Epilepsy Diagnosis Using EEG Image Analysis

Seeba Doddmani, Sana Mulla, Dilipsingh Rajpurohit, Rajashri Khanai and Prema T. Akkasaligar
Department of Computer Science and Engineering, KLE Technological University, Belagavi, India

Keywords: Epilepsy Detection, EEG Image Analysis, Machine Learning (ML), Convolutional Neural Network (CNN), Scalogram Images, VGG-16 Model, Deep Learning, Feature Extraction, Transfer Learning, Bern-Barcelona EEG Dataset, Seizure Classification, Signal Processing, Time-Frequency Analysis, Neural Networks, Supervised Learning, Data Preprocessing, Binary Classification, Medical Diagnosis, Graph Spectral Features, Long Short-Term Memory (LSTM), Real-Time Detection.

Abstract: Epilepsy is a neurological disorder that affects millions of people around the world and is characterized by recurrent seizures caused by abnormal brain activity. Electroencephalograms (EEG) are the primary diagnostic tool, but traditional manual analysis is time-intensive and prone to errors. This project leverages machine learning techniques to automate epilepsy detection using scalogram images generated from EEG signals. A custom Convolutional Neural Network (CNN) model was developed and trained on the Bern-Barcelona EEG dataset, achieving a training precision of 74.07% and a testing precision of 73.22%. The model demonstrates good training performance and testing accuracy. An implemented VGG-16 gave a training accuracy of 81.13% and a testing accuracy of 80.04%. This study aims to help clinicians improve diagnostic accuracy and provide a scalable, real-time solution for epilepsy detection, particularly in underserved regions.

1 INTRODUCTION

Epilepsy, a chronic neurological disorder, affects more than 50 million people worldwide, causing significant health and social challenges. Diagnosis of epilepsy traditionally relies on manual analysis of EEG recordings, a process that is not only time-consuming but also susceptible to errors due to the complexity of EEG waveforms. With advances in machine learning, there is growing interest in automating this process to improve diagnostic accuracy and efficiency. This project focuses on converting EEG signals into scalogram images, which capture both temporal and frequency domain features, and applying a custom CNN model for seizure detection. The proposed system aims to address the limitations of manual analysis, offering a reliable and scalable solution for clinicians, particularly in resource-limited areas where access to specialized neurologists is scarce.

The high prevalence of epilepsy, coupled with the challenges of timely and accurate diagnosis, motivates the need for automation in the detection of epilepsy. Automated systems can significantly reduce the burden on neurologists, improve diagnostic accuracy, and ensure better access to healthcare in underserved areas. Using machine learning, these systems can offer reliable and real-time diagnoses, facilitating earlier intervention and better management of epilepsy.

Several approaches have been explored for epilepsy detection, combining traditional machine learning and deep learning techniques: Krishnasamy et al. proposed supervised learning algorithms, including Support Vector Machines (SVM) and CNNs, achieving accuracy rates up to 99.7% but requiring large datasets and computational resources. Pattnaik et al. utilized scalogram images and transfer learning with ResNet50, achieving a classification accuracy of 95.23% with high sensitivity and specificity. Sesha Sai et al. combined CNNs with SVMs for automatic feature extraction, achieving 94.48% accuracy but heavily dependent on data quality. Other studies, such as by Wang et al., explored hybrid CNN-LSTM (Long Short Term Memory) models for capturing both spatial and temporal features in EEG data.

Despite their promising results, these approaches face several limitations: Heavy reliance on large, high-quality datasets, making them less effective in diverse clinical settings. Computational complexity, particularly for deep architectures like Residual

Doddmani, S., Mulla, S., Rajpurohit, D., Khanai, R. and T. Akkasaligar, P.
Epilepsy Diagnosis Using EEG Image Analysis.
DOI: 10.5220/0013602300004664

In *Proceedings of the 3rd International Conference on Futuristic Technology (INCOFT 2025) - Volume 2*, pages 776-783
ISBN: 978-989-758-763-4

Network (ResNet) and hybrid models, which hinders real-time application. Limited generalization to unseen data, with significant performance degradation when applied to different patient demographics or new datasets. Challenges in adapting models to resource-limited environments due to high hardware requirements.

Epilepsy detection faces several challenges: High variability in seizure patterns across patients complicates the creation of a universally robust model. EEG data often contains noise and artifacts, requiring sophisticated pre-processing. Balancing model complexity with computational efficiency is critical for real-time applications. Addressing class imbalance in datasets is essential to improve model generalization.

To address these challenges, this project proposes a machine learning pipeline that converts EEG signals into scalogram images to capture time-frequency domain features. A custom CNN is designed to classify these images into seizure and non-seizure categories. The pipeline incorporates preprocessing techniques like normalization and leverages training strategies to improve model generalization.

Developed a novel pipeline to transform EEG signals into scalogram images for time-frequency analysis. Designed and implemented a custom CNN architecture optimized for binary classification. Evaluated model performance on the Bern-Barcelona EEG dataset, achieving a training accuracy of 97.77%. Identified challenges in generalization through testing accuracy analysis, providing insights for future improvement. Highlighted the potential for real-time deployment of the proposed system in resource-limited settings.

The rest of paper is organized as follows. Section II provides a brief review of the recent works. Section III details the problem statement, background and the proposed methodology. Section IV describes the experiment details along with the results and discussions. Finally, the paper concludes in Section V.

2 LITERATURE SURVEY

The methodology for identifying and selecting studies for the literature survey involved several stages to ensure relevance and quality. Initially, 42 papers were identified from reputable sources such as IEEE Conference, Springer, Frontiers, and BioMed Central(BMC). During the screening phase, 10 papers were excluded—6 for being older than five years and 4 for having incorrect titles. Figure. 1 shows the flowchart illustrating this process. After screening, 32 papers were assessed for eligibility, and 16 were further excluded—3 were review papers, and 13 focused on signals rather than images. Ultimately, 16 studies were included in the literature survey, providing a robust foundation for the research.

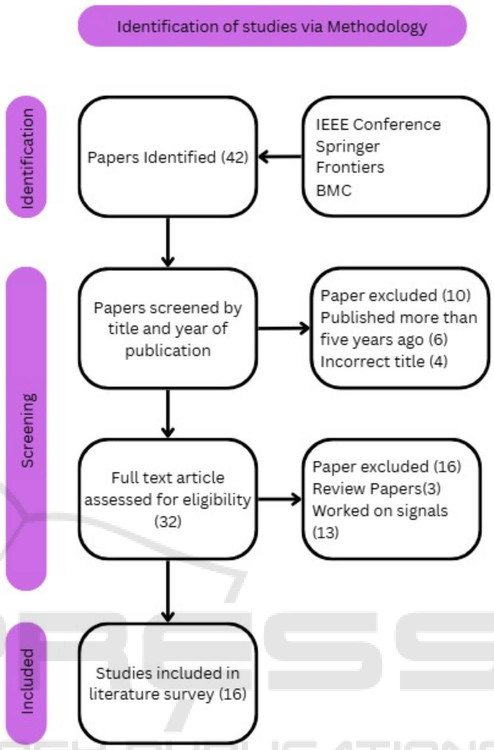

Figure 1: Flowchart of literature survey paper selection

(Krishnasamy et al., 2024) explored supervised learning algorithms for epileptic seizure detection, achieving high accuracy rates (98.5% with SVM and 99.7% with fuzzy classifiers). Their system automates seizure detection effectively, reducing complexity for experts, but struggles with real-time deployment due to computational demands and limited generalizability across datasets. Similarly, (Hafeez and Shakil, 2024) utilized EEG-based brainwave images to classify stress levels using LSTM (70.67% accuracy) and CNN (90.46% accuracy), demonstrating CNN's capability for spatial feature extraction but facing challenges with noise in EEG data and small participant pools. (Pattnaik et al., 2024) achieved 95.23% classification accuracy using scalogram images analyzed with a pre-trained ResNet50 model, showcasing the utility of transfer learning for EEG signal analysis.

(Sadam and Nalini, 2024) combined CNN and SVM to achieve 94.48% accuracy, highlighting the advantages of hybrid models for automated feature extraction but showing sensitivity to noisy datasets. (Georgis-Yap et al., 2024) compared supervised and unsupervised approaches like CNN, CNN-LSTM, and TCN, finding comparable results for patient-specific seizure prediction. While their methods reduce preprocessing needs, they require substantial preictal data for optimal performance. (Krishnan et al., 2024) applied GASF to convert EEG signals into image representations, achieving up to 96% accuracy. However, this approach is computationally intensive, especially during image transformation and feature extraction.

(Shankar et al., 2023) utilized CNNs with phase synchronization matrices to classify seizures with 83.3% accuracy, demonstrating effective spatial feature extraction. However, intensive preprocessing and the need for large datasets remain barriers to real-time applications. (Khasawneh et al., 2022) achieved 99.8% precision using Faster R-CNN and transfer learning for K-complex detection, with adaptability across datasets. Nonetheless, overlapping image generation methods may introduce data leakage, compromising model reliability. (Hu et al., 2020) proposed a hierarchical neural network (HNN) using transfer learning, attaining 98.97% accuracy on the CHB-MIT dataset, though the reliance on large pre-trained DNNs limits real-time applicability.

(Sharma and Meena, 2024) introduced a model integrating GFT and DWT for feature extraction, achieving over 98% accuracy across datasets. While graph spectral features enhance detection accuracy, the method's complexity increases computational demand, complicating real-time deployment. (Kunekar et al., 2024) used LSTM networks to achieve 97% accuracy, effectively handling temporal dependencies in EEG data. However, dataset class imbalance could hinder generalizability. (Jridi et al., 2024) employed deep ResNet for multi-disorder detection, reaching 100% accuracy for epilepsy detection on the UBonn dataset, but high computational requirements pose challenges for implementation in resource-limited environments.

(Saleem et al., 2023) combined CNN with traditional classifiers, achieving 98.49% accuracy in seizure detection. The hybrid model effectively identifies subtle EEG patterns but requires larger, more diverse datasets to ensure generalizability. (Majzoub et al., 2023) used AlexNet for multi-channel EEG signal classification, achieving 98.25% accuracy in binary classification. However, its accuracy drops to 92.98% with new patient data, highlighting the need for diverse training samples. (Wang et al., 2023) developed a CNN-LSTM hybrid model, achieving 98% accuracy in ternary classification and 100% in binary classification. This model excels in capturing both spatial and temporal features but demands high computational resources.

Finally, (Supriya et al., 2021) adopted graph-theory-based methods, such as VG and HVG, for feature extraction, achieving accuracies above 95%.

These methods are advantageous for their processing speed and classification efficacy but are limited by their sensitivity to dataset size and threshold dependencies. Collectively, while these studies demonstrate significant advancements in EEG-based seizure detection and classification, the major limitations across models include dependency on large datasets, computational demands, and challenges with generalization across diverse patient populations.

This literature survey reviews advancements in EEG-based epileptic seizure detection using supervised learning, deep learning, and hybrid models. Techniques like CNN-SVM hybrids (94.48% accuracy), transfer learning with scalograms (95.23%), and GASF image transformations (96%) demonstrate high accuracy but face challenges with computational demands and noisy data. Deep learning models like CNNs and hybrid CNN-LSTM architectures achieve exceptional spatial-temporal feature extraction but require large datasets. Despite progress, issues such as dataset dependency, generalizability, and real-time applicability remain significant hurdles.

3 PROBLEM STATEMENT AND SYSTEM MODEL

Epilepsy is a chronic neurological disorder marked by recurrent seizures resulting from abnormal electrical activity in the brain. Diagnosing epilepsy relies heavily on EEGs, which capture electrical patterns and help identify abnormalities. Traditional approaches involve manual inspection, a labor-intensive task requiring specialized expertise.

EEG, or electroencephalography, plays a significant role in medical diagnosis as a non-invasive and cost-effective method for recording brain activity. It is particularly useful in identifying abnormal brain wave patterns, which can indicate potential seizures or other neurological disorders. By capturing electrical activity in the brain, EEG provides valuable insights into the functional state of the brain, aiding clinicians in the diagnosis and management of various conditions.

Traditional methods of diagnosing epilepsy present several challenges. The manual analysis of long EEG recordings is a time-consuming process, requiring significant effort from trained specialists. Furthermore, the variability in seizure patterns across individuals adds complexity to achieving a consistent and accurate diagnosis. These issues are compounded by the limited availability of trained neurologists, particularly in rural or underserved areas, making timely and effective diagnosis even more difficult.

Epilepsy is a neurological disorder that impacts over 50 million people worldwide, characterized by abnormal brain activity leading to seizures. Traditional epilepsy diagnosis using EEGs involves manually analyzing extensive recordings to detect epileptic events, a process prone to errors and inefficiencies. This project aims to automate the detection of epilepsy by employing Machine Learning (ML) techniques to classify epileptic and non-epileptic events using 2D scalogram images derived from EEG signals. The goal is to create a robust, accurate system that can enhance diagnostic efficiency, particularly in resource-limited areas.

The primary objective of this work is to develop a machine learning pipeline capable of automatically classifying EEG-derived scalogram images as "seizure" or "non-seizure." This automated approach aims to improve diagnostic accuracy by increasing sensitivity and specificity, thereby assisting clinicians in making reliable diagnoses. Additionally, the solution is designed to be deployable in real-time systems, enabling rapid analysis and decision-making in clinical settings. To ensure versatility, the proposed solution is scalable and adaptable to diverse datasets and hardware setups, including mobile and low-power devices.

Figure.2. shows the system model that outlines a process for classifying EEG signals, into two categories: seizure (focal) and non-seizure. The process begins with signal input, followed by data preprocessing to transform the signal into a scalogram image, which visually represents the frequency content over time. The dataset is then split into training and testing subsets. Two models are employed for classification: a custom CNN and a pre-trained VGG16 model. Both models undergo training and validation to learn patterns from the data. Finally, model evaluation is conducted to assess performance, leading to the classification of the input signal as either seizure (focal) or non-seizure.

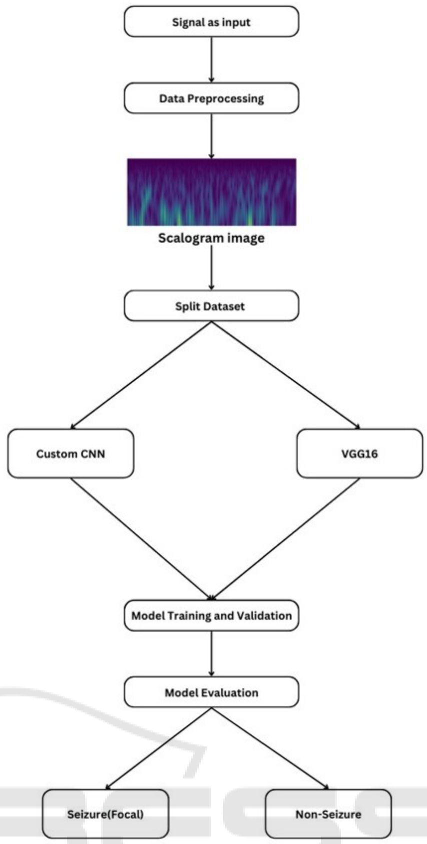

Figure 2: System model for EEG detection

4 PROPOSED METHODOLOGY

The proposed work is executed on device is powered by an AMD Ryzen 7 4800H with Radeon Graphics processor, operating at 2.90 GHz, and is equipped with 16 GB of RAM. It runs a 64-bit version of Windows 11 Home and the NVIDIA GTX 3050 GPU. For implementation, we have used several Python libraries are used for efficient data analysis and visualization. NumPy supported numerical computations, Pandas handled data manipulation. Matplotlib are employed for creating visualizations, tensorflow used for building the CNN Architecture.

The dataset originally consisted of EEG signals represented by both X and Y components. These signals were processed and converted into scalogram images, with separate scalograms generated for the X and Y components. Following this transformation, normalization was applied to the images to standardize pixel intensity values, ensuring they fell within a consistent range to enhance model training.

4.1 Custom CNN

CNNs are a class of deep learning models specifically designed for processing structured data like images. A CNN consists of multiple layers, including convolutional layers that extract spatial and hierarchical fea-

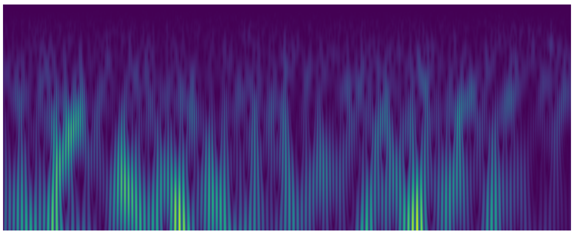

Figure 3: Focal Scalagram image (Seizure)

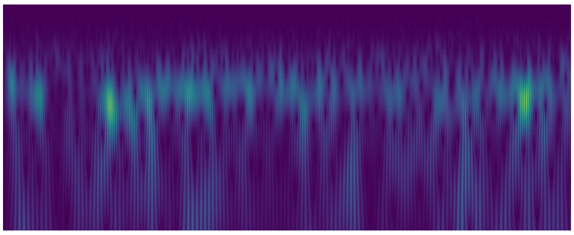

Figure 4: Non-Focal Scalagram image (Non-Seizure)

tures, pooling layers that reduce dimensionality, and fully connected layers for final classification. The training process involves optimizing a loss function, typically using stochastic gradient descent and back-propagation, to adjust weights and minimize classification errors. Regularization techniques like dropout are often applied to prevent overfitting. CNNs are highly effective in capturing spatial patterns, making them well-suited for tasks such as image recognition, object detection, and, in this case, epilepsy detection from EEG-based scalogram images.

The custom CNN for a multi-class classification problem consists of sequential convolutional blocks with increasing filters (32, 64, 128, and optionally 256), each containing a Conv2D layer with ReLU activation, batch normalization for stability, and MaxPooling2D for down-sampling. To prevent overfitting, L2 regularization is applied to both the convolutional and dense layers, and dropout is used in the dense layers for further regularization. The fully connected layers include a flatten layer to convert feature maps into a 1D vector, followed by dense layers with ReLU activation to learn complex representations. The model uses the Adam optimizer, an adaptive method suitable for deep learning tasks.

After training the CNN model for 50 epochs with early stopping at 11, it achieved a training accuracy of 74.07%, indicating that the model effectively learned patterns from the training data and the testing accuracy is 73.22%.

Figure.5. This graph depicts the training and testing accuracy of a model over several epochs (labeled on the x-axis). The training accuracy (blue line) starts higher and maintains relatively stable values with slight fluctuations, while the testing accuracy (orange line) begins lower but rises sharply, eventu-

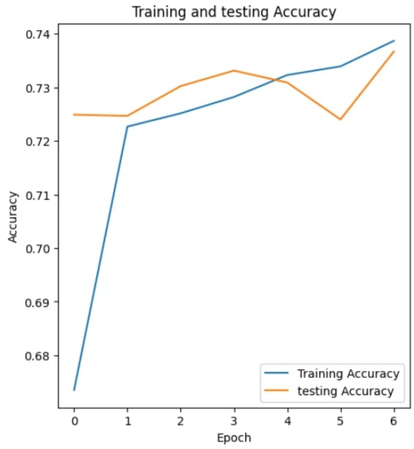

Figure 5: Training vs testing accuracy for custom CNN

Figure 6: Training and testing error vs Epoch for Custom CNN

ally fluctuating and surpassing the training accuracy in later epochs. This pattern may suggest overfitting at the final epochs, as the testing accuracy's instability could indicate sensitivity to the validation set. The testing accuracy surpassing training at times suggests randomness in the dataset split. However, the convergence of both accuracies around epoch 4 indicates balanced model performance during this phase.

Figure.6. The graph shows the training and testing loss of a machine learning model over six epochs. The training loss (blue line) starts at a moderate level, gradually decreasing as the model learns from the training data. The testing loss (orange line), which measures the model's performance on unseen data, begins at a higher value but initially drops sharply, indicating improvement in generalization. After a few epochs, the testing loss increases slightly, where the model begins to perform better on training data.

4.2 VGG 16

VGG16 is a pre-trained model and is well-suited for study due to its ability to effectively extract spatial features from images. EEG signals converted into visual representations, such as scalogram images, contain patterns indicative of epileptic activity. VGG16's architecture, with its small 3x3 convolutional filters, is adept at capturing these fine-grained spatial details. Additionally, using a pre-trained VGG16 model through transfer learning allows leveraging its general feature extraction capabilities while fine-tuning it for the specific task of epilepsy detection. This approach is particularly advantageous when working with limited data, as it reduces the need for extensive training from scratch and ensures robust performance in classifying complex patterns in EEG images.

Learning Rate: 0.001, adjusted dynamically using schedulers. Batch Size: 32 for balanced efficiency and performance. Epochs: 25 with early stopping to prevent overfitting. Optimizer: Adam for faster convergence and adaptive learning. Image Resolution: 128x128 for consistent input size. Dropout layer to prevent overfitting and also applied the L2 regularization Loss Function : Binary cross entropy for binary classification Activation Function : We used both Relu and Sigmoid

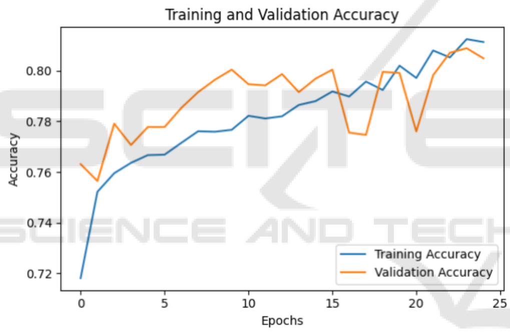

Figure 7: Training vs testing accuracy for VGG-16

Figure.7. illustrates the training and validation accuracy over epochs. Initially, both training and validation accuracy increase, indicating that the model is learning effectively. Around epoch 10, the validation accuracy starts to fluctuate, showing signs of overfitting as the training accuracy continues to improve steadily while the validation accuracy varies. Despite these fluctuations, the validation accuracy remains relatively close to the training accuracy, suggesting that the model is performing reasonably well.

Figure.8. shows the training and validation loss over epochs. Both losses decrease steadily, indicating that the model is learning and improving its predictions. However, the validation loss consistently remains lower than the training loss after a few epochs, which might suggest differences in how the training and validation datasets are handled. The smooth decline in loss for both suggests that the training process is stable, and there are no significant issues like overfitting or divergence in this range of epochs.

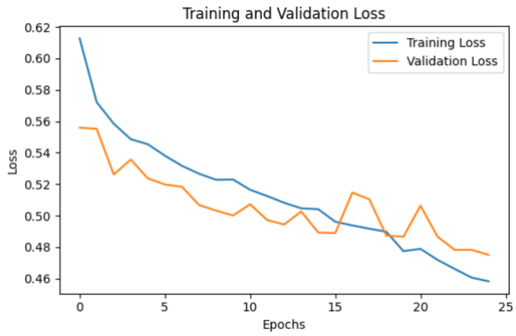

Figure 8: Training vs testing loss for VGG-16

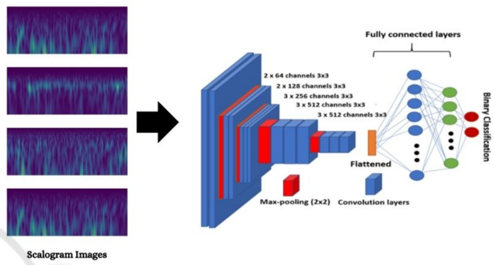

Figure 9: The architecture of VGG16

Figure.9. illustrates the convolutional neural network (CNN) architecture used for binary classification of scalogram images into focal and non-focal categories. The model processes input scalograms through a series of convolutional layers, each extracting increasingly complex features, with max-pooling layers reducing the spatial dimensions. The extracted features are then flattened and passed through fully connected layers, culminating in a sigmoid-activated output layer for binary classification. This architecture is adapted from the sea ice classification of SAR imagery presented by Khaleghian et al. (2021) (Khaleghian et al., 2021).

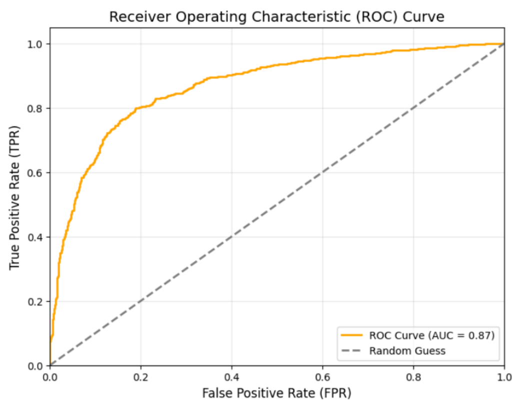

Figure 10: ROC curve of VGG-16

Figure.10. shows the ROC curve demonstrates the performance of the model in distinguishing between classes, with an Area Under the Curve (AUC) of 0.87. This indicates that the model has a high ability to discriminate between focal and non-focal classes, performing significantly better than random guessing (represented by the diagonal line). The curve's proximity to the top-left corner suggests a good balance between the true positive rate (sensitivity) and false positive rate, making the model reliable for classification tasks.

Table 1: Comparison results of custom CNN and VGG16

Parameter	VGG16	Custom CNN
Training Accuracy	81.13%	74.07%
Testing Accuracy	80.04%	73.22%

5 CONCLUSIONS

This study highlights the potential of machine learning in automating epilepsy detection using EEG-based scalogram images. The proposed custom CNN model achieved a training accuracy of 74.07% and a testing accuracy of 73.22%, demonstrating its ability to learn meaningful patterns from the data. Additionally, the VGG-16 model outperformed the custom CNN, achieving a training accuracy of 81.13% and a testing accuracy of 80.04%.

The CNN model has 9 layers but yet works good in comparison with the 16 layers of VGG-16. These results underscore the utility of advanced image-based analytics in healthcare while also emphasizing the importance of optimizing models for enhanced performance and generalization. This project lays a foundation for developing scalable, real-time systems for epilepsy diagnosis, with future work focusing on improving model robustness, leveraging larger and more diverse datasets, and exploring deployment strategies for real-world applications.

REFERENCES

Georgis-Yap, Z., Popovic, M. R., and Khan, S. S. (2024). Supervised and unsupervised deep learning approaches for eeg seizure prediction. *Journal of Healthcare Informatics Research*, 8(2):286–312.

Hafeez, M. A. and Shakil, S. (2024). Eeg-based stress identification and classification using deep learning. *Multimedia Tools and Applications*, 83(14):42703–42719.

Hu, D., Cao, J., Lai, X., Wang, Y., Wang, S., and Ding, Y. (2020). Epileptic state classification by fusing hand-crafted and deep learning eeg features. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 68(4):1542–1546.

Jridi, A., Djemal, R., and Belwafi, K. (2024). Neurological disorder diagnosis through deep residual network-based eeg signal analysis. In *2024 IEEE 7th International Conference on Advanced Technologies, Signal and Image Processing (ATSIP)*, volume 1, pages 144–149. IEEE.

Khaleghian, S., Ullah, H., Kræmer, T., Hughes, N., Eltoft, T., and Marinoni, A. (2021). Sea ice classification of sar imagery based on convolution neural networks. *Remote Sensing*, 13(9):1734.

Khasawneh, N., Fraiwan, M., and Fraiwan, L. (2022). Detection of k-complexes in eeg waveform images using faster r-cnn and deep transfer learning. *BMC Medical Informatics and Decision Making*, 22(1):297.

Krishnan, P. T., Erramchetty, S. K., and Balusa, B. C. (2024). Advanced framework for epilepsy detection through image-based eeg signal analysis. *Frontiers in Human Neuroscience*, 18:1336157.

Krishnasamy, L., Sriwastav, Y. K., Bharat, S. P., and Ganachar, S. R. (2024). Epilepsy detection using supervised learning algorithms. In *2024 IEEE International Conference on Contemporary Computing and Communications (InC4)*, volume 1, pages 1–6. IEEE.

Kunekar, P., Gupta, M. K., and Gaur, P. (2024). Detection of epileptic seizure in eeg signals using machine learning and deep learning techniques. *Journal of Engineering and Applied Science*, 71(1):21.

Majzoub, S., Fahmy, A., Sibai, F., Diab, M., and Mahmoud, S. (2023). Epilepsy detection with multi-channel eeg signals utilizing alexnet. *Circuits, Systems, and Signal Processing*, 42(11):6780–6797.

Pattnaik, S., Rao, B. N., Rout, N. K., and Sabut, S. K. (2024). Transfer learning based epileptic seizure classification using scalogram images of eeg signals. *Multimedia Tools and Applications*, pages 1–15.

Sadam, S. S. P. and Nalini, N. (2024). Epileptic seizure detection using scalogram-based hybrid cnn model on eeg signals. *Signal, Image and Video Processing*, 18(2):1577–1588.

Saleem, A., Khan, M. A., and Yousaf, H. M. (2023). Advancing epilepsy disease classification through machine learning and deep learning models utilizing eeg data. In *2023 17th International Conference on Open Source Systems and Technologies (ICOSST)*, pages 1–8. IEEE.

Shankar, A., Chakraborty, D., Saikia, M. J., Dandapat, S., and Barma, S. (2023). Seizure type detection using eeg signals based on phase synchronization and deep learning. In *2023 IEEE 19th International Conference on Body Sensor Networks (BSN)*, pages 1–5. IEEE.

Sharma, R. and Meena, H. K. (2024). Enhanced epileptic seizure detection through graph spectral analysis of eeg signals. *Circuits, Systems, and Signal Processing*, pages 1–21.

Supriya, S., Siuly, S., Wang, H., and Zhang, Y. (2021). Epilepsy detection from eeg using complex network techniques: A review. *IEEE Reviews in Biomedical Engineering*, 16:292–306.

Wang, X., Wang, Y., Liu, D., Wang, Y., and Wang, Z. (2023). Automated recognition of epilepsy from eeg signals using a combining space–time algorithm of cnn-lstm. *Scientific Reports*, 13(1):14876.