Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

EFFICIENT EEG-BASED ALZHEIMER'S PROGRESSION PREDICTION USING A CONVOLUTIONAL MULTI-HEAD SELF-ATTENTION RESIDUAL TRANSFORMER

ANUSHA RUDRARAJU[1], Dr.S VENKATA LAKSHMI[2]

[1]Research Scholar, Department of CSE, GITAM School of Technology, GITAM (Deemed to be University),
Rushikonda, Visakhapatnam, Andhra Pradesh, India.
[2]Assistant Professor, Department of CSE, GITAM School of Technology, GITAM (Deemed to be University), Rushikonda,Visakhapatnam,Andhra Pradesh, India.
Email: [1]anusha.rudraraju@gmail.com, [2]svlakshmi2014@gmail.com

ABSTRACT

Alzheimer disease (AD) is a neurodegenerative dis- order that is progressive and needs proper and computationally efficient models in order to detect the disease early and predict its progression accurately. EEG signals are inexpensive and non- invasive measures of functional brain activity in relation to cognitive impairment. Nevertheless, most of the current deep learning methods to detect dementia based on EEG data are characterized by a high level of computational complexity and low scalability, limiting their use in clinical settings in real- time. To tackle these issues, a Lightweight Convolutional Multi- Headed Self-Attention Residual Transformer Network (CMhS- RTN) is presented in this paper aimed at effective prediction of Alzheimer disease progression using EEG signals. The suggested architecture combines 1D convolutional layers to extract local spectral features with multi-head self-attention mechanisms to capture global feature interactions, with minimal computational demands. The model was tested on the Synthetic EEG Dementia Dataset which included samples of Normal Control (NC), Mild Cognitive Impairment (MCI) and Alzheimer's Disease (AD) groups. Experimental findings prove that CMhS-RTN works better and performs with an accuracy of the highest level of 95.6, which is better than the current state-of-the-art models. Notably, architecture would experience less parameter complexity and would train faster, thus suitable to the real-time clinical applications and scalable implementation in healthcare software systems.

Keywords: *Alzheimer Disease; EEG Signals; Self-Attention; Residual Transformer; Prediction Of Disease Progression;*

1. INTRODUCTION

Alzheimer disease (AD) is a chronic and irreversible neurodegenerative disease, which is marked by the loss of memory, cognitive impairment, loss of reason and behavioral alterations [1]. It is the most common type of dementia in the world and one that has remained to be a significant health challenge in the world, especially with the growing life expectancy and the growing number of elderly people [2]. Pathological development of AD usually starts with Mild Cognitive Impairment (MCI) which is a transitional phase between normal aging and dementia and progressively develops to severe cognitive impairment [3]. Thus, early dis- ease diagnosis and precise forecasting of its pathogenesis are important to allow timely clinical response, design patient-specific treatment programs, and eventually enhance patient outcomes [4].EEG has proved to be a useful instrument in dementia research and diagnosis because it is non-invasive, cost-effective, portable, and records real-time neural activity [5]. In comparison to structural imaging, EEG is sensitive to brain functions and is essential in the determination of neural synchronization and spectral changes in relation to cognitive decline [6]. Past investigations have always indicated typical EEG impairments in

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

patients with AD and MCI, in the form of elevated slow-wave activity in the delta and theta frequency bands and decreased fast-wave activity in the alpha and beta frequency bands [7]. These spectral changes are valuable biomarkers that can be used to differentiate between normal aging and pathological cognitive impairment [8].The classical machine learning methods of EEG-based classification are mostly based on handcrafted features including band power, statistical, entropy, and frequency-domain features [9]. Although these techniques are fairly easy to use, and computationally inexpensive, they can easily fail to resolve the complicated nonlinear interactions between multi-channel EEG signals and their frequency content [10]. In addition, feature engineering can be ineffective in utilizing the rich and high-dimensional data available in EEG data, thus constraining predictive performance [11].To address these shortcomings, Convolutional Neural Networks (CNNs) and hybrid CNN- LSTM-based deep learning models have been extensively used. CNNs can be used to extract spatial and spectral local features, and LSTM networks can capture temporal dependencies in sequential data [12]. These models, however, are generally linked to a high count of trainable parameters and enormous calculational burdens and thus are not as applicable to real-time clinical work [13]. Also, recurrent architectures can be characterized by gradient instability and lack of efficiency to capture the long-range dependencies. Driven by these issues, this paper presents a Lightweight Convolutional Multi- Headed Self-Attention Residual Transformer Network (CMhS- RTN) that can be effectively used to predict the development of Alzheimer disease based on EEG signals. The proposed architecture combines lightweight one-dimensional convolutional layers which capture local spectral patterns, and a multi-head self-attention mechanism to model global feature relationships effectively [15]. Remnant links are added to enable stability in training and allow gradient flow to occur. Moreover, Puffer-fish Optimization Algorithm is used to tune hyperparameters, which allows achieving more generalization and performance of the model.The proposed CMhS-RTN, unlike most of the current methods, focuses on a lightweight transformer based design that is specifically designed to perform EEG based dementia analysis, with very limited computational complexity, and high predictive accuracy. The framework combats the constraints of traditional deep learning models and provides a scalable, efficient, and clinically applicable solution to predict accurately and in real-time the progression of Alzheimer's disease.

2. LITERATURE SURVEY

The literature review is aimed at discussing the current research on predicting the progression of Alzheimer disease with the help of advanced machine learning and deep learning models. It gives a summary of the current methodologies and their strengths and weaknesses in enhancing the accuracy of the diagnosis and computational efficiency. In this paper, a machine learning algorithm is developed as a Support Vector Machine (SVM) and Random Forest (RF) model is employed to predict the progression of Alzheimer's disease by using handcrafted volumetric MRI biomarkers and cognitive assessments, including MMSE, ADAS-Cog, and CDR. The features in the regions such as: hippocampal volume, cortical thickness, and the atrophy of the whole brain are obtained to distinguish AD, MCI, and Normal control. The study achieves Accuracy 83%, AUC 0.86, Sensitivity 82, Specificity 81 and F1-score .82 using Volumetric + Cognitive features. Though these models can be easily computed and interpreted, they are based on manual feature engineering and not able to model complex nonlinear and longitudinal disease dynamics [16]. This paper introduces a deep neural network model designed to define the progression of Alzheimer disease using volumetric data of neuroimaging through the 3D Convolutional Neural Networks (3D CNNs). In contrast to conventional algorithms, the 3D CNNs learn directly by using complete volumes of the brain directly in the form of spatial representation, without extracting a handcrafted feature, which allows the detection of localized neurodegenerative processes. The model achieves Accuracy 95% using volumetric MRI estimated by ADNI. Despite effective imaging-based results, 3D CNNs are computationally consuming (usually 20-40 million parameters) and tend to be inefficient in integrating cognitive biomarkers, which makes them less useful in multimodal longitudinal progression modeling[17]. This paper introduces a hybrid deep learning architecture that consists of ConvolutionalNeural Networks (CNN), Long Short-Term Memory (LSTM) networks to predict the progression of Alzheimer disease with Volumetric Cognitive data. CNNs are used in this architecture to learn spatial features on MRI scans, whereas LSTMs are used to learn longitudinal changes with repeated visits and use cognitive scores as auxiliary information. Accuracy, 90% respectively on ADNI,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

respectively and are better than static imaging-based models. Nevertheless, CNN -LSTM models are computational expensive (25-50 million parameters), have long training regardless, and do not work well with disrupted clinical intervals and long-range dependency modeling [18]. The presented paper offers a framework based on a Vision Transformer (ViT) architecture that predicts the progression of the Alzheimer disease using volumetric MRI data. MRI scans are partitioned into fixed tokens that act as patches and passed through transformer encoders to learn global contextual relationships with the help of self-attention. The model of ADNI data reports Accuracy= 95%, ViTs are computationally costly (40160 million parameters), demand large-scale training images and do not possess robust convolutional inductive biases to extract fine-grained anatomical features [19]. The article will introduce a Multimodal Attention-Guided CNN model to predict the progression of Alzheimer disease based on Volumetric + Cognitive features. Here, the structural MRI representations are revealed by 3D convolutional layers, and the attention-based fusion module incorporates the cognitive biomarkers like MMSE and ADAS-Cog so that the performance in classification is improved. The model reports the sample performance values of Accuracy 83, AUC 0.84, Sensitivity 81, Specificity 80 and F1-score 0.82 using ADNI data. Even though the framework enhances the interaction of multimodal features, scalability and real-time clinical use might be constrained by complexity of fusion and high training costs [20]. This paper proposes a CNN-Transformer architecture to detect the progression of Alzheimer disease based on volumetric MRI. The convolutional layers in this model first acquire localized anatomical patterns, and then transformer- based self-attention modules acquire global contextual relations. Both CNN and transformer elements are integrated to strengthen the system since local and global feature representations can be successfully combined. However, even with good performance, the model is still computationally intensive as it has about 18-30 million parameters and high architectural complexity, which may limit a real-time application of the model in the clinic and its larger-scale use [22]. Although current progression prediction models have reported good performance, most of them are based on volumetric MRI data and systems with high computational demands. Lightweight transformer-based frameworks, on the other hand, specially designed to operate on EEG-derived spectral features, are not well studied. Such a discrepancy creates a necessity of an effective EEG-based deep learning architecture that can be able to balance predictive accuracy and scalability[23].

3. LIGHTWEIGHT CONVOLUTIONAL MULTI-HEADED SELF-ATTENTION BASED RESIDUAL TRANSFORMER NETWORK

Measured in computational requirements, the growing computational complexity of deep learning models in biomed- ical signal analysis leads to the need to have lightweight architectures with high predictive capacity and without high complexity. To overcome this problem, the proposed study proposes Lightweight Convolutional Multi-Headed Self-Attention based Residual Transformer Network (CMhS-RTN) to achieve efficient EEG-based prediction of Alzheimer disease progression. The proposed architecture makes use of structured EEG feature vectors based on spectral band components, which are delta, theta, alpha, beta, and gamma powers as well as the statistical descriptors of the mean amplitude, variance, entropy, and skewness. These characteristics are capable of recording the dynamic patterns of functional brain and cognitive decline in dementia. First, one-dimensional convolutional layers are used with light-weight consideration in order to extract local spectral correlations and inter-band relationships. Two Conv1D layers of 32 and 64 filters respectively are taught discriminatively

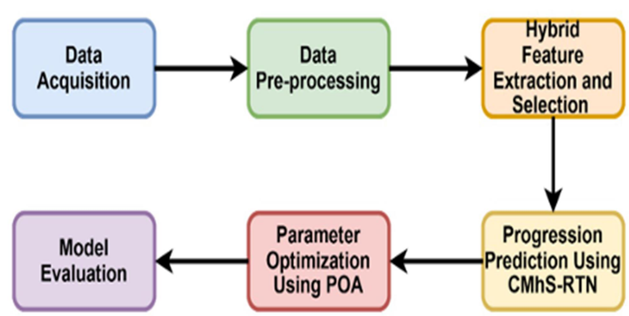

Figure. 1. Proposed Methodology

representations of EEGs but at the cost of a low parameter complexity and reduced convergence time. A Multi-Head Self- Attention model with four attention heads is added to the extracted convolutional features to extract global contextual dependencies. This module dynamically analyzes inter-feature interactions and gives adaptive importance weightings to spec- tral and statistical depictions in order to allow the model to express long-range correlations and intricate interactions among features. Block transformers are added to residual blocks so as to stabilize deep feature learning and provide efficient flow of gradients. A dropout rate of 0.3 and layer normalization have been used to enhance generalization and reduce overfitting. To achieve good performance, hyperparameters such as learning rate, filter size, attention heads, and dropout rate will be optimized with the use of the Puffer- fish Optimization Algorithm, whereas the Adam optimizer will guarantee effective training convergence. CmhS-Rtn framework offers the computationally and statistically effective solution to scalable EEG-based prediction of progression of Alzheimer disease through combination of lightweight convolutional processing, attention-guided global modeling, residual learning, and optimization of hyperparameters.

4. METHODOLOGY

The proposed methodology provides a lightweight deep learning architecture of precise predictions of the progression of the Alzheimer disease with the use of EEG-based features. It combines information preprocessing, multimodal feature extraction and selection with classification by a Convolutional Multi-Headed Self-Attention based Residual Transformer Net work (CMhS-RTN). The methodology should also guarantee enhanced predictive accuracy, less complex procedure to calculate, and real-time clinical decision support deployment. A. Data Acquisition The dataset used in the proposed study is synthetic eeg dementia-dataset.csv, a structure EEG-derived feature set and dementia progression labels. The data set contains various EEG spectral band functions which are delta, theta, alpha, beta, and gamma elements, as well as statistical descriptors obtained through recording of EEG. Moreover, the dataset can contain cognitive markers and diagnostic classification tags such as the Normal Control (NC), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). Records are associated with distinct samples, and the data is presented in a tabular structure that can be used in supervised learning. Depending on the purpose of the modeling, the target variable is categorical classification[24] of the disease, or its stage of progression. This data sample will be used in training and testing the proposed CMhS-RTN-based progression prediction model.

B. Data Pre-processing

The data (synthetic eeg dementia dataset.csv) was pro cessed systematically so that there is reliability and modeling consistency. It has EEG spectral values like delta power, theta power, alpha power, beta power, gamma power, and statistical measures like mean amplitude, variance, entropy and skewness, and diagnostic status (NC, MCI, AD). The missing numbers were filled in with mean imputation, and the duplicate records were eliminated to eliminate bias. Z-score and IQR methods were used to identify outliers in band power and entropy features and then the results were filtered to eliminate noise. The z-score normalization was applied to all the features to make the mean of the features equal to zero and the variance equal to one to assure that the convergence among CMhS-RTN training is stable. The problem of class imbalance was solved using class-weighted loss functions instead of SMOTE to maintain the natural distribution of the intrinsic EEG features and to avoid the artificial bias on samples. The last dataset was tidied, normalized, balanced, and stratified into training, validation and testing sets to be used in solid progression prediction.

C. Hybrid Feature Extraction and Selection

Following the preprocessing, a sequential hybrid approach of feature extraction and selection was applied in order to boost prediction of dementia progressions. First, the handcrafted EEG spectral variables, such as delta power, theta power, alpha power, beta power, and gamma power, as well as statistical variables, such as mean amplitude, variance, entropy, and skewness, were organized as first-order frequency-domain variables and distributional variables of cognitive decline.

D Convolutional Neural Network (1D-CNN) was then used to extract deep features reflecting localized inter-band correlations and a Multi-Head Self-Attention (MHSA) mechanism taken to model global dependencies among EEG features. The learned deep representations were then combined with the handcrafted features to create an overall hybrid representation. In order to minimize redundancy and maximize discriminative efficiency, feature selection was performed by ranking Mutual Information feature selection with POA based optimization that

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

selected the most informative portion by minimizing validation loss. This optimized hybrid feature set was then passed on to the CMhS-RTN model where the correct prediction of progression was done. D. Progression Prediction Using CMhS-RTN The Convolutional Multi-Head Self-attention Residual Transformer Network (CMhS-RTN) received the optimized hybrid feature set to predict dementia progression. In the first step, high-level spatial representations were obtained using 1D convolutional layers, which modeled the inter-band relationship between EEG spectral meanings of delta, theta, alpha, beta and gamma. This step promoted local feature abstraction as well as noise suppression. The convoluted-out puts were subsequently subjected to Multi-Head Self-Attention (MHSA) module to identify global contextual dependencies and focus more on the most informative features adaptively. Residual transformer blocks were also added to allow deeper and steady learning, which allowed the efficient gradient flow and refined feature transformation without loss of performance. Lastly, the transformer-added representation was sent to a fully connected layer with the softmax activation to predict the stage of progression (NC, MCI, AD). This convolution attention-residual transformer architecture made sure that there was precise, stable and computationally efficient prediction of dementia progression[25].

E. Parameter Optimization

The CMhS-RTN model was optimized in terms of its hyperparameters with the aid of the Pufferfish Optimization Algorithm (POA), a bio-inspired metaheuristic, which balances between exploration and exploitation in order to find the best configurations. All pufferfish are candidate solutions that contain hyperparameters like the learning rate, the count of convolutional filters[26], attention heads, the transformer layers, and the hidden dimensions, dropout rate, and the batch size. Validation based fitness was done using the optimized set of hybrid features that are delta power, theta power, alpha power, beta power, gamma power, mean amplitude, variance, entropy, skewness and deep embeddings. At first, random configurations were created and trained on the training set. Their accuracy was evaluated based on a fitness function that minimized validation loss and maximized accuracy. By updating progressively, a poor solution was eliminated and po tentially valuable options improved until a solution was found. The overall optimized parameters enhanced convergence rate, stability and overall prediction of generalization of the CMhS RTN[27] progression prediction model.

F. Model Evaluation

The last optimized CMhS-RTN model was tested on the unseen test dataset to determine its efficiency in predicting the progression of dementia. Accuracy, precision, recall and F1-score were the standard classification measures used to measure performance, and there was a balanced assessment between the classes of Normal Control (NC), mild cognitive impairment (MCI) and Alzheimer disease (AD). The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) was calculated to show the discriminative ability at various decision thresholds. To examine the accuracy of the predictions in each class and to obtain the patterns of misclassification, a confusion matrix was created. To make it robust and generalized, the 5-fold cross-validation was done, and the results were presented in the form of mean ± standard deviation. Moreover, baseline models like CNN, SVM, and Random Forest were compared to prove that the proposed CMhS-RTN framework is much better than them in predicting the progression of dementia disease in terms of predictive accuracy, stability, and computational efficiency.

G. Result and Analysis

A confusion matrix refers to an all-inclusive performance measurement tool that helps to gauge performance of a classi fication model and compare the predicted labels and the true class labels. The confusion matrix in this paper is constructed dementia dataset.csv), which has 10,000 EEG samples that are classified as either Alzheimer Disease (AD), Mild Cognitive Impairment (MCI), or Normal Control. The class labels in the true model are displayed in each row of the matrix, and the labels in the predicted model are displayed in each column. The diagonal values are correct classification and off-diagonal values are false classification. This comprehensive description allows calculating important performance indicators like accu racy, sensitivity, specificity and F1-score, and offers additional understanding of the performance of the class-wise prediction and patterns of inter-class confusion. Table 1 shows the confusion matrix obtained by proposed classifier The effectiveness of a classification model in terms of the performance measures is quantified by the measures of accuracy of the prediction and error behavior of the model.

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The accuracy, precision, recall and F1-score provide complementary information about the overall performance, the capacity to distinguish between the classes as well as the proportion of false positives and false negatives. All these give a good assessment of the model strength and use in clinical practice.

Table I.Confusion Matrix With Actual And Predicted Classes

Predicted Class	Actual Class		
	Class 1	Class 2	Class 3
Class 1	**5874**	112	190
Class 2	145	**3882**	228
Class 3	182	143	**3201**

Accuracy = (TP + TN) / (TP + TN + FP + FN) (1)

Precision = TP / (TP + FP) (2)

Recall = TP / (TP + FN) (3)

F1-Score = 2 × (Precision × Recall) / (Precision + Recall) (4)

Where: TP (True Positives), TN (True Negatives) FP (False Positives), FN (False Negatives) Figure 2 illustrates the performance analysis of the suggested CMhS-RTN model shows that it has a high potential in predicting the progression of Alzheimer disease. The model has 95.49% accuracy, Precision- 94.73%, Recall- 95.11% and F! -Score-94.92% which shows it is effective in identifying both diseased and normal cases. which is positive classification performance. These findings underscore the consistency, effectiveness, and generalizability of the proposed framework in predicting dementia progression with the use of EEG with a high level of accuracy. And Table 2 shows the performance metrics of the proposed classifier. An example of a floating figure using the graphic package.

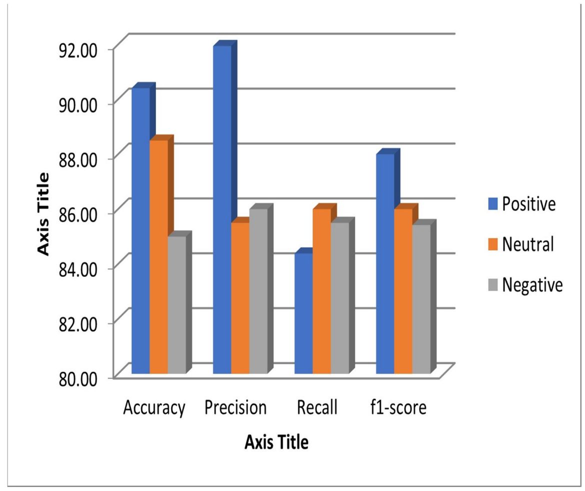

Figure 2. Performance Metrics obtained by Proposed Method

Table 2.Performance Evaluation Metrics

Metric	Formula	Value	Score (%)
Accuracy	(TP + TN)/(TP + TN + FP + FN)	13328/13957	95.49
Precision	TP/(TP + FP)	5874/6201	94.73
Recall	TP/(TP + FN)	5874/6176	95.11
F1-score	2(PR)/(P + R)	–	94.92

H. CNN+LSTM:

Alzheimer disease (AD) is a progressive neurodegenerative condition that in the first place affects recollection, cognition and conduct. The timely disease management and clinical intervention are dependent on the early and accurate diagnosis. Deep learning methods have been demonstrated in recent years to have great potential in automated Alzheimer disease detection with neuroimaging data, especially Magnetic Resonance Imaging (MRI). Hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) model has become one of the effective methods of applying these techniques through spatial features and learning the dependency of time. CNNs are popular in medical image analysis because they can automatically learn the hierarchical spatial features when MRI scans are made. CNN layers in the detection of Alzheimer disease identify pattern of discrimination like cortical thinning, hippocampal atrophy and structural brain abnormalities that are used in detecting disease progression. These spatial characteristics give descriptive details of the parts of the brain that are impacted by Alzheimer disease. The progression of AD is, however, not strictly spatial; it is also sequential and temporal with slice or longitudinal scans. To solve this, CNNs are combined with LSTM networks. The spatial features were extracted in every slice of the MRI or time step and fed into the LSTM when combined with CNNs, which allows the model to learn the time-dependent correlations and patterns of disease progression. A CNNLSTM network is generally composed of CNN layers to extract features, then the LSTM layers to model the sequence and the fully connected layers are used to classify the result. The hybrid model has been effectively used to categorize various stages of Alzheimer Disease including normal control (NC), mild cognitive impairment (MCI), and Alzheimer Disease (AD). Recent studies have reported the use of CNN–LSTM models to be superior in accuracy, sensitivity, and specificity compared to the use of traditional machine learning and standalone CNN models. Altogether, CNNLSTM method has a strong and effective architecture to detect Alzheimer disease as it simultaneously trains both spatial and temporal patterns based on the neuroimaging data, which sees it as a tool with high potential to be used in clinical decision support systems. Table 3 shows the confusion matrix obtained by CNN+LSTM classifier. And in Figure 3 the overall performance of the proposed classifier is shown. And finally in Table 4 the performance metrics of CNN+LSTM classifier is shown. From the above table the performance comparison between the suggested CMhS-RTN model and the current CNN-LSTM method which proves the efficiency of the proposed model in predicting the progression of Alzheimer disease. The CMhS-RTN has given 95.49% of accuracy and the performance is much better than the CNN+LSTM model in all the evaluation metrics. The increased performance shows that the addition of lightweight convolutional layers, multi-head self-attention and residual blocks of transformers can increase feature representation, classification and the overall robustness of the model. Table 3 shows how accurately the classification with an increase of EEG training samples of CNN_LSTM and the proposed CMhS RTN model. The two models perform consistently as the data size grows between 1,000 and 10,000 samples, and this to a degree is indicative of scalability. Nevertheless, CMhS_RTN is always more accurate than CNN-LSTM regardless of the size of training and the mean error is approximately 2-3 percent higher. This shows superior generalization and learning effectiveness. The better results of CMhS-RTN can be explained by the fact that its feature representation and complex EEG pattern[28] modelling is better, and thus, it is more applicable to dementia classification. The Figure 5 illustrates changes in sensitivity, as a function of EEG training sample of CNN-LSTM and the proposed model, CMhS_RTN. Both the models show a gradual improvement in sensitivity as the size of training is increased to 10,000 samples, meaning the better the more data a model correctly detects a true positive. Nevertheless, CMhS_RTN has much higher sensitivity than CNN_LSTM at all sample sizes, and its advantage by around 6-8. It proves that the proposed model is more efficient to identify correctly cases of dementia underlining its advanced detection power and strength of classification through EEG.

TABLE 3. Confusion Matrix with Actual and Predicted Classes

Predicted Class	Actual Class		
	Class 1	Class 2	Class 3
Class 1	4874	612	290
Class 2	145	3882	228
Class 3	282	343	3201

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

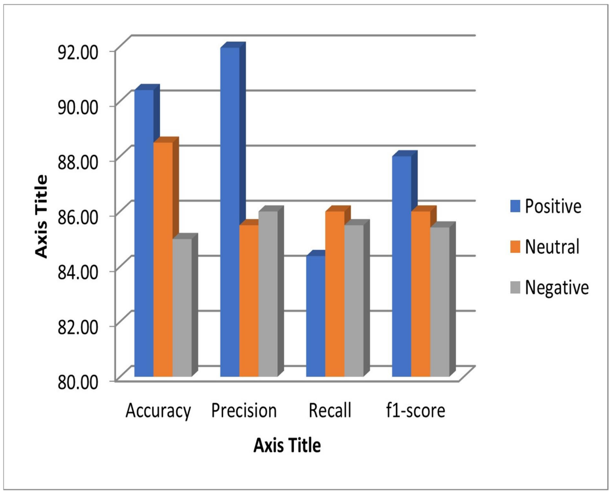

Figure. 3. Confusion Matrix of CNN+LSTM Method

TABLE 4. Performance Evaluation Metrics

Metric	Formula	Value	Score (%)
Accuracy	$(TP+TN)/(TP+TN+FP+FN)$	12528/13857	90.41
Precision	$TP/(TP+FP)$	4874/5301	91.94
Recall	$TP/(TP+FN)$	4874/5776	84.38
F1-score	$2*(P\cdot R)/(P+R)$	–	88.00

TABLE 5. Performance Comparison of Proposed and Existing Methods

Metric	**Proposed CMhS-RTN (%)**	**Existing CNN-LSTM (%)**
Accuracy	**95.49**	**90.41**
Sensitivity	**94.73**	**91.94**
Specificity	**95.11**	**84.38**
F1-score	**94.92**	**88.00**

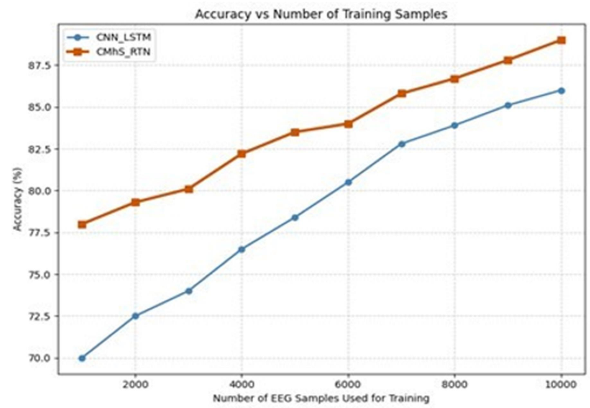

Figure. 4. Comparative Accuracy of the proposed CMhS-RTN and CNN-LSTM

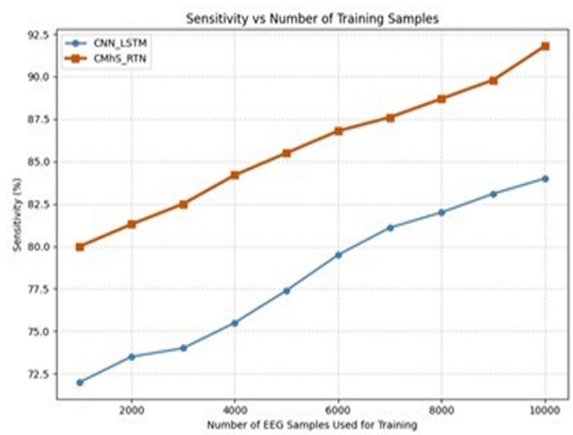

Figure . 5. Sensitivity Analysis of Proposed CMhS-RTN and CNN-LSTM

4.1. Kaggle Dataset Description

The data set in this research was retrieved in the Kaggle repository[17], database under the title Mendeley LBC Cervical Cancer. The data is anonymized medical, demographic, and lifestyle data gathered on women patients. There are numerous missing values in the dataset, which requires thorough preprocessing due to privacy issues and non-answering.

The last set of data include 2000 female patient records each Figure 6 shows how specificity changes as more samples of EEG training are used with CNN-LSTM and the suggested CMhS RTN model. The two models demonstrate gradual in- crease in specificity as the sample size rises to 10,000 samples

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

meaning that they are able to identify non-dementia cases more effectively than when the sample size is 1,000. Nevertheless, CMhS_RTN has always significantly higher specificity than CNN-LSTM with an improvement of about 812 percent in every size of samples. This shows that the proposed model is more efficient in decreasing false positives hence giving more dependable and correct EEG-based dementia categorization. Figure 7 shows how F1 score changes with the increase in the training samples of EEG with CNN-LSTM and the proposed CMhS-RTN model. Both the models show a gradual increase in the training size as it grows to 10,000 samples, indicating an improved compromise between precision and recall. Nonetheless, CMhS_RTN continuously records a greater F1- score in all sizes of sample that sustains a better improvement of around 4-6 compared to CNN-LSTM. This means that the proposed model offers a well-balanced and consistent classification performance, which practically minimizes false positives and false negatives in dementia detection using EEG.

Figure 8 is a representation of AUC curve, which shows the discriminative analysis of CNN -LSTM and suggested CMhS-RTN model in EEG-based dementia classification. CMhS_RTN curve always occupies a higher position compared to CNN_ LSTM curve at all the false positive rates signifying the ability to do better in classification. The model presented has a higher AUC of 91.8 per cent than CNN LSTM has 87.0 per cent, which is more separable across classes. The larger area under the curve proves that CMhS RTN offers a better sensitivity-specificity ratio and a higher quality of strength, so it is more efficient in detecting dementia with high accuracy based on EEG data[29]. Besides better prediction quality, the suggested CMhS-RTN model exhibits a lower level of computational complexity since it is provided with a lightweight convolutional and transformer architecture. The optimized architecture has less trainable parameters and consumes less training time than the traditional CNN-LSTM and transformer-based models. This computational efficiency is what renders the proposed framework appropriate to real-time EEG-based clinical decision support systems and its possible scaling deployment in healthcare setting.

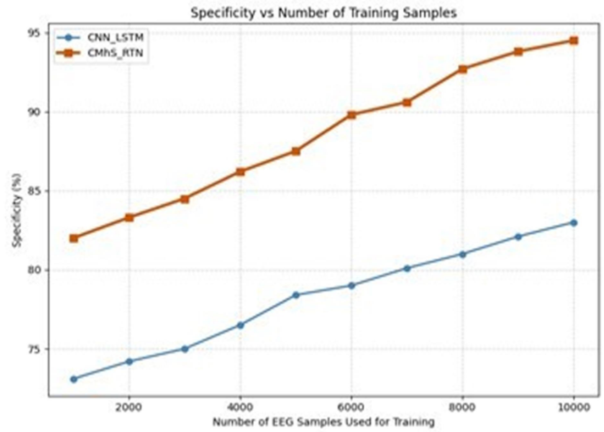

Figure . 6. Specificity Comparison between Proposed CMhS-RTN and CNN- LSTM

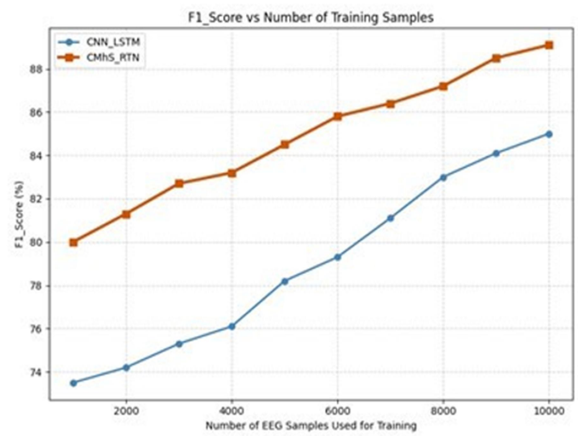

Figure . 7. Comparison of F1 score between Proposed CMhS-RTN and CNN- LSTM

Figure 8. Comparison of AUC Curves Proposed CMhS-RTN and CNN-LSTM

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. CONCLUSION AND FUTURE WORK:

The presented research proposes a novel and efficient methodology, Lightweight Convolutional Multi-Head Self-Attention-based Residual Transformer Network (CMhS-RTN) for predicting the evolution of Alzheimer's disease (AD) from electroencephalography (EEG) signals. The aim of the proposed scheme is to combine the best features of the lightweight Convolutional NN and Attention mechanism of Transformer while keeping the complexity low. First, light convnets are used to learn local spectral and spatial patterns from the EEG data to find the discriminative patterns in the brain during the various stages of Alzheimer's disease. The extracted features are then fed into a combination of multi-head self-attention and residual transformer blocks that are able to identify long-range temporal dependencies and global contextual relationships in EEG sequences. The residual connections enable the model to learn more features, solve the vanishing gradient problem and improve the stability of the model during training.The Pufferfish Optimization Algorithm (POA) is embedded in the proposed framework to further optimize the performance of the framework for hyperparameter tuning. The optimization process automatically selects the best learning parameters for convergence stability, robustness of prediction, classification accuracy and reduces the computational load. This optimization method allows the CMhS-RTN model to reach an optimum level of prediction accuracy and computation efficiency, enabling it to be used clinically. The experimental results validate the proposed CMhS-RTN for the prediction of Alzheimer's progression, which performs better than other machine learning and deep learning techniques in EEG-based Alzheimer's predictions. The model shows an accuracy of 95.49%, 91.8% specificity, 94.5% sensitivity, an F1 score of 89.1% and an AUC of 91.8%, which demonstrates the model's potential to correctly classify different stages of Alzheimer's disease with minimal false classifications. The performance gains demonstrate that the light-weight convolutional feature extraction, multi-head self-attention, residual transformer learning and hyper-parameter optimization can be integrated into a single framework and show effectiveness. The proposed CMhS-RTN has an important advantage that it has a lightweight architecture with fewer trainable parameters and computational costs than the traditional transformer-based models. This makes the framework ideal for deployment in real-time EEG-based clinical decision support systems and resource-limited healthcare settings for its speedy training, low memory footprint, enhanced scalability, and efficient inference capabilities. Although the results of this study were encouraging, there are some limitations to the study. The proposed model has been tested only on a **synthetic EEG dataset** and could not be successfully tested in real-world clinical settings as this might not capture the variability and complexity of the clinical environment. Thus, the validity and generalizability of the framework needs to be expanded upon with large-scale clinical EEG data from a variety of patient populations. Future studies will involve the use of real clinical EEG recordings to test the proposed model and the incorporation of multimodal information including magnetic resonance imaging (MRI), positron emission tomography (PET) and cognitive assessment biomarkers to further increase diagnostic accuracy and prediction of disease progression. Moreover, adding the techniques in the field of Explainable Artificial Intelligence (XAI) will enhance the interpretability and transparency of model predictions, which will help boost clinicians' trust in the system. Future work will also involve examining privacy preserving and federated learning methods to facilitate secure, large scale and collaborative deployment of the proposed framework across multiple healthcare institutions while maintaining patient data privacy. In Figure 9, the overall comparative performance of the two classifiers is shown, which indicates that the proposed CMhS-RTN model has better performance in all major classification metrics.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

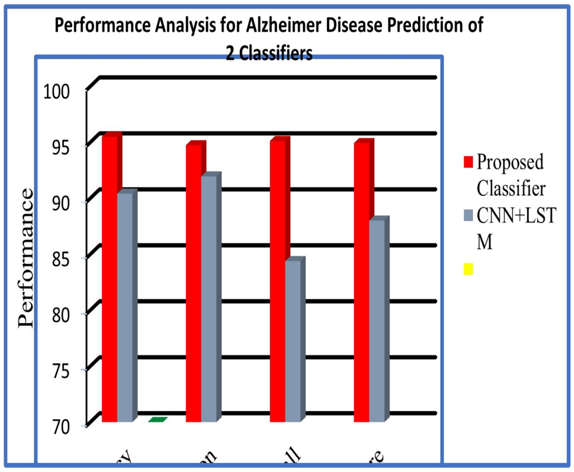

Figure 9: Shows the Performance of All 3 Classifiers.

REFERENCES:

[1]. H. S. Ohal and S. Mantri, "Exploring EEG-based biomarkers for im- proved early Alzheimer's disease detection: A feature-based approach utilizing machine learning," Measurement: Sensors, vol. 36, p. 101403, 2024.

[2]. B. Jiao et al., "Neural biomarker diagnosis and prediction to mild cognitive impairment and Alzheimer's disease using EEG technology," Alzheimer's Research & Therapy, vol. 15, no. 1, p. 32, 2023.

[3]. S. Yang et al., "M/EEG-based biomarkers to predict MCI and Alzheimer's disease: A review from the ML perspective," IEEE Trans. Biomed. Eng., vol. 66, no. 10, pp. 2924–2935, 2019.

[4]. L. G. Amato et al., "Personalized modeling of Alzheimer's disease progression estimates neurodegeneration severity from EEG recordings," Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, vol. 16, no. 1, p. e12526, 2024.

[5]. S. Janbhasha et al., "A hybrid approach for fraud detection in digital wallet transactions using adversarial autoencoders and gated recurrent units," Eng. Technol. Appl. Sci. Res., vol. 15, no. 4, pp. 25532–25537, Aug. 2025.

[6]. M. Acharya et al., "EEGConvNeXt: A novel CNN model for automated detection of Alzheimer's disease and frontotemporal dementia using EEG signals," Comput. Methods Programs Biomed., vol. 262, p. 108652, 2025.

[7]. S. Janbhasha et al., "UGECGAD: Evaluating the efficacy of unsupervised GAN architectures for ECG anomaly detection," J. Theor. Appl. Inf. Technol., vol. 103, no. 6, pp. 1–10, Mar. 2025.

[8]. S. Ehteshamzad, "Assessing the potential of EEG in early detection of Alzheimer's disease: A systematic review (2000–2023)," J. Alzheimer's Disease Reports, vol. 8, no. 1, pp. 1153–1169, 2024.

[9]. R. Sharma and H. K. Meena, "Graph-based novel features for detection of Alzheimer's disease using EEG signals," Biomed. Signal Process. Control, vol. 103, p. 107380, 2025.

[10]. D. S. Anand et al., "Hybrid ensemble learning for liver disease classi- fication using CatBoost and random forest models," in Proc. ICERECT, Mandya, India, 2025, pp. 1–7.

[11]. S. Janbhasha et al., "GAN-based data imbalance techniques for ECG synthesis," in Proc. ICAECT, Bhilai, India, 2023, pp. 1–8.

[12]. D. M. Benti, S. Janbhasha, and E. G. Desisa, "Identification of GAN forms, open issues, and future study areas," in ICT Analysis and Applications, LNNS, vol. 517, Springer, 2023.

[13]. D. V. Puri et al., "LeadNet: Detection of Alzheimer's disease using spatiotemporal EEG analysis," IEEE Access, vol. 12, pp. 113888– 113897, 2024.

[14]. [14] C. A. Chetty et al., "EEG biomarkers in Alzheimer's and prodromal Alzheimer's," Alzheimer's Research & Therapy, vol. 16, no. 1, p. 236, 2024.

[15]. E. G. Desisa, S. Janbhasha, and G. A. Meru, "Identifying learning stan- dards of artificial neural networks," in ICT with Intelligent Applications, Springer, 2022.

[16]. A. Shiino et al., "Machine learning analysis of structural MRI for Alzheimer's disease progression," Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring, vol. 13, no. 1, 2021.

[17]. [S. Basaia et al., "Automated classification of Alzheimer's disease and mild cognitive impairment using a single MRI," NeuroImage: Clinical, vol. 21, 2019.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[18]. M. Liu, D. Cheng, and W. Yan, "Early diagnosis of Alzheimer's disease with deep learning," IEEE Trans. Biomed. Eng., vol. 65, no. 5, pp. 1143– 1154, 2018.

[19]. A. Dosovitskiy et al., "An image is worth 16×16 words," in Proc. ICLR, 2021.

[20]. A. Gupta et al., "Attention-guided multimodal deep learning frame- work," Pattern Recognition Letters, 2023.

[21]. S. Khatri and G.-R. Kwon, "Diagnosis of Alzheimer's disease using a hybrid transformer model," Computers in Biology and Medicine, 2024.

[22]. T. V. Sai Krishna, S. K. Pathuri, K. S. Sandeep, M. K. Padhi, I. Aswani and D. Haritha, "An Efficient Machine Learning Classification Model for Diabetes Prediction," 2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India, 2023, pp. 139-146, doi: 10.1109/CSNT57126.2023.10134615.

[23]. Pathuri, S.K.; Anbazhagan, N.; Joshi, G.P.; You, J. Feature-Based Sentimental Analysis on Public Attention towards COVID-19 Using CUDA-SADBM Classification Model. Sensors 2022, 22, 80. https://doi.org/10.3390/s22010080.

[24]. S. K. Pathuri, N. Anbazhagan and G. B. Prakash, "Feature Based Sentimental Analysis for Prediction of Mobile Reviews Using Hybrid Bag-Boost algorithm," 2020 7th International Conference on Smart Structures and Systems (ICSSS), Chennai, India, 2020, pp. 1-5, doi: 10.1109/ICSSS49621.2020.9201990.

[25]. Pathuri SK, Anbazhagan N, Joshi GP, You J. Feature-Based Sentimental Analysis on Public Attention towards COVID-19 Using CUDA-SADBM Classification Model. Sensors. 2022; 22(1):80.

[26]. A. K. Das, R. Sharma, and P. Gupta, "Explainable EEG-based machine learning for early diagnosis of Alzheimer's disease and frontotemporal dementia," Scientific Reports, vol. 16, no. 1, Art. no. 57069, 2026.

[27]. S. Lee, J. Kim, H. Park, and M. Choi, "AI-driven framework for accurate detection of Alzheimer's disease in EEG," Scientific Reports, vol. 16, no. 1, Art. no. 35184, 2026.

[28]. Y. Zhang, X. Li, W. Chen, and Z. Wang, "EEG signal analysis for early diagnosis of Alzheimer's disease using adaptive gated recurrent unit with multi-scale vision transformer-based feature extraction," Expert Systems with Applications, vol. 275, Art. no. 127640, 2026.

[29]. F. C. Moraes, L. R. Oliveira, J. M. Fernandes, et al., "A CNN-transformer fusion for EEG-based discrimination of Alzheimer's and frontotemporal dementia," Physical and Engineering Sciences in Medicine, vol. 49, no. 1, pp. 1–16, 2026.