SCHIZOPHRENIA DETECTION USING EEG: A COMPARATIVE MODEL EVALUATION

Submitted in partial fulfilment for the award of the degree of

B. Tech Computer Science & Engineering

by:

AMEEN MEHVISH SANA 20BCD7089

CHIRAG SATISH NAIR 20BCE7073

SATYAM SAHA 20BCE7238

Near Vijayawada, - 522237, Andhra Pradesh, India.

Under the guidance of:

Dr. D. SUMATHI

SCOPE

May, 2023

CERTIFICATE

This is to certify that the CSE4006 Deep Learning Project titled **"Schizophrenia Detection Using EEG: A Comparative Model Evaluation"** that is being submitted by **Ameen Mehvish Sana(20BCD7089)**, **Chirag Satish Nair(20BCE7073)** and **Satyam Saha(20BCE7238)** is in partial fulfilment of the requirements for the award of Bachelor of Technology, is a record of bonafide work done under my guidance. The contents of this Project work, in full or in parts, have neither been taken from any other source nor have been submitted to any other Institute or University for award of any degree or diploma and the same is certified.

Dr. D. Sumathi
Guide

The thesis is satisfactory / unsatisfactory

ABSTRACT

Schizophrenia is a complex mental disorder that necessitates early detection and accurate diagnosis for effective treatment. This research paper proposes an approach for comparison of performance of deep learning models in schizophrenia detection. The three different models used: VGG16, CNN, and LSTM. Each model will be evaluated separately to assess its performance on the task. The VGG16 architecture will be utilized to extract high-level features from neuroimaging data, while the CNN model will focus on learning spatial patterns. The LSTM layers will capture temporal dependencies in the data. Each model will undergo individual training and evaluation on a large dataset, allowing for a comprehensive analysis of their performance in accurately distinguishing between individuals with schizophrenia and healthy controls. By comparing the results obtained from each model, valuable insights into their relative strengths and weaknesses can be gained. This comparative study will provide a comprehensive understanding of the capabilities and limitations of each model in the context of schizophrenia detection, contributing to the advancement of diagnostic approaches in the field.

CONTENTS

Chapter 1

Introduction

1.1 SCHIZOPHRENIA: *a gist*

Schizophrenia is a complex and chronic mental disorder that affects individuals worldwide. It is characterized by a range of symptoms, including hallucinations, delusions, disorganized thinking, and social withdrawal. Early detection and accurate diagnosis of schizophrenia are crucial for providing effective treatment and improving outcomes for patients. Hence it is a complex mental disorder that requires early detection and an accurate diagnosis for effective treatment. A comprehensive analysis of learned traits will provide insight into the neurobiological mechanisms associated with schizophrenia and contribute to our understanding of the disorder. We have explored the neurobiological mechanisms underlying schizophrenia to enhance our understanding of the disorder.

1.2 PROPOSED MODELS

This study presents an evaluation of the accuracy and performance of different models, namely VGG16-CNN, Basic Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM), for the detection of schizophrenia using EEG data from adolescents. The dataset consists of EEG records from two groups: healthy controls and individuals with symptoms of schizophrenia. The aim is to systematically test and compare the performance of each model in accurately distinguishing between the two groups. By conducting a comparative study, this study seeks to gain insights into the strengths and limitations of each model in detecting schizophrenia based on EEG data. The findings will contribute to a better understanding of the suitability and effectiveness of these models in the context of schizophrenia detection, which will facilitate future research and the development of improved diagnostic approaches for this mental disorder, with the potential to aid early detection and intervention by clinicians and ultimately improve outcomes for patients with schizophrenia.

1.3 BLUEPRINT BRIEF

As the dataset used in this study comprises EEG records from two distinct groups: healthy controls and individuals exhibiting symptoms of schizophrenia. The primary objective is to systematically test and compare the performance of each model in accurately distinguishing between these two groups. By conducting a comparative analysis, the study aims to uncover the strengths and limitations of each model in detecting schizophrenia based on EEG data. The findings from this research have the potential to significantly contribute to our understanding of the suitability and effectiveness of these computational models in the context of schizophrenia detection. Such insights can facilitate early detection and aid in the development of improved diagnostic approaches for this mental disorder.

Chapter 2

Background

2.1 DEEP LEARNING AND MEDICAL INSIGHTS

2.1.1 NEURO BIOLOGICAL INSIGHTS

Schizophrenia, a complex psychiatric disorder, is characterized by perceptual, cognitive, and behavioural disturbances. Neurobiological insights into the disorder have revealed key findings. The dopamine hypothesis suggests excessive dopamine activity as a contributor to positive symptoms, while glutamate dysfunction, particularly involving NMDA receptors, may contribute to cognitive deficits. Neurodevelopmental abnormalities, disrupted white matter integrity, and genetic factors have been implicated, along with GABAergic dysfunction and altered synaptic function. These insights shed light on the complex interplay of genetic, neurochemical, and environmental factors underlying schizophrenia and provide a foundation for further understanding its pathophysiology.

2.1.2 DEEP LEARNING MODELS

Schizophrenia, a complex psychiatric disorder characterized by perceptual, cognitive, and behavioural disturbances, necessitates early detection and accurate diagnosis for effective treatment. Deep learning techniques have exhibited promise in various medical applications, including mental health disorders. This study aims to evaluate the performance and accuracy of different models, namely VGG16-CNN sub architecture, Basic convolutional neural networks (CNN), and long short-term memory (LSTM), for detecting schizophrenia. The neuroimaging data is pre-processed, extracting high-level features from brain images using the VGG16 architecture. These features are then inputted into the basic CNN model to learn spatial patterns and enhance their representation. LSTM layers are employed to capture temporal dependencies. By utilising a comprehensive dataset of brain images from schizophrenia patients and healthy controls, the study conducts a comparative analysis of the models' accuracy, surpassing individual models' performance. Furthermore, in-depth analysis of learned features provides valuable insights into the underlying neurobiological mechanisms associated with schizophrenia, unveiling potential biomarkers and patterns that contribute to our understanding of the disorder's pathophysiology. Ultimately, this research contributes to a possible robust deep learning framework for integrating VGG16, Basic CNN, and LSTM, for facilitating early detection and intervention by clinicians, thereby potentially improving outcomes for patients with schizophrenia; achieved by testing each of the above-mentioned model performances and then evaluating the results and then integrating the perk-based limitations and vice versa of each model and the subsequent counterpart fix in the other model.

2.2 ELECTROENCEPHALOGRAPHY (EEG)

2.2.1 EEG FEATURES AND SCHIZOPHRENIA

Electroencephalography (EEG) has emerged as a valuable non-invasive tool for studying brain activity and identifying potential biomarkers associated with schizophrenia. Analysis of a resting EEG reveals spectral power anomalies such as: increasing theta and delta power or decreasing alpha power. This indicates an impairment of neural oscillations and a potential cognitive impairment. Event-related potentials (ERPs) provide insight into cognitive processes, and a schizophrenic patient exhibits a reduced response to her P300 amplitude and mismatch negativity (MMN), indicating high levels of cognitive function and sensory processing deficits. suggesting. Changes in connectivity patterns derived from EEG signals, including coherence and graph theory analyses, reveal dyssynchronization and abnormal functional networks in schizophrenia patients. Machine learning approaches applied to EEG data enable the detection and classification of schizophrenia, contributing to early detection and personalized treatment. Overall, EEG function plays a key role in understanding the neurobiological mechanisms of schizophrenia and improving patient outcomes through improved detection and treatment strategies.

2.2.1 EEG INFORMATION

- <u>Abnormal EEG patterns:</u>

Patients with schizophrenia often exhibit different EEG patterns during recordings at rest, including: increasing theta and delta power or decreasing alpha power. These abnormal spectral performance patterns serve as potential biomarkers for the presence of schizophrenia.

- <u>Event-Related Potential (ERP):</u>

ERPs are time-limited brain responses to specific stimuli or events and can provide insight into cognitive processes. In schizophrenia, ERP abnormalities such as decreased P300 amplitude or impaired mismatch negative (MMN) responses indicate deficits in attention, memory, and perception. These ERP abnormalities may indicate the presence of schizophrenia.

- <u>Connectivity measures:</u>

EEG-based connectivity measures such as coherence, phase synchrony, and graph theory analysis may reveal impairments in coordination and communication between different brain regions in schizophrenia patients. Changes in functional connectivity patterns, characterized by synchronization disruptions and network anomalies, have been observed and may indicate disruption. [9]

By leveraging the rich information embedded in EEG features, the algorithms can learn discriminative patterns and build predictive models. These models can aid in the early detection of schizophrenia and contribute to personalized treatment strategies.

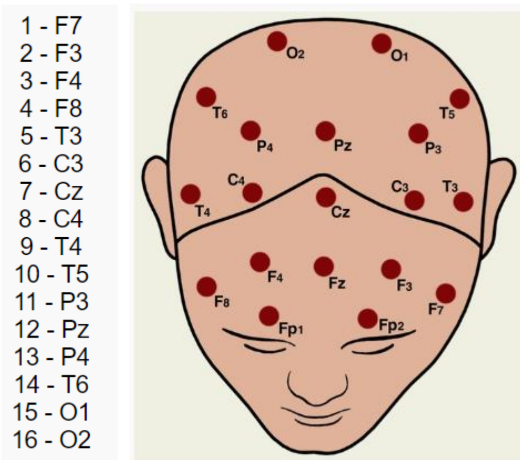
Chapter 3

Problem Definition

The objective of this study is to address the problem of accurately detecting and classifying schizophrenia using EEG data from adolescents. The study aims to compare and evaluate the performance of three different models, namely VGG16, CNN, and LSTM [14], in accurately distinguishing between healthy controls and individuals with symptoms of schizophrenia based on EEG recordings. The goal is to assess the strengths and limitations of each model in detecting schizophrenia and gain insights into their effectiveness in analysing EEG data for diagnostic purposes. By providing a comprehensive analysis, this research seeks to contribute to the development of improved diagnostic approaches and facilitate early detection and intervention for patients with schizophrenia.

3.1 DATASET

The dataset used is EEG data for adolescents with symptoms of schizophrenia and healthy controls [6].



The dataset comprises EEG records from two groups of adolescents: a healthy group (n = 39) and a group with symptoms of schizophrenia (n = 45). Each TXT file represents one subject's EEG recording and contains columns with EEG samples from 16 electrode channels. The columns correspond to specific channels, and each number represents the EEG amplitude (in µV) at a given sample. With a sampling rate of 128 Hz, 7680 samples correspond to a 1-minute segment of the EEG record. The dataset includes the topographical positions of the electrode channels, which aid in understanding the spatial distribution of the recorded EEG signals and their relationship to different brain regions. This dataset offers an opportunity to investigate the EEG characteristics and potential biomarkers associated with schizophrenia in adolescents, facilitating further research in the field.

Chapter 4

Objectives

The objective of this research is to evaluate and compare the performance of VGG16, CNN, and LSTM models for the detection of schizophrenia using EEG data from adolescents. Specific objectives include:

1. Pre-processing the EEG data
2. Implementing the VGG16 model
3. Implementing the CNN model
4. Implementing the LSTM model
5. Comparative analysis: Conduct a comprehensive comparative analysis of the accuracy and performance metrics of the VGG16, CNN, and LSTM models. Assess their strengths and limitations in detecting schizophrenia based on the EEG data.
6. Interpretation of results Interpret the results of the comparative study to gain insights into the strengths and weaknesses of each model. Identify the model that demonstrates the highest accuracy and the most promising performance for schizophrenia detection using EEG data.

Chapter 5

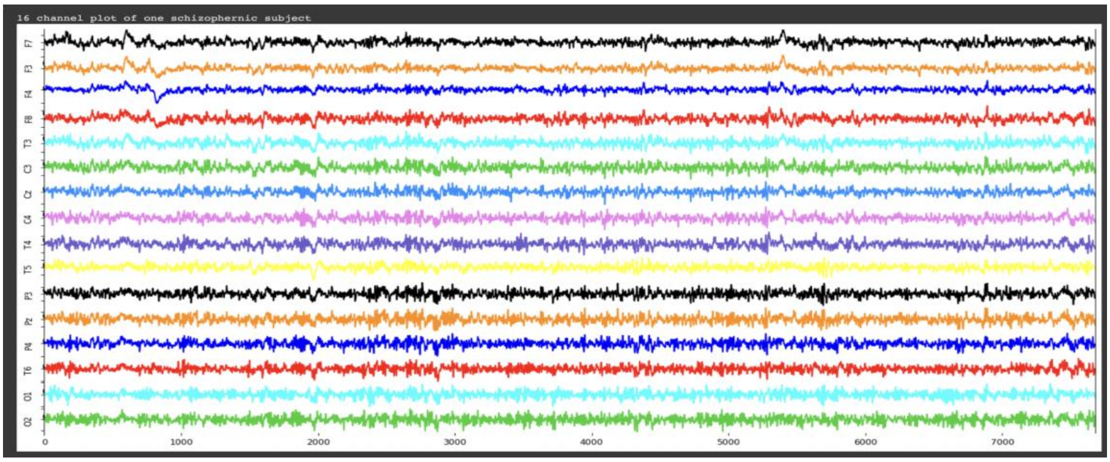
Methodology

5.1 WORKFLOW

- Collected EEG data on schizophrenia from credible sources.
- Applied suitable filtering techniques to remove noise and artifacts from the EEG data.
- Partitioned the data into shorter epochs [1] for simplifying analysis.
- Now, generate spectrogram pictures from pre-processed EEG data using the Short-Time Fourier Transform (STFT).
- Divide the dataset into training, validation, and testing sets.
- Train deep learning models such as Basic CNN, VGG16, and LSTM using the spectrogram pictures and fine-tune hyperparameters.
- Evaluate the performance of each trained model using the validation set and relevant performance indicators (accuracy, precision, recall, and F1-score).
- Select the model with the highest accuracy as the best performer.
- Assess the chosen model's performance using the testing set.
- Analyse the outcomes and draw inferences based on model training and testing.
- Identify any study limitations and provide recommendations for further research.

5.2 DATA PREPROCESSING

1. <u>Signal Filtering:</u> The EEG data was filtered using bandpass filters to eliminate noise and unwanted frequencies outside the range of interest, ensuring that the neural activity relevant to schizophrenia was retained.[10]



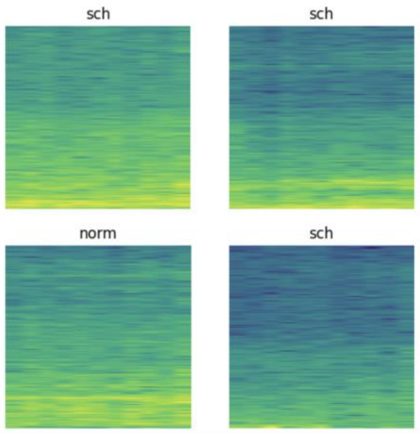
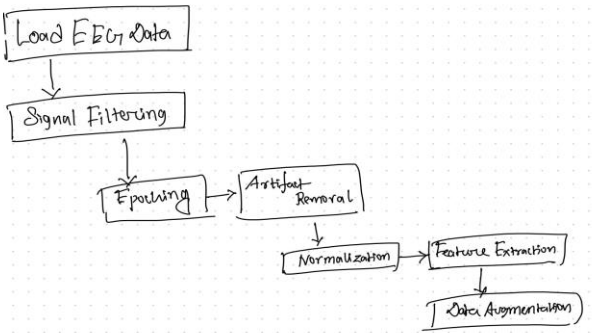


Fig.a) 16-channel plot of EEG signals of a schizophrenic subject

Fig b) The spectrogram plot of the schizophrenic and healthy person in comparison

2. <u>Epoching:</u> The EEG data was segmented into smaller epochs based on triggers or events that occurred at the start and end of each epoch. This segmentation allowed for a focused analysis of brain activity during specific time periods, facilitating the identification of patterns related to schizophrenia.
3. <u>Artifact Removal</u>: Independent Component Analysis (ICA) and regression-based methods were employed to identify and remove artifacts in the EEG signals, such as eye blinks, muscle movements, or electrode drift. Hence, we separated the artifacts from the underlying brain signals associated with schizophrenia.
4. <u>Normalization:</u> Normalization of the EEG data was performed by removing the mean of the signal from each epoch. This normalization step reduced inter-subject variability and enhanced comparability across different participants.



5. <u>Feature Extraction:</u> Various feature extraction techniques were applied to the pre-processed EEG data to capture relevant information. Time-frequency analysis was used to identify variations in phase coherence with time and frequency, which gave insights into the dynamic characteristics of brain activity associated with schizophrenia.[13]
6. <u>Data Augmentation:</u> Within our future scope, we wish to augment the available dataset and improve generalization [7]; hence, data augmentation techniques can be applied. These techniques may introduce variations such as random shifts or rotations to the EEG data, which will expand the diversity of the training dataset.

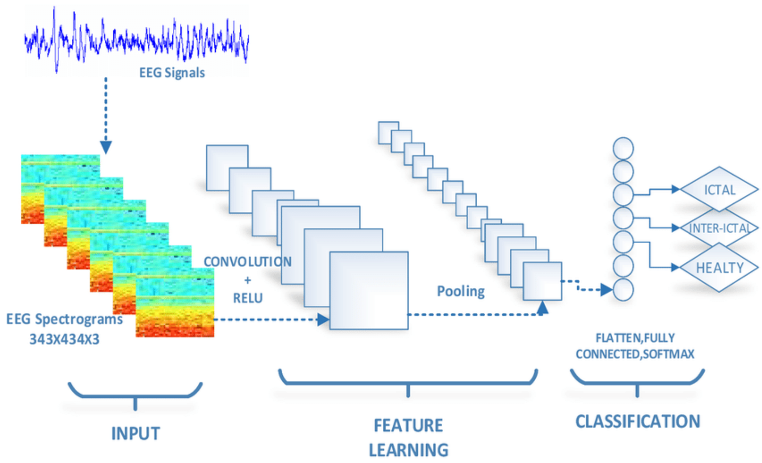
5.3 MODEL TRAINING

Basic CNN, VGG16, and LSTM are trained using the spectrogram pictures generated from the pre-processed EEG data. The models are trained to learn patterns and features that are indicative of schizophrenia. During the training process, hyperparameters such as learning rate, number of layers, filters, and other architectural configurations are fine-tuned to optimize the performance of the models. This involves experimenting with different combinations of hyperparameter values and evaluating their impact on the model's accuracy and convergence. By adjusting these hyperparameters, the models can better capture the underlying patterns in the spectrogram pictures and improve their ability to accurately classify EEG data from individuals with schizophrenia and healthy controls.

5.4 MODEL ARCHITECTURES

5.4.1 CONVOLUTIONAL NEURAL NETWORKS (CNN)

A CNN model architecture for EEG data analysis involves several key components. The input layer receives the EEG data, which is typically represented as a 2D or 3D matrix. Convolutional layers apply filters to capture spatial patterns, followed by activation functions to introduce non-linearity and pooling operations to down sample the features. Fully connected layers aggregate the learned features and capture global relationships. Finally, the output layer produces the classification or regression output.[11] Pre-processing steps and data augmentation techniques can be included to improve performance. It's important to customize the architecture based on data characteristics and task requirements.

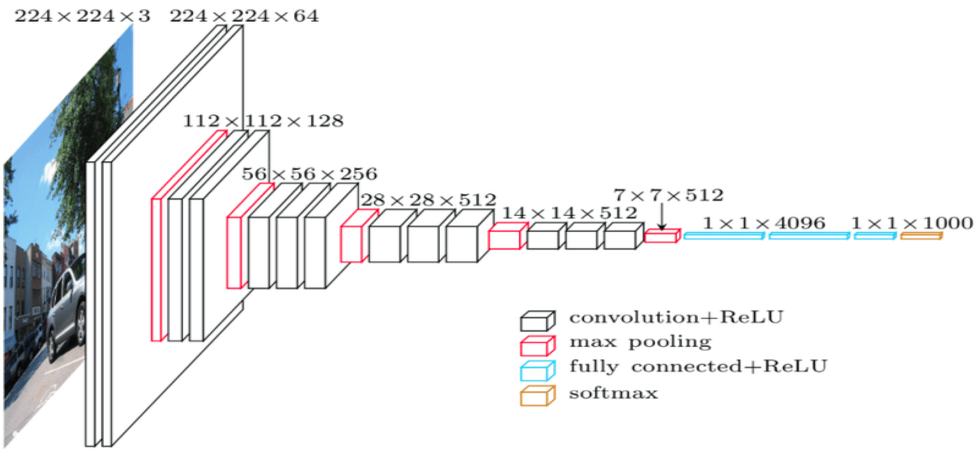


Benefits of CNN for EEG data in schizophrenia detection:

1. Localized Pattern Extraction: CNNs are effective at capturing localized patterns in EEG data, such as spatial and temporal features associated with specific brain activities. They can automatically learn and extract relevant features from the raw data.
2. Translation Invariance: CNNs are invariant to translations, meaning they can detect patterns regardless of their exact position in the EEG data. This property is advantageous when analysing brain signals that may vary in their timing or spatial location.
3. Hierarchical Feature Learning: CNNs learn features in a hierarchical manner, with early layers capturing low-level features (e.g., edges or simple patterns) and deeper layers capturing more complex and abstract features. This allows for the detection of subtle and higher-level patterns associated with schizophrenia.
4. Parameter Sharing: CNNs use parameter sharing across the spatial dimensions of the input data, which reduces the number of parameters to be learned. This makes them computationally efficient for analysing large-scale EEG datasets.
5. Robust to Noise: EEG data is often noisy due to various artifacts and biological sources. CNNs can handle noisy data by learning to focus on the most relevant features while disregarding noise.

5.4.2 VISUAL GEOMETRY GROUP (VGG16)

The VGG16 architecture, originally designed for image classification, can be adapted for EEG data analysis with certain modifications. To apply VGG16 to EEG data, adjustments need to be made to the input layer to accommodate the dimensions of the EEG data. The convolutional layers, activation functions, and pooling layers can be used, treating each channel as a separate "image" in the input. The fully connected layers at the end of the architecture may need to be modified or replaced to suit the specific EEG task, such as schizophrenia detection. However, it's important to note that the performance of VGG16 on EEG data may not be optimal due to the different characteristics of EEG data compared to images.



Therefore, experimentation, customization, and incorporation of EEG-specific pre-processing and augmentation techniques are recommended for better results.

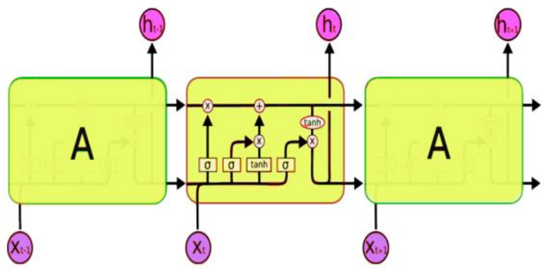
The VGG network is a CNN that was first launched in 2014. It is well-known for its deep architecture (up to 19 layers) and performance on image classification tasks.

The VGG network is similar to a conventional CNN, but it contains more convolutional layers and fewer pooling layers, allowing it to capture finer-grained characteristics in the input data.

VGG16 is a 16-layer deep neural network, as the name implies. VGG16 is thus a rather large network with 138 million parameters—a massive network even by today's standards. The simplicity of the VGGNet16 design, on the other hand, is its key selling point.

5.4.3 LONG SHORT-TERM MEMORY (LSTM)

The LSTM network is a recurrent neural network (RNN) intended to handle sequential data such as time series or natural language. It is particularly beneficial for activities requiring the retention of knowledge over a longer period of time, as it may learn to selectively recall or forget information as needed.



The LSTM network is made up of memory cells that are linked together by gates that govern the flow of information in and out of the cells. Based on the input data, these gates are trained to selectively recall or forget information. A typical LSTM [12] network is made up of multiple memory blocks known as cells. The cell state and the concealed state are the two states that are passed to the following cell. Memory blocks are in charge of remembering things, and modifications to this memory are carried out via three primary processes known as gates. Each of these is described more below.

Chapter 6

Results and Discussion

6.1 PERFORMANCE EVALUATION:

Performance evaluation on deep learning models implemented (VGG16, CNN, and LSTM models for schizophrenia detection) was done using metrics: accuracy, precision, recall, and F1-score.

- Accuracy = (number of true positives + number of true negatives) / (total number of samples)
- Precision = number of true positives / (number of true positives + number of false positives)
- Recall = number of true positives / (number of true positives + number of false negatives)
- F1-score = 2 * precision * recall / (precision + recall)

```
2/2 [==============================] - 0s 26ms/step - loss: 0.2916 - accuracy: 0.9028
simple CNN's accuracy: 90.28%
2/2 [==============================] - 1s 73ms/step - loss: 0.4604 - accuracy: 0.9444
VGG-16's accuracy: 94.44%
```

```
3/3 [==============================] - 1s 34ms/step - loss: 0.7469 - accuracy: 0.7500
model's accuracy: 75.0%
```

6.2 CONCLUSIVE DISCUSSION

The evaluation metrics provided of the performance of the above models: VGG16-CNN gave the highest accuracy of 94.44%, while Basic CNN model gave an accuracy of 90.28% and the least was of the LSTM model of as low as 75.0%.
VGG is a complex network architecture with more layers, hence allows it to capture more complex patterns in the EEG signals. The basic CNN was also able to capture some of these patterns, but may not be as effective as VGG due to its simpler architecture. The LSTM, is primarily designed for sequential data and may not be as well-suited to this particular dataset. Moreover, the relatively low accuracy of the LSTM model indicates that spectral information may not be as informative as temporal information in detecting schizophrenia leading to a lower accuracy.

Chapter 7

Conclusion and Future Scope

- The results obtained by the implementation of the above basic CNN, VGG16 are quite promising and suggest that these models have the potential to be powerful tools in detecting schizophrenia from EEG data.
- The high accuracy of the VGG16 and CNN models demonstrates the effectiveness of using convolutional neural networks for feature extraction from EEG signals. It also suggests that the models have effectively captured the discriminative patterns present in the EEG data that are characteristic of schizophrenia.
- While the LSTM model's performance shows the importance of considering the temporal dynamics of the signals; additionally, the relatively low accuracy of the LSTM model indicates that spectral information may not be as informative as temporal information in detecting schizophrenia.
- Overall, these findings demonstrate the potential of usage of CNN and its architectures for
- Early detection of schizophrenia using EEG data and highlights the importance of considering both spectral and temporal information in such analyses. Future studies could consider using a combination of these models to improve detection accuracy further.
- In future work, it would be interesting to explore the combination model built with the cumulative architectural dependencies of the three models, tested in this paper. Moreover, metric evaluations can be implemented using different statistical analogies.

Appendices

CNN classification with a simple model

```
num_classes = len(class_names)

simple_model = tf.keras.Sequential([
  layers.Conv2D(32, 3, activation='relu'),
  layers.Conv2D(32, 3, activation='relu'),
  layers.MaxPooling2D(),
  layers.Conv2D(32, 3, activation='relu'),
  layers.Conv2D(32, 3, activation='relu'),
  layers.MaxPooling2D(),
  layers.Conv2D(64, 3, activation='relu'),
  layers.Conv2D(64, 3, activation='relu'),
  layers.MaxPooling2D(),
  layers.Conv2D(128, 3, activation='relu'),
  layers.Conv2D(128, 3, activation='relu'),
  layers.Conv2D(128, 3, activation='relu'),
  layers.MaxPooling2D(),
  layers.Flatten(),
  layers.Dense(512, activation='relu'),
  layers.Dense(512, activation='relu'),
  layers.Dense(num_classes, activation='softmax')
])

opt = keras.optimizers.Adam(learning_rate=0.0001)
simple_model.compile(optimizer=opt, loss=tf.keras.losses.SparseCategoricalCrossentropy(), metrics=['accuracy'])


simple_model.fit(
  train_ds,
  validation_data=val_ds,
  epochs=60)
```

CNN classification with a VGG-16 model

```
num_classes = len(class_names)

vgg16_model = keras.Sequential()
vgg16_model.add(layers.Conv2D(input_shape=(224,224,3),filters=64,kernel_size=(3,3),padding="same", activation="relu"))
vgg16_model.add(layers.Conv2D(filters=64,kernel_size=(3,3),padding="same", activation="relu"))
vgg16_model.add(layers.MaxPool2D(pool_size=(2,2),strides=(2,2)))
vgg16_model.add(layers.Conv2D(filters=128, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.Conv2D(filters=128, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.MaxPooling2D(pool_size=(2,2),strides=(2,2)))
vgg16_model.add(layers.Conv2D(filters=256, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.Conv2D(filters=256, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.Conv2D(filters=256, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.MaxPooling2D(pool_size=(2,2),strides=(2,2)))
vgg16_model.add(layers.Conv2D(filters=512, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.Conv2D(filters=512, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.Conv2D(filters=512, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.MaxPooling2D(pool_size=(2,2),strides=(2,2)))
vgg16_model.add(layers.Conv2D(filters=512, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.Conv2D(filters=512, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.Conv2D(filters=512, kernel_size=(3,3), padding="same", activation="relu"))
vgg16_model.add(layers.MaxPooling2D(pool_size=(2,2),strides=(2,2)))
vgg16_model.add(layers.Flatten())
vgg16_model.add(layers.Dense(4096,activation="relu"))
vgg16_model.add(layers.Dense(4096,activation="relu"))
vgg16_model.add(layers.Dense(num_classes, activation="softmax"))
vgg16_model.summary()

opt = keras.optimizers.Adam(learning_rate=0.0001)
vgg16_model.compile(optimizer=opt, loss=tf.keras.losses.SparseCategoricalCrossentropy(), metrics=['accuracy'])

vgg16_model.fit(train_ds, validation_data=val_ds, epochs=60)
```

CNN classification with a LSTM model

```
model = Sequential()

model.add(TimeDistributed(Conv2D(32, (3, 3), padding='same',activation = 'relu'), input_shape = (12, 224, 224, 3)))
model.add(TimeDistributed(MaxPooling2D((2, 2), strides=(2, 2))))
model.add(TimeDistributed(Conv2D(64, (3, 3), padding='same',activation = 'relu')))
model.add(TimeDistributed(MaxPooling2D((2, 2), strides=(2, 2))))
model.add(TimeDistributed(Conv2D(128, (3, 3), padding='same',activation = 'relu')))
model.add(TimeDistributed(MaxPooling2D((2, 2), strides=(2, 2))))
model.add(TimeDistributed(Flatten()))

model.add(LSTM(64))
model.add(Dense(512, activation = "relu"))
model.add(Dense(2, activation = "softmax"))

model.summary()

optimizer = keras.optimizers.Adam(learning_rate=0.00009)
model.compile(loss = 'sparse_categorical_crossentropy', optimizer=optimizer, metrics=["accuracy"])

model_history = model.fit(X_train, y_train, epochs=60, batch_size=8 ,validation_split=0.2, callbacks = [checkpoint])
```

REFERENCES

[1] Neuhaus A, Popescu FC, Grozea C, Hahn E, Hahn C, Opgen-Rhein C, et al., Single-Subject Classification of Schizophrenia by Event-Related Potentials during Selective Attention, NeuroImage, 2011;55(2): p 514–521 < For medical parameter understanding>

[2] Yandong L, Zhongwei M, Lu W, Li Y. Automatic Removal of the Eye Blink Artifact from EEG using an ICA-based Template Matching Approach, Physiological Measurement, 2006;27: p 425–436. <For EEG spectrogram reference>

[3] Peled A, Geva A, Kermen W, Blankfeld H, Estfandiarfard R, Nordahl T. Functional Connectivity and Working Memory in Schizophrenia: An EEG Study, International Journal of Neuroscience, 2001; p 106: 47–61. < For EEG pre-processing>

[4] Charlson, F.J.; Ferrari, A.J.; Santomauro, D.F.; Diminic, S.; Stockings, E.; Scott, J.G.; McGrath, J.; Whiteford, H.A. Global Epidemiology and Burden of Schizophrenia: Findings from the Global Burden of Disease Study 2016. Schizophr. Bull. 2018, 44, p 1195–1203.
< For medical parameter understanding>

[5] Kendler, K.S. Phenomenology of Schizophrenia and the Representativeness of Modern Diagnostic Criteria. JAMA Psychiatry 2016, 73, p 1082–1092. < For medical parameter understanding>

[6] Makarov V. V., Romanov A. N. (2017). EEG data for adolescents with symptoms of schizophrenia and healthy controls [Dataset]. Moscow State University. < For EEG parameter understanding>

[7] Borisov, S.V., Kaplan, A.Y., Gorbachev kaya, N.L. et al. Analysis of EEG Structural Synchrony in Adolescents with Schizophrenic Disorders. Hum Physiol 31,p 255–261 (2005).
< For EEG parameter understanding>

[8] Benes, F.M., Emerging Principles of Altered Neural Circuitry in Schizophrenia, Brain Res. Rev., 2000, vol. 31, p. 251.< For medical parameter understanding>

[9] Tauscher, J., Fischer, P., Neumeister, A., et al., Low Frontal Electroencephalographic Coherence in Neuroleptic-Free Schizophrenic Patients, Biol. Psychiatry, 1998, vol. 44, no.6, p. 438.< For medical EEG understanding>

[10] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 2015, pp. 730-734, < For VGG16 CNN reference understanding>

[11] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 2017, pp. 1-6, < For CNN understanding>.

[12] Sepp Hochreiter, Jürgen Schmidhuber; Long Short-Term Memory. Neural Comput 1997; 9 (8): 1735–1780. < For LSTM understanding>

[13] W. A. W Azlan and Y. F. Low, "Feature extraction of electroencephalogram (EEG) signal - A review," 2014 IEEE Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, Malaysia, 2014, pp. 801-806,< For medical EEG understanding>

[14] C. Shi, X. Yu, E. Cheung, D. Shum and R. Chan, "Revisiting the therapeutic effect of rTMS on negative symptoms in schizophrenia: A meta-analysis", Psychiatry research, vol. 215, no. 3, pp. 505-513, 2014. < For medical understanding>

[15] L. Kuncheva and J. Rodriguez, "Interval feature extraction for classification of event-related potentials (ERP) in EEG data analysis", Progress in Artificial Intelligence, vol. 2, no. 1, pp. 65-72, 2013. < For deep learning & EEG understanding>