentially slowing the progression of the disease.

Conclusion

The importance of early detection and intervention for Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI) cannot be stressed enough. Since AD is the most common neurodegenerative disorder, and MCI can be a sign of future cognitive decline, accurate and timely diagnosis is essential. This study shows that EEG data, combined with advanced computer models, can be effective in identifying AD. We used a specific type of model called a Convolutional Learning and Attention-based Bi-LSTM (ATBiLSTM) and achieved impressive accuracy in detecting both MCI and AD. This demonstrates the effectiveness of our proposed method. Our findings also highlight the versatility and robustness of the ATBiLSTM model. It's not just useful for diagnosing AD; it can also be used to analyze EEG signals from people with other neurodegenerative disorders, as shown by its success with "other AD datasets" (although the specific datasets aren't mentioned here). By showing that this model can be used for a wider range of conditions, our research contributes to the development of more general-purpose systems for making decisions in neurological diagnostics. Furthermore, our model achieved exceptional accuracy, which was confirmed through rigorous testing. This suggests that it could be a valuable tool in clinical settings to help identify potential biomarkers that are important for diagnosing AD. Moreover, our study presents a significant advancement in the field of diagnosing neurodegenerative diseases. We propose a new computer-aided diagnostic model that analyzes EEG data. This model achieved high accuracy in identifying both MCI and AD, suggesting it has the potential to improve patient outcomes by enabling earlier detection and personalized treatment plans. As we continue to refine and validate our approach, we hope to see it integrated into clinical practice, ultimately leading to better diagnosis and management of neurodegenerative disorders.

Data availability

All datasets used in this study are freely available through the open repositories on the web. All data generated or analyzed during this study are included in this published articles and its supplementary information files from Figshare[49]. https://figshare.com/articles/dataset/dataset_zip/5450293.

Received: 14 June 2024; Accepted: 25 October 2024
Published online: 29 October 2024

References

1. Perez-Valero, E., Lopez-Gordo, M. Á., Gutiérrez, C. M., Carrera-Muñoz, I. & Vílchez-Carrillo, R. M. <ArticleTitle Language="En">A self-driven approach for multi-class discrimination in Alzheimer's disease based on wearable EEG. *Comput. Biol. Med.* **220**, 106841. https://doi.org/10.1016/j.cmpb.2022.106841 (2022).
2. Petersen, R. C. Mild cognitive impairment. *CONTIN Lifelong Learn. Neurol.* **22**, 404. https://doi.org/10.1212/CON.0000000000000313 (2016).
3. Petersen, R. C. et al. Mild cognitive impairment: Clinical characterization and outcome. *Arch. Neurol.* **56** (3), 303–308. https://doi.org/10.1001/archneur.56.3.303 (1999).
4. Dugger, B. N., Tu, M., Murray, M. E., Dickson, D. W. & University of Washington Alzheimer's Disease Research Center. Neuropathological comparisons of amnestic and nonamnestic mild cognitive impairment. *BMC Neurol.* **15** (1), 1–8. https://doi.org/10.1186/s12883-015-0344-0 (2015).
5. Jessen, F. et al. The characterisation of subjective cognitive decline. *Lancet Neurol.* **19** (3), 271–278. https://doi.org/10.1016/S1474-4422(19)30368-0 (2020).
6. Venkata Phanikrishna, B., Prakash, J., Suchismitha, C. & A., & Deep review of machine learning techniques on detection of drowsiness using EEG signal. *IETE J. Res.* **69** (6), 3104–3119. https://doi.org/10.1080/03772063.2021.1913070 (2023).
7. Gawel, M., Zalewska, E., Szmidt-Sałkowska, E. & Kowalski, J. The value of quantitative EEG in differential diagnosis of Alzheimer's disease and subcortical vascular dementia. *J. Neurol. Sci.* **283** (1–2), 127–133. https://doi.org/10.1016/j.jns.2009.02.332 (2009).

8. Oltu, B., Akşahin, M. F. & Kibaroğlu, S. A novel electroencephalography based approach for Alzheimer's disease and mild cognitive impairment detection. *Biomed. Signal Process. Control*. **63**, 102223. https://doi.org/10.1016/j.bspc.2020.102223 (2021).
9. Azami, H. et al. Beta to theta power ratio in EEG periodic components as a potential biomarker in mild cognitive impairment and Alzheimer's dementia. *Alzheimers Res. Ther.***15** (1), 1–12. https://doi.org/10.1186/S13195-023-01280-Z/FIGURES/5 (2023).
10. Li, R. X., Ma, Y. H., Tan, L., Yu, J. T. & Prospective biomarkers of Alzheimer's disease: A systematic review and meta-analysis. *Ageing Res. Rev.***81**, 101699. https://doi.org/10.1016/J.ARR.2022.101699 (2022).
11. Khan, P. et al. Machine learning and deep learning approaches for brain disease diagnosis: Principles and recent advances. *IEEE Access.***9**, 37622–37655. https://doi.org/10.1109/ACCESS.2021.3062484 (2021).
12. Safi, M. S. & Safi, S. M. M. Early detection of Alzheimer's disease from EEG signals using Hjorth parameters. *Biomed. Signal Process. Control*. **65**, 102338. https://doi.org/10.1016/J.BSPC.2020.102338 (2021).
13. Gong, S., Xing, K., Cichocki, A. & Li, J. Deep learning in EEG: Advance of the last ten-year critical period. *IEEE Trans. Cogn. Dev. Syst.***14** (2), 348–365. https://doi.org/10.1109/TCDS.2021.3079712 (2021).
14. Miltiadous, A. et al. Alzheimer's Disease and Frontotemporal Dementia: A Robust Classification Method of EEG Signals and a Comparison of Validation Methods. *Diagnostics*. **11** (8), 1437. https://doi.org/10.3390/DIAGNOSTICS11081437 (2021).
15. Alsharabi, K., Salamah, B., Abdurraqeeb, Y., Aljalal, A. M., Alturki, F. A. & M., & EEG Signal Processing for Alzheimer's Disorders Using Discrete Wavelet Transform and Machine Learning Approaches. *IEEE Access.***10**, 89781–89797. https://doi.org/10.1109/ACCESS.2022.3198988 (2022).
16. Tautan, A. M. et al. Preliminary study on the impact of EEG density on TMS-EEG classification in Alzheimer's disease. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS (Vol. 2022-July, pp. 394–397). (2022). https://doi.org/10.1109/EMBC48229.2022.9870920
17. Alvi, A. M., Siuly, S. & Wang, H. A long short-term memory based framework for early detection of mild cognitive impairment from EEG signals. *IEEE Trans. Emerg. Top. Comput. Intell.***7** (2), 375–388. https://doi.org/10.1109/TETCI.2022.3186180 (2022).
18. Alessandrini, M. et al. EEG-Based Alzheimer's Disease Recognition Using Robust-PCA and LSTM Recurrent Neural Network. *Sensors*. **22** (10), 3696. https://doi.org/10.3390/S22103696 (2022).
19. Amini, M., Pedram, M. M., Moradi, A. R. & Ouchani, M. Diagnosis of Alzheimer's Disease by Time-Dependent Power Spectrum Descriptors and Convolutional Neural Network Using EEG Signal. Computational and Mathematical Methods in Medicine, 2021. (2021). https://doi.org/10.1155/2021/5511922
20. Senturk, U., Polat, K. & Yucedag, I. A non-invasive continuous cuffless blood pressure estimation using dynamic Recurrent Neural Networks. *Appl. Acoust.***170**, 107534. https://doi.org/10.1016/J.APACOUST.2020.107534 (2020).
21. Sharma, G., Parashar, A. & Joshi, A. M. DepHNN: A novel hybrid neural network for electroencephalogram (EEG)-based screening of depression. *Biomed. Signal Process. Control*. **66**, 102393. https://doi.org/10.1016/J.BSPC.2020.102393 (2021).
22. Klepl, D., He, F., Wu, M., Blackburn, D. J. & Sarrigiannis, P. G. Apr. Adaptive Gated Graph Convolutional Network for Explainable Diagnosis of Alzheimer's Disease using EEG Data. Accessed: May 26, 2023. [Online]. Available: (2023). https://arxiv.org/abs/2304.05874v1
23. Shikalgar, A. & Sonavane, S. Hybrid Deep Learning Approach for Classifying Alzheimer Disease Based on Multimodal Data. *Adv. Intell. Syst. Comput.*, **1025**, 511–520. https://doi.org/10.1007/978-981-32-9515-5_49/COVER (2020).
24. Fouladi, S., Safaei, A. A., Mammone, N., Ghaderi, F. & Ebadi, M. J. Efficient Deep Neural Networks for Classification of Alzheimer's Disease and Mild Cognitive Impairment from Scalp EEG Recordings. *Cogn. Comput.***14** (4), 1247–1268. https://doi.org/10.1007/S12559-022-10033-3 (2022).
25. Huggins, C. J. et al. Deep learning of resting-state electroencephalogram signals for three-class classification of Alzheimer's disease, mild cognitive impairment and healthy ageing. *J. Neural Eng.***18** (4), 046087. https://doi.org/10.1088/1741-2552/AC05D (2021).
26. Ambeth Kumar, V. D. et al. An Internet of Medical Things-Based Mental Disorder Prediction System Using EEG Sensor and Big Data Mining. *J. Circuits Syst. Computers*, 2450197. https://doi.org/10.1142/S0218126624501974. (2024).
27. Dahan, F. et al. A smart IoMT based architecture for E-healthcare patient monitoring system using artificial intelligence algorithms. *Front. Physiol.***14**, 1125952. https://doi.org/10.3389/fphys.2023.1125952 (2023).
28. Imani, M. Alzheimer's diseases diagnosis using fusion of high informative BiLSTM and CNN features of EEG signal. *Biomed. Signal Process. Control*. **86**, 105298. https://doi.org/10.1016/J.BSPC.2023.105298 (2023).
29. Nobukawa, S. et al. Classification Methods Based on Complexity and Synchronization of Electroencephalography Signals in Alzheimer's Disease. *Front. Psychiatry*. **11**, 511787. https://doi.org/10.3389/FPSYT.2020.00255/BIBTEX (2020).
30. Yu, H., Lei, X., Song, Z., Liu, C. & Wang, J. Supervised Network-Based Fuzzy Learning of EEG Signals for Alzheimer's Disease Identification. *IEEE Trans. Fuzzy Syst.***28** (1), 60–71. https://doi.org/10.1109/TFUZZ.2019.2903753 (2020).
31. Bi, X. & Wang, H. Early Alzheimer's disease diagnosis based on EEG spectral images using deep learning. *Neural Netw.***114**, 119–135. https://doi.org/10.1016/J.NEUNET.2019.02.005 (2019).
32. Miltiadous, A., Gionanidis, E., Tzimourta, K. D., Giannakeas, N. & Tzallas, A. T. DICENet: A Novel Convolution-Transformer Architecture for Alzheimer Detection in EEG Signals. *IEEE Access.***11**, 71840–71858. https://doi.org/10.1109/ACCESS.2023.3294618 (2023).
33. Lopes, M., Cassani, R. & Falk, T. H. Using CNN saliency maps and EEG modulation spectra for improved and more interpretable machine learning-based alzheimer's disease diagnosis. Computational Intelligence and Neuroscience, 2023. (2023). https://doi.org/10.1155/2023/3198066
34. Fouad, I. A., El-Zahraa, F. & Labib, M. Identification of Alzheimer's disease from central lobe EEG signals utilizing machine learning and residual neural network. *Biomed. Signal Process. Control*. **86**, 105266. https://doi.org/10.1016/J.BSPC.2023.105266 (2023).
35. kumar Ravikanti, D. & Saravanan, S. EEGAlzheimer'sNet: Development of transformer-based attention long short term memory network for detecting Alzheimer disease using EEG signal. *Biomed. Signal Process. Control*. **86**, 105318. https://doi.org/10.1016/j.bspc.2023.105318 (2023).
36. Xie, J. et al. A transformer-based approach combining deep learning network and spatial–temporal information for raw EEG classification. *IEEE Trans. Neural Syst. Rehabil Eng.***30**, 2126–2136. https://doi.org/10.1109/TNSRE.2022.3194600 (2022).
37. Ferri, R. et al. Stacked autoencoders as new models for an accurate Alzheimer's disease classification support using resting-state EEG and MRI measurements. *Clin. Neurophysiol.***132** (1), 232–245. https://doi.org/10.1016/j.clinph.2020.09.015 (2021).
38. You, Z., Zeng, R., Lan, X., Ren, H. & Guo, Y. Alzheimer's disease classification with a cascade neural network. *Front. Public. Health*. **8**, 584387. https://doi.org/10.3389/fpubh.2020.584387 (2020).
39. Rad, E. M. et al. Diagnosis of mild Alzheimer's disease by EEG and ERP signals using linear and nonlinear classifiers. *Biomed. Signal Process. Control*. **70**, 103049. https://doi.org/10.1016/j.bspc.2021.103049 (2021).
40. Morabito, F. C., Ieracitano, C. & Mammone, N. An explainable artificial intelligence approach to study MCI to AD conversion via HD-EEG processing. *Clin. EEG Neurosci.* 15500594211063662. https://doi.org/10.1177/15500594211063662 (2021).
41. Araújo, T., Teixeira, J. P. & Rodrigues, P. M. Smart-data-driven system for Alzheimer disease detection through electroencephalographic signals. *Bioengineering*. **9** (4), 141. https://doi.org/10.3390/bioengineering9040141 (2022).
42. Nour, M., Senturk, U. & Polat, K. A novel hybrid model in the diagnosis and classification of Alzheimer's disease using EEG signals: Deep ensemble learning (DEL) approach. *Biomed. Signal Process. Control*. **89**, 105751. https://doi.org/10.1016/j.bspc.2023.105751 (2024).

43. Siddiqui, M. K., Huang, X., Morales-Menendez, R., Hussain, N. & Khatoon, K. Machine learning based novel cost-sensitive seizure detection classifier for imbalanced EEG data sets. *Int. J. Interact. Des. Manuf. (IJIDeM)*. **14**, 1491–1509. https://doi.org/10.1007/s12008-020-00715-3 (2020).
44. Siddiqui, M. K., Morales-Menendez, R., Huang, X. & Hussain, N. A review of epileptic seizure detection using machine learning classifiers. *Brain Inf.***7** (1), 5. https://doi.org/10.1186/s40708-020-00105-1 (2020).
45. Ieracitano, C., Mammone, N., Hussain, A. & Morabito, F. C. A novel multi-modal machine learning based approach for automatic classification of EEG recordings in dementia. *Neural Netw.***123**, 176–190. https://doi.org/10.1016/j.neunet.2019.12.006 (2020).
46. Song, Z., Deng, B., Wang, J. & Yi, G. An EEG-based systematic explainable detection framework for probing and localizing abnormal patterns in Alzheimer's disease. *J. Neural Eng.***19** (3), 036007. https://doi.org/10.1088/1741-2552/ac697d (2022).
47. Zhang, D., Jin, X., Shi, P. & Chew, X. Real-time load forecasting model for the smart grid using bayesian optimized CNN-BiLSTM. *Front. Energy Res.***11**, 1193662. https://doi.org/10.3389/fenrg.2023.1193662 (2023).
48. Li, H. et al. Automatic electrocardiogram detection and classification using bidirectional long short-term memory network improved by Bayesian optimization. *Biomed. Signal Process. Control*. **73**, 103424. https://doi.org/10.1016/j.bspc.2021.103424 (2022).
49. Cejnek, M., Vysata, O., Valis, M. & Bukovsky, I. Novelty detection-based approach for Alzheimer's disease and mild cognitive impairment diagnosis from EEG. *Med. Biol. Eng. Comput.***59** (11), 2287–2296. https://doi.org/10.1007/s11517-021-02427-6 (2021).
50. Fiscon, G. et al. Combining EEG signal processing with supervised methods for Alzheimer's patients classification. *BMC Med. Inf. Decis. Mak.***18** (1), 1–10. https://doi.org/10.1186/s12911-018-0613-y (2018).
51. Trambaiolli, L. R., Spolaôr, N., Lorena, A. C., Anghinah, R. & Sato, J. R. Feature selection before EEG classification supports the diagnosis of Alzheimer's disease. *Clin. Neurophysiol.***128** (10), 2058–2067. https://doi.org/10.1016/j.clinph.2017.06.251 (2017).
52. Amezquita-Sanchez, J. P., Mammone, N., Morabito, F. C., Marino, S. & Adeli, H. A novel methodology for automated differential diagnosis of mild cognitive impairment and the Alzheimer's disease using EEG signals. *J. Neurosci. Methods*. **322**, 88–95. https://doi.org/10.1016/j.jneumeth.2019.04.013 (2019).
53. Triggiani, A. I. et al. Classification of healthy subjects and Alzheimer's disease patients with dementia from cortical sources of resting state EEG rhythms: a study using artificial neural networks. *Front. NeuroSci.***10**, 604. https://doi.org/10.3389/fnins.2016.00604 (2017).
54. Cassani, R. et al. Towards automated electroencephalography-based Alzheimer's disease diagnosis using portable low-density devices. *Biomed. Signal Process. Control*. **33**, 261–271. https://doi.org/10.1016/j.bspc.2016.12.009 (2017).
55. Morabito, F. C. et al. Deep convolutional neural networks for classification of mild cognitive impaired and Alzheimer's disease patients from scalp EEG recordings. In 2016 IEEE 2nd International Forum on Research and Technologies for Society and Industry Leveraging a better tomorrow (RTSI) (pp. 1–6). (2016)., September https://doi.org/10.1109/RTSI.2016.7740576
56. Bevilacqua, V. et al. Advanced classification of Alzheimer's disease and healthy subjects based on EEG markers. In 2015 International Joint Conference on Neural Networks (IJCNN) (pp. 1–5). (2015)., July https://doi.org/10.1109/IJCNN.2015.7280463
57. Kanda, P. A. M. et al. Clinician's road map to wavelet EEG as an Alzheimer's disease biomarker. *Clin. EEG Neurosci.***45** (2), 104–112. https://doi.org/10.1177/1550059413486272 (2014).
58. Siddiqui, M. K., Islam, M. Z. & Kabir, M. A. Analyzing performance of classification techniques in detecting epileptic seizure. InAdvanced Data Mining and Applications: 13th International Conference, ADMA 2017, Singapore, November 5–6, 2017, Proceedings 13 2017 (pp. 386–398). Springer International Publishing. https://doi.org/10.1007/978-3-319-69179-4_27
59. Cong, G., Peng, W. C., Zhang, W. E., Li, C. & Sun, A. (eds). Advanced Data Mining and Applications: 13th International Conference, ADMA 2017, Singapore, November 5–6, 2017, Proceedings (Vol. 10604). Springer. (2017). https://doi.org/10.1007/978-3-319-69179-4_27
60. Saab, K. et al. Towards trustworthy seizure onset detection using workflow notes. *npj Digit. Med.***7** (1), 42. https://doi.org/10.1038/s41746-024-01008-9 (2024).
61. Khosravi, M. et al. EEG signal-based machine learning approaches for Alzheimer's disease: a review of methodological analysis, EICEEAI 2023, Jordan, Dec. (2023). https://doi.org/10.1109/EICEEAI60672.2023.10590088
62. Khosravi, M. et al. Dec., A novel EEG-based deep approach for diagnosing Alzheimer's disease using attention-time-aware LSTM, EICEEAI 2023, Jordan, (2023). https://doi.org/10.1109/EICEEAI60672.2023.10590201

Acknowledgements
We sincerely thank the National Elites Foundation.

Author contributions
M.K., H.P., and K.R. have designed the methodology, and provided the tests, and the initial manuscript. H.P., and M.H. have managed and reviewed the research and its results. All authors reviewed the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Ethical approval for animals and human participants
The authors conform that in this research, they have used the publicly available dataset (for free download, refer to: https://figshare.com/articles/dataset/dataset_zip/5450293). Thus, no new consent for ethical approval is needed. Based on the guide provided by the editorial office, this clarification is required.

Additional information
Correspondence and requests for materials should be addressed to M.K. or H.P.