SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I1P107

Volume 13 Issue 1, 74-88, January 2026
© 2026 Seventh Sense Research Group®

Original Article

An Innovative Frontier in Healthcare: Optimizing Alcoholism Detection with an Assorted Convolutional Neural Network and Long Short-Term Memory System

G. Usha[1], K. Narasimhan[2]

[1]Department of Electronics and Communication Engineering, Srinivasa Ramanujan Centre, SASTRA University, Kumbakonam, Tamil Nadu, India.
[2]School of Electrical and Electronics Engineering, SASTRA University, Thanjavur, India.

[2]Corresponding Author : knr@ece.sastra.edu

Received: 05 November 2025 Revised: 07 December 2025 Accepted: 06 January 2026 Published: 14 January 2026

Abstract *- Alcohol is considered an intense-affective agent in brain functions that can cause abrupt health issues. Hence, the predominant approach for detecting alcohol consumption is to perform an alcohol diagnosis. Although clinical applications for determining these causes have widely evolved, classical processes have several drawbacks for generating favorable facilities. Clinicians have increasingly developed methods for determining consumption on a technological basis in recent years. Basically, EEG tools aid in exhibiting brain activity through EEG signals. Currently, instigated technology is relatively dependent on ML techniques; however, it has major defects, such as poor spatial resolution evaluation and high computational requirements for precise outcomes. Therefore, the proposed model utilizes a DL approach in which both Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) are used to analyze spatial data to further define model enhancement. The proposed research utilized ten types of pretrained neural network models for image classification. For comprehending the data samples, the proposed system used the EEG-Alcohol dataset to evaluate model performance and efficiency on classification, and was examined by the accuracy metric. In essence, comparative analysis is conducted through the respective models and their comparison with existing research that applied DL-based techniques, unifying LSTM to CNN custom, which led to prominence in the classification of alcoholic and nonalcoholic EEG signals, where the highest accuracy attained by the proposed model is VGG19+LSTM, with an accuracy rate of 96.72%. Furthermore, the model intends to contribute to therapeutic services while pivoting as an impactful intervention in patient care.*

Keywords *- Deep Learning, CNN, LSTM, EEG Signals, Alcoholism.*

1. Introduction

Alcoholism is a harmful binary disease worldwide, and its causes include both physical and mental illness [1]. Predicting alcoholism in an early stage could diminish the mental illness [2]. Global statistics indicate that alcoholic cases account for more than 25% of psychiatric disorders, with excessive alcohol consumption accounting for 7.1% of male and 2.2% of female deaths globally [3, 4].

Chronic alcohol abuse leads to numerous impairments, especially neurotoxic effects on the brain, resulting in long-term mental disorders such as dementia, brain syndromes, and panic disorders. Therefore, early detection of alcoholism is crucial for reducing the risk of mental illness and improving patient outcomes [5]. Traditionally, clinicians have relied on Electroencephalogram (EEG) signals to assess brain activity and diagnose related disorders. However, EEG-based methods are often hampered by inaccuracies and are both time- and resource-intensive [6, 7]. The test results have been represented in EEG records [8, 9]. Primarily, EEG signals detect negative effects on actual brain functions; although they predict alcoholism predispositions, the EEG method still recognizes added spikes such as inaccuracies, is time-intensive, and is resource-intensive [10]. To resolve these limitations, many studies have focused on Artificial Intelligence (AI) technologies for generating effective treatment plans in EEG facilities [11]. Recent advancements in various fields have been relatively enhanced by machine learning [12] and Deep Learning (DL) techniques.

Accordingly, DL algorithms are capable of human brain resemblance, which is improbable in ML methods since an extensive feature of DL is determined from its neural network models [13]. Several researchers [14, 15] have contributed to the classification of alcoholic and nonalcoholic alcohol consumption through the use of ML and DL models. It was found that there is a significant difference in quantifying its accuracy rate, high spatial resolution, and low cost. For instance,

This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)

existing research [16] utilized neural networks and ML methods of regression, Support Vector Machines (SVMs), Decision Trees (DT), Random Forest (RF), and K-Nearest Neighbor (KNN). The system was tested using 38 demographic and sample variables. Among the utilised models, the RF model obtained an ROC of 78% with 15 variables. Another study supported this relative case, which employed Principal Component Analysis (PCA) for feature extraction and LSTM for Classification and achieved a feasible outcome [17]. While these advancements are promising, a critical limitation persists: existing DL models either predominantly analyze EEG signals by extracting either spatial features or temporal features, but they fail to integrate these two dimensions simultaneously, which leads to insufficient accuracy. In addition, the conversion of EEG to a spectrogram builds rich features; however, existing studies have not applied it in their architecture for better representation learning.

To address such constraints and enhance the screening process for alcohol detection, the proposed work employs a hybrid Convolutional Neural Network (CNN) with Long Short-Term Memory (LSTM) architecture. In this approach, ten pre-trained CNNs are utilized to extract spatial features from the EEG spectrogram image. The input EEG signals are preprocessed using noise filtering, and STFT is used to convert the EEG signal into a time-frequency-based spectrogram image representation. Then, the preprocessed image is fed into the CNN for spectral analysis, followed by an LSTM for sequential pattern recognition. By leveraging spatial-temporal feature hierarchies, the model enables automated detection of alcoholism. This dual-modality approach achieves high accuracy by capturing both localized brain activity and long-term neurological trends. The model optimizes efficiency through regularization and data augmentation, ensuring clinical applicability.

The foremost contribution of the proposed model is examined in sequential cases:

- To propose a hybrid CNN–LSTM framework that captures both spatial and temporal characteristics from the EEG spectrum data. The proposed framework achieves better robustness and accuracy over the traditional ML and standalone DL approaches for the classification task.
- To evaluate ten different pre-trained neural network models for image classification and unify them with LSTM, the research identifies the VGG19+LSTM configuration as the top performer, achieving a high classification accuracy of 96.72% for distinguishing alcoholic from nonalcoholic EEG signals.
- To validate the clinical acceptability of the model, performance assessment is performed through various key parameters.

This paper is arranged based on the effective techniques involved in the classification of alcohol use. In addition, techniques applied to the existing research undertaken on the analysis of similar fields are reviewed in Section II. Section III describes the methodologies used to carry out the proposed research. Consequently, the findings contributed by the respective models are demonstrated in Section IV. Finally, the conclusions and future implications of the proposed model are specified in Section V.

2. Review of Literature

Alcohol Use Disorder (AUD) is one of the most common neuropsychiatric disorders affecting the brain. Electroencephalography (EEG) is fast becoming one of the most widely used neurophysiological modalities for the diagnosis of AUD. Many researchers have implemented ML and DL approaches to automatic detection of alcoholism from EEG signals. This section provides a comprehensive review of existing approaches, synthesizing their contributions and critically analyzing their limitations.

2.1. Traditional and Classical ML Models in Alcoholism Prediction

Traditional alcoholism detection methods and analysis of EEG signals to evaluate brain activity involve monitoring specific brainwave patterns and other physiological indicators. It was highly time-consuming and required extensive clinical resources [18]. To mitigate this, initially, classical ML models such as SVM, RF, KNN, MLP (multilayer perceptron), and Bayesian classifiers were utilized to automate the detection process. For instance, in [19], an RF-based ML classifier was employed for alcoholic classification through EEG samples. The model has yielded an average accuracy of 69%. In addition to this case, another researcher introduced a method that utilizes absolute gamma band power as a feature and employs an ensemble subspace K-NN classifier to distinguish between alcoholics and normal subjects. An Improved Binary Gravitational Search Algorithm (IBGSA) selects important EEG channels, achieving a detection accuracy of 92.50% [20]. Though ML models achieved good performance, these techniques require a manual feature engineering process and fail to capture the complex, non-linear patterns inherent in EEG data [21].

2.2. Deep Learning Models For Alcoholism Prediction

The advent of DL approaches has ensured automatic feature extraction, which can effectively learn the complex patterns from large-scale data. Correspondingly, the existing system has been applied to neural networks for classifying alcoholic and control groups based on prevailing predictions of EEG signals. Recently, prevailing work has utilized Convolutional Neural Networks (CNNs) based on the VGG-16 method to enhance accurate representations in the data extraction process, particularly when working with spatial resolution data [22]. Another framework has implemented both Machine Learning (ML) approaches and neural networks as classifiers [23]. In this NN model, showcased improvements in performance over

Classical ML models. It has been perceived that a prominent alcoholic classification was undertaken from EEG signals to attain feasible results. The author has presented a Graphical User Interface (GUI) that extracts features such as sample entropy and standard deviation from EEG signals to detect alcoholism status. A Quadratic SVM classifier for detection and achieved 95% accuracy [24].

In addition, distinct CNNs have been implemented to improve the efficiency of the activation functions of alcoholic classifications from EEG signals [25]. The author in [26] has incorporated Transfer Learning with various CNNs as feature extractors, combined with classical classifiers like SVM and RF. The combination of the MobileNet CNN with an SVM classifier achieved 95.33 % performance in point of accuracy. Another existing research [27] Hybridising Fast Fourier Transform (FFT) and three classification models were utilized. Among those, the FFT+SVM model achieved better performance than the others. Research also imposed a dissimilar approach in regard to vector functions for enhancing the alcoholic classification of EEG signals [28, 29]. The author in [30] has employed a cascaded process starting with LASSO regression for initial clustering and feature extraction, followed by meta-heuristic algorithms (PSO, BCHS, BDA) for feature minimization, and finally classifying using various models, including SVM, random forests, ANN, EANN, and LSTM. In this study, LASSO regression combined with BDA-based enhanced ANN obtained remarkable performance. Though DL models performed well, the presence of missing values, feature redundancy, noisy datasets, and imbalanced data can adversely affect the efficacy of these prediction models.

2.3. Hybrid Model for Alcoholism Prediction

To address the aforementioned problem, many researchers implemented hybrid models subsequently to enrich the prediction accuracy [31]. Accordingly, in [32], dimensionality reduction using PCA (Principal Component Analysis) and classification with ML models KNN, SVM, and XGBoost were compared with CNN, RNN, and LSTM, and CNN combined with LSTM to improve model performance. The model was examined, and the demonstrated hybrid model achieves better performance than standalone models. In [33], the authors explored the use of CNN and bidirectional LSTM networks to classify EEG signals for alcoholism detection, achieving a diagnostic accuracy of 95.32%.

Another study [34] utilized the Multichannel Pyramidal Convolutional Neural Network (MP-CNN), which has been assessed with 61 channels of EEG signals covering over five brain regions for alcoholic predictions of EEG signal-based data. The results are extremely precise, but these conventional monologues have been reduced to larger sample sizes. Moreover, the CNN combined activation on the model has been estimated with better performance on the list of performance metrics in alcoholic classification. Another research [35] has implemented three DL models: CNN, LSTM, and CNN+LSTM. The framework yielded 92.77%, 89%, and 91% efficiency in accuracy on test samples. The suggested work [36] has converted 64-channel temporal data into images using FFT, ICA, and SAX techniques, then employed an ensemble model, LSTM, and EfficientNet models. Testing on a public dataset, their method achieved 85.52% accuracy, outperforming the state-of-the-art EEG-NET's 81.19% accuracy, demonstrating the potential of multi-perspective EEG analysis with ensemble methods for alcoholism detection.

Furthermore, researchers have looked at attention mechanisms, transformer-based models, and multi-modal learning to make EEG classification more robust and generalizable [37, 38]. In spite of these advances, many current models are computationally heavy, require large quantities of training data, or depend on extensive pretraining, limiting their use in clinical settings. Similarly, none of the approaches rely on complex ensemble frameworks, which can impact interpretability and deployment. Nevertheless, spatial CNNs and temporal LSTMs have demonstrated considerable promise individually; their joint potential remains underexplored in lightweight architectures specifically tailored for EEG-based alcoholism detection.

Table 1. Comparative analysis of proposed method with existing state-of-the-art techniques

Study	Method(s)	Dataset	Accuracy	Key Takeaway
Zhang et la., (2020) [26]	MobileNet CNN with an SVM	EEG	95.33 %	Achieved a better outcome
Mukhtar et al. (2021) [10]	Regularized Deep CNN	EEG	92.3%	CNN with dropout reduces overfitting
Nor et al., (2022) [27]	FFT+SVM, FFT+KNN and FFT+ANN	EEG	91 %	The SVM model performed better than other models
Cohen et al. (2023) [36]	LSTM + EfficientNet + Linear Nets	EEG	85.52%	Ensemble improves robustness
Pain et al. (2023) [9]	GNN + Connectivity	EEG	93.28%	GNN is effective for activations + connectivity
Proposed Model	**Assorted CNN + LSTM**	**EEG (UCI)**	**96.72%**	**Temporal modeling improves the pre-trained CNN**

2.4. Problem Identification

While reviewing the existing studies, the following aspects were found,

- High computational complexity in ensemble, wavelet-based, and transformer-based models causes problems in their use [23, 39].
- Traditional methods have unified more ML-based algorithms for the classification of alcoholic states with EEG signal images, but lack methods for enhancing the alcoholic predictions of EEG signals [29].

3. Proposed Methodology

Alcoholism is regarded as a common addiction disorder that substantially malfunctions the human immune system. ML techniques are confined to providing a sufficient level of precision for alcoholic and nonalcoholic detection. Consequently, traditional screening methods are limited in terms of time consumption, processing, and the need for excessive resources.

Several researchers have attempted to use ML and DL technologies for effective screening of alcohol, but conventional models are limited by inadequate mechanism incorporation in the model paradigm. For better screening for alcoholic classification, the proposed research works with a DL approach that integrates LSTM with a CNN by using an EEG-alcohol dataset. Figure 1 represents the workflow of the CNN-LSTM system.

Figure 1 illustrates the respective procedures carried out to perform the proposed CNN-LSTM technique for the classification of alcoholic and nonalcoholic features using the EEG-Alcohol dataset. This technique involves data collection, preprocessing, data splitting, training, classification, and prediction. The following sections depict a broad demonstration of each stage.

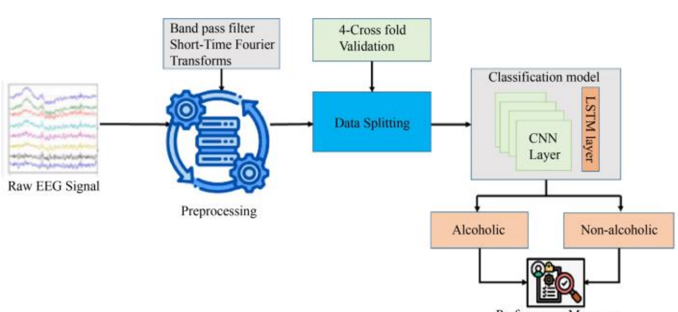


Fig. 1 Flow diagram of the proposed method

3.1. Data Collection

The proposed research utilized the EEG-Alcohol dataset for the binary classification of alcoholics and non-alcoholics. The respective dataset is extracted from the Kaggle website, which is a cloud-based database composed of various AI datasets with data scales ranging from small to medium and large. Primarily, the design modality of the proposed dataset covers the EEG signal denotation of alcoholism. The classification mechanism is processed through stimulus and two-stimulus specifications. The significant features of the assumed dataset are depicted in Table 2.

Table 2. Significant aspects of the EEG-Alcohol dataset

S. no	Aspects	Aspect Description
1.	In-specific	It is particularly useful for analyzing EEG signal images.
2.	Standardized	Each dataset subset can be preprocessed.
3.	Predictive	The sensor value measurements are estimated in microvolts.

The employed EEG-Alcohol dataset was acquired from the following link [40]:

Subsequently, to ensure that the model is effective, the MATLAB 2022a software is used for execution and for tuning the parameter value control. For a reliable source requisite, the deep learning toolbox delivers functions and tools for designing and implementing the proposed CNN networks. To evaluate the data structure analysis, a set of modules was employed with the PCA toolbox.

3.2. Preprocessing

It is a crucial process in data preparation, essential for transforming raw data into a suitable format for analysis and modeling. The quality of the image significantly influences the accuracy of the data, so image preprocessing enables the lessening of distortions. Moreover, it is crucial to perform this process before importing the dataset to the classification model. The respective research imposed kernel matrices for processing the collected EEG signal sample images, wherein each kernel is designed to detect specific features by employing these matrices on local regions of the input data.

The output of this operation is known as a feature map, which highlights the presence of the particular features in the input. Applying the kernel function permits measuring the relationship between different segments of EEG data without explicitly mapping them into high-dimensional space. With the imparted mask, the extracted feature understanding of images is attained by computing convolutional functions.

Properly preprocessed data improves the performance by providing accurate and relevant inputs. Streamlining the data can significantly reduce the computational resources and time required for model training. Figure 2 demonstrates the depiction of the data processing process carried out in the present research to modify and obtain a clean dataset collection regardless of the presence of noisy or blurred series of data samples.

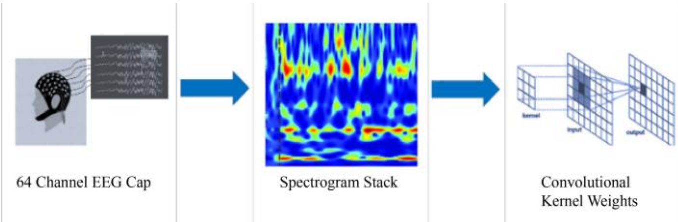


Fig. 2 Data-preprocessing flow of the proposed model

3.3. Data Splitting

Data splitting is an essential process in determining the flow of any AI model. In the proposed model, a fourfold cross-validation method is applied to validate performance and to reduce the bias issue in the dataset. As per Figure 3, the data is divided into four equal subgroups. Here, during every iteration, a single subgroup is designated as the test dataset, while the remaining subgroups are utilized as the training dataset. In every iteration, the model undergoes training and is subsequently assessed using the test dataset. Then, all iterations undergo metric evaluation, after which their averages are computed to assess the overall performance of the model. This approach effectively tackles the overfitting issue and improves the performance of the model.

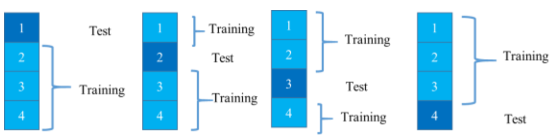


Fig. 3 4-Fold Cross Validation

3.4. Classification: Hybrid CNN-LSTM

The proposed research uses a CNN-LSTM model to classify the alcoholic and control (nonalcoholic) groups via the TensorFlow framework on the utilized sample images from the EEG-Alcohol dataset. The respective model conducts the training phase beforehand; the proposed model is constructed with the use of neural networks as a pre-trained model. The employed images depart from the EEG-Alcohol dataset. This section illustrates the methods and algorithms used for alcoholic classification in the proposed research.

The proposed model contains two key components, such as CNN and LSTM. In which a component CNN is used for feature extraction and an LSTM for sequence modeling purposes. Initially, the process takes spectrogram images of the EEG alcohol signal as an input for the CNN layer, which involves a kernel method for data processing, to extract relevant features and similarity data from the raw data, hence improving the quality of the EEG data. Subsequently, the respective model employs a pooling operation for noise reduction, and to avoid overfitting issues, the dropout classifier adapts to the proposed architecture to increase model capacity.

Following that, the outcome of the conventional layer is fed into the LSTM layers that are utilized to learn durable dependencies between individual time series from a set of data sequences. The LSTM network has memory cells that recall the stored data over time, and it also includes gates that regulate the flow of data into and out of these cells. The structure enables the model to learn and remember patterns in the data that span extended periods of time. In addition, before training the model, samples of EEG data are converted to a spectrogram stack that produces a series of segments for preprocessed data. Sequentially, linear layers, called fully connected layers, are also included in the configuration that applies a linear transformation using a weight matrix, followed by a non-linear activation function. This task allows the model to learn complex patterns and interconnect the entire feature map into a single vector of data. This vector is utilized to detect the alcoholism state signal. The obtained data is used to generate output data via binary classification of alcoholic and nonalcoholic features.

The classification is automated by using the DL method of CNN and LSTM to improve the accuracy of the proposed model. Together, these strategies can contribute to a more effective DL framework that maximizes classification performance. The next section illuminates the CNN mechanism in the classification task.

3.4.1. Convolutional Neural Network

CNN is one of the most widely used artificial neural networks in AI models; it is processed on the basis of designated filters. The predominant work of CNNs involves features. Moreover, it delivers better accuracy in images and classification. Feature detectors work at identifying the intended image and progress throughout CNN layers [29]. Basically, CNN operation is carried out in three-layer patterns: convolutional layer, pooling layer, and Fully Connected (FC) layer [30].

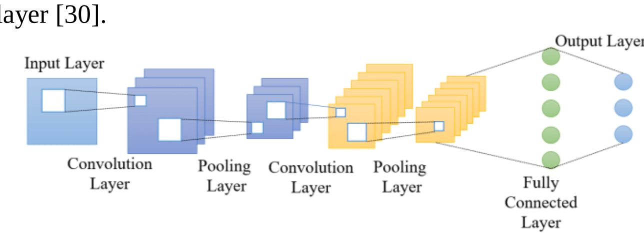


Fig. 4 The architecture of the CNN model

Figure 4 depicts the working process of the CNN model. In this, the convolutional layer transforms the image into a standard numerical representation, where it prepares the data, and neural networks enable the interpretation of the conclusions. In the next step, the pooling operation bends the output to arrays produced from aggregated functions. The mathematical expression for the convolutional action is formulated accordingly, and the integral operation produces a convolution function.

$$f(t) * g(t) = \int_0^t \frac{f(T)g(t-T)dT}{(f*g)(t)} \tag{1}$$

The pooling layer has one of two paths: max pooling or average pooling. The max pooling layer breaks down larger classes, where average pooling is performed when neutral classes are used. The chief use of CNNs for deducting complexes enhances accuracy and overfits the results. The fully connected layer works as an interconnected system where each node is connected to every other node of perceived outcomes. FC classifies appropriate functions, but it cannot predict high-dimensional vectors. The elaborate action of FC is an increasingly time-consuming process and requires expensive computational models during the training phase. Occasionally, massive interconnections may lead to data loss.

The CNN design is categorized into variant structures, and different models are represented by different activation functions that determine the performance of the trained model [32]. AlexNet is one of the most widely used architectures, and it is the preferred architecture for large-scale datasets. The model is composed of five convolutional layers, a single max pooling layer, and three FC layers [33]. Finally, another layer is divided into two norms, each of which specializes in promoting graphical-centered approaches. Another system called ResNet also leverages large sizes, yet not much more than AlexNet, but features such as skips or shortcuts are possible. This model is configured in three versions: ResNet50, ResNet101, and ResNet152. ResNet models are also widely recognized as consensus-based and standardized models for obtaining energy efficiency results. Ideally, mutual development is also implicit before layer addition. VGGNet has smaller feature maps for its functions. VGGNet has two versions: VGG16 and VGG19. GoogLeNet also generates a wide range of performance metrics for classification [34]. Furthermore, the GoogleNet model aids in comprehensive learning of image feature representations, and it is an idealized pre-trained model for classifying images. However, other named SqueezeNet models are faster than the GoogLe model. The added updates with the GoogLeNet model included a new model, an enhanced model known as the InceptionV3 model.

The proposed CNN-LSTM system employs the respective pre-trained neural networks according to their idealized selection of layer architectures. In this case, the method utilized ten standardized pre-trained models to impart a specialized approach for determining the efficacy of the heightening classification process. These neural network models have previously departed from the typical requirement for classifying image data; however, the use of the proposed method enables the reconstruction of model optimism in the classification of representative alcoholic and nonalcoholic EEG images.

3.4.2. Long Short-Term Memory

LSTM is boxed in an Recurrent Neural Network (RNN). The operation of LSTM is slightly similar to that of the FC layer, but the specific features of LSTM determine the added effectiveness to model performance. The optimist verification on each layer possibly detects the actions of the data sequences. The memory is recorded with all sequential connections, which can handle large sample sizes [35]. LSTM works with time series conditions, and the recorded set of memory enables the evaluation of enumerations. Enumerations may differ with model purposes. The application of these procedures precisely detects different states in the sequences where there are no unlikely missing data problems.

3.4.3. Proposed Classification Model—Hybrid CNN+LSTM

This section explicitly describes the specifications of the classification methods used for model construction. The proposed research employs the CNN-LSTM model to classify alcoholic and nonalcoholic states in humans. The respective technique is applied to the image of the EEG signals. To train the model, the EEG-Alcohol dataset is used for alcoholic identification from EEG records.

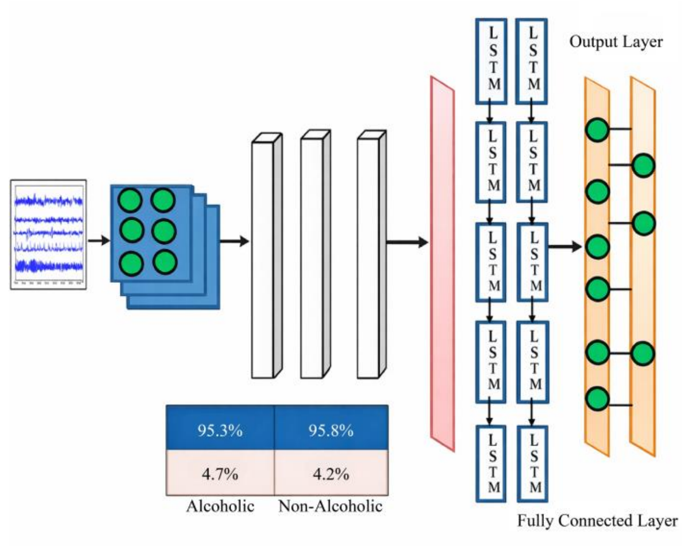


Fig. 5 Architecture of the CNN-LSTM classification model

As Figure 5 illustrates, the predominant use of the max pooling layer from the CNN network layer executes relevant data values by detecting image fluctuations.

Remarkably, LSTM has an elongated memory of data information, and it considers all the particulars of memory blocks. The feature of the forget gate on LSTM works to prevent inappropriate data reflection on summarizing outcomes because it primarily relies on yielding continuous predictions. This can avoid complexity in the estimation of predictions, so LSTM promotes the effective analysis of state conditions. This serialized recognition of sequential EEG signals improves learning tasks more effectively in the training phase. Essentially, the proposed CNN-LSTM model disables the conditions of complexity, inapt storage, and vanishing gradients. These constraints limit the performance of classical CNN architectures. The combination with LSTM can generate a precise evaluation of testing data. Eventually, the efficacy of the trained model is enhanced compared with that of the normal CNN model. The framework of Figure 5 depicts the CNN-LSTM classification model. The model starts by receiving a preprocessed EEG spectrogram image of EEG signals, which are time-series data representing brain activity collected from scalp electrodes. These images are first processed by CNN layers, which extract spatial and local temporal features, such as frequency components and characteristic patterns, applying convolutional filters across the input. The resulting feature maps are then flattened into 1D vectors to prepare them for sequential modeling. This flattened output is reshaped into a sequence format suitable for LSTM input, ensuring the temporal structure of the data is preserved.

Next, the LSTM layers (sometimes arranged in parallel or as bidirectional streams) process these sequences to learn long-range temporal dependencies, effectively capturing how EEG features evolve. The high-level temporal features produced by the LSTM are then passed through fully connected layers, which integrate the information and perform the final classification. The output layer maps these features to specific classes, such as different mental states, motor imagery tasks, or seizure detection categories.

This hybrid CNN-LSTM model significantly enhances Alcoholic and nonalcoholic classification by effectively capturing both spatial features and long-range temporal dependencies within the data. By leveraging the strengths of both architectures, it achieves improved accuracy and robustness in distinguishing complex brain activity patterns for various clinical and cognitive applications.

Table 3. Training parameters with the LSTM hybrid models

Model	Layer	Optimizer	No. of Epochs	Elapsed Time	Hardware Resource	Learning rate schedule	Learning rate
Alexnet+LSTM	pool5	adam	500	10 min 43 sec	Single CPU	Constant	0.001
DenseNet201+ LSTM	avg_pool	adam	2000	9 min 5 sec	Single CPU	Constant	0.001
GoogleNet+LSTM	pool5-7x7_s1	adam	2000	9 min 44 sec	Single CPU	Constant	0.001
InceptionV3+LSTM	avg_pool	adam	1500	3 min 13 sec	Single CPU	Constant	0.001
ResNet18+LSTM	pool5	adam	4000	10 min 25 sec	Single CPU	Constant	0.001
ResNet 50+LSTM	avg_pool	adam	2000	16 min 42 sec	Single CPU	Constant	0.001
ResNet101+LSTM	pool5	adam	2000	14 min 32 sec	Single CPU	Constant	0.001
VGG16+LSTM	pool5	adam	150	8 min 3 sec	Single CPU	Constant	0.001
VGG19+LSTM	pool5	adam	200	10 min 11 sec	Single CPU	Constant	0.001
SqueezeNet+LSTM	pool10	adam	7000	78 min 54 sec	Single CPU	Constant	0.001

Table 3 shows the training parameters of the proposed neural network models, which are summed with their layers formed in the CNN architecture, and the accuracy obtained through the EEG-alcohol dataset for predicting the model performance. Then, to minimize the loss functions, the proposed system applied an optimizer named Adam. The Adam optimizer was applied to all successive neural network models to execute the accelerated optimization procedures for model training. The subsequent addition of pooling layers to the respective CNN models causes computational complexity and results in adequate memory requirements.

Hence, the performance of the proposed model allows for the exploration of improvements in image classification performance. The elapsed time calculated at InceptionV3 was less than that of the other proposed neural network models. However, other mere models have procured more time, whereas the corresponding number of epochs is determined in accordance with resolving overfitting issues. Significantly, to alleviate overfitting and increase the generalization ability of the neural networks, the optimal range of the number of epochs is employed. Furthermore, to avoid waver or overrun, the research employed learning rates with a default value of 0.001. A constant learning rate is formulated for the adopted models to update in the optimizer.

4. Experimental Analysis

This section demonstrates the outcome attained by the proposed model. In addition, Explanatory Data Analysis (EDA), performance metrics, and a comparative analysis of the respective model with existing mechanisms are illustrated.

4.1. Explanatory Data Analysis (EDA)

The EDA method is further used to analyze and perceive dataset automation based on model features, and the characteristic features of the dataset are also identified in this section.

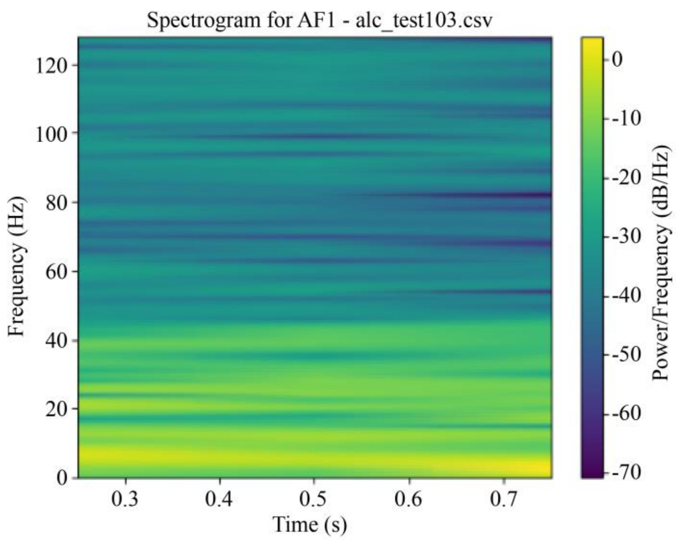


Fig. 6 Training sample

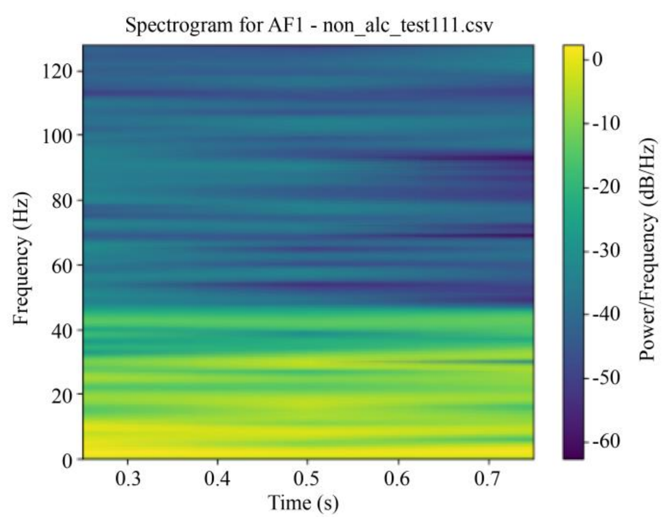


Fig. 7 Testing sample

Figures 6 and 7 represent the training and testing samples used to analyze the classification performance of the CNN-LSTM model, respectively.

This testing sample is processed with a trained model to determine the efficacy of the proposed model. Training sample of the CNN-LSTM model on the EEG-Alcohol dataset.

4.2. Performance Metrics

The respective models were tested with standard performance metrics to evaluate the performance of the proposed approach. This research used accuracy metrics for evaluating model performance since accuracy enhancement is crucial for model progression in alcoholic classification.

The accuracy metrics are determined to be significant for determining the most appropriate prediction for the classification. It is evaluated by the correct predictions as a proportion of the sum of predictions. The formula for determining the accuracy is given below in Equation (2)-(5):

$$Accuracy(Acc) = \frac{True\ P + True\ N}{Trye\ P + False\ P + True\ N + False\ N} \tag{2}$$

$$F1\ score\ = 2 * \frac{RC \times Pc}{Rc + Pc} \tag{3}$$

$$Precision\ = \frac{True\ P}{(True\ P + False\ P)} \tag{4}$$

$$Recall\ = \frac{True\ P}{(True\ p + False\ N)} \tag{5}$$

Where,
$True\ P$ -The model correctly predicts an alcoholic
$True\ N$- The model correctly predicts not an alcoholic
$False\ P$- The model inaccurately predicts an alcoholic
$False\ N$- The model fails to predict an alcoholic

4.3. Performance Analysis

The outcomes for the CNN-LSTM model are demonstrated in this section. The outcomes from the operation of the proposed model were compared with those from the use of the EEG-Alcohol dataset.

Table 4 illustrates the ten proposed pre-trained model outcomes using various metrics.

Table 4. Pre-trained model performance analysis

Model	ACC %	Precision %	Recall %	F1 Score %
Alexnet	95.43	98.7	92.1	95.27
DenseNet201	95.55	95.77	95.33	95.55
GoogleNet	90.86	87.08	96.03	91.34
InceptionV3	91.44	90.19	92.57	91.28
ResNet18	92.38	93.46	91.51	92.47
ResNet 50	91.68	86.21	96.85	91.22
ResNet101	90.74	86.91	94.18	90.5
VGG16	95.63	95.33	95.33	95.33
VGG19	96.72	95.8	97.62	96.7
SqueezeNet	85.95	98.72	83.48	86.43

Table 4 illustrates the performance outcomes obtained in the proposed ten pretraining models. From Table 4, observing that pre-trained models Alexnet, DenseNet, VGG16, and VGG19 have shown more accuracy in the prediction, of which VGG19 has obtained the highest accuracy of 96.72% in detecting the alcoholic state of the patients. Figure 8 depicts the graphical representation of the pre-trained model's performance with the accuracy metrics.

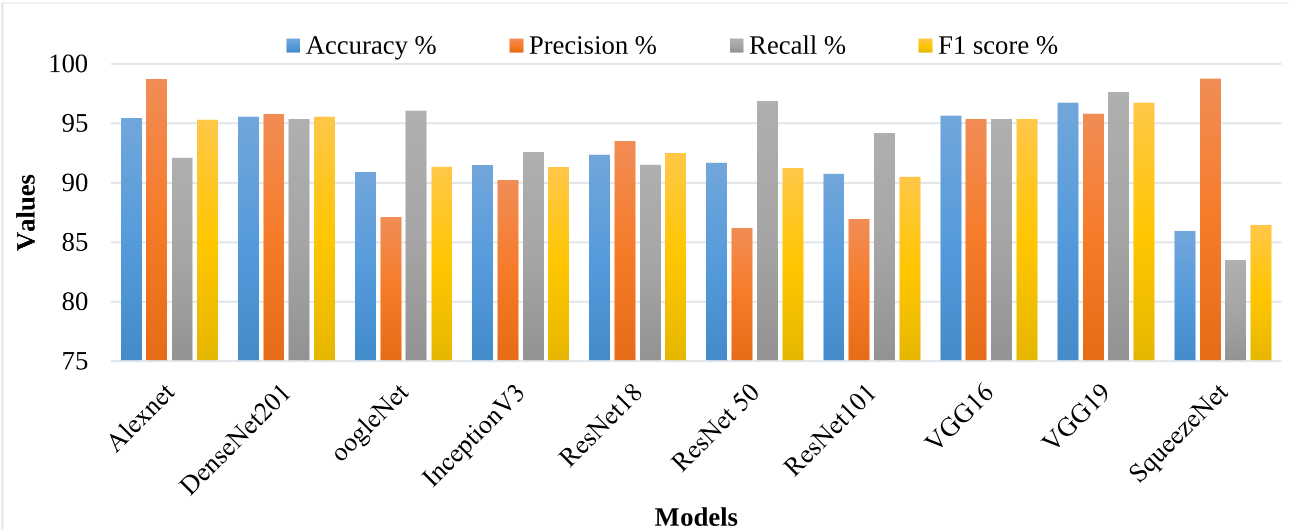


Fig. 8 Performance analysis of the pre-trained model

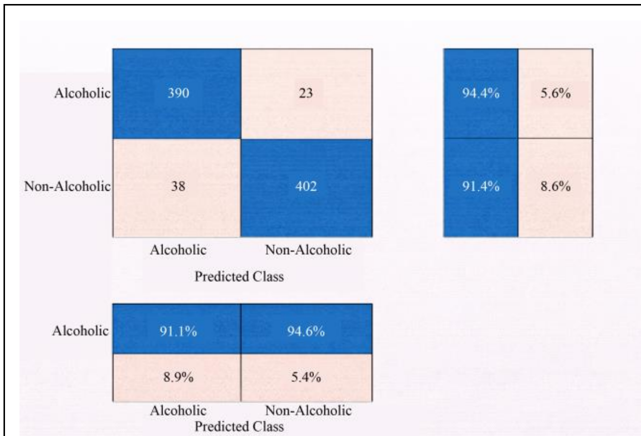
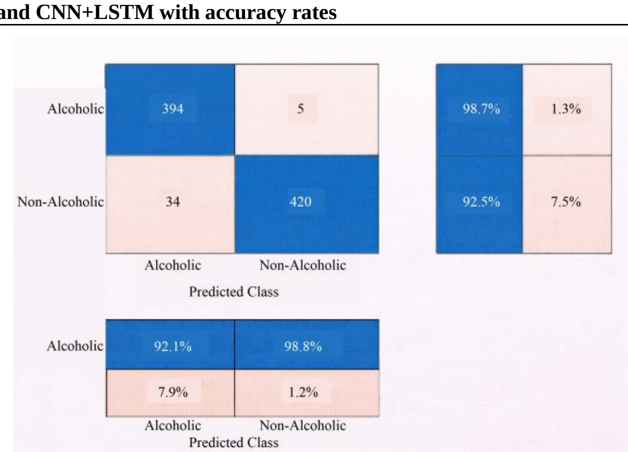
Table 5. Comparative analysis of the CNN-LSTM model with the pretrained model

Pre-trained Models	ACC %	Proposed Model	ACC %
AlexNet	92.85	AlexNet+LSTM	95.43
DenseNet-201	86.87	DenseNet-201+LSTM	95.55
GoogleNet	84.06	GoogleNet+LSTM	90.86
InceptionV3	88.98	InceptionV3+LSTM	91.44
ResNet18	82.65	ResNet18+LSTM	92.38
ResNet50	86.40	ResNet50+LSTM	91.68
ResNet101	82.65	ResNet101+LSTM	90.74
VGG16	85.35	VGG16+LSTM	95.63
VGG19	86.40	VGG19+LSTM	96.72
SqueezeNet	84.88	SqueezeNet+LSTM	85.95

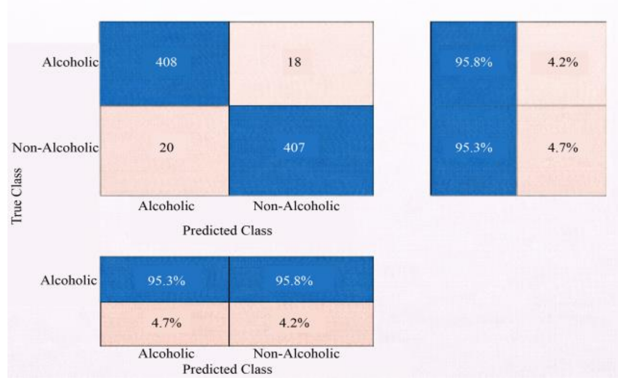
4.4. Comparative Analysis

This section analyses the outcomes attained by the proposed CNN-LSTM model with existing models, which determines the efficacy of the present framework with CNN models. Table 5 depicts the comparative analysis of the CNN-LSTM method with the existing model. The present research evaluated the existing CNN models with the CNN-LSTM system model to estimate the efficacy of the proposed method.

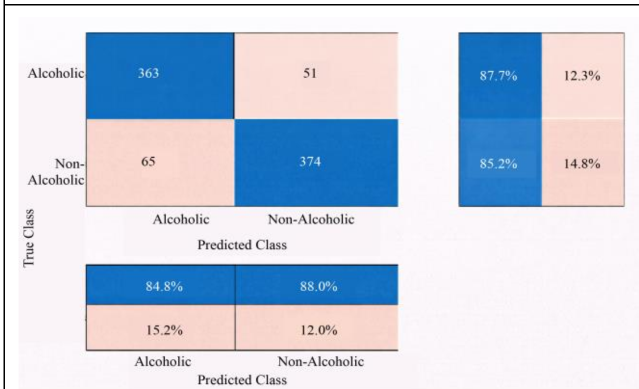
Table 6. Confusion matrices of CNN and CNN+LSTM with accuracy rates

 **Alexnet**	 **Alexnet+LSTM**

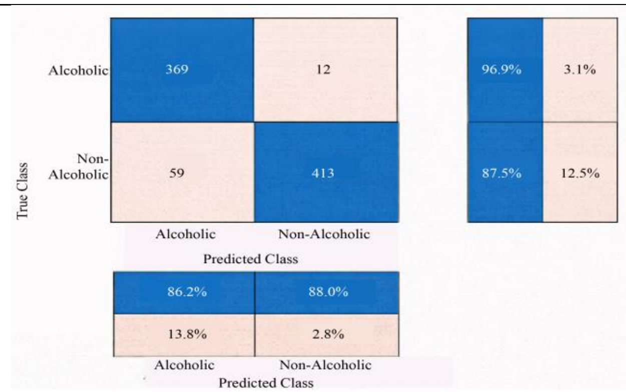
Densenet201



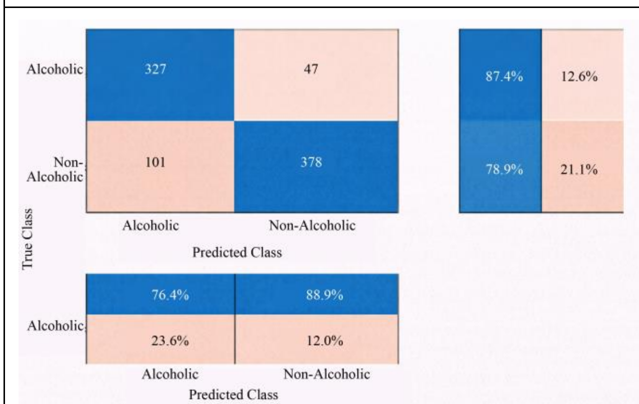
Densenet201+LSTM



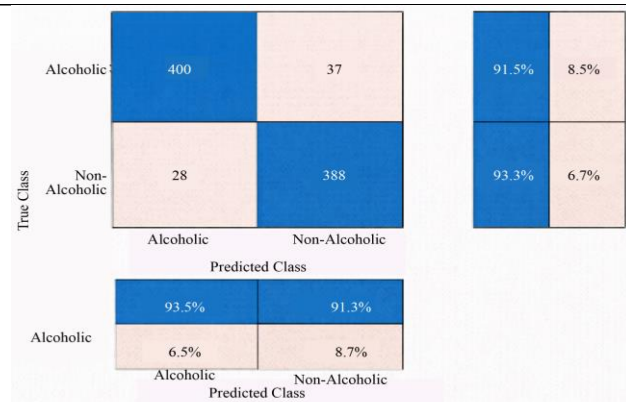
Resnet50



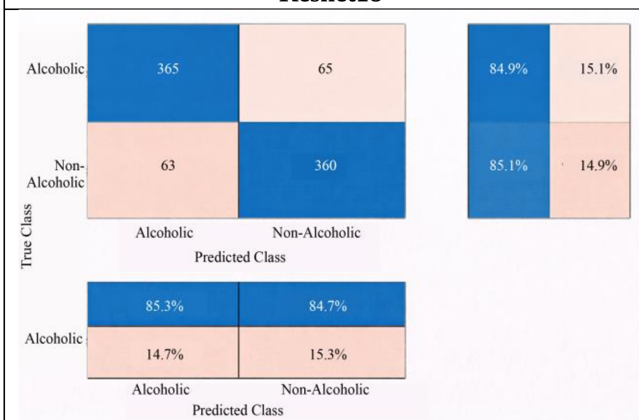
Resnet50+LSTM



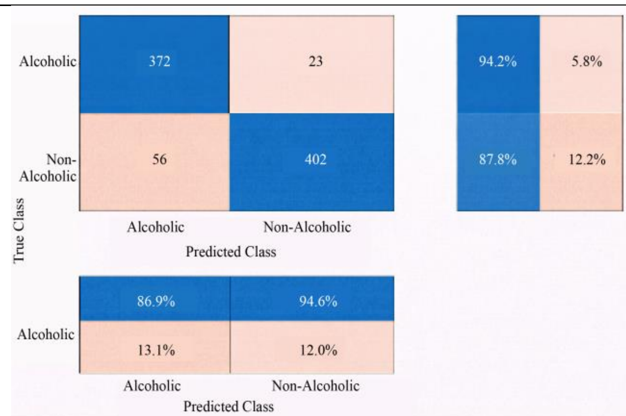
Resnet18



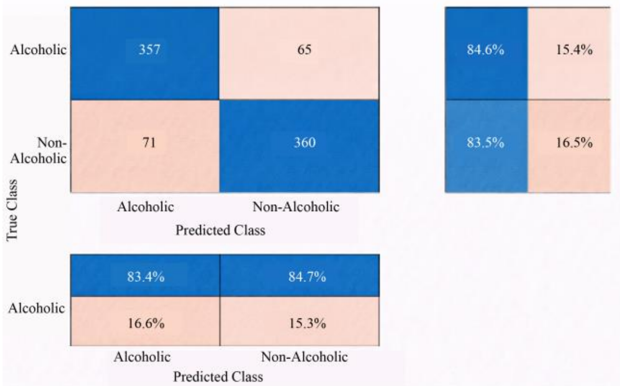
Resnet18+LSTM



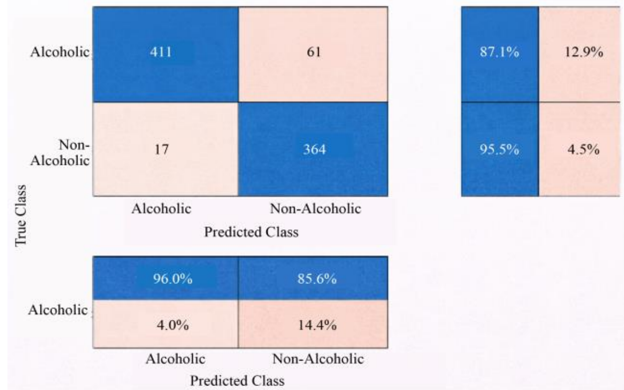
Resnet 101



Resnet 101 +LSTM

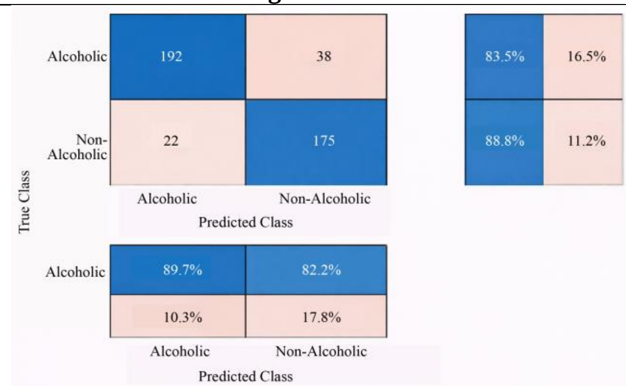


Googlenet

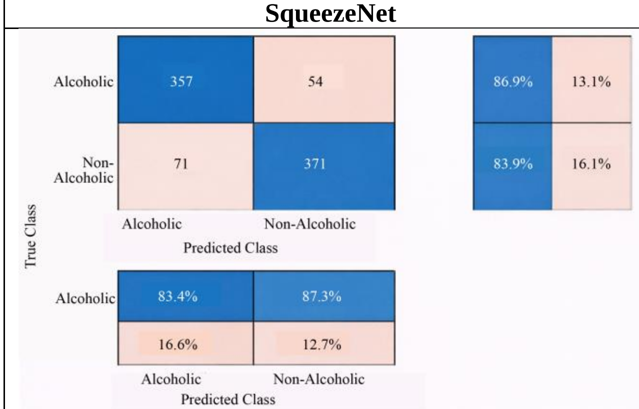


Googlenet+LSTM

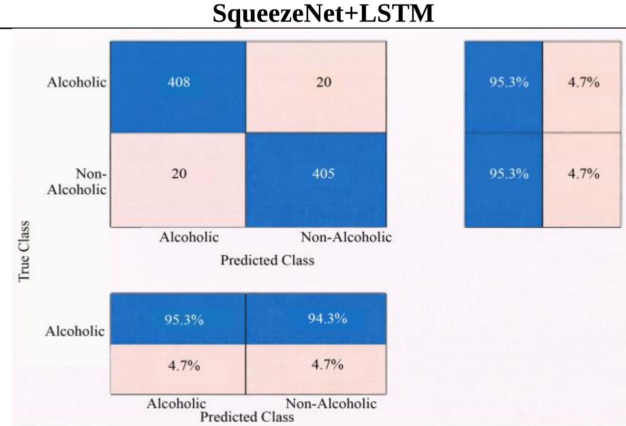
SqueezeNet



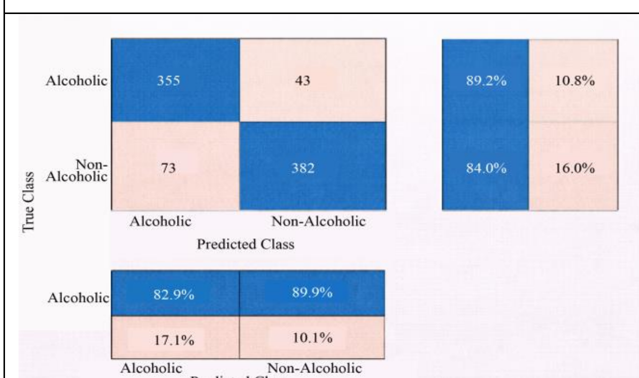
SqueezeNet+LSTM



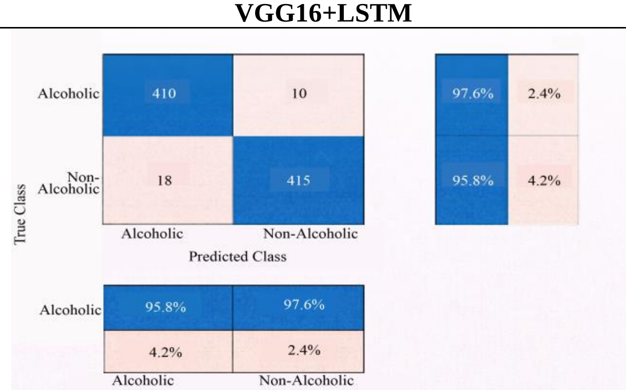
VGG16



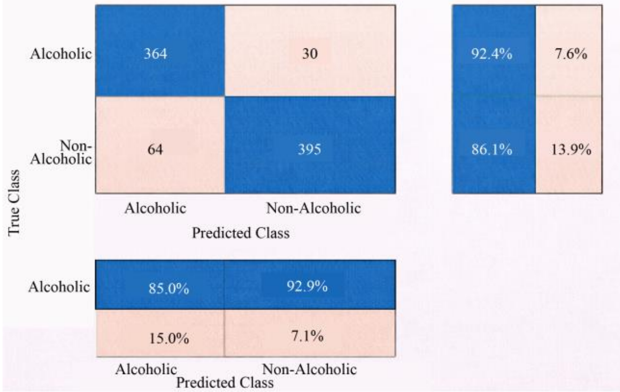
VGG16+LSTM



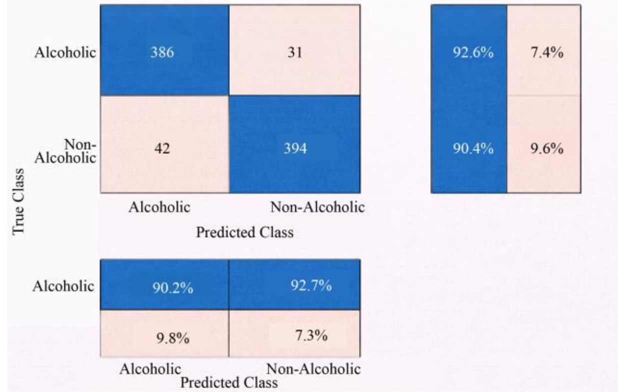
VGG19



VGG19+LSTM

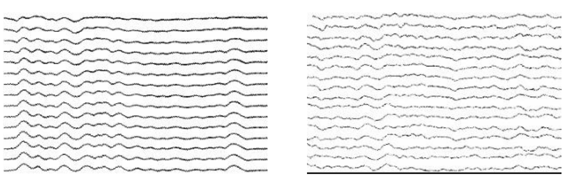


InceptionV3



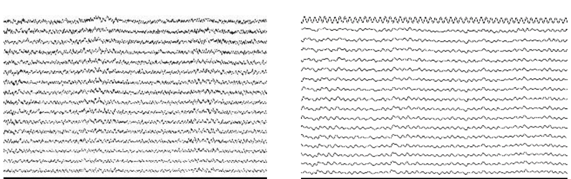
InceptionV3+LSTM

Table 6 illustrates the outcome of the proposed model in which LSTM was cascaded with existing CNN models, which were examined with 10 CNN-based pre-trained models utilized in this present research, and the accuracy improvement of all 10 models was evaluated. The highest accuracy rate is identified from the VGG19 model combination, with an accuracy of 96.72%. The accuracy of the confusion matrix collections of the alcoholic classification of all ten pretrained CNN models with the proposed CNN-LSTM models is shown in Table 6, which shows the accuracy rates of the CNN models with and without the LSTM combination. After LSTM is incorporated into the mechanism, the accuracy increases for all ten models. Henceforth, it is evident that LSTM is a major feature for enhancing the prediction of alcoholic classification.



a) Alcoholic **b) Nonalcoholic**

Fig. 9 Classification Results of the CNN System



a)Alcoholic **b) Nonalcoholic**

Fig. 10 Classification Results of the CNN-LSTM System

Figures 9(a), 9(b), and 10(a), 10(b), illustrate the alcoholic classification on the CNN and CNN+LSTM models using the EEG-Alcohol dataset. Figures 9 and 10 explicitly show the precise outcomes of the EEG signals and a clear classification of Alcoholic and Nonalcoholic indications. The efficacy of EEG signal detection on alcoholic classification is enhanced to a greater extent in the proposed CNN-LSTM system. Therefore, further accurate results for classifying alcoholic and nonalcoholic states are needed.

4.5. Discussion

The conventional model [25] employs a generalized CNN architecture to evaluate the classification performance of the model, and the accuracy rates of the models reached a maximum of 95%. In addition to evaluating the efficacy of the proposed model through the assessment of the accuracy rates of the pretrained models, typical research [37] has also evaluated the proposed methods via external considerations. Although research has prevailed in image detection, accuracy attainment reached a minimal extent of 94.47%, rather than the results of the proposed work. Subsequently, another recent study [21] also explored feature extraction. The hybrid of LSTM to CNN is used for improving model performance, but the marked accuracy rate decreases by only 93% when using the CNN-LSTM approach. However, the existing models have achieved better accuracy and improved model training, as further improvements beyond accuracy rates have not been achieved. According to the proposed research, the performance of CNN-LSTM reached 96.72% for the VGG19 model, which was computed with LSTM. Regardless of the LSTM, the model achieved an accurate rate of 86.40%.

5. Conclusion

Alcoholism is regarded as a global disease that predominantly affects brain functionality, and its causes lead to malfunctioning human brain functions. It is crucial to determine and anticipate ways to reduce the risk factors that contribute to these causes. The EEG model depicts brain activity and, moreover, monitors functionality and detects alcohol consumption. In summary, hospitalists' involvement in increasing the efficiency of EEG signal specifications is considered a greater demand for generating accurate results to prevent risks. Contemporary methods, such as ML techniques, are preferred during the initial process; nevertheless, ML model outcomes have been shown to yield better results than manual procedures. However, there are still possibilities for improving the efficacy of EEG signals; however, the ML mechanism is not yet available for alcoholic classification. Relatively, the neural network learning task is a more relevant application for analyzing alcohol detection, where the proposed system utilizes the EEG-Alcohol dataset for alcoholic classification using Deep learning CNN and LSTM, to enhance the efficacy of the classification. The respective research used LSTM with its grouping on CNN models such as AlexNet, DenseNet201, GoogleNet, InceptionV3, ResNet18, ResNet50, ResNet101, VGG16, VGG19, and SqueezeNet. Each CNN model combination increased the accuracy of the model. A comparative analysis of the CNN and CNN-LSTM models revealed that the accuracy percentages obtained by the proposed model are increased. The accuracies of the CNN models AlexNet, DenseNet201, GoogLeNet, InceptionV3, ResNet18, ResNet50, ResNet101, VGG16, VGG19, and SqueezeNet were 92.85%, 86.87%, 84.06%, 88.98%, 82.65%, 86.40%, 82.65%, 85.35%, 86.40%, and 84.88%, respectively. The accuracies of the CNN+LSTM, AlexNet+LSTM, DenseNet201+LSTM, GoogLeNet+LSTM, InceptionV3+LSTM, ResNet18+LSTM, ResNet50+LSTM, ResNet101+LSTM, VGG16+LSTM, VGG19+LSTM and SqueezeNet+LSTM models were 95.43%, 95.55%, 90.86%, 91.44%, 92.38%, 91.68%, 90.74%, 95.63%, 96.72% and 85.95%, respectively. The proposed approach achieved higher accuracies than the pre-trained neural network models. Hence, the outcomes of the proposed models elucidate the significance of LSTM for alcoholic and nonalcoholic image classification. Therefore, the accomplished percentage rates explicitly represent the proposed approach's certainty on magnified classification rather than the execution on prevailing AI models. Furthermore, the proposed research considers traditional screening, which has various drawbacks, and consequently, high computation time is required for existing models to achieve qualified performance. Numerous studies have shown that the DL approach can be used for the effective classification of alcohol use based on EEG signals. As a slight difference in incorporation, the proposed system employed LSTM with pre-trained neural networks, which are combined to enhance the accuracy metrics for the classification. The Results section identifies the prominent difference in the accuracy of model performance.

The proposed hybrid-structured model integrates LSTM with CNN architecture, although the model improved the ability to classify alcoholic and nonalcoholic detection. The respective model is still confined to some contemplation. For training, a large range of noise limits is imposed when screening attack instances. The proposed model processes the data using the functionality of 1D-CNN configurations. Therefore, denotation, such as weight sharing and computational cost, will still be limited in the present application.

References

[1] J. Morris et al., “The "Alcoholic Other": Harmful Drinkers Resist Problem Recognition to Manage Identity Threat,” *Addictive Behaviors*, vol. 124, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Palak Mahajan et al., “Ensemble Learning for Disease Prediction: A Review,” *Healthcare*, vol. 11, no. 12, pp. 1-21, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Daniel Q. Huang, “Global Epidemiology of Alcohol-Associated Cirrhosis and HCC: Trends, Projections and Risk Factors,” *Nature Reviews Gastroenterology & Hepatology*, vol. 20, pp. 37-49, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Xavier Noël, “Leveraging Memory Suppression from a Goal-Directed Perspective to Regain Control over Alcohol Consumption,” *Alcoholism, Clinical and Experimental Research*, vol. 48, no. 12, pp. 2242-2245, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] Kaloso M. Tlotleng, and Rodrigo S. Jamisola, “An Analysis of the Severity of Alcohol Use Disorder Based on Electroencephalography using Unsupervised Machine Learning,” *Big Data and Cognitive Computing*, vol. 9, no. 7, pp. 1-27, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[6] Niyozov Shuxrat Toshtemirovich, Turayev Bobir Temirpulotovich, and Mamurova Mavludaxon Mirxamzayevna, “Modern Methods of Diagnosing and Treating Neurological Changes Observed in Alcoholism,” *Journal of Medical Genetics and Clinical Biology*, vol. 2, no. 1, pp. 107-116, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[7] Maytal Wolfe et al., “Alcohol and the Central Nervous System,” *Practical Neurology*, vol. 23, no. 4, pp. 273-285, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Muhammad Tariq Sadiq et al., “Alcoholic EEG Signals Recognition based on Phase Space Dynamic and Geometrical Features,” *Chaos, Solitons & Fractals*, vol. 158, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Subrata Pain et al., “Detection of Alcoholism by Combining EEG Local Activations with Brain Connectivity Features and Graph Neural Network,” *Biomedical Signal Processing and Control*, vol. 85, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Hamid Mukhtar, Saeed Mian Qaisar, and Atef Zaguia, “Deep Convolutional Neural Network Regularization for Alcoholism Detection using EEG Signals,” *Sensors*, vol. 21, no. 16, pp. 1-19, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[11] Said Abenna, Mohammed Nahid, and Hamid Bouyghf, “Alcohol use Disorders Automatic Detection based BCI Systems: A Novel EEG Classification based on Machine Learning and Optimization Algorithms,” *International Journal of Information Science and Technology*, vol. 6, no. 1, pp. 14-25, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Salman Ahmed et al., “An Empirical Analysis of State-of-Art Classification Models in an it Incident Severity Prediction Framework,” *Applied Sciences*, vol. 13, no. 6, pp. 1-27, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] D. Merlin Praveena, D. Angelin Sarah, and S. Thomas George, “Deep Learning Techniques for EEG Signal Applications–A Review,” *IETE Journal of Research*, vol. 68, no. 4, pp. 3030-3037, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Muhammad Tariq Sadiq et al., "Efficient Novel Network and Index for Alcoholism Detection from EEGs," *Health Information Science and Systems*, vol. 11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[15] Hamid Mukhtar, *Chapter 11 - EEG Signal Processing with Deep Learning for Alcoholism Detection*, Artificial Intelligence and Multimodal Signal Processing in Human-Machine Interaction: Artificial Intelligence Applications in Healthcare and Medicine, pp. 211-226, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[16] Levi N. Bonnell et al., "A Machine Learning Approach to Identification of Unhealthy Drinking," *The Journal of the American Board of Family Medicine*, vol. 33, no. 3, pp. 397-406, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] L. Farsi et al., "Classification of Alcoholic EEG Signals Using a Deep Learning Method," *IEEE Sensors Journal*, vol. 21, no. 3, pp. 3552-3560, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[18] Qasem Abu Al-Haija, and Moez Krichen, "A Lightweight in-Vehicle Alcohol Detection Using Smart Sensing and Supervised Learning," *Computers*, vol. 11, no. 8, pp. 1-18, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[19] Dennis Flathau et al., "Early Detection of Alcohol Use Disorder Based on a Novel Machine Learning Approach using EEG Data," *2021 IEEE International Conference on Big Data (Big Data)*, Orlando, FL, USA, pp. 3897-3904, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[20] Sandeep Bavkar, Brijesh Iyer, and Shankar Deosarkar, "Rapid Screening of Alcoholism: An EEG based Optimal Channel Selection Approach," *IEEE Access*, vol. 7, pp. 99670-99682, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] Jardel das C. Rodrigues et al., "Classification of EEG Signals to Detect Alcoholism Using Machine Learning Techniques," *Pattern Recognition Letters*, vol. 125, pp. 140-149, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[22] Gaowei Xu et al., "A Deep Transfer Convolutional Neural Network Framework for EEG Signal Classification," *IEEE Access*, vol. 7, pp. 112767-112776, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[23] Fatima Faraz et al., *Classification of Normal and Alcoholic EEG Signals Using Signal Processing and Machine Learning*, Artificial Intelligence Enabled Signal Processing based Models for Neural Information Processing, CRC Press, pp. 33-50, 2024. [Google Scholar] [Publisher Link]

[24] Liu Jiajie et al., "Clinical Decision Support System for Alcoholism Detection Using the Analysis of EEG Signals," *IEEE Access*, vol. 6, pp. 61457-61461, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[25] Mingkan Shen et al., "Detection of Alcoholic EEG Signals Based on Whole Brain Connectivity and Convolution Neural Networks," *Biomedical Signal Processing and Control*, vol. 79, 2023, [CrossRef] [Google Scholar] [Publisher Link]

[26] Hongyi Zhang et al., "Bi-Dimensional Approach based on Transfer Learning for Alcoholism Pre-Disposition Classification via EEG Signals," *Frontiers in Human Neuroscience*, vol. 14, pp. 1-13, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[27] Nur Shahirah Md Nor et al., "Automated Classification of Eight Different Electroencephalogram (EEG) Bands using Hybrid of Fast Fourier Transform (FFT) with Machine Learning Methods," *Neuroscience Research Notes*, vol. 5, no. 1, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[28] Yue Zhao et al., "E3GCAPS: Efficient EEG-based Multi-Capsule Framework with Dynamic Attention for Cross-Subject Cognitive State Detection," *China Communications*, vol. 19, no. 2, pp. 73-89, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[29] Tatyana Yakovleva, and A.V. Krysko, "Processing Alcoholism EEG Signals Using Neural Networks," *Russian Journal of Biomechanics*, vol. 28, no. 1, pp. 110-126, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[30] Gowri Shankar Manivannan et al., "Detection of Alcoholic EEG signal using LASSO Regression with Metaheuristics Algorithms based LSTM and Enhanced Artificial Neural Network Classification Algorithms," *Scientific Reports*, vol. 14, no. 1, pp. 1-24, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[31] Neeraj et al., "Augmenting Common Spatial Patterns to Deep Learning Networks for Improved Alcoholism Detection using EEG Signals," *Computers in Biology and Medicine*, vol. 193, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[32] Supreyo Chakravorty Pretom et al., "Alcoholism Detection from Electroencephalogram Using Machine Learning and Deep Learning," *2025 International Conference on Electrical, Computer and Communication Engineering (ECCE)*, Chittagong, Bangladesh, pp. 1-6, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[33] Savarapu Chandra Shekar et al., "Detection of Alcoholic using LSTM and Enhanced ANN Classification Algorithms," *2025 6th International Conference for Emerging Technology (INCET)*, Belgaum, India, pp. 1-6, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[34] Emad-ul-Haq Qazi, Muhammad Hussain, and Hatim A. AboAlsamh, "Electroencephalogram (EEG) Brain Signals to Detect Alcoholism Based on Deep Learning," *Computers, Materials & Continua*, vol. 67, no. 3, pp. 3329-3348, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[35] Nandini Kumari, Shamama Anwar, and Vandana Bhattacharjee, "A Deep Learning-Based Approach for Accurate Diagnosis of Alcohol Usage Severity using EEG Signals," *IETE Journal of Research*, vol. 69, no. 11, pp. 7816-7830, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[36] Seffi Cohen et al., "Ensemble Learning for Alcoholism Classification using EEG Signals," *IEEE Sensors Journal*, vol. 23, no. 15, pp. 17714-17724, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[37] Sagnik De, Anurag Singh, and Ashish Kumar Bhandari, "A Novel Vision Transformer based Multimodal Fusion Approach for Clinical MDD Diagnosis Using EEG and Audio Signals," *IEEE Transactions on Computational Biology and Bioinformatics*, vol. 22, no. 6, pp. 3399-3409, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[38] Gourav Siddhad et al., "Efficacy of Transformer Networks for Classification of EEG Data," *Biomedical Signal Processing and Control*, vol. 87, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[39] Yuji Bai, and Lin Yu, "Evaluating a Deep Learning Model for EEG Categorization of Alcoholic and Non-Alcoholic Subjects," *Journal of Ambient Intelligence and Humanized Computing*, vol. 16, pp. 523-532, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[40] EEG-Alcohol, Kaggle. [Online]. Available: https://www.kaggle.com/datasets/nnair25/Alcoholics/data