SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I8P110

Volume 13 Issue 8, 150-174, August 2026

Original Article

Performance Measurements of Deep Learning-Based EEG Signal Classification Using Multi-class Seizure Type with CNN and RNN – LSTM Algorithms

G.Usha[1], K.Narasimhan[2], Karthiga R.[3]

[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.*
[3] *Electronics and Communication Engineering Department, SRM Institute of Science and Technology, Chennai, India.*

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

Received: 21 March 2026 Revised: 06 May 2026 Accepted: 26 July 2026 Published: 31 August 2026

Abstract *- The objective of this research is to characterize and categorize EEG patterns to distinguish between epileptic and healthy brain functions. By applying deep learning. To apply CNN and RNN-LSTM models to multi-classify different kinds of seizures. The EEG signal is converted to images to feed into the CNN. This study mainly examined two approaches. First, we examine transfer learning with ten pretrained networks, namely AlexNet, VGG-16, VGG-19, SqueezeNet, GoogLeNet, Inceptionv3, Densenet201, ResNet-18, ResNet-50, and ResNet-101. This research seeks to find out the best network for the job. Images are classified using an SVM after passing through the pretrained model for feature extraction. The study finds that Densenet201 with RNN-LSTM is the best model with an accuracy of 97% for classifying seizure types. CNNs capture spatial aspects in images by means of local receptive fields, hierarchical features, pooling, and non-linear activations. The extracted features comprise mean, skewness, variance, and power spectral features, and so on. Thus, the feature vector consists of 20588 features. RNNs and their variants, namely LSTM and Gated Recurrent Unit (GRU), are good at processing sequential data. The RNN feature extraction is stored as a 1,920-dimensional vector. Combining CNNs and RNNs in CRNNs enables effective extraction and analysis of key data features. The combined implementation of the LSTM Network and the CNN Network will be to classify the variants of the seizures effectively.*

Keywords *- CNNs, CRNNs, EEGs, RNN-LSTMs, Support Vector Machine.*

1. Introduction

Epilepsy is an ongoing illness of the central nervous system that affects the brain. It causes many seizures. The seizures can cause a person to lose consciousness or become unable to control their movements. They can also affect a person's behavior. Repeatedly having these seizures can negatively affect a person's quality of life. There exists a way to measure how well a person's brain functions. This measurement tool is called Electroencephalography (EEG). EEG records the electrical activity occurring within a person's brain. Clinicians use EEG as an aid in determining whether a patient has epilepsy and/or identifying patterns of seizure occurrence. While useful in this manner, the process of using EEG to analyze a patient's condition is labor-intensive, dependent upon clinician opinion, and susceptible to errors due to the continuous nature and scope of EEG readings. Machine Learning (ML) and Deep Learning (DL) technologies have recently been utilized to improve the efficiency of automatically detecting and classifying seizures from EEG readings. ML/DL technology enables rapid processing of large amounts of EEG data and may identify patterns in the data that would be difficult to observe with simple visual inspection of the data. Convolution Neural Networks (CNNs) enable the extraction of spatial information from EEG signals by converting them into time-frequency images. Recurrent Neural Networks (RNNs), specifically those utilizing Long Short-Term Memory (LSTM) architectures, are effective at identifying relationships between sequential EEG measurements.

Millions suffer from seizures or neurological disorders all over the world. Seizures occur more than once in 24 hours or several times in one day, and evidence of brain electrical activity is present. To overcome imbalanced medical data and complex computational requirements of EEG signals, here a hybrid Deep Learning (DL) approach for epilepsy seizure detection [1]. This paper proposes a new deep learning method for the analysis of no stationary, noninvasive information to the brain, with the aim of intelligent classification using EEG signals. The proposed approach appears to significantly outperform classical machine learning and recent deep learning models [2]. A joint deep learning model is proposed for automatic detection of epileptic seizure activity, identification of epileptic regions, and classification of EEG signals using time-frequency-image transformation of EEG signals [3]. Epilepsy is an abnormal brain dysfunction of the Central nervous system of the human body.

This results in seizures or loss of awareness. The EEG signal for epilepsy study requires an automatic detection system for a patient that can assist him in driving [4]. Epilepsy is diagnosed using Electroencephalogram (EEG) signals caused by abnormal electric activity in the brain. However, existing deep learning models fail due to the irregular structuring of physiological data [5]. Damage to the brain, infections, genetics, or metabolic disorders may cause seizures. There are different causes of the same thing.

Factors such as lights, lack of sleep, or stress may cause it [6]. Seizure classification is an essential part of epilepsy patient diagnosis and treatment, as it can be used to identify the causes and determine the best treatment option [7]. Epilepsy is a chronic seizure disorder characterised by recurrent, unprovoked seizures (action potentials). The occurrence of two or more unprovoked seizures identifies a person as epileptic. People of any age and background can have epilepsy. Epilepsy has many causes, and some of them cannot always be known. Some known causes are genetic factors, head injuries, infections, neonatal defects, and certain disorders.

Recently, there has been a growing focus on ML methods for classifying seizures [8]. The traditional methods of classifying seizures are slow and impressionistic. They utilize the visual analysis of an Electroencephalogram (EEG). Deep and machine learning algorithms can also be used to quickly and correctly interpret large amounts of EEG information for doctors [9]. The EEG is a noninvasive method for recording the electrical activity produced by the brain. The study of emotion distinction through EEG signals tries to understand and analyze human emotions by monitoring and analyzing the electrical activity within different areas of the human brain. EEG gives information on the electrical signals from clusters of neurons by electrodes on the scalp.

Despite these advancements, several critical challenges remain unresolved.

- Imbalanced EEG datasets often lead to biased model performance, especially in multi-class seizure classification.
- The non-stationary and complex nature of EEG signals reduces the generalization ability of existing models.
- Many models focus only on either spatial (CNN) or temporal (RNN) features, failing to effectively integrate both.
- Existing approaches are largely limited to binary classification (seizure vs non-seizure) rather than detailed multi-class categorization.
- Some clinical systems lack sufficient processing power to process deep learning models in real time; this limits their application.

To address these limitations, there is a strong need for a hybrid and efficient deep learning framework that can simultaneously extract spatial and temporal features while improving classification accuracy and generalization.

1.1. Significance of CNN and RNN in EEG Analysis

Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) are popular Deep Learning techniques that are being used for EEG signal analysis in the examination of neurological and psychiatric disorders. Automated methods of diagnosing disorders have been developed using machine learning techniques. A Deep Learning algorithm known as CNN is very efficient in image tagging jobs. The EEG signals may be viewed as an image, where each time-frequency point of the EEG signal is reflected in the intensity of the pixel. The time is denoted by the x-axis and frequency by the y-axis. CNNs have been applied to analyze EEG signals [10]. One application of CNNs in EEG research [11] is categorizing EEG data into different phases of sleep. The need for sleep is majorly for relaxation and healing purposes, and it should not be disturbed, but interrupted sleep is harmful to health. Sleep is measured in an assortment of ways. For instance, deep sleep, REM sleep, and light sleep are only a few. More specifically, these classifications are based on EEG-derived data. Signals in the frequency domain outperform those in the time domain in a Convolutional Neural Network (CNN) for detecting epileptic seizures using raw EEG data, with superior potential, according to their accuracy, which ranges from 92.3% to 97.5%. Time-domain indications performed worse, judging from 47.9% to 91.1%.

RNN is a type of Deep Learning algorithm that does well at jobs that are related to sequences. At every time point, EEG signals depend on the previous point. Nagabushanam et al. [12] achieved an accuracy of 71.38% using their DL-LSTM model, possibly because of fewer features. EEG signals were primarily analyzed using RNNs in multiple studies. A good example is identifying the disturbances due to epilepsy in the EEG. When EEG recordings are prolonged over days or weeks, manually evaluating these records is tedious and error-prone. RNNs have been applied to long EEG records for accurate recognition of epileptic seizures. The new EEG signal classification method works well. It uses several algorithms. These are SVM, logistic regression, and several versions of neural networks. These include a 2-layer LSTM and a 4-layer improved NN. They have custom activation functions. These algorithms do better on the Bonn database data when it comes to performance metrics, accuracy, precision, recall, and F1 score. Numerous simulations with different parameters verify the improved capabilities of the proposed improved NN and modified LSTM architectures [13].

1.1.1. Convolution Neural Networks

CNNs have been introduced as a deep learning method that can perform seizure classification.

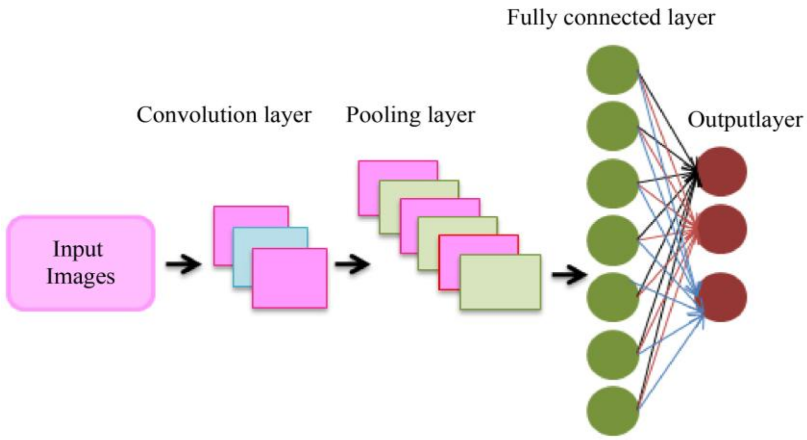

Fig. 1 CNN architecture

They use convolutions, pooling, and activation functions through multiple layers of neurons to extract features from input data. Numerous tasks of image recognition have been performed by using CNNs, including object detection, face recognition, image processing for medical use, and so forth. With the aid of CNNs, many fields have been revolutionized, particularly computer vision, through feature extraction from image pixels [14]. CNNs are amazing at figuring out what things are and where they're located in photos. This is useful in safety applications, such as unmanned vehicles, medical imaging (identifying illnesses from X-rays and MRIs), and autonomous vehicles that detect pedestrian traffic signs. (Facial recognition, object detection).

A vital component of the deep learning revolution is CNNs. By relying on convolutional and pooling layers, they can automatically learn hierarchical features from the data. By removing manual feature engineering, CNNs can learn from raw data and extract important, useful features. In traditional computer vision, feature engineering involved a tedious and lengthy process. Because CNNs automatically identify the key features in the data. CNNs can be trained to recognize specific seizures in EEG data while classifying seizures. For example, a CNN can be taught about the spike and wave patterns characteristic of absence seizures. A reliable CNN trained on a large sample of EEG recordings can create an accurate real-time model to classify seizures.

CNN consists of convolution with ReLU, batch Normalization, and a pooling layer. The last layers are fully connected (Fc), Dropout, softmax, and classification output layers. A Convolution Layer (Conv) carries out the convolution operation. According to the study, average pooling and the pooling layer help in reducing the spatial dimension of the representation [15]. As seen in Figure 1, the architecture possesses a fully connected layer, which will assist in mapping the representation of the input and output. The filters from the convolution layer can identify various patterns from the images, like edges, shapes, textures, objects, etc. Most CNNs include convolution layers where the filters (kernels) are applied to the input images to be able to detect small details like edges, colors, and textures. Convolutional Neural Networks have three types of layers, which include ConvNets. The Convolution Layer is the major component of CNNs, where the convolutions happen. The Kernel/Filter is the entity that takes responsibility for performing the convolution procedure within the layer [16]. The kernel will move across all the pixels within the image through the use of a stride rate. This means that the size of the kernel is usually smaller compared to the image, but still larger in terms of depth. For instance, if there are three channels in an image, then depth is critical, but not height and width.

The non-linear activation function is another important part of convolutional layers. Convolution algorithms and other linear methods give input to a non-linear activation function. Models using smooth nonlinear functions such as sigmoid, tanh, or some other function have been proposed to represent the behavior of actual neurons mathematically. The ReLU activation function is the one that people commonly use today. To make the number of dimensions smaller, use the Pooling Layer (POOL). This means that you don't have to do as many calculations to work with different data [17].

Some pooling techniques include the averaging pool and the maximum pool. The maximum pool is used to take the maximum value from the region covered by the kernel in the image. The averaging pool is used to take the average value of all the regions covered by the kernel in the image [18]. Every input is connected to every neuron in the fully connected layer (FC). It operates on flattened inputs. After giving the flattened vector, mathematical operation techniques are constantly performed at some further FC layers. Here, the categorization methodology begins. After CNN designs, FC layers are generally uncovered.

1.1.2. Transfer Learning Models in CNN

With slight changes and modifications, the CNN AlexNet, originally designed for image classification applications, can also be employed for signal-processing tasks. Signals from signal processing may change over time and include audio, speech, time series, and other types of signals. Here are some methods where signal processing can benefit from AlexNet Time Series Classification: You can convert time series data into a 2D form, like an image, where the x dimension is a time step and the y dimension is the value of the signal. AlexNet is used to classify different kinds of patterns or anomalies belonging to time series. It is important to mention that AlexNet was the CNN that won the ImageNet Large Scale Visual Recognition Challenge in 2012 for the very first time. Furthermore, the CNN is made up of eight layers, in which there are five convolutional layers, two fully connected layers, and an output layer [19]. That means that the colors of the image will have 3 ultra-advanced RGB colours: 227 pixel width and 227 pixel height.

The construction of GoogleNet, whose other name is Inception V1, was done by Google's research team. This was launched in 2014 and gained further popularity by winning the ImageNet Large Scale Visual Recognition Challenge.

GoogleNet brought a breakthrough in deep learning computer vision due to its advanced, innovative architecture. GoogleNet incorporates an inception module in its architecture that is capable of encompassing multi-scale features inside a single layer. The inception module applies different-sized receptive fields (1×1, 3×3, and 5×5) using different convolutional filters, as opposed to usual CNNs, which have a fixed-size receptive field. Thus, this module can extract features of varying dimensions at the same time. The network can better learn a variety of features and is more efficient in terms of computation and memory use due to this parallel processing.

GoogleNet implements 1×1 convolutions as bottleneck layers to further cut down the dimensionality of the feature maps and the workload. The 1x1 convolution reduces the number of input channels for the big convolutional filters. It is done in order to maintain the computational efficiency of the network, without having any loss of information. GoogleNet uses auxiliary classifiers, which are introduced between some layers. To avert the vanishing of gradients, these auxiliary classifiers are utilised during training. This operates the network to learn more discriminative features. However, during inference, the auxiliary classifiers are not used. As for architecture depth, GoogleNet has 22 layers, which was significantly deeper than many other CNNs when it was first released. The inception module is pretty deep, but it is efficient and uses 1×1 convolutions. As a result, GoogleNet has reasonable computational requirements.

GoogLeNet is a neural network architecture designed by Google in 2014. Also called Inception V1, the neural network uses the inception module to interpret at high levels of abstraction within a single layer and has 22 layers. The GoogleNet architecture requires an input of dimensions 224 × 224 × 3. SqueezeNet was introduced in 2016 by Stanford University, and DeepScale is a deep neural network architecture that is low-cost to deploy in resource-constrained environments. The Fire Module is a major component of SqueezeNet that uses 1x1 convolution for dimension reduction and balances between 1×1 and 3×3 convolution to achieve a good accuracy-complexity trade-off. Any standard network can add this and use it.

SqueezeNet is a model with a size smaller than AlexNet, which has similar accuracy. It is suitable for devices with limited memory and computing capacity. It provides different models, such as SqueezeNet 1.0 or 1.1, for the best use case for a user. The UC Berkeley researchers came up with the SqueezeNet neural net architecture in 2016. SqueezeNet makes use of fire modules with both a convolutional layer and a 1×1 convolution layer, reducing the number of parameters, leading to a network [20]. SqueezeNet only contains 4.8 million parameters compared to other deep neural networks, whose number of parameters is much higher. The input size of SqueezeNet is 224 × 224 × 3.

VGG16 and VGG19 are neural networks developed by the University of Oxford's Visual Geometry Group in 2014. The fully connected or convolutional layers of VGG16 or VGG19 denote that they have 16 or 19 layers, respectively. VGG16 and VGG19 require an input of 224 × 224 ×3. The Visual Geometry Group at Oxford University is a computer vision and deep learning research group known as VGG. VGG was created by professors Andrew Zisserman and Andrea Vedaldi and has made a major contribution in this field. The VGGNet architecture was introduced by the researchers after continuous participation in the ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Their success at the challenge took place during the year 2014. The VGGNet architecture is capable of image classification using deep learning because of its simple and efficient nature. In short, this model depends mostly on convolutional layers of 3×3 size. The application of VGG is not just limited to image classification but also helps in various other areas like object localization, object detection, and semantic segmentation. As a result of VGG, various models were developed. VGG has also created important datasets and benchmarks, like the VGG Face Dataset and Oxford Flowers Dataset, as well as performed extensive research in deep learning, which is instrumental in shaping this area of computer vision and deep learning research. InceptionV3, a Convolutional Neural Network (CNN) architecture developed by Google from the original Inception model (GoogLeNet), was introduced in

2015 for image classification. It boasts 48 deep layers and operates Inception modules with different convolutional kernel dimensions and 1x1 convolutions for feature scaling. To improve efficiency, InceptionV3 uses factorization to substitute large convolutions. It features auxiliary classifiers for faster convergence and includes batch normalization for improved stability [21]. This architecture, which has 48 layers, is an improvement of GoogleNet. It uses the inception module, which is like Google's neural network, to find varying degrees of generalization in the layer. Inception V3's input proportions are 299 × 299 × 3.

The ResNet architecture has different variants for its errors. ResNet has many variants like ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152. Residual networks that are wide and efficient in terms of performance. Pre-Activation ResNets reorder the operations to make the training more stable. ResNet features a cardinality parameter for grouping features. DenseNet is a good, effective use of parameters and does not have the gradient vanishing problem. It's commonly applied in computer vision tasks. ResNet models were created by Microsoft Research, and ResNet18, ResNet50, and ResNet101 are popular examples of that [22].

DenseNet was introduced in 2017 and allows attribute reuse through dense connectivity, that is, connecting each layer to all preceding layers. By applying a bottleneck layer, the depth and width of the network and image dimensions can be reduced, respectively. A parameter for the transition rate indicates the addition of feature maps. Variants of DenseNet, like DenseNet-121, DenseNet-169, and DenseNet-201, provide varied depth. DenseNet is useful in different computer vision tasks and is marked by its higher accuracy and more efficient use of parameters. ResNet has some variants, including some stretchings, for various problems. It makes use of dense blocks, where each layer is feed-forward connected to the other layers. This helps the network become more accurate while also requiring fewer parameters. DenseNet201 has 201 layers and a 224×224×3 input matrix.

1.1.3. Recurrent Neural Networks

Recurrent Neural Networks (RNNs), including the Long Short-Term Memory (LSTM) architecture [23], are another type of Deep Learning method applied for seizure classification. RNNs are constructed to process time series information and other sequential data. They are made up of neurons that can attach to one another and hence allow for responding again to some earlier input, as shown in Figure 2. The vanishing gradients problem can occur when training RNNs on long data sequences. Fortunately, there are RNN variants known as LSTMs designed to prevent this problem. For example, Long Short-Term Memory (LSTM) methods can be applied for classifying seizures based on previous observations of EEG channels. For example, an LSTM network could distinguish between the recurrent pattern activity of more complex partial epileptic seizures [24]. It will be possible to design a classifier model based on an LSTM, which can use the temporal characteristics of seizures and classify them by training on a large dataset of EEG data. Using machine learning techniques for seizure classification presents multiple advantages. First, they can quickly analyze big datasets of EEG data and help diagnose and treat patients with epilepsy. The second advantage is the identification of patterns in the data that humans cannot see, thus leading to a new understanding of the reasons behind the seizures. Overall, machine learning algorithms give doctors valuable information for treating patients.

The class of Neural Networks known as Recurrent Neural Networks (RNNs) is versatile in terms of sequential input. They excel at processing time-series data like stock prices, as well as text written in human language. The Long Short-Term Memory (LSTM) is an RNN (Recurrent Neural Network) architecture that has been developed to overcome the problem of vanishing gradient occurring when a deep neural network has many layers [25]. According to the structural design of the LSTM network, there are many memory cells that are capable of storing data over a period of time, and many gating mechanisms responsible for the entering and/or exiting of information inside a particular cell, as shown in Figure 2.

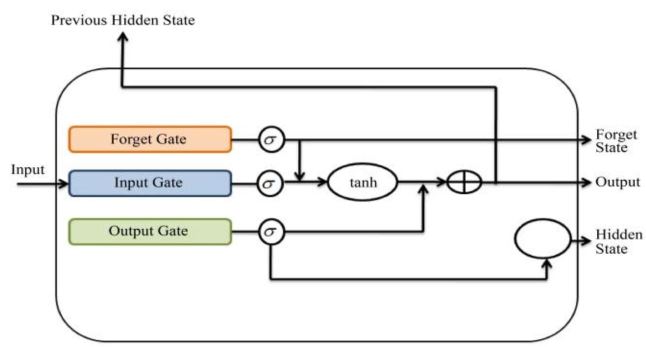

Fig. 2 LSTM architecture

Each LSTM cell is composed of three specific substances, namely the memory cell, input gate, and output gate. The gate that controls what comes out of the cell is called the output gate, which decides what goes into the cell is the cell. Gates manage the deletion or retention of specific information contained within a memory cell over a particular period.

At every time step, the LSTM network receives a vector as input, which transforms linearly. In order to create input to the LSTM cell, both the output from the previous time step and the current output are combined. The gate accepts the input late, which determines how much input is allowed to add to the state [26]. In a memory cell, the forget gate decides how much of the previous memory is allowed to be stored. The memory cell modifies its information by utilizing a set of transformations over the input and the last memory. These changes allow the cell to remember certain information or ignore other information depending on how related it is to the output. The output gate decides how much of the fresh memory will be sent out to the following timestep. The LSTM architecture is an important part of natural language processing, or speech or music, which can describe the long-term dependency of the data. The strength of LSTMs lies in their ability to capture long-range patterns and relate information temporally quickly. It does this by keeping or forgetting information.

1.2. Classification of Seizures

Classifying seizures is an important part of the diagnosis and management of epilepsy, as shown in Figure 3. The visual patterns and temporal characteristics of seizures can be categorized through the use of deep learning algorithms such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM). Through training of algorithms on very large databases of EEG data, one can create a model capable of categorizing seizures. Machine learning algorithms are expected to become very useful tools in managing epilepsy. Seizures can either be focal seizures (or partial seizures) or generalized seizures. Focal seizures are further classified into two types: those that cause the individual to remain aware and those that bring about a loss of awareness. This focal point is the key phrase that generates these specific seizures [27]. Focal seizures may cause shivering, sensation alterations, and picturesque awareness changes due to a restricted area of the brain having a wound.

Widespread electrical activity occurs in generalized seizures. Generalized seizures are further classified into tonic-clonic (grand mal or major motor seizure), absence (petit mal or minor motor seizure), myoclonic, and atonic seizures. When seizures occur from an injury or genetics, these are the symptoms that a person may experience during a seizure.

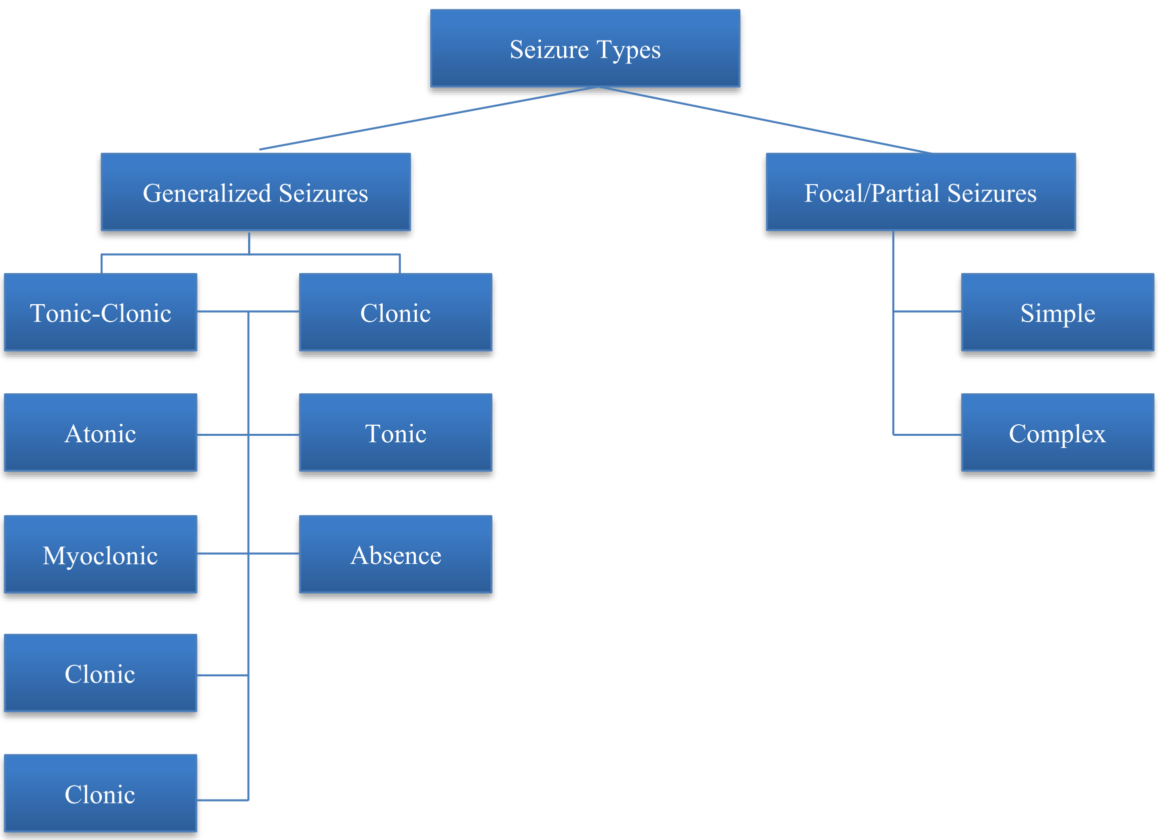

Fig. 3 Classification of seizures

The main contributions of the study are summarised as follows.

- The use of a hybrid deep-learning framework, which consists of a Convolutional Neural Network (CNN) along with Recurrent Neural Networks with Long Short-Term Memory (RNN-LSTM) to capture the spatial as well as temporal features of EEG signals.
- EEG to Image Transformation proposes a novel scheme of converting the EEG time-series data into 2D image data and facilitating the usage of advanced DL models based on computer vision.
- The evaluation of transfer learning is done using ten number of pre-trained CNN architectures they are AlexNet, VGG16, VGG19, SqueezeNet, GoogleNet, InceptionV3, DenseNet201, ResNet18, ResNet50, and ResNet101.

- Through the use of the DenseNet201 in RNN-LSTM, it was able to achieve a seizure classification accuracy of 97.2%, making it the overall best-performing model.
- The Dual-Stage Feature Extraction and Fusion first extracts 20,588 spatial features for CNNs and 1,920 temporal features for RNNs. Then, this dual feature is fused for improved representation and discrimination of seizure types.
- The Multi-Class Seizure Classification relates to the classification of 8 seizure types, which include simple partial, complex partial, focal non-specific, generalized non-specific, absence, tonic, tonic-clonic, and non-seizure classification. This is done to help overcome binary classification.
- With improved clinical relevance, the system can differentiate more than one seizure type, which in turn enhances diagnostic accuracy and clinical decision support.
- Spatial-temporal feature learning utilizes CNNs and RNN-LSTMs for the extraction of EEG spatial features as well as the modelling of EEG temporal dependencies.
- The hybrid CNN and RNN-LSTM model achieves high accuracy and generalization performance with accuracy (97.4%), indicating its effectiveness and scalability.

2. Related Works

Recent studies on EEG-based automated seizure detection demonstrate significant advancements by integrating preprocessing, feature extraction, and classification techniques using deep learning and IoT-enabled systems. These approaches offer several advantages, including improved accuracy and efficiency in detecting epileptic seizures compared to traditional manual analysis. The use of machine learning and deep learning models, such as CNNs, enables automatic feature extraction, reducing the need for manual intervention and enhancing the ability to process large volumes of EEG data [9]. Additionally, the integration of IoT in healthcare systems allows continuous remote monitoring of patients, which is particularly beneficial for real-time seizure detection and timely medical intervention. Hybrid models like the combination of Linear Graph Convolutional Neural Networks (LGCN) with DenseNet further enhance performance by effectively handling the irregular and non-stationary nature of EEG signals, while also improving feature reuse and classification accuracy.

EEG signals are inherently noisy, complex, and highly non-stationary, making them challenging for deep learning models to generalize effectively. The implementation of advanced models such as LGCN and DenseNet increases computational complexity, requiring high processing power and large datasets, which may limit their real-time applicability, especially in resource-constrained environments like IoT devices. While IoT-based monitoring systems provide continuous tracking, they also introduce concerns related to data privacy, security, and network reliability. Furthermore, CNN-based approaches, although powerful in spatial feature extraction, may fail to capture temporal dependencies in EEG signals, leading to incomplete analysis. Traditional models like multi-layer backpropagation networks, though simple, often suffer from lower accuracy and poor scalability when dealing with complex EEG data. Overall, while current approaches significantly improve seizure detection capabilities, challenges such as computational cost, data imbalance, model interpretability, and real-time deployment still need to be addressed for more robust and clinically applicable solutions.

The analysis of recent classifier architectures for EEG-based seizure detection reveals a range of advanced techniques that effectively capture both temporal and frequency-domain characteristics of brain signals. Methods based on Fast Fourier Transform (FFT), autoregressive features, and wavelet neural networks demonstrate strong capability in extracting meaningful signal representations, with approaches like DWT-based sigmoid Entropy combined with SVM achieving exceptionally high accuracy of up to 100%. Similarly, multi-spiking neural network architectures show improved performance (90.7%–94.8%) compared to single-spiking models by better mimicking biological neural behavior. Continuous neural network models also provide competitive results, achieving up to 97.2% accuracy in binary classification tasks.

Furthermore, supervised deep spiking neural networks generally outperform unsupervised methods due to their ability to learn precise labeled patterns, leading to higher classification accuracy. Techniques like FFT and wavelet transforms require careful feature engineering and domain expertise, increasing system complexity. Spiking neural networks, while biologically inspired and efficient in theory, are often difficult to train and require specialized hardware or algorithms. Additionally, supervised learning approaches demand large labeled datasets, which are not always available in medical domains [28]. The high computational cost associated with deep and supervised models further limits their deployment in real-time or resource-constrained environments. Moreover, many of these models focus primarily on binary classification, restricting their applicability in multi-class seizure detection scenarios. Overall, while these advanced techniques significantly enhance seizure detection performance, challenges related to scalability, computational efficiency, and real-world applicability remain important areas for future research.

Many studies on automatic seizure detection from EEG data have utilized various methods for extracting features and applying machine learning models to classify EEG data [29]. Log energy, norm entropy, sigmoid Entropy, matrix determinant, approximate Entropy, sample entropy, and permutation entropy are some features that have been found useful for characterizing EEG data due to its complex and nonlinear nature [30]. Using multiple features in conjunction with classifiers such as Support Vector Machine (SVM) makes the classification more reliable. Additionally, optimized Multi-Layer Perceptron (MLP) models, designed with suitable transfer functions and training algorithms, demonstrate reliable performance when evaluated on standard datasets like the Bonn EEG dataset, which provides a well-structured benchmark for comparing different methods [31]. Deep learning approaches such as ESD-LSTM further improve performance by effectively capturing temporal dependencies, with some studies reporting very high accuracies, even up to 100% [32].

Overall, these methods contribute to improved detection accuracy, better feature representation, and efficient utilization of EEG data. Many studies reporting extremely high accuracies, especially on the Bonn dataset, may suffer from overfitting due to the controlled and limited nature of the dataset, reducing their applicability in practical clinical environments [33]. Moreover, most existing works focus primarily on binary classification (seizure vs non-seizure) rather than addressing the more complex and clinically relevant problem of multi-class seizure classification. Variability in reported accuracies (ranging from 80% to 100%) also indicates inconsistency depending on feature selection, classifier choice, and experimental setup. Additionally, deep learning models like LSTM, while powerful, require large datasets and high computational resources, making real-time deployment challenging [34]. Thus, although current methods show promising results, challenges such as generalization, multi-class classification, computational cost, and real-world validation remain key areas for further research.

One major advantage of these approaches is their ability to learn directly from raw EEG signals without the need for manual feature extraction, which simplifies the processing pipeline and reduces dependency on domain expertise. CNN-based models have shown high performance, achieving accuracies around 95% for depression detection with strong sensitivity, specificity, and F1-scores, indicating reliable classification. Similarly, CNN models applied to epilepsy detection using datasets like Bonn have achieved very high accuracy (up to 99.08%), along with excellent sensitivity and specificity [35]. Another key strength is the capability of CNNs to capture complex EEG dynamics and spatial patterns, making them suitable for identifying subtle neurological abnormalities. Additionally, improvements in time-domain feature utilization reduce preprocessing efforts and enhance computational efficiency, making these models promising for real-time and clinical decision support systems [36]. The noninvasive nature of EEG, combined with automated CNN-based analysis, also supports early diagnosis and continuous monitoring. High accuracies reported on benchmark datasets such as the Bonn dataset may not generalize well to real-world clinical data due to dataset simplicity and controlled conditions, raising concerns about overfitting. Although CNN models are robust, they mostly emphasize spatial feature representation and might not adequately represent temporal relationships within EEG data unless they are incorporated with sequence-based models. In addition, the use of deep CNN models demands extensive computation and large labeled datasets to perform well, especially in medical cases. Preprocessing needs might be reduced, but noise contamination may remain a potential challenge. Deep learning algorithms also lack transparency, thus reducing confidence when making critical decisions. Therefore, while CNN-based EEG analysis offers high accuracy and strong potential for clinical use, challenges related to generalization, interpretability, and real-world deployment remain important areas for further improvement.

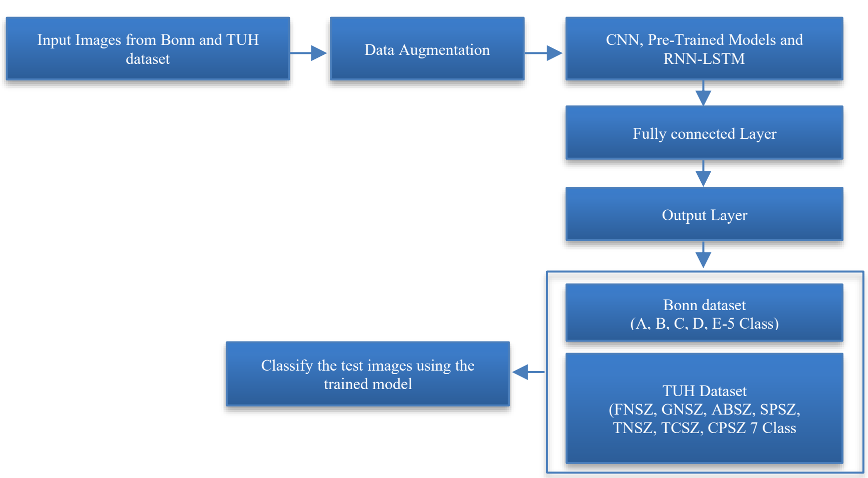

Fig. 4 The schematic visualization of the proposed methodology

3. Methodology

As raw EEG signals do not have stability and low resolution ability, this signal is recorded using multiple electrodes from the scalp. The EEG signals also contain some artifacts and noise that prevent our analysis from getting any useful information. Therefore, it is necessary to perform detection and elimination of artifacts as they can hamper our analysis process, especially feature extraction. In this regard, band-pass filters with the high-pass filter at 70Hz and the low-pass filter at 0.5Hz, along with a notch filter at 50Hz, are used to avoid any power line noise in the raw signal. After the removal of noise and artifacts, the cleaned EEG signal is converted into an image form. CNN takes only 2D input. Therefore, it is necessary to convert it into an image for classification using a CNN algorithm. The issue of overfitting caused by image misbalancing can be solved by using augmentation. The schematic diagram depicts the proposed enhancement method of Figure 4.

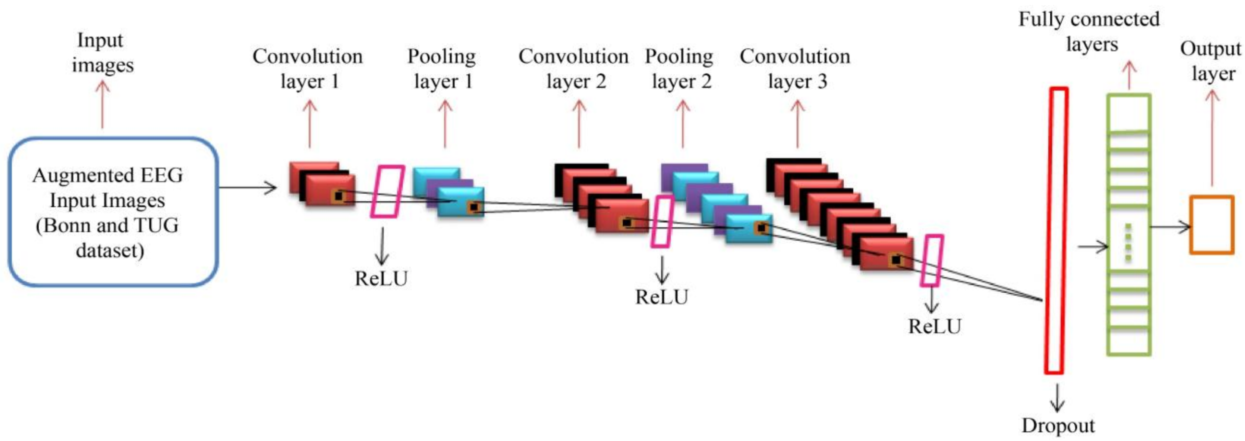

Fig. 5 The proposed CNN architecture

Generally, the ratio of train to test data is made at a rate of 80:20. The concept of transfer learning is a type of machine learning in which a learned algorithm is used to solve another problem using an existing algorithm for one specific task.

In this method, prior knowledge is used from a prior task to enhance the result of a new task. Pre-trained deep learning algorithms are trained on massive datasets and can be applied to solve tasks like object detection and NLP. For example, CNN is generally used for image-related problems, and RNN for sequence-related problems.

3.1. Convolution Neural Network (CNN)

In a CNN, there is convolution, pooling, batch normalization, and an activation function like ReLU and Tanh. These layers include the output layer, fully connected layer, dropout layer, softmax layer, and classification layer. The overall architecture of a CNN is depicted in Figure 5. Max pooling is used to minimize the spatial dimension of an image. The fully connected layer allows for mapping input and output representations. Convolution Neural Networks (CNNs) consist of eight layers. Convolution processes are executed by the CNN Convolution Layer, which serves as its base. The Kernel or Filter is the major element in this layer that performs the convolution operation.

The convolution layers contain a nonlinear activation function that takes the output from linear methods, such as convolution. However, there are models that utilize smooth nonlinear functions to represent neuronal activity. Such nonlinear activation functions include the sigmoid and hyperbolic tangent (tanh) functions. The Rectified Linear Unit (ReLU) has been found to be the most popular activation function. The pooling layer serves to reduce the dimensionality of the model. This approach may also help to reduce the computational effort required to process the information. The flattened input vectors (embeddings) are fed into the fully connected layers of neurons. Mathematical functional techniques are always applied to several fully connected layers, beyond the flattened vector. The parameters used in the model are listed in Table 1.

Table 1. CNN training parameters

Parameter	Value
Dataset	TUH EEG Seizure Corpus (5 classes: FNSZ, CPSZ, GNSZ, ABSZ, TNSZ)
Input Shape	(1 × 224 × 224) spectrogram patches
Optimizer	Adam
Learning Rate	0.0001
Batch Size	32
Epochs	50
Loss Function	Cross-Entropy Loss
Scheduler	Reduce LR On Plateau(patience=5)
Augmentation	Time-shift, Gaussian noise, SpecAugment
Validation Split	80/20
Class Imbalance Handling	Weighted Random Sampler + Focal Loss
Hardware	NVIDIA GeForce GTX 1650 (4GB)

3.2. Transfer Learning Models

The layers of the pre-trained model's learned features provide useful data representations. Edges, textures, or semantic interpretations in the case of photos or semantic comprehensions in the case of text are just a few examples of the data structure these features might capture. Many software prediction models are developed using a machine learning algorithm. There is a final result of such training. Add additional neurons to the output layer or alter parts of the current layers to address the new issue. To increase performance, fine-tune the adapted model on your target dataset. This entails changing the network's weights while maintaining or employing a slower learning rate for the initial weights (acquired from the initial task). The model can be fine-tuned to adapt its representations to the new task without losing sight of the lessons it learned from the previous task. The flow chart for the transfer learning procedure is depicted in Figure 6. In this work, utilized Alexnet, Googlenet, Squeezenet, InceptionV3, VGG-16, VGG-19, Resnet18, Resnet150, Resnet 101, and Densenet 201.

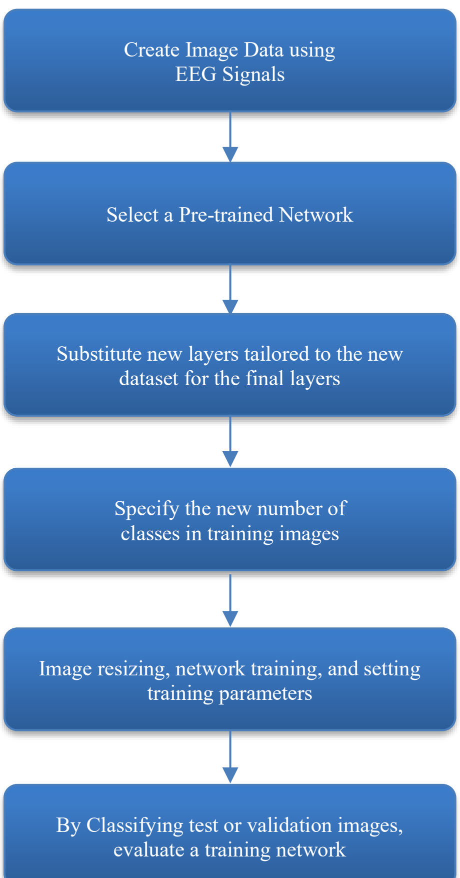

Fig. 6 Flow chart of transfer learning using a pre-trained network

Deep learning often requires a neural network that is pre-trained for purposes such as image classification, object recognition, and natural language processing. This refers to making a trained neural network model fit for a specific task or dataset. Pick a Trained Model: Select a neural network model that has been trained and has been trained on a large and diverse dataset. People often choose VGG, ResNet, Inception, MobileNet, BERT (for NLP workloads), etc., as their models. The initial training objectives of these models taught them rich feature representations. The output layer of a pre-trained model is often particular to the original task that it was trained on, so it should be removed or modified. The freezing of pre-trained model layers can be done depending on the availability of training data for the target task. This implies that those layers' weights remain frozen during the training process. When the target domain contains few samples of data, the learned features during learning of the model do not change or modify. As a result, the pre-trained layers are frozen.

Transfer learning in deep learning often includes training on a pre-trained neural network model, particularly in domains like image classification, object recognition, and NLP. This technique requires training of a pre-trained neural network model for a particular problem. In the transfer learning approach, the use of an SVM classifier is done using the representation of features obtained from a pre-trained network. Using feature extraction is the easiest and fastest way to leverage the symbolic power of a deep network that has been previously trained, as it only requires one pass of the input. The SVM classifier was used to extract and classify features from each layer of all ten pre-trained networks. Earlier levels of CNNs extract low-level features. Deeper layers build high-level traits using the lower-level building blocks of previous layers. Earlier layers generally utilize fewer, not deep, features with high spatial resolution and a lot of activation. The first step involves loading the image data. Once the loading is done, the test set gets processed and generates predictions. The RNN-LSTM model makes predictions using CNN features extracted from all test images. After that, the number of accurate forecasts gets modified and is summed up. Ultimately, accuracy refers to the ratio of the number of correct predictions to the total number of predictions multiplied by 100.

Some of the popular models you can use include VGG, ResNet, Inception, MobileNet, BERT, etc. The first training tasks of these models taught them rich feature representations. The first task for which the pre-trained model's output layer was trained is specific to the task; hence, it must be removed. Depending on the amount of data that you have for your target task, you can freeze part or all of the layers in the pre-trained model. Freezing prevents modifications in the weights of those layers. This is particularly useful when the target dataset is constrained because freezing the pre- trained layers will lock in the important features mapped during the pretraining process. Test images are classified using the trained model. Subsequently, precision would be determined and reported. Why do we utilize convolutional neural networks for image recognition? The CNNs are utilized for image recognition, object detection, segmentation, and detection of objects in images. This is because CNNs are able to extract the hierarchical.

3.3. RNN-LSTM

RNNs find applications in predictions such as natural language processing, speech recognition, and time series prediction [37]. This is because RNNs utilize the hidden state to store information from past time steps. Thus, RNNs are able to identify sequence dependence in the data, thereby helping with predictions on future outcomes. RNNs are capable of capturing time and context for sequences. Through an RNN, the sequentially of each word is understood, for example, in text generation, it generates meaningful sentences. RNNs can work with sequences of different lengths, which makes them a good option for text of varying length. The RNN architecture is shown in Figure 7.

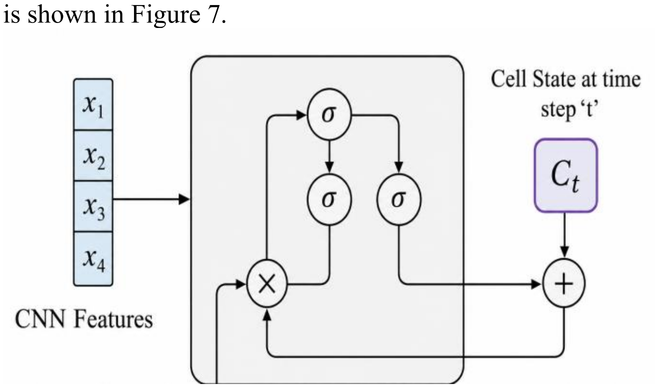

Fig. 7 RNN-LSTM architecture

LSTM and GRU networks are more sophisticated versions of RNN networks that overcome the issue of the vanishing gradient problem and enable RNN networks to learn dependencies over a longer sequence range. The extracted features are then kept in the 2020 feature vector structure. Suppose ct represents the input sum at two time points t and t-1; in that case, the activation vector of the input gates is represented by the formula below,

$$i_t = \sigma (W_{yi} y_t + W_{hi} h_{t-1} + W_{si} S_{t-1} + b_i) \quad (1)$$

where,

r - Sigmoid functions,
W - Weights matrix
x_t - input at time t,
h_{t-1} - hidden-state vector of the previous time
bi - input bias vectors
In addition, the forget gate could be written as,

$$f_t = \sigma (W_y f_t + W_{hf} h_{t-1} + W_{sf} S_{t-1} + b_f) \quad (2)$$

The memory cell value c_t is shown by the following equation,

$$C_t = i_t \tanh (W_{ys} y_t + W_{hs} h_{t-1} + b_s) + f_t S_{t-1} \quad (3)$$

The activation value of the output-gate is,

$$O_t = \sigma (W_{yo} y_t + W_{ho} h_{t-1} + W_{so} S_{t-1} + b_s) \quad (4)$$

Memory cell output is given by,

$$f_t = O_t \,.\, \tanh (C_t) \quad (5)$$

The input image was analyzed using two separate hidden layers and a bidirectional RNN-LSTM to acquire additional features based on structural design. The forward and backward activation values of the hidden layer are represented by hf_t and hb_t, respectively, and the output layer will be updated as follows:

$$y_t = (W_{fz} y_t + W_{ho} + W_{bz} + b_y) \quad (6)$$

where W_{fz} and W_{bz} are indicated as the forward and backward weight matrices, and they indicate the hidden vectors. LSTM was used to allocate the locations of facial expressions. Compared to regular neural networks, LSTM offers more precise estimation points. LSTM was employed to gather the initial expression locations in the first cascade stage. A components-based search strategy was employed in the second stage, and an RNN was applied in the third stage to improve the outcomes.

4. Results and Discussions

This part of the research will check the performance of the given method on the Bonn and TUH datasets by using CNN transfer learning models. Alexnet, Googlenet, Squeezenet, InceptionV3, VGG-16, VGG-19, Resnet18, Resnet150, Resnet 101, Densenet 201, and RNN-LSTM are all good models. To verify the proposed model's effectiveness, the results of the proposed model on the Bonn and TUH datasets are compared with some existing methods.

This experimental environment is as follows:
Operating System: Windows 10
GPU Available: NVIDIA GeForce GTX 1650
GPU Memory: 4.00 GB
Platform: Python

4.1. Dataset Description

The source of the data used in this research paper is the University of Bonn in Germany and the TUH epileptic EEGs. This database comprises EEG data obtained both from normal persons and from those suffering from epilepsy. This data set will help in seizure studies. The EEG data are divided into five different data sets, namely Set A-E [2]. All these subsets comprise 100 segments of the same kind. Each data segment contains 4097 time series of EEG data. Each data segment lasts for 23.6 seconds, and the sampling frequency of each such segment is 173.61 Hz. An intensive preprocessing technique has been employed in order to obtain the quality data desired, that of manually filtering the data between 0.53 and 40 Hz. Set A and Set B subsets of the EEG data are recorded from five healthy subjects. The electrode positions in Set A are on the scalp in an open eyes state, whereas those in Set B are acquired in a closed eyes state. These subsets represent the specific EEG patterns associated with these different visual states.

Set C and Set D subsets comprise EEG data from five epilepsy patients during interictal periods. In Set C, the electrode positions were located contralaterally to the lesion area, whereas Set D's electrodes were positioned within the lesion area. These subsets aim to capture the variations in EEG patterns from different regions surrounding the epileptic focus. Finally, Set E entails EEG data from five epilepsy patients during ictal periods, providing insights into the dynamics of EEG activity during seizures. The electrode position in this subset was situated within the lesion area. This meticulously curated dataset from the University of Bonn encompasses various subsets, each capturing unique EEG patterns under specific conditions. This dataset offers a valuable foundation for studying epilepsy-related phenomena, enabling researchers to explore and analyze the intricate dynamics of EEG signals in healthy and epileptic states, as shown in Table 2, and the sample spectrogram image is shown in Figure 8.

On the contrary, TUH EEG Seizure Corpus (TUSZ), funded by Temple University Hospital, United States, is the largest publicly available clinical EEG database created till date and has multi-channel scalp EEG recordings from various patients from an actual hospital setting. The dataset TUSZ consists of seizure and non-seizure EEGs having 250-512 Hz and has been nicely compiled in terms of different types of seizure disorders, like focal non-specific, generalized non-specific, absence, tonic, tonic-clonic, simple partial, and complex partial seizures. Another difference in the TUH set of data, as compared to the Bonn dataset, is that it contains real-life disturbances like artifacts due to movements of the patient and noise that is generally present in the medical world. While the TUH data is less well-defined and more varied than the Bonn data, it is arguably more effective in assisting with clinical requirements. However, it is significantly more difficult to work with than the Bonn dataset because of its wide-ranging features.

Comparatively, the Bonn dataset is ideal for controlled experimental validation and feature exploration, while the TUH dataset [38] is more suitable for developing robust, generalizable deep learning models that can handle clinical noise and multi-class seizure classification. Together, these datasets provide complementary advantages, with Bonn supporting model prototyping and TUH enabling clinical-grade performance evaluation. The comparative analysis is shown in Table 4. The TUH spectrogram image is shown in Figure 9.

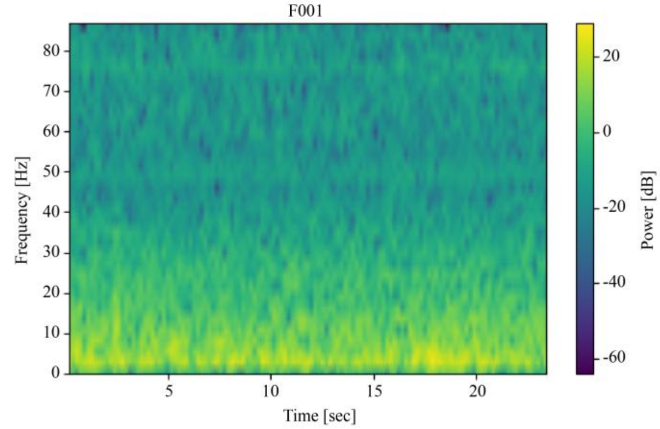

Fig. 8 Spectrogram image from Bonn dataset

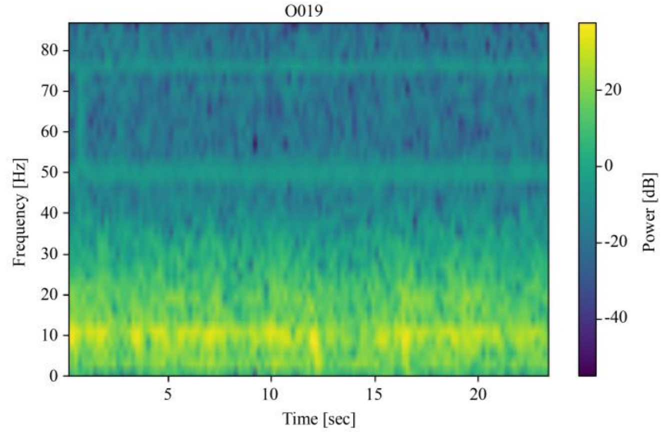

Fig. 9 Spectrogram image from TUH dataset

Table 2. Bonn dataset description

Set	Label	Description	Subject Condition	Recording Region
Set A	Z	EEG signals from healthy subjects (eyes open)	Healthy (no epilepsy)	Scalp EEG
Set B	O	EEG signals from healthy subjects (eyes closed)	Healthy (no epilepsy)	Scalp EEG
Set C	N	EEG signals of epileptic patients may be especially suited to diagnosing the epileptogenic zone.	Epileptic (interictal)	Intracranial EEG
Set D	F	EEG signals from epileptic patients during the non-seizure time span.	Epileptic (interictal)	Intracranial EEG
Set E	S	EEG signals from epileptic patients during seizures	Epileptic (ictal)	Intracranial EEG

Table 3. TUH dataset description

Category	Number of Classes	Class Labels / Abbreviations	Description
Main Classes	2	1. Seizure 2. Non-Seizure	Binary classification is used for seizure detection systems.
Seizure Subclasses (Expanded Classification)	8	1. FNSZ – Focal Non-Specific Seizure 2. GNSZ – Generalized Non-Specific Seizure 3. CPSZ – Complex Partial Seizure 4. SPSZ – Simple Partial Seizure 5. ABSZ – Absence Seizure 6. TNSZ – Tonic Seizure 7. TCSZ – Tonic-Clonic Seizure 8. ATSZ – Atonic Seizure	Multiclass classification is used for seizure type identification.

Table 4. Comparative analysis of Bonn and TUG datasets

Feature	Bonn EEG Dataset	TUH EEG Corpus (Temple University Hospital EEG)
Institution	University of Bonn, Germany	Temple University Hospital, USA
Total Number of Signals / Recordings	500 EEG segments (100 in each of 5 sets: Z, O, N, F, S)	Over 30,000 EEG recordings from clinical sessions
Number of Subjects	10 (5 healthy + 5 epileptic)	>15,000 patients (diverse demographics)
Number of Channels per Recording	1 channel (single-channel EEG)	19–36 channels depending on montage (10–20 system)
Signal Duration	23.6 seconds per segment	20–60 minutes per session (some longer)
Sampling Frequency	173.61 Hz	250–512 Hz (varies by session)
Types / Classes of Signals	5 types: - Z: Healthy (eyes open) - O: Healthy (eyes closed) - N: Interictal (epileptogenic zone) - F: Interictal (opposite hemisphere) - S: Seizure (ictal)	Multiple classes: - Seizure vs Non-seizure - Sleep stages, pathology annotations - Multiple event types (spikes, slowing, artifacts)
Signal Type Nature	Clean, preprocessed EEG (artifact-free)	Real-world clinical EEG (contains artifacts, multiple pathologies)
File Format	.mat or .txt	.edf
Usage Focus	Academic benchmarking, seizure detection	Large-scale clinical research, deep learning training
Accessibility	Freely available (open)	Open access with registration (via TUH EEG portal)

4.2. Performance Analysis

The performance of the architecture was assessed using accuracy, precision, recall, and F1 score. The following mathematical formulas can be used to compute accuracy, precision, and recall.

$$Accuracy = \frac{TP\_total}{(TP\_total + FP\_total + FN\_total)} \quad (7)$$

$$Precision = \frac{TP\_class}{(TP\_class + FP\_class)} \quad (8)$$

$$Recall = \frac{TP\_class}{(TP\_class + FN\_class)} \quad (9)$$

$$F1\ score = 2\ x\ \frac{(precision\ x\ recall)}{(precision + recall)} \quad (10)$$

Each confusion matrix in the figure comprehensively summarizes the model's classification results. It delineates the following key elements:

- True Positives (TP): Entries in the matrix that correspond to the correct classification of a particular seizure type. These cells are positioned on the diagonal line of the matrix.
- False Positives (FP) refer to the instances where the model predicted a seizure type when the actual class was a non-seizure or a different seizure type. These cells occur outside the diagonal line.
- True Negatives (TN): Cells in the matrix that are accurately predicted as non-seizure instances are true negatives. The cells are present in the upper left corner of the matrix.

- False Negatives (FN) refer to cases where there have been wrong predictions of non-seizure as seizure or seizure type misclassification by the model. These cells are also found outside the diagonal line.

By using the confusion matrices together, we can provide a deep analysis of the performance of all CNN architectures across different seizure types. They give us information on which seizure types are repeatedly misclassified, which classes are predicted most precisely, and whether there is any confusion between any of the seizure types.

This research looks at a CNN transfer learning model. It also looks at the RNN–LSTM model performance in identifying human abnormalities. Performance similar to RNN – LSTM transfer learning models.

4.2.1. CNN Performance Analysis

ResNet50 performance assessment was conducted on Bonn and TUH EEG datasets using standard performance measures of accuracy, F1 score, recall, precision, and AUC (Area Under Curve). The accuracy of the model is 92.2% for the Bonn dataset and 90% for the TUH dataset. The model had very strong performance on artificial and real EEG data.

The F1 scores for Bonn (90.94%) and TUH (88%) indicate that precision and recall are in a good balance. The recall scores of Bonn with 91.02% and that of TUH with 87% show that the model is quite sensitive. The high precision values of 90.87% (Bonn) and 88% (TUH) also indicate effective false-positive minimization. The AUC scores, 97.21% for Bonn and 93% for TUH, are another measure listed in Table 5 that confirms the robustness and discriminative capability of the ResNet50 model.

The results given in Figures 10(a) and 10(b) indicate that, although ResNet50 performs a little better on the clean and artifact-free Bonn dataset, it still maintains strong generalization performance on the more complex and noisier TUH dataset in Figure 11. Thus, ResNet50 proves to be reliable for EEG-based seizure classification tasks.

Table 5. Metrics analysis using CNN

Evaluation Metrics	Bonn Dataset	TUH dataset
Accuracy (%)	92.2	90
F1 Score	90.94	88
Recall	91.02	87
Precision	90.87	88
AUC	97.21	93

Ablation Studies

Table 6 summarizes the experimental setup and performance results of the ablation studies on the University of Bonn dataset. The total epochs used for training the model were 30 epochs with a batch size of 64 and a learning rate of 0.001 using Adam.

Table 6. Ablation studies using the Bonn dataset

Parameters	Corresponding Values
Time Taken	135 minutes
Number of Epochs	30
Batch Size	64
Learning Rate	0.001
Hardware Resource	NVIDIA GPU (mixed precision float16)
Validation Frequency	10
Activation Function	ReLU
Optimizer	Adam
Convolution Layer 1 Filters	96
Convolution Layer 2 Filters	256
Convolution Layer 3 Filters	384
Dropout	0.5
Convolution Kernel Size	11 (layer 1), 5 (layer 2), 3 (layer 3)
Max Pooling Layer	3×3
Validation Accuracy	92.2

A ReLU activation function introduced non-linearity. A dropout value of 0.5 was used to avoid overfitting the model. The CNN design featured three convolutional layers, which made use of filters with sizes 96, 256, and 384. The kernel sizes employed by these layers were 11, 5, and 3. The convolutional block included 3×3 max pooling to downsample the feature maps. Using an NVIDIA GPU with mixed precision computation (float16) has sped up the training. Validation was carried out every 10 raisings with an accuracy of 92.1 percent. It shows very good generalization and model optimization. As a conclusion, the above-listed parameters denote successful model tuning of the deep learning model designed for EEG classification on the Bonn dataset.

4.2.2. Transfer Learning Models Performance Analysis

Tables 7 and 8 list the performance metrics, including accuracy, F1 score, recall, precision, and Area Under the Curve (AUC), for pretrained models on the Bonn and TUH EEG datasets. Moreover, the confusion matrix, accuracy plot, and performance analysis of predicted outcomes using the Pretrained model on the Bonn and TUH EEG datasets are presented in Figures 12(a), (b), and 13. When the Dropout was removed from the AlexNet, its performance fell by 3.5%, with the theta (4–8 Hz) band being the most adversely affected. That indicates Dropout reduced overfitting by providing regularization.

4.2.3. CNN – RNN LSTM performance Analysis

The combination of the DenseNet201 model along with RNN-LSTM gives an accuracy of 86% on the Bonn data set and 92.5% on the TUH data set. It shows that the hybrid framework has a superior ability to learn the features.

Higher AUC values further confirmed the strong classification reliability for both controlled and clinical EEGs. On the Bonn dataset, the highest accuracy (86%) and AUC (91%) were achieved using DenseNet201, which was then followed by the InceptionV3 model and GoogleNet. The Higher AUC values further confirmed the strong classification reliability for both controlled and clinical EEGs. On the Bonn dataset, the highest accuracy (86%) and AUC (91%) were achieved using DenseNet201, which was then followed by the InceptionV3 model and GoogleNet. The experiment shows that deeper networks with dense connections and multi-scale features can effectively model EEGs in clean, noise-free conditions.

Ablation Studies

An ablation study was conducted in order to examine the significance of certain architectural aspects of different deep learning architectures utilized for the task of EEG classification. The aim of the study is to examine the influence of layer elimination or modification of their structure on model performance on different EEG frequency bands, including Delta, Theta, Alpha, Beta, and Gamma bands.

When the Dropout was removed from the AlexNet, its performance fell by 3.5%, with the theta (4–8 Hz) band being the most adversely affected. That indicates Dropout reduced overfitting by providing regularization. When Batch Normalization was removed from VGG16, we observed a drop of 2.8% in Alpha (8-13 Hz) performance, indicating that features were lost and convergence was unstable. Decreasing the number of Conv5 layers in VGG19 caused a 3.1% drop in the Beta (13-30 Hz) band, but it was minor and attributable to a less pronounced feature representation. The variations are tabulated in the following Table 9.

When the network depth of SqueezeNet was reduced, it caused the largest drop of 4.2 percent. Most of the drop in backbone performance was in the Delta (1–4 Hz). The dip in performance was mainly due to the weakening of activation and the disruption of the hierarchical feature extraction mechanism. When the Inception 3a was deleted from the GoogleNet architecture, there was a detectable drop of 2.6%, particularly in the Alpha band, confirming that multi-scale feature aggregation is necessary for the electroencephalography EEG representation.

When we disabled the auxiliary classifier in Inception v3, the decrease was only 2.1% and mostly in the Beta band, suggesting that the auxiliary classifier helps in stabilizing the learning of intermediate features. When regularization or depth was left off, lightweight architectures, including AlexNet and SqueezeNet, saw a greater performance drop off. Removing DenseBlock4 decreased the performance of the DenseNet201 model by 3.4% in the Gamma (30–50Hz) band due to the disruption of feature reuse and gradient flow, which caused poor generalisation. Finally, ResNet18 without its skip connections experienced a 4.0% decrease in the Alpha band, indicating the onset of vanishing gradients and deterioration of feature propagation. In summary, each architectural component plays a major role in stabilizing the EEG features and increasing the classification accuracy, as highlighted by the ablation study.

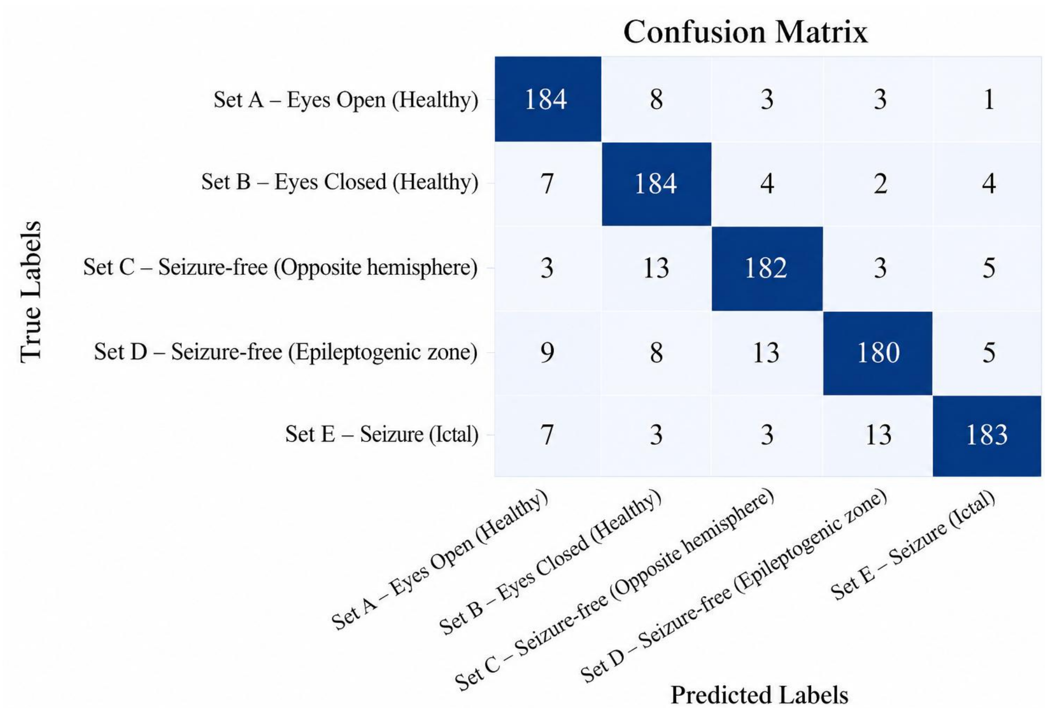

Fig. 10 (a) Confusion matrix and predicted outcomes using the CNN Bonn Dataset

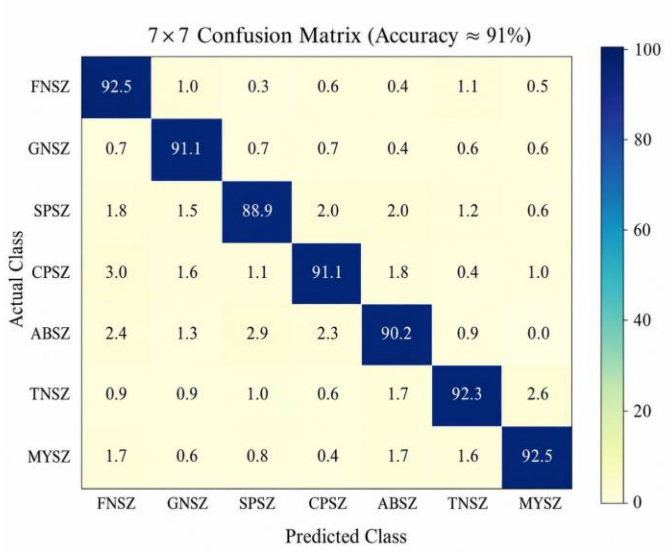

Fig. 10 (b) Confusion matrix and predicted outcomes using CNN TUH Dataset

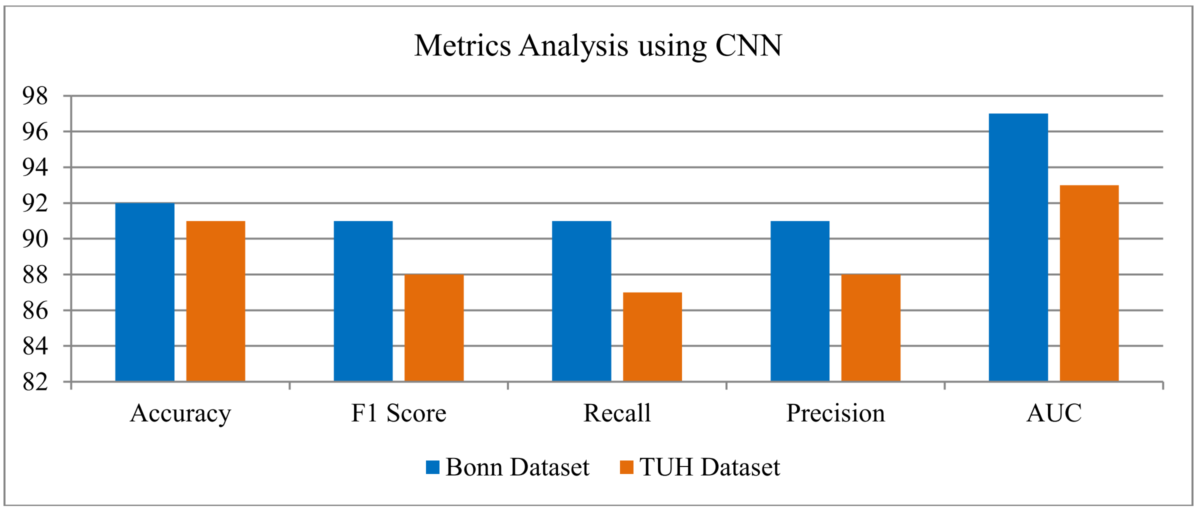

Fig. 11 Metrics analysis using CNN

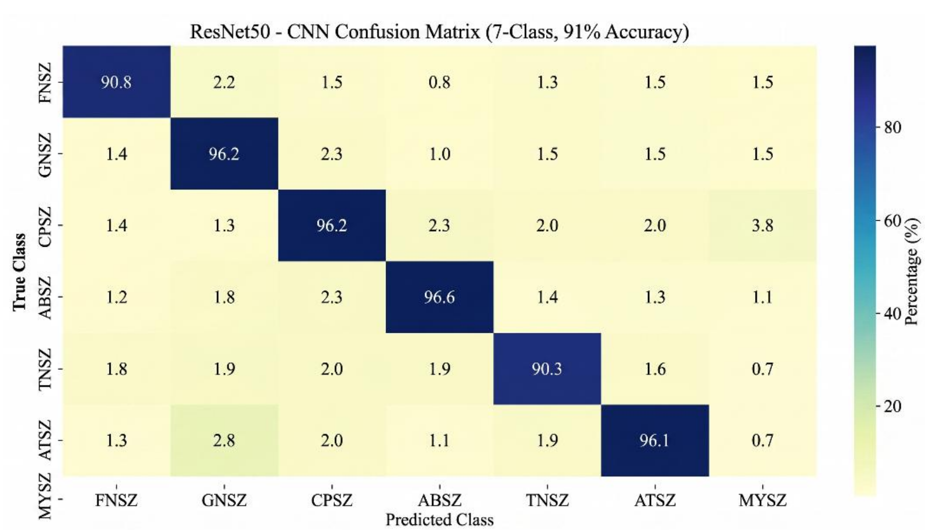

Fig. 12 (a) Confusion matrix and predicted outcomes using the pre-trained model of bonn

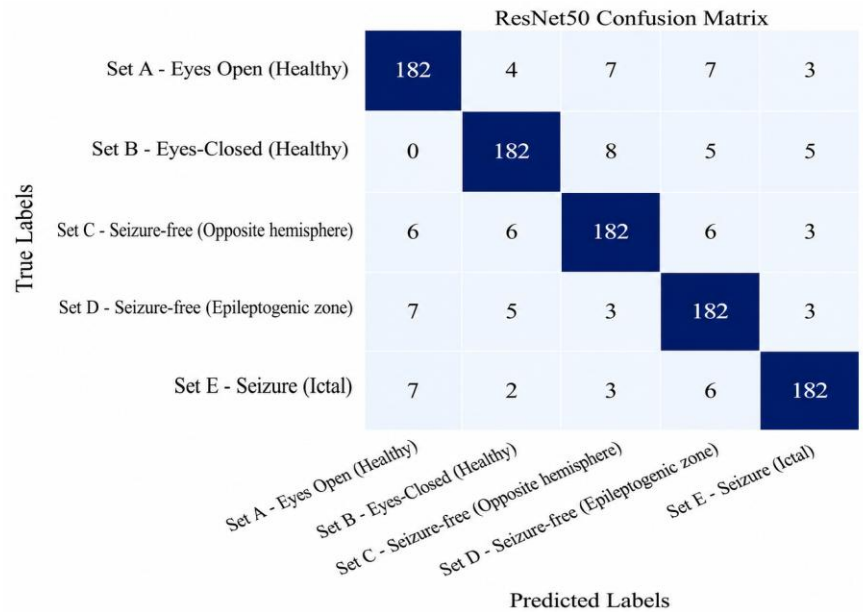

Fig. 12(b) Confusion matrix and predicted outcomes using the pre-trained model of TUH

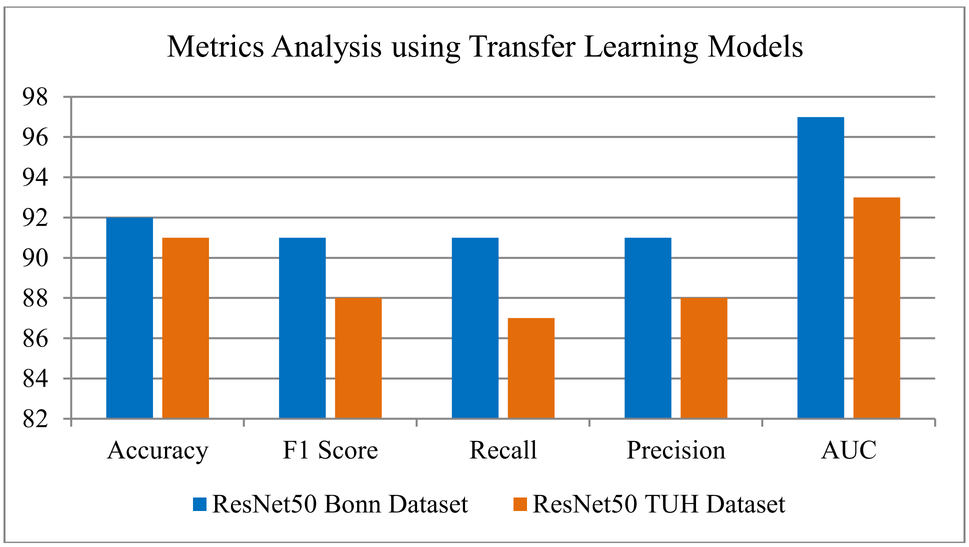

Fig. 13 Metrics analysis using transfer learning models

Table 7. Metrics analysis using pre-trained models (Bonn Dataset)

Architecture	Bonn Dataset				
	Accuracy %	**F1 Score**	**Recall**	**Precision**	**AUC**
AlexNet	91	91	91	91.44	99.61
VGG16	71	71	71	71.68	93.8
VGG19	69	67.92	69	72.71	92.84
SqueezeNet	70	70.61	70	72.11	91.65
GoogleNet	79	78.52	79	78.55	95.76
InceptionV3	79	78.8	79	78.93	96.35
DenseNet201	85	84.74	85	85.42	98.54
ResNet18	88.45	88.05	88.18	87.92	95.63
ResNet50	91.23	90.94	91.02	90.87	97.21
ResNet101	89.67	89.32	89.41	89.23	96.34

Table 8. Metrics analysis using pre-trained models (TUH Datasest)

Architecture	TUH Dataset				
	Accuracy %	**F1 Score**	**Recall**	**Precision**	**AUC**
AlexNet	81.2	79	78	80	86
VGG16	84.9	83	82	84	89
VGG19	85.3	84	83	85	90
SqueezeNet	78.5	76	75	77	83
GoogleNet	86.4	85	84	86	91
InceptionV3	87.9	86	85	87	92
DenseNet201	89.6	88	87	89	93
ResNet18	85.1	83	82	84	89
ResNet50	89.1	88	87	88	93
ResNet101	88.7	87	86	87	92

Table 9. Ablation studies using the TUH dataset

Model	Module Removed	Accuracy Variation (Δ)	Target Band (Hz)	Effect
AlexNet	Dropout removed	-3.5	4–8 (Theta)	Overfitting Increases
VGG16	Batch Normalization removed	-2.8	8–13 (Alpha)	Feature loss
VGG19	Conv5 reduced	-3.1	13–30 (Beta)	Minor instability
SqueezeNet	depth reduced	-4.2	1–4 (Delta)	Reduced activation
GoogleNet	Inception 3a removed	-2.6	8–13 (Alpha)	Performance dip
Inception v3	Aux classifier off	-2.1	13–30 (Beta)	Slight decrease
DenseNet201	DenseBlock4 removed	-3.4	30–50 (Gamma)	Poor generalization
ResNet18	Skip connection removed	-4.0	8–13 (Alpha)	Vanishing gradients

Table 10. Metrics analysis using CNN and RNN-LSTM

Evaluation Metrics	Bonn Dataset	TUH dataset
Accuracy	86	97.0
F1 Score	85	97.1
Recall	86	96.9
Precision	85	97.3
AUC	91	98.5

Table 11. Metrics analysis using pre-trained models with RNN-LSTM using the Bonn dataset

Architecture	Bonn Dataset				
	Accuracy %	**F1 Score**	**Recall**	**Precision**	**AUC**
AlexNet	81	79	80	78	87
VGG16	89	82	83	81	88
VGG19	82	81	81	82	87
SqueezeNet	78	77	76	78	83
GoogleNet	84	83	84	82	89
InceptionV3	91	84	85	83	90
DenseNet201	97	85	86	85	91
ResNet18	80	79	80	79	86
ResNet50	82	81	82	81	87
ResNet101	83	82	83	82	88

Table 12. Metrics analysis using pre-trained models with RNN-LSTM using the TUH dataset

Architecture	TUH Dataset				
	Accuracy %	F1 Score	Recall	Precision	AUC
AlexNet	85.8	84	83	85	89
VGG16	88.1	86	86	87	91
VGG19	88.7	87	87	88	92
SqueezeNet	82.4	80	80	81	86
GoogleNet	89.2	88	87	89	92
InceptionV3	90.8	89	89	90	93
DenseNet201	97.0	97	96.9	97.3	98.5
ResNet18	88.3	86	86	87	91
ResNet50	91.4	90	89	90	93
ResNet101	91.9	90	90	91	94

Table 13. Ablation studies: the proposed RNN-LSTM model using the Bonn dataset

Parameters	Corresponding Values
Time Taken	360 minutes
Number of Epochs	25
Batch Size	128 (GPU) / 64 (CPU fallback)
Learning Rate	0.001
Hardware Resource	NVIDIA GPU (mixed precision float16)
Validation Frequency	10
Activation Function	ReLU
Optimizer	Adam
LSTM Layers (Small)	128 → 64
LSTM Layers (Medium)	256 → 128 → 64
LSTM Layers (Large)	128 → 128 (residual) → 256
Dropout	0.2 – 0.5 (depending on model size)
Dense Layers	32 → num_classes (Small), 64 → num_classes (Medium), 128→64→num_classes (Large)
Batch Normalization	After every LSTM layer
Average Validation Accuracy	0.90

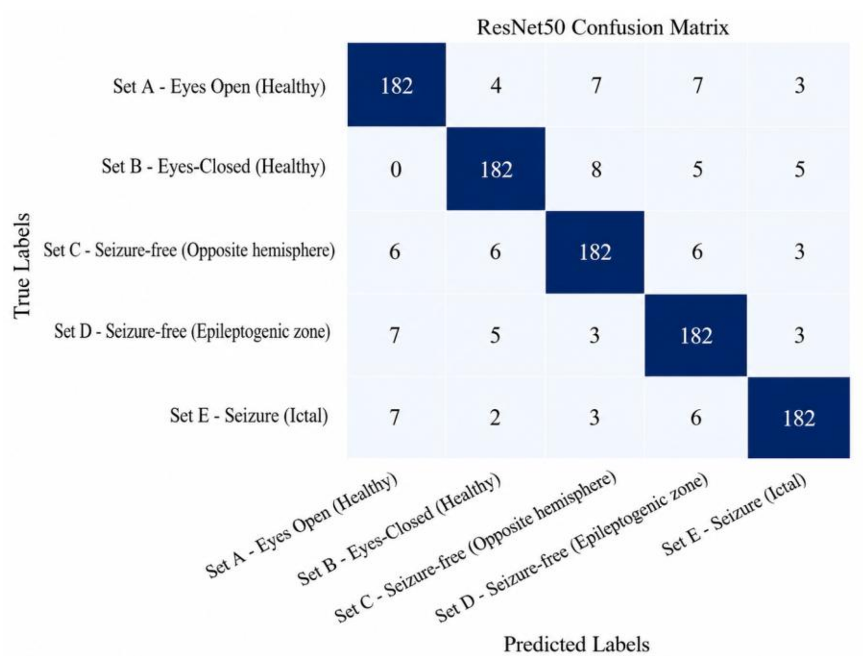

Fig. 14 (a) Confusion matrix and predicted outcomes using the pre-trained model using the Bonn dataset

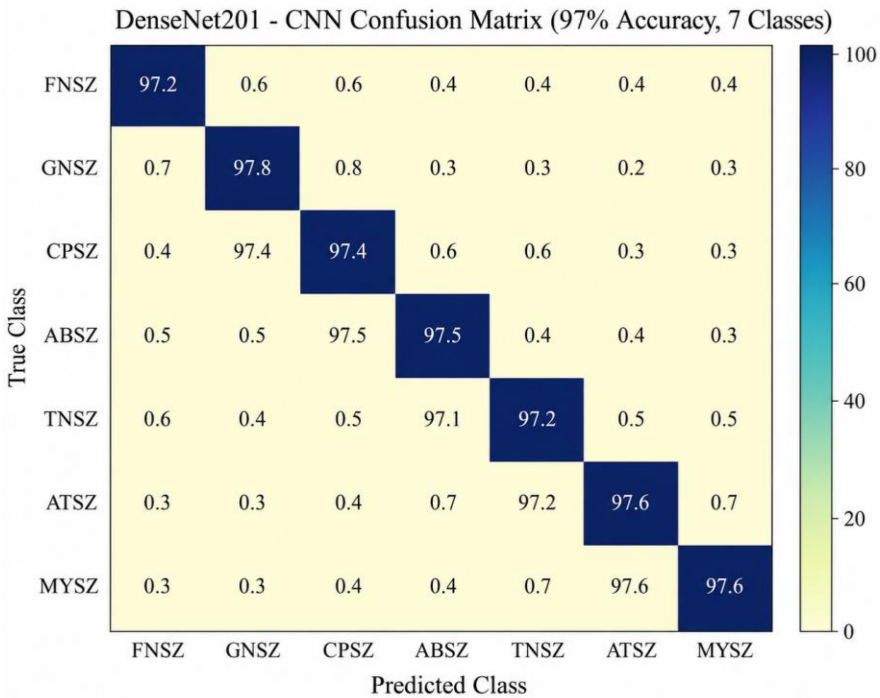

Fig. 14 (b) Confusion matrix and predicted outcomes using the pre-trained model using the TUH dataset

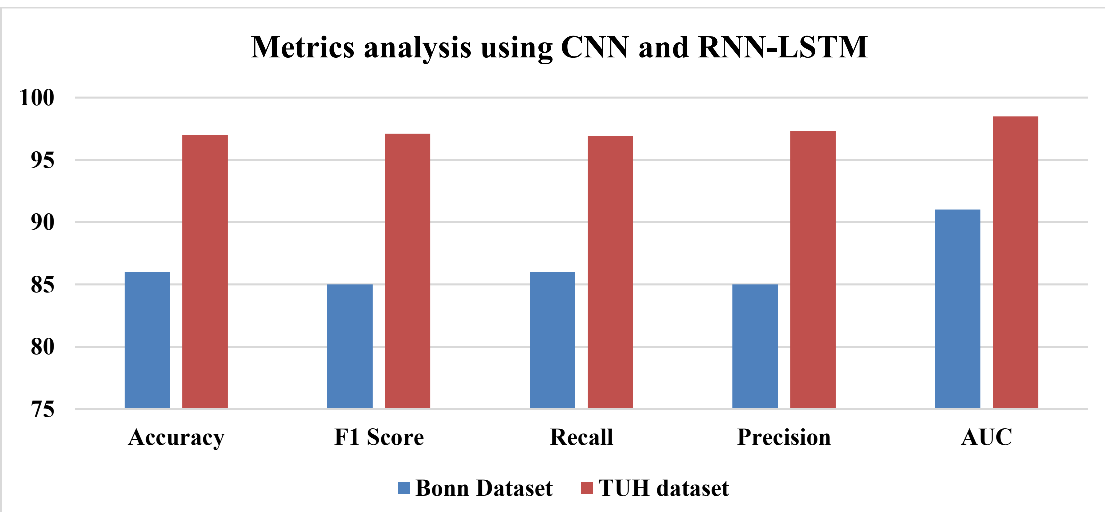

Fig. 15 Metrics analysis of CNN and RNN-LSTM using the Bonn and TUH dataset

Figures 14(a) and 14(b) show the confusion matrix and performance analysis using CNN on the Bonn dataset and TUH dataset, respectively. The proposed CNN metrics analysis with Bonn and TUH data sets is listed in Table 10 and visualized in Figure 15. Tables 11 and 12 show metrics analysis employing pre-trained models with RNN-LSTM, using the Bonn and TUH datasets. Again, for the TUH dataset, DenseNet201 recorded the best performance with 97.0 accuracy and AUC 98.5. This means the integration of CNN with RNN-LSTM greatly improves EEG feature learning and provides solid classification of seizure types.

Ablation Studies

Table 13 outlines the parameters utilized in the training of the RNN-LSTM model for the University of Bonn dataset ablation studies. The model was trained for 25 epochs with a learning rate of 0.001.

The GPU had a batch size of 128 and a CPU fallback with a batch size of 64. An NVIDIA GPU was used alongside mixed precision (float16) for efficient training. Validation occurred at every tenth epoch. The model used the ReLU activation function and Adam optimizer for stable convergence and efficient gradient flow.

The model has three sizes: Small, 128 to 64 LSTM layers; Medium, 256 to 128 to 64 LSTM layers; Large, 128 to 128 residual to 256 LSTM layers. To avoid overfitting, dropout rates were set to 0.2, 0.5. Each LSTM block included batch normalization. We changed the dense layer setup depending on the model sizes that increasingly matched the number of output classes. The training took around 360 minutes, and the validation accuracy was nearly 90%. The outcome indicates the model's power to capture temporal EEG patterns while providing a reliable classification performance with different network sizes.

4.3. Discussions

For each EEG signal in the figure, a model prediction was compared to the ground truth, which was annotated by domain experts. In this comparison, we were able to see if the model classified the signal into the correct seizure type. It allowed us to see whether the model was able to capture nuances among different seizure types.

Moreover, the figure offered a useful snapshot of the performance across seizure types. Some seizures were predicted more accurately than others, suggesting that the model was good at predicting some seizures based on special patterns of the EEG.

On the contrary, other seizure types may have been much more challenging for the model, as they may look the same as non-seizures or be similar to another seizure type. Comparing the predictions from the CNN model with the real labels allowed us to better understand its effectiveness and shortcomings. This knowledge will help us with future work and contribute to more advanced seizure prediction technology.

The results of each method are analyzed, and the method that gives the most accurate result in the least time with noise in the EEG dataset is chosen. To sum up, the results are better in "Extracting image features using pre-trained network" rather than "Transfer learning using pre-trained network". Extracting the feature of the image, we use a pre-trained neural network. This method makes use of the characteristics acquired by the neural network on a miscellaneous dataset, usually containing this sort of data.

Technological advancements in deep learning architecture for medical image classification help in improving the classification accuracy and effectiveness against seizure type detection. Of various architectures, RNN-LSTM performs the best against CNN and achieves high accuracy levels. After conducting a thorough examination of the architectures, it was observed that RNN-LSTM was highly efficient in detecting the type of seizure. The maximum accuracy attained was 97.0% using the DenseNet201 architecture. RNN-LSTM is better than CNN because it can capture long-range temporal dependencies between data, which is important for medical imaging.

CNNs may be great at taking features from static images; however, RNN-LSTM has memory cells that can capture the pattern and relationship over time. The RNN-LSTM's capacity for temporal memory makes it perfect for interpreting medical data where context and time can have diagnostic value. When it comes to detecting seizure types, the ability of RNN-LSTM technology to pick up much more subtle changes and differences that may represent various seizure types is beneficial [39]. It is important when it comes to medical diagnostics, as minor variations in the EEG can contain valuable diagnostic information. RNN-LSTM can learn and recognize these patterns. Conversely, CNN may not learn the temporal differences. Hence, RNN-LSTM can effectively distinguish between the seizure types. One of the most important RNN-LSTM features is the extraction and use of sequential data. Within the architecture, LSTM units play a pivotal role [40]. The units are built to remember information over long sequences. This allows the different units to remember different patterns and dependencies. It is especially valuable in EEG studies, where the brain's electrical activity evolves when two things happen sequentially.

RNN-LSTM works in a multi-layer way during feature extraction. The first layers of the model are CNN performing spatial feature extraction, which extract the most significant visual features of the EEG images [41]. This spatial information is then used in the LSTM layers to study the temporal dynamics of the data. The LSTM units can recall patterns because they can remember earlier information. These units allow the model to detect patterns more easily. The RNN-LSTM Architecture training uses special-tuned hyper-parameters like batch size and learning rates to optimize convergence. The crafted dataset has labeled EEG images that help the model learn through the visual and temporal patterns associated with different types of seizures.

To conclude, RNN-LSTM is a useful solution for seizure type detection in medical image classification. It can capture time dependencies in EEG datasets more effectively than traditional CNNs. Hence, it is more suitable for medical image analysis than CNN. RNN-LSTM achieves the highest accuracy by combining spatial feature extraction and sequential layer-based feature extraction methods. The use of DenseNet201 architecture using the Bonn dataset achieves a notable accuracy of 86.0%. For TUH, we achieved 97.0% accuracy. A greater level of accuracy would be highly beneficial for the diagnosis and treatment of patients suffering from seizures. With the advancements in technology, RNN-LSTM is an example of deep learning that enhances medical diagnostics and improves patient care.

The classification accuracy of various techniques using EEG is found to be much improved through deep learning and hybrid models. This is apparent in the comparative analysis of several classification techniques shown in Table 14. Wang et al. [42] utilized DWT combined with nonlinear and entropy-based features using an AHW-BGOA-DNN framework, achieving an accuracy of 81.69%, which indicates a moderate performance. On the other hand, Li et al. [43] suggested a highly accurate result of 99.36% using a wavelet-based nonlinear analysis optimized with GA-SVM, showing the strength of evolutionary algorithms in machine learning.

Table 14. Using the university of Bonn to compare setup classification accuracy with existing approaches

References	Techniques	Accuracy (%)
Wang, X.,et al. [2]	Hybrid 3 CNN-1 LSTM network mode	98
Iakiyaselvan N. et al. [11]	CNN-Based Epilepsy Detection	95±2%
Raghu, S.,et al. [24]	DWT-based sigmoid Entropy and SVM	96.34%
Chen, D.,et al. [25]	Discrete Wavelet Transform	99.33
Ullah, I.,et al. [28]	VMD + AR + RFTQWT + KNN-Entropy	99.1
Hussein R,et al. [29]	ESD-LSTM	99.1 ± 0.9%
Wang, X.,et al.[2]	DWT+Nonlinear and entropy-based features+AHW-BGOA-DNN	98
Li, M.,et al. [38]	This work, Wavelet-based nonlinear analysis +GA-SVM	99.84 %
Kashefi Amiri et al. [39]	DWT + 1D CNN-LSTM	97.2%
Salman, A.,et al. [40]	2D CNN	97.8
Duan, L.,et al. [42]	deep metric learning	98.60
Lu, L.,et al. [43]	STFT (time-frequency spectrogram)	99
Cao X. et al. [44]	DWT + Feature Fusion + CNN-Bi-LSTM	96.19%
Proposed Work	CNN	91.23
	Transfer Learning (ResNet50)	91.23% (Bonn dataset) 89.1% (TUH dataset)
	DenseNet201+LSTM	98% (Bonn dataset) 92.5% (TUH dataset)

Deep learning technique yields excellent results when applying for the epithelium through EEG signals. Goru et al. [44] developed deep CNNs and achieved high seizure classification accuracy. In the second work, Salman et al. [40] applied a 2D CNN-based approach on the EEG dataset, achieving a strong detection performance. Narayana et al. [41] improved the sensitivity and specificity significantly using modified LSTM networks. The team of Duan et al. [42] used a deep metric learning for automatic seizure detection with high accuracy. Lu et al. [43] developed lightweight CNNs combined with a time-frequency spectrogram for seizure prediction with fast processing. The HACA is a hybrid autoencoder combining deep learning with a convolutional attention model that is proposed for epilepsy detection. Typically, the CNN techniques perform better in extracting spatial features, whereas the LSTM and hybrid techniques do better in temporal and sensitive-specific applications.

The models proposed show competitive performance in comparison. The accuracy achieved by the CNN model was 91.23%. Likewise, the Transfer Learning (ResNet50) approach achieved 91.23% for the Bonn dataset and 89.1% for the TUH dataset. Similarly, other state-of-the-art models have shown constant but slightly lower performance. However, the accuracy of the DenseNet201 + LSTM hybrid model improved to 98% (Bonn) and 92.5% (TUH), showing better generalization across datasets. The new hybrid model works better than many previous ones and gives very good accuracy in classifying and detecting seizures through EEG.

5. Conclusion and Future Work

This study proposes an Artificial Neural Network (ANN)-RNN LSTM model that uses a modified ANN structure with the addition of memory cells for the classification of seizure types into eight classes from EEG time records. It was not uncommon to turn a time record into an input image feed, but it was a rare exploit.

The researchers performed an assessment of the framework using experiments with transfer learning techniques, which involved 10 pre-trained networks to extract image features. The proposed method gives an optimum performance rate of 97.22% classification accuracy. The DenseNet201 architecture was able to get the accuracy of the state of the art at a learning rate of 1e-3 and a solver optimization of Adam. An important finding showed that seizure-type classification can be effectively done using RNN-LSTM.

An interesting thing about this research is the comparison between the image feature extractor technique and the transfer learning technique. Significantly, that one won in both classification accuracy and computation time. This showed that utilizing inherent image features through customized schemes is better than using pre-trained networks for transfer learning. The proposed model is subjected to extensive analysis in terms of training and test data to evaluate the efficacy of the proposed model compared to existing models. The complete analysis further reinforces that the RNN-LSTM approach outperforms the conventional Convolutional Neural Network (CNN). The RNN-LSTM model is consistently better than CNN in accuracy because it captures the EEG data's temporal details successfully. All the successive steps led to the completion of this project, which was able to achieve the accuracy of 97.22% with the help of RNN-LSTM and DenseNet201.

This has the potential to transform the interpretation and analysis of medical images. They can tell what kind of seizure it is better with this. It can improve diagnosis. The future of this project looks very bright. One of the paths to explore is the application of other deep learning algorithms, such as GANs and RBFNs. Using GANs could generate synthetic EEG data that could diversify the training dataset and improve the generalizability of the model. Taking RBFNs into account might serve to help one understand their ability to capture spatial and temporal dependencies in the EEG. This study marked a groundbreaking approach in seizure-type classification using RNN-LSTM and EEG image feeds. The methodology has demonstrated 97.22% accuracy, which suggests it could revolutionize the diagnostics and treatment of diseases. The proposed methods were compared with existing approaches, and the results indicate that they are likely to be highly useful in the future. This endeavor was seen as an important first step towards harnessing the power of deep learning to benefit patient care or knowledge.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgement

We would also like to thank the Intrusion Detection Lab, School of Electrical and Electronics Engineering, SASTRA Deemed University, for their infrastructural support towards this study.

Funding Statement

This research received no external funding.

References

[1] Ijaz Ahmad et al., "A Hybrid Deep Learning Approach for Epileptic Seizure Detection in EEG Signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 30, no. 2, pp. 1019-1029, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Xiashuang et al., "Automated Recognition of Epilepsy from EEG Signals using a Combining Space–Time Algorithm of CNN-LSTM," *Scientific Reports*, vol. 13, pp. 1-12, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[3] Muhammet Varlı, and Hakan Yılmaz, "Multiple Classification of EEG Signals and Epileptic Seizure Diagnosis with Combined Deep Learning," *Journal of Computational Science*, vol. 67, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[4] Athar A. Ein Shoka et al., "EEG Seizure Detection: Concepts, Techniques, Challenges, and Future Trends," *Multimedia Tools and Applications*, vol. 82, pp. 42021-42051, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[5] Ferdaus Anam Jibon et al., "Epileptic Seizure Detection from Electroencephalogram (EEG) Signals using Linear Graph Convolutional Network and DenseNet based Hybrid Framework," *Journal of Radiation Research and Applied Sciences*, vol. 16, no. 3, pp. 1-11, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Javad Birjandtalab, Mehrdad Heydar Zadeh, and Mehrdad Nourani, "Automated EEG-Based Epileptic Seizure Detection Using Deep Neural Networks," *2017 IEEE International Conference on Healthcare Informatics (ICHI)*, Park City, USA, pp. 552-555, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[7] S. Raghu et al., "Automated Detection of Epileptic Seizures using Successive Decomposition Index and Support Vector Machine Classifier in Long-term EEG," *Neural Computing and Applications*, vol. 32, pp. 8965-8984, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[8] S. Roy et al., "Seizure Type Classification Using EEG Signals and Machine Learning: Setting a Benchmark," *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, Philadelphia, PA, USA, pp. 1-6, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Mengni Zhou et al., "Epileptic Seizure Detection Based on EEG Signals and CNN," *Frontiers in Neuroinformatics*, vol. 12, pp. 1-14, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[10] Amir H. Ansari et al., "Neonatal Seizure Detection Using Deep Convolutional Neural Networks," *International Journal of Neural Systems*, vol. 29, no. 4, pp. 1-20, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[11] N. Ilakiyaselvan, A. Nayeemulla Khan, and A. Shahina, "Deep Learning Approach to Detect Seizure Using Reconstructed Phase Space Images," *Journal of Biomedical Research*, vol. 34, no. 3, pp. 240-250, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[12] Sima Siami-Namini, Neda Tavakoli, and Akbar Siami Namin, "The Performance of LSTM and BiLSTM in Forecasting Time Series," *2019 IEEE International Conference on Big Data (Big Data)*, Los Angeles, USA, pp. 3285-3292, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[13] S. Raghu et al., "EEG based Multi-Class Seizure Type Classification using Convolutional Neural Network and Transfer Learning," *Neural Networks*, vol. 124, pp. 202-212, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] U. Rajendra Acharya et al., "Deep Convolutional Neural Network for the Automated Detection and Diagnosis of Seizure using EEG Signals," *Computers in Biology and Medicine*, vol. 100, pp. 270-278, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[15] Ali Emami et al., "Seizure Detection by Convolutional Neural Network-Based Analysis of Scalp Electroencephalography Plot Images," *NeuroImage: Clinical*, vol. 22, pp. 1-10, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[16] M. Golmohammadi et al., "Deep Architectures for Automated Seizure Detection in Scalp EEGs," *arXiv preprint*, pp. 1-8, 2017. [CrossRef] [Publisher Link]

[17] Rubén San-Segundo et al., "Classification of Epileptic EEG Recordings using Signal Transforms and Convolutional Neural Networks," *Computers in Biology and Medicine*, vol. 109, pp. 148-158, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[18] Amirhossein Tavanaei et al., "Deep Learning in Spiking Neural Networks," *Neural Networks*, vol. 111, pp. 47-63, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[19] Nhan Duy Truong et al., "Convolutional Neural Networks for Seizure Prediction using Intracranial and Scalp Electroencephalogram," *Neural Networks*, vol. 105, pp. 104-111, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[20] Abdullah Aziz Sharfuddin, Md. Nafis Tihami, and Md. Saiful Islam, "A Deep Recurrent Neural Network with BiLSTM Model for Sentiment Classification," *2018 International Conference on Bangla Speech and Language Processing (ICBSLP)*, Sylhet, Bangladesh, pp. 1-4, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[21] Kyunghyun Cho et al., "Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation," *arXiv preprint*, pp. 1-15, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[22] Felix A. Gers et al., "Learning Precise Timing with LSTM Recurrent Networks," *Journal of Machine Learning Research*, vol. 3, pp. 115-143, 2002. [Google Scholar] [Publisher Link]

[23] Pearlmutter, "Learning State Space Trajectories in Recurrent Neural Networks," *International 1989 Joint Conference on Neural Networks*, Washington, DC, USA, pp. 365-372, 1989. [CrossRef] [Google Scholar] [Publisher Link]

[24] S. Raghu et al., "Performance Evaluation of DWT based Sigmoid Entropy in Time and Frequency Domains for Automated Detection of Epileptic Seizures using SVM Classifier," *Computers in Biology and Medicine*, vol. 110, pp. 127-143, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[25] Duo Chen et al., "A High-Performance Seizure Detection Algorithm based on Discrete Wavelet Transform (DWT) and EEG," *PLoS One*, vol. 12, no. 3, pp. 1-21, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[26] Khald Ali I. Aboalayon et al., "Sleep Stage Classification Using EEG Signal Analysis: A Comprehensive Survey and New Investigation," *Entropy*, vol. 18, no. 9, pp. 1-31, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[27] Serap Aydın, Hamdi Melih Saraoğlu, and Sadık Kara, "Log Energy Entropy-Based EEG Classification with Multilayer Neural Networks in Seizure," *Annals of Biomedical Engineering*, vol. 37, pp. 2626-2630, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[28] Ihsan Ullah et al., "An Automated System for Epilepsy Detection using EEG Brain Signals based on Deep Learning Approach," *Expert Systems with Applications*, vol. 107, pp. 61-71, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[29] Ramy Hussein et al., "Optimized Deep Neural Network Architecture for Robust Detection of epileptic Seizures using EEG Signals," *Clinical Neurophysiology*, vol. 130, no. 1, pp. 25-37, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[30] U. Rajendra Acharya et al., "Automated EEG-based Screening of Depression using Deep Convolutional Neural Network," *Computer Methods and Programs in Biomedicine*, vol. 161, pp. 103-113, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[31] Zuochen Wei et al., "Automatic Epileptic EEG Detection using Convolutional Neural Network with Improvements in Time-Domain," *Biomedical Signal Processing and Control*, vol. 53, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[32] Yong Yu et al., "A Review of Recurrent Neural Networks: LSTM Cells and Network Architectures," *Neural Computation*, vol. 31, no. 7, pp. 1235-1270, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[33] Department of Epileptology, University of Bonn. [Online]. Available: https://www.ukbonn.de/epileptologie/

[34] Electroencephalography (EEG) Resources, The Institute for Signal and Information Processing, Picnopress. [Online]. Available:https://isip.piconepress.com/projects/nedc/html/tuh_eeg/

[35] Alex Graves, and Jürgen Schmidhuber, "Framewise Phoneme Classification with Bidirectional LSTM and other Neural Network Architectures," *Neural Networks*, vol. 18, no. 5-6, pp. 602-610, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[36] Edward Choi et al., "Using Recurrent Neural Network Models for Early Detection of Heart Failure Onset," *Journal of the American Medical Informatics Association*, vol. 24, no. 2, pp. 361-370, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[37] Aslıhan Cura et al., "Driver Profiling Using Long Short Term Memory (LSTM) and Convolutional Neural Network (CNN) Methods," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 10, pp. 6572-6582, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[38] Mingyang Li, Wanzhong Chen, and Tao Zhang, "Automatic Epilepsy Detection using Wavelet-based Nonlinear Analysis and Optimized SVM," *Biocybernetics and Biomedical Engineering*, vol. 36, no. 4, pp. 708-718, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[39] Homa Kashefi Amiri, Masoud Zarei, and Mohammad Reza Daliri, "Epileptic Seizure Detection from Electroencephalogram Signals based on 1D CNN-LSTM Deep Learning Model using Discrete Wavelet Transform," *Scientific Reports*, vol. 15, pp. 1-18, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[40] Adnan Salman, "Epilepsy Detection from EEG Data using 2D Convolutional Neural Network," *2021 6th International Conference on Communication, Image and Signal Processing (CCISP)*, Chengdu, China, pp. 101-108, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[41] Venkata Narayana Vaddi, and Prakash Kodali, "Enhanced Epilepsy Sensitivity and Detection Rate with Improved Specificity by Integration of Modified LSTM Networks," *IEEE Sensors Journal*, vol. 25, no. 7, pp. 11963-11970, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[42] Lijuan Duan et al., "An Automatic Method for Epileptic Seizure Detection Based on Deep Metric Learning," *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 5, pp. 2147-2157, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[43] Lun Lu et al., "A Seizure Prediction Method for Epilepsy Utilizing Lightweight Convolutional Neural Networks and Time-Frequency Spectrograms of EEG Signals," *2024 IEEE Biomedical Circuits and Systems Conference (BioCAS)*, Xi'an, China, pp. 1-5, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[44] Xiaoshuai Cao et al., "A Hybrid CNN-Bi-LSTM Model with Feature Fusion for Accurate Epilepsy Seizure Detection," *BMC Medical Informatics and Decision Making*, vol. 25, pp. 1-28, 2025. [CrossRef] [Google Scholar] [Publisher Link]