EEG Biometric: Rapid Identification across varied Electrode Configurations and Limited Recording Times

Submitted in partial fulfillment of the requirements for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

INFORMATION TECHNOLOGY

Submitted by

Akash Biswas , Ritej Dhamala, Aashsih Agarwal

(IIT2020001, IIT2020006, IIT2020009)

With the supervision of

Dr. Anupam Agrawal

(Associate Professor)

Department of Information Technology

Indian Institute of Information Technology - Allahabad

Prayagraj-211015

May 2024

CANDIDATE S' DECLARATION

This is hereby declare that the work which we presenting here in this report entitled "EEG Biometric: Rapid Identification across varied Electrode Configurations and Limited Recording Times ", submitted towards fulfillment of THESIS that will be report of Bachelor of Technology in IT at Indian Institute of Information Technology, Allahabad,Prayagraj that it is an authenticated for the record of our original work carried out under the guidance of **Dr. Anupam Agrawal**. I am profoundly grateful to Mr. GC Jana(Ph.D) for his expert guidance and continuous encouragement throughout to see that this project meets its target since its commencement to its completion. With the Due acknowledgements which have been made in this text to all of which other material used.That the project was done in the fully compliance with all the requirements and constraints of the prescribed curriculum.

Signature of Students: ————————————-

Name: **Akash Biswas, Ritej Dhamala, Aashsih Agarwal**

Enrolment No: **IIT2020001, IIT2020006, IIT2020009**

Department: **Information Technology**

Indian Institute of Information Technology - Allahabad

CERTIFICATE FROM SUPERVISOR

This is to certify that the statement made by the candidate is all correct to the best of our knowledge and belief. That the project titled ad "EEG Biometric: Rapid Identification across varied Electrode Configurations and Limited Recording Times" is the record for the candidates' work carried out by him under my guidance and supervision. I do hereby recommend that it should be accepted in the fulfilment of the requirements of the BACHELOR'S THESIS at IIIT Allahabad.

Signature of the Supervisor: ——————————-

Name of the Supervisor: **Dr. Anupam Agrawal**

Designation: **Professor**

Department: **Dept. of Information Technology**

Indian Institute of Information Technology - Allahabad

CERTIFICATE OF APPROVAL

The foregoing thesis is hereby approved as a creditable study in Information Technology and its allied areas. It is carried out and presented in a satisfactory manner to warrant its acceptance as a prerequisite to the degree for which it has been submitted. It is understood that by this approval the undersigned do not necessarily endorse or approve any statement made, opinion expressed or conclusion drawn therein but the thesis only for the purpose for which it is submitted.

Committee Members for Evaluation of the Thesis for Final Examination:

....

....

....

....

ACKNOWLEDGEMENTS

First of all, I would like to thank my institute **Indian Institute of Information Technology, Allahabad**. This institute gave a unique platform to learn new techniques and enhance my technical skills. I am highly grateful to the honorable Director, **IIIT Allahabad, Dr. Mukul Sutaone** for his helping attitude and encouragement to excel in studies.

I extend my thanks to my brilliant thesis supervisor, **Prof. Anupam Agrawal, Associate Professor, IIIT Allahabad**, for his excellent ideas and insights. This report would not have been possible without his many important contributions, and I am heartily grateful for his open and honest feedback. Having a good thesis supervisor is key when studying for a higher degree, and I could simply not have wished for a better guide and work partner than **Dr. Anupam Agrawal**.

Akash Biswas, Ritej Dhamala, Aashsih Agarwal
IIT2020001, IIT2020006, IIT2020009
B.Tech (IT)

Abstract

The use of biometric recognition for identifying individuals has become common in modern secure systems. The growing need for this technology and the simplicity with which it can be utilised have, however, also led to an increase in the accompanying security vulnerabilities. Therefore, it is essential to discover biometric characteristics that are more reliable than the ones that are currently being utilised. A viable biometric alternative has recently been found as the brain signal that is acquired using the electroencephalogram (EEG) approach. The EEG technique possesses an excellent level of distinctiveness, consistency, and global application capability.

The purpose of this study is to demonstrate an innovative method of biometric identification that makes use of EEG data. Feature extraction is accomplished through the utilisation of a short-time Fourier transform (STFT), which is then followed by the channel-wise division of the data being processed. This is followed by the implementation of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models for each channel on an individual basis. The final step involves the integration of these models using a hard-voting ensemble classifier. Through this ensemble approach, the system achieves improved accuracy and robustness by combining the predictions from multiple models.

Contents

List of Figures

List of Tables

Abbreviations

STFT	Short-time Fourier transform
CNN	Convolution Neural Network
RNN	Recurrent Neural Network
LSTM	Long short-term memory
FC	Fully Connected
Ind	Individual

Chapter 1

Introduction

Biometric identification systems are becoming an essential part of modern civic life that totally has applications, ranging from unlocking the mobile phone to person identification for security surveillance applications and stuff. All biometric systems are like multiclass classification systems, where each class corresponds to an individual. We can retrieve the feature vectors for each candidate from the candidate's physical features, behavioural patterns, and cognitive responses. Facial recognition, fingerprint, and iris scanners are among the most widely used modern biometric systems. However, they do not guarantee liveness detection and are susceptible to deception. Man's exposure to the environment makes these biometric traits vulnerable to spoofing attacks. Most surfaces the candidate touches will imprint their fingerprints, and even a long-range camera can capture the candidate's face. Researchers are focusing more and more on cognitive biometric authentication based on electroencephalograms as an alternative to these biometric authentication methods that can't be faked.

EEG technology detects the electrical activity of brain waves by placing electrodes on the scalp.[1]. These electrodes measure the synchronous activity of millions or even trillions of neurons with similar spatial orientation, providing valuable data on cognitive functions and neural reactions. EEG recordings, measured in microvolts, are specific to individuals and cannot be replicated. Due to its excellent temporal resolution, mobility, affordability, and ease of use, EEG is the most commonly used non-invasive method for collecting brain signals. EEG waves have unique biometric qualities, making EEG a valuable tool in various applications. With the increasing use of wireless and portable headsets like the Epoc[2], EEG data collection and management have become easier. Thus, it is not possible to replicate or mimic the EEG patterns of another person. Since brain waves have special biometric qualities, EEG has been used extensively in the past ten years. In traditional scalp EEG, electrodes are placed on different parts of the scalp according to a standardized method. Recorded EEGs provide information about the brain's specific reactions to internal or external stimuli. Information on the brain's particular reaction to the given external or internal stimuli can be found in recorded EEGs. Event-related potentials (ERPs) are the associated potential changes in the EEG that occur in response to various stimuli [3]. Event-related potentials (ERPs) are changes in the EEG associated with responses to various stimuli, such as visual or auditory stimuli [4], or during the brain's rest state.

In this paper, we work on identification using auditory evoked potentials (AEPs) and visual evoked potentials (VEPs).The AEPs are neuroelectrical potentials generated by applying an auditory stimulus to the ear while VEPS are the ones generated as response to some visual elements like images or videos. As the case for biosignals, this signalis more secure than face and fingerprints. Moreover, the AEPs and VEPs are stimulus-dependent, i.e. they change by varying the auditory and visual stimulus [5]. This adds an extra feature to the AEP and VEP signals over conventional biometrics and other bio signals allowing these signals to be cancellable. So, even if the AEP signal is breached or hacked,it still can be reused for biometric authentication by varying the respective stimulus.

Biometric authentication utilising AEP and VEP signals can be advantageous in specialised applications such as tele-medicine and situations where protective equipment is employed. Telemedicine applications utilise AEPs for a range of medical purposes, such as detecting emotions and assessing mental health [6]. Advances in wearable technology and biomedical instrumentation have made it possible to conveniently and remotely record AEPs. Integrating the advanced biometric feature into AEP/VEP will streamline the process of gathering data for mental health assessment and simultaneously identify the patient using the same signal, eliminating the need for further data collection for identification purposes. For this reason, the authers have used a dataset containing such signals to build a biometric identification system.

The rest of this paper is divided into six sections. Section II describes the Related Work. Section III provides a detailed description of the data set, including the stimuli used, the equipment employed, and the acquisition setting and protocol the proposed framework and methodology. Section IV describes the experimental setup of the model, followed by Section V contains an analysis of result, compares our work with previously existing works and further discussion. Finally, Section VI summarizes the major conclusions that can be drawn from this work.

Chapter 2

Literature survey

2.1 Literature survey for EEG based Biometric Identification

Many works have been proposed in the papers about the EEG biometric methods so far. After going through various papers, we found out that the BED dataset and the Physionet Auditory Potential Evoked dataset were used the most. In their first paper, they proposed a Convolutional Neural Network (CNN) approach for biometric identification using the BED dataset. The proposed CNN achieved a correct classification rate (CCR) of 72.71%. However, by utilising only 4 channels, the CCR improved to an average of 83.51%. Additionally, the training time was significantly reduced from 626 to 306 seconds. It demonstrates the effectiveness of CNN architectures in processing EEG data for biometric identification tasks. The improvement in classification accuracy and reduction in training time highlight the potential for optimising EEG-based biometric systems using CNN models [7].

In the other paper, they investigated the performance of multi-channel and single-channel approaches using the BED dataset. The multi-channel approach achieved an impressive accuracy of 95%. On the other hand, the single-channel approach yielded training and testing accuracies of 98.27% and 97.13%, respectively. The study compares the efficacy of utilising multiple EEG channels versus focusing on individual channels for biometric identification. It demonstrates that both approaches can achieve high accuracy, with the single-channel approach exhibiting slightly superior performance [8].

Finally, for auditory evoked potential Dataset (AEP), the research examines how auditory stimuli (AS) affect EEG-based biometric authentication to improve performance. It examines how AS affects authentication accuracy, how AS conduction methods affect it, and how AS language variance affects it. Used MLP, KNN, and XGBoost classifiers on EEG epochs and cluster map features. They concluded that auditory stimuli boost authentication accuracy by 9.27% over resting states [9].

Chapter 3

Problem definition and Objectives

Conventional EEG-based biometric identification systems often struggle to meet the demands of real-world applications due to their reliance on specific electrode setups and lengthy recording periods. This dependence makes it challenging to create a reliable and effective system for quick biometric identification, as the number of electrodes and recording times can vary significantly. Moreover, the extended recording times inconvenience users and limit the practicality of EEG-based identification systems, especially in time-critical scenarios. To address these limitations, the project aims to develop a robust EEG-based biometric identification system with several key objectives.

These objectives include creating a solution independent of electrode count, achieving accurate identification results within a 2-second recording window, prioritizing a seamless user experience, and exploring Ensemble Learning methods to ensure high accuracy across various electrode setups and limited recording durations. By combining innovative approaches and cutting-edge technologies, the project seeks to revolutionize EEG-based biometric identification, making it more accessible and practical for diverse real-world applications.

Chapter 4

Methodology

Our proposed methodology contains a completely comprehensive framework designed to effectively address the big challenges which are inherently present in EEG-based biometric identification systems which are variable electrode configuration and inconsistent data acquisition time period. Our main is to design a biometric identification system using electroencephalogram irrespective of the number of electrodes used during eeg data reading and keep the duration of data acquisition as short as 2 seconds for fast and efficient identification of subjects. Thus prepared methodology encompasses several key steps:

4.1 Dataset

For our project, we utilize two publicly available datasets, namely the **Auditory Evoked Potential EEG-Biometric dataset (AEP datatset)** and the **Biometric EEG dataset (BED dataset)**. These datasets provide valuable EEG recordings obtained from human subjects under various experimental conditions.

4.1.1 Auditory Evoked Potential EEG-Biometric dataset:

This dataset consists of over 240 two-minute EEG recordings from 20 volunteers, including both resting-state and auditory stimuli experiments. EEG signals were recorded using the 10/10 worldwide EEG system, with electrodes at specific locations such as T7, F8, Cz, and P4. Reference and ground electrodes were inserted in the left and right ears for calibration. The stimuli for data acquisition included three-minute sessions of resting-state with eyes open and closed, listening to songs in native and non-native languages, and listening to neutral music, using both in-ear and bone-conducting headphones.

4.1.2 BED (Biometric EEG Database) dataset:

The BED dataset consists of EEG responses from 21 subjects to various stimuli across three chronologically disjointed sessions. Stimuli presented to subjects include images designed to elicit specific emotions, mathematical computations (2-digit additions), resting-state conditions with eyes closed, and visual evoked potentials at different frequencies. EEG recordings are segmented, structured, and annotated according to the presented stimuli, facilitating analysis and interpretation. The dataset also includes features extracted from each EEG segment, as described in the associated publication.

4.1.3 Proposed Methodology

Our proposed methodology contains a completely comprehensive framework designed to effectively address the big challenges which are inherently present in EEG-based biometric identification systems which are variable electrode configuration and inconsistent data acquisition time period. Our main is to design a biometric identification system using electroencephalogram irrespective of the number of electrodes used during eeg data reading and keep the duration of data acquisition as short as 2 seconds for fast and efficient identification of subjects. Thus prepared methodology encompasses several key steps:

4.1.3.1 Data Preprocessing

In this stage, the EEG data undergoes several preprocessing steps to enhance its quality and usability for subsequent analysis. A first-order Butterworth filter with cutoff frequencies between 1 Hz and 40 Hz is applied to remove noise and retain relevant frequency components. A notch filter at 50 Hz is employed to eliminate electrical interference, such as power line noise. The preprocessed data is divided into short time windows, typically around 2 seconds in duration. This segmentation aids in managing the data and allows for the analysis of temporal dynamics.

FIGURE 4.1: Unfiltered eeg signal depicting one channel of a subject in Auditory evoked potential EEG Biometric dataset

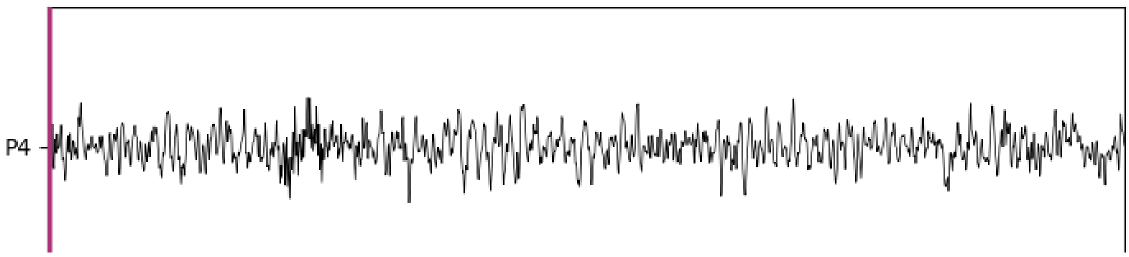

FIGURE 4.2: Filtered eeg signal depicting one channel of a subject in Auditory evoked potential EEG Biometric dataset

4.1.3.2 Feature Extraction

The Short-Time Fourier Transform (STFT) is a time-frequency analysis technique used to extract features from EEG signals. The continuous EEG signal is divided into short overlapping segments, often using a windowing function like the Hamming window. Fourier Transform is applied separately to each segment to analyze the frequency content over time. The magnitude of the resulting STFT coefficients is squared to obtain the spectrogram, a visual representation of frequency content versus time. Spectrograms provide insights into how the spectral characteristics of the EEG signals evolve over time.

$$S(m,k) := \sum_{n=0}^{N-1} x(n+mH)\omega(n)e^{(-i2\pi kn/N)} \tag{4.1}$$

Equation 1. [12]

where, K is the frequency index w.r.t Nyquist frequency. N is the section's duration. hop size (H) is the step size of the window to be shifted. w: [0, N-1] -¿ R is sampling frequency, m: [0,M] and M is maximum frame index mathematically M = (L-N)/H.

4.1.3.3 Channel Selection:

Channel selection involves choosing a subset of EEG channels for further analysis based on predefined criteria. The 10/10 worldwide EEG system was followed in our instance when choosing the electrode locations, which were T7, F8, Cz, and P4 [13–15].The 10/10 worldwide EEG system was followed in our instance when choosing the electrode locations, which were T7, F8, Cz, and P4 [13–15]. The AEP dataset however had only these 4 electrodes as original set of electrodes. It has been observed through multiple previous research that these selected electrodes are more relevant for the identification task they have high signal-to-noise ratio (SNR) and minimal artifact contamination. We have done a comparative study between the model that incorporates all available channels and a model that ingest data from selected channel only in the result section of this paper.

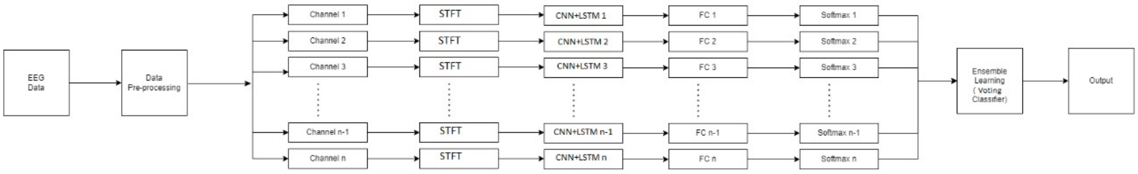

FIGURE 4.3: Complete training workflow diagram of the model

4.1.3.4 Individual Channel Processing:

In our methodology, each selected EEG channel undergoes a comprehensive processing pipeline to extract meaningful features and facilitate accurate biometric identification. The processing pipeline consists of CNN followed by LSTM and Fully Connected layer. We have also introduced attention layer for explerimentation purpose.

The CNN (Convolutional Neural Network) is a type of neural network architecture that combines the convolutional layers to learn spatial and temporal features of sequential data. It is widely used in the field of image and eeg processing for classification, detection, and localization tasks. CNNs are powerful in handling spatial features in data, while LSTM are capable of modeling temporal dependencies. The CNN model's mix of these two neural network types makes it perfect for examining time-series data with intricate structures, such as egg signals. The convolutional layers, pooling layers, and fully connected layer make up the three primary components of the CNN model. By applying convolutional operations to the input data, the convolutional layers extract features from the audio stream. You can think of these layers as an eeg signal feature map extractor. The convolutional and recurrent layers make up the majority of the CNN+LSTM model. It creates dependencies on past or future outcomes by combining the functions of recurrent and convolutional layers. We employed a CNN+LSTM model with two detection layers and one classification layer for this assignment. A LSTM (Long Short-Term Memory) layer and a dense layer make up the classification layer, while a 2D convolutional layer and a max-pooling layer make up the detection layers. A series of eeg spectrograms produced from the egg recordings using the Short-Time Fourier Transform (STFT) serve as the CNN model's input. The frequency domain of the eeg signal at various time intervals is represented by the spectrograms. With time, the CNN model recognizes patterns in the spectrograms and categorizes them into target individual.

Convolutional Neural Network (CNN)

Convolutional layers are responsible for extracting local features from the input data. These layers use filters to convolve over the input data and learn spatially localized patterns. Convolutional layers are the backbone of most convolutional neural networks (CNNs) and are responsible for learning the spatial features of input data.

In a convolutional layer, a set of filters are applied to the input data. Each filter is a small window that slides over the input data, and at each position, it performs a dot product between the filter weights and the corresponding input values within the window. The result of this operation is a single value that represents a feature at that particular position.

Convolutional layers have several parameters that can be adjusted to control the behavior of the layer. These parameters include:

Filters: The number of filters is the number of output channels produced by the layer. Each filter detects a particular feature of the input data. Increasing the number of filters increases the model's ability to detect more complex features.

Kernel size: The kernel size is the size of the filter applied to the input data. Increasing the kernel size increases the area of the input data that each filter sees, allowing the model to detect larger features.

Stride: The stride is the step size that the filter takes as it moves across the input data. Increasing the stride reduces the spatial resolution of the output and can lead to faster computation.

Padding: Padding is used to preserve the spatial dimensions of the input data as it passes through the layer. Padding can be 'valid' (no padding) or 'same' (padding the input so that the output has the same dimensions as the input). Activation function: The activation function is applied to the output of the layer to introduce non linearity into the model. Common activation functions include ReLU, sigmoid, and tanh.

By adjusting these parameters, convolutional layers were used to build models that were capable of detecting a wide range of features in the input data. Convolutional layers have a few important parameters, including the number of filters, filter size, stride, and padding. The number of filters determines how many features the layer will learn, while the filter size specifies the width and height of the window. The stride parameter determines the amount by which the window shifts at each step, and padding can be used to add zeros around the input data to preserve its size.

In a CNN, convolutional layers are usually followed by activation functions (such as ReLU) to introduce non-linearity, and pooling layers to reduce the spatial dimensionality of the output. Pooling Layers are used to reduce the spatial size of the input feature maps and help to extract important features from them. The pooling operation is typically performed after a convolutional layer and involves down-sampling the feature maps using a fixed-size window with a certain stride.

There are two types of pooling operations commonly used in CNNs: Max Pooling: This operation takes the maximum value from each window of the feature map to obtain a down-sampled output. The size of the window and the stride are two important parameters that need to be specified while creating the max pooling layer.

Average Pooling: This operation takes the average value from each window of the feature map to obtain a down-sampled output. Again, the size of the window and the stride are two important parameters that need to be specified while creating the average pooling layer.

In addition to these parameters, there are several other hyper parameters that can be set while creating a pooling layer in CNNs. Some of these include: Padding: This is used to add zeros around the edges of the input feature maps to ensure that the pooling operation does not reduce the spatial size too much.

Pooling Mode: This determines whether max or average pooling is used. Stride: This determines the amount of shift applied to the window during the pooling operation.

Pool Size: This determines the size of the window used for pooling.

Equation for convolutional network:

Output_shape = (input_shape - kernel_size + 2*padding) / stride + 1

In a convolution layer, the input is convolved with a set of learnable filters, also called kernels, to produce a set of output feature maps. Each filter learns to detect a specific pattern or feature in the input. The output shape of the feature maps is determined by the kernel size, padding, and stride used. The activation function used in the convolution layer is usually ReLU (Rectified Linear Unit).

Long Short Term Memory (LSTM)

Recurrent layers are a type of layer in neural networks that are designed to work with sequential or time-series data, where the order of inputs matters. They can be used in various types of tasks, including language modeling, speech recognition, and music analysis.

Recurrent layers have a few parameters that can be set to control their behavior:

Units: This parameter controls the number of memory cells in the layer. The more units, the more complex patterns the layer can learn.

Activation: This parameter controls the activation function used on the output of each memory cell. Common activation functions include sigmoid, tanh, and ReLU.

Recurrent Activation: This parameter controls the activation function used on the recurrent connections between the memory cells. Common choices include sigmoid and tanh.

Return Sequences: This parameter controls whether the layer should return the output for each time step (i.e., a sequence) or just the output for the final time step. Return State: This parameter controls whether the layer should return the final internal state of the memory cells in addition to the output.

Dropout: This parameter controls the amount of regularization applied to the outputs of the memory cells and the recurrent connections.

LSTM networks are a type of RNN that are specifically designed to capture temporal dependencies and patterns in sequential data, making them well-suited for analyzing EEG signals over time. Recurrent neural networks (RNNs) with Long Short-Term Memory (LSTM) architecture were developed to address the difficulties associated with encoding long-term relationships in sequential input. An LSTM is primarily composed of memory cells, which serve as the network's memory by maintaining a cell state throughout the whole sequence.

LSTM utilizes three gates to control the flow of information into and out of the memory cell:

- Forget Gate: Determines what information to discard from the cell state.
- Input Gate: Decides which new information to store in the cell state.
- Output Gate: Controls what information to output from the cell state.

A sigmoid neural network layer and a pointwise multiplication operation make up each gate. The information flow is managed by the sigmoid layer, which produces values between 0 and 1.

In a recurrent layer, the output from the previous time step is fed back into the network as input for the current time step. This allows the network to capture the

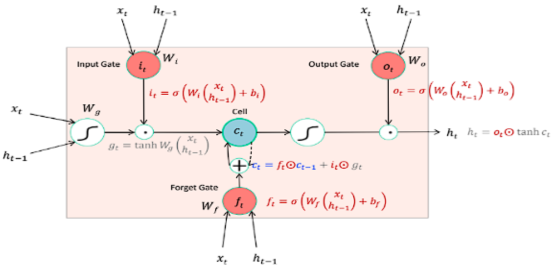

FIGURE 4.4: Diagram of a LSTM cell. [16]

temporal dynamics of the input sequence. The LSTM uses a gating mechanism to selectively forget or remember information from previous time steps. The activation function used in the recurrent layer is typically tanh or ReLU.

Attention Model:

For the purpose of experiment an attention mechanism was incorporated into the architecture, right after LSTM to dynamically weight the importance of different regions in the data. The attention mechanism allowed the model to focus on salient features in the EEG data, enhancing its discriminative power by giving more weight to relevant parts of the input.An attention model, in the context of neural networks, is a mechanism that allows the model to focus on specific parts of the input data, giving them more weight during processing. This mechanism is inspired by human attention, where certain elements in a scene or a piece of information are prioritized over others based on their relevance or importance. to see if the mechanism that allows the model to focus on specific parts of the input data, giving them more weight during processing.

The process of incorporating attention into an LSTM-based architecture typically involves the following steps:

Attention Score Calculation: First, attention scores are calculated for every element in the input sequence. These scores, which measure each element's importance in relation to the LSTM layer's present state, are generated using a trainable method that is frequently carried out using neural networks.

Normalisation of Attention scores: Next, an appropriate activation function (such as softmax) is used to normalise the attention ratings across the input sequence. A probabilistic interpretation of the attention distribution is made possible by this normalisation, which guarantees that the attention weights add up to unity overall.

Weighted Summation of Input Features: Next, using the normalised attention scores, a weighted total of the input features is calculated, with each feature vector being weighted according to its matching attention weight. By using the attention mechanism to decide what information is most salient within the input sequence, this procedure efficiently creates a context vector.

Integration with LSTM Computation: Lastly, the LSTM layer's computation incorporates the context vector that was retrieved from the attention mechanism. The internal state dynamics and output generation process of the LSTM are impacted by this integration, which normally takes place at each time step. The LSTM improves its ability to identify long-range dependencies and concentrate on pertinent information within the input sequence by integrating contextual information obtained via the attention mechanism.

Fully Connected Layers (Dense Layers):

FC layers (also called dense layers) are used for classification or regression tasks. They take the output from the convolutional and recurrent layers and produce a final output.

In a CNN model, the FC layer is used to perform the final classification based on the extracted features. The FC layer takes in the output of the previous layer, which is typically a sequence of feature vectors, and outputs a single vector that represents the class probabilities.

The dense layer in a CNN+LSTM model typically has a large number of parameters, and its output size is determined by the number of classes in the classification task.

The number of units in the dense layer can be adjusted based on the complexity of the classification task and the amount of available training data.

Fully connected layers connect every neuron in the current layer to every neuron in the next layer. They are used for classification and regression tasks. In a FC layer, each neuron receives input from all the neurons in the previous layer, and its output is calculated as a weighted sum of these inputs, followed by a non-linear activation function.

The parameters of a dense layer include:

Number of neurons: This is the number of neurons in the layer. It determines the output size of the layer and is specified when defining the layer. Activation function: This is the non-linear function applied to the output of each neuron. It introduces non-linearity to the network and helps to model complex relationships between the inputs and outputs.

Kernel initializer: This is the method used to initialize the weights of the layer. The weights are initialized randomly, and the kernel initializer determines the distribution from which the random values are drawn. Popular kemel initializers include uniform, normal, and glorot_uniform.

Bias initializer: This is the method used to initialize the bias values of the layer. The bias is an additional parameter added to each neuron that allows the network to shift the activation function to the left or right. Popular bias initializers include zeros and ones.

Regularization: This is a technique used to prevent overfitting in the network. Regularization methods include Ll and L2 regularization, which penalize large weight values and encourage sparsity, and dropout, which randomly drops out some neurons during training to encourage the network to learn more robust features.

Other optional parameters: Other optional parameters of dense layers include whether to use bias, whether to use batch normalization, and whether to use kernel constraint.

The dense layer is followed by an activation function, which is used to introduce non-linearity into the output of the dense layer. In the CNN+LSTM model, the activation function used is usually a softmax function, which is commonly used in multi-class classification tasks to convert the output of the dense layer into class probabilities.

The parameters of the dense layer can be optimized during training using back-propagation and gradient descent. The goal is to minimize the loss function, which measures the difference between the predicted class probabilities and the true class labels. The optimization process adjusts the weights and biases of the dense layer to minimize the loss function and improve the accuracy of the model on the training data.

Equation for fully connectd layer

y = activation_function(Wx+b)

In a dense layer, all neurons are fully connected to the input from the previous layer. Each neuron computes a weighted sum of the inputs and applies an activation function to produce the output. The activation function used in the dense layer is typically softmax for multiclass classification or sigmoid for binary classification.

Softmax Layer:

The final layer of the processing pipeline is a softmax layer, responsible for generating class probabilities for each segment of EEG data. Softmax activation transforms the raw output of the network into probability distributions over the possible classes, facilitating interpretation and decision-making.

4.1.3.5 Ensemble Learning with Voting Classifier:

A voting classifier is an ensemble learning method that combines the predictions of multiple base classifiers and predicts the class label by majority vote (or by averaging probabilities). It aggregates the predictions of each base classifier and selects the class label that receives the most votes. There are different types of voting classifiers:

• Hard Voting: In hard voting, the predicted class label is determined by a simple majority vote among the base classifiers. The class with the most votes is selected as the final prediction.

• Soft Voting: In soft voting, the predicted class label is determined by averaging the probabilities predicted by each base classifier for each class, and then selecting the class with the highest average probability. This method often performs better than hard voting, especially when the base classifiers produce well-calibrated probabilities.

We have consider voting classifiers as they are effective when the base classifiers have diverse characteristics and make different types of errors, as the ensemble can compensate for individual weaknesses. We employ hard voting among the output of our trained models to get the final prediction.

We however would like to make a small change in the classifier logic. We have added a condition for classifier that if the support for a class is less than 75% i.e. there is no particular class which was predicted by 75% of the total models then the test subject will be considered unidentified which means the test subject is not authenticated.

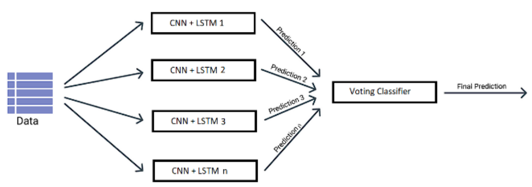

FIGURE 4.5: Architecture of a stacking classifier.

4.1.3.6 Output:

The output of the proposed methodology is a robust EEG-based biometric identification system capable of accurately identifying individuals within a short recording window. Evaluation metrics such as accuracy, precision, recall, and F1-score are used to assess the performance of the system. The model's effectiveness is validated across various electrode setups and recording durations to ensure its practicality and reliability in real-world scenarios. By meticulously following each step of this detailed methodology, we aim to create a sophisticated EEG-based biometric identification system that addresses existing limitations and delivers reliable performance in diverse settings.

4.1.3.7 Testing:

At test time. we would follow the similar data acquisition, data preprocessing and feature extraction process as in the training time. We will feed thus generated data to our trained model which will in turn identify the subject.

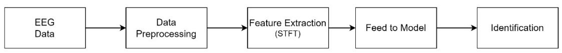

FIGURE 4.6: Workflow of the testing phase of the proposed model

Anytime a new user comes into our system we will have to retrain the model with new incoming data. Hence continuous retraining is a hindrance that has to be beared with this biometric system.

Chapter 5

Experiments

The training phase of our model start with data being molded into a consistent input shape which is fed to the starting layer of our model. We have find out two convolution layers each followed by max pooling layer to be best performing for our case. We have used relu activation function and experimented with different strides convoluted with 3,3 filters. The output from this layer is subjected to dropout layer for regularization and reduce number of neurons.Thus follows two LSTM layer each accompanied by dropout layer with drop factor of 0.2 before being fed to fully connected layers. We have used Dense layer with 64 neurons at the end of all models which gets fed to final Dense layers that outputs an array of 20 dimensions representing the predicted output of each input eeg signal. The training of the model has been optimized with Adam optimizer and loss function is chosen as categorical crossentropy.

For the purpose of experiment an attention mechanism was incorporated into the architecture, right after LSTM to dynamically weight the importance of different regions in the data. The attention mechanism allowed the model to focus on salient features in the EEG data, enhancing its discriminative power by giving more weight to relevant parts of the input.An attention model, in the context of neural networks, is a mechanism that allows the model to focus on specific parts of the input data, giving them more weight during processing. This mechanism is inspired by human attention, where certain elements in a scene or a piece of information are prioritized over others based on their relevance or importance. to see if the mechanism that allows the model to focus on specific parts of the input data, giving them more weight during processing.

```
Layer (type)                 Output Shape
==================================================
conv2d (Conv2D)              (None, 199, 116, 32)

max_pooling2d (MaxPooling2   (None, 99, 58, 32)
D)

conv2d_1 (Conv2D)            (None, 97, 56, 64)

max_pooling2d_1 (MaxPoolin   (None, 48, 28, 64)
g2D)

dropout (Dropout)            (None, 48, 28, 64)

flatten (Flatten)            (None, 86016)

repeat_vector (RepeatVecto   (None, 4, 86016)
r)

lstm (LSTM)                  (None, 4, 256)

dropout_1 (Dropout)          (None, 4, 256)

lstm_1 (LSTM)                (None, 128)

dropout_2 (Dropout)          (None, 128)

dense (Dense)                (None, 64)

dropout_3 (Dropout)          (None, 64)

dense_1 (Dense)              (None, 20)
```

FIGURE 5.1: Model summary

Evaluation metrics:

Evaluation metrics that offer insights into the predictive power of classification models across many classes, such as accuracy, precision, recall, and F1 score, are essential instruments for evaluating the performance of these models.

- Accuracy: It is simply the proportion of correctly classified samples to the total number of samples. You can use this to evaluate the overall performance of your model.

Accuracy = Total Number of Predictions / Number of Correct Predictions X 100%

- Precision: It is a metric used to assess how well the model can distinguish true positive labels from all positive labels. When a precision score is high, it implies a low false positive rate, which implies that the model is more likely to be accurate when it predicts a positive class. Precision is calculated as the ratio of true positives (TP) to the sum of true positives and false positives (FP).

  Precision = TP / (TP + FP)

- Recall: Recall gauges the model's accuracy in identifying positive cases among the real positive instances; it is sometimes referred to as sensitivity or true positive rate. It emphasises how comprehensive the optimistic forecasts are. A low rate of false negatives is shown by a high recall, which means that the majority of positive events can be captured by the model. The ratio of true positives (TP) to the total of true positives and false negatives (FN) is used to compute recall.

  Recall = (TP + FN) / TP

- F1-Score: A statistic called the Fl-score aggregates recall and precision into a single number. It offers a fair assessment of a model's efficacy by taking precision and recall into account at the same time. The Fl-score assigns equal weight to precision and recall by taking the harmonic mean of these two metrics. When there is an imbalance between the classes or when recall and precision are both crucial, the Fl-score comes in handy.

  F1= 2 X (Precision + Recall) / (Precision X Recall)

Chapter 6

Result and Discussion

In our first study, we utilized the Auditory Evoked Potential EEG-Biometric dataset for biometric identification purposes. Our suggested methodology's performance was compared to a baseline model that used a straightforward CNN+LSTM architecture that took in all channels simultaneously. Notably, the baseline model achieved an accuracy of 90%. Adding attention model on top of that model didn't excceed the accuracy as it only could be accuracy 89% of the time. However, our method, which leveraged individual channel processing and ensemble learning, demonstrated a significantly higher accuracy of 98%. This substantial improvement underscores the efficacy of our approach in enhancing biometric identification accuracy. Through the use of hard voting, the ensemble learning technique aggregated predictions from several models trained on different channels, enabling robust decision-making. This allowed us to achieve better performance than the baseline model by utilizing the distinct information that each channel captured. More performance metrics has been compared and listed in Table I.

In our second study, we experimented with the BED: Biometric EEG dataset and assessed the performance of two distinct approaches for biometric identification.First off, analysing all 14 EEG channels at once with a simple CNN and LSTM architecture yielded an accuracy of 89%. Introducing attention model on top of that couldn't

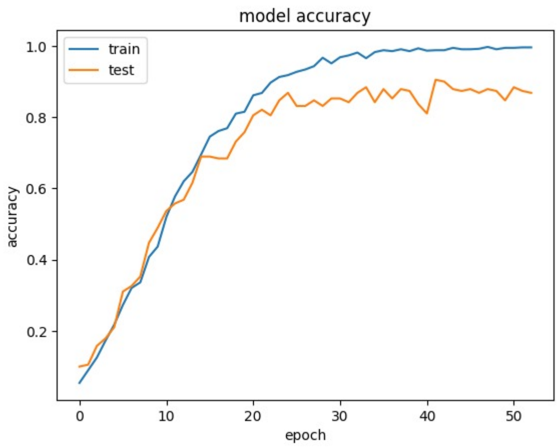

FIGURE 6.1: Accuracy curve for CNN+LSTM training with all 4 channel on AEP dataset

TABLE 6.1: Metrics comparison between model with all channels taken at once vs ensemble of channel independent models (AEP dataset)

Metrics	All channels	Individual(ensemble)
Accuracy	0.905	0.980
Precision	0.896	0.995
Recall	0.900	0.988
F1 score	0.892	0.992

TABLE 6.2: Accuracy comparison between pre-existing models vs our proposed model for (AEP)

Models	Accuracy
k Nearest Neighbours (KNN) [9]	60.3%
Multilayer perceptron (MLP) [9]	61.9%
eXtream Gradient Boosting (XGBoost) [9]	85.4%
CNN+LSTM	90%
CNN+LSTM with attention	89.1%
CNN+LSTM with Ensemble learning	98.0%

make the model perform any better with stagnent accuracy of 89% only. These two outcome offered a performance metric baseline for comparison. We then looked into a more sophisticated approach that made use of ensemble learning strategies. We used ensemble learning to combine predictions from several models trained on separate EEG channels, initially concentrating on just four of them. This method produced a 96% accuracy rate, which is a significant improvement. On using all of the 14 channels, i.e modelling them indiviually and building an ensemble of them through hard voting yielded 99% accuracy, which is a tremendous increase in performance. Similar improvement can be seen in other performance metrics like precision, recall and F1 score as shown in Table III

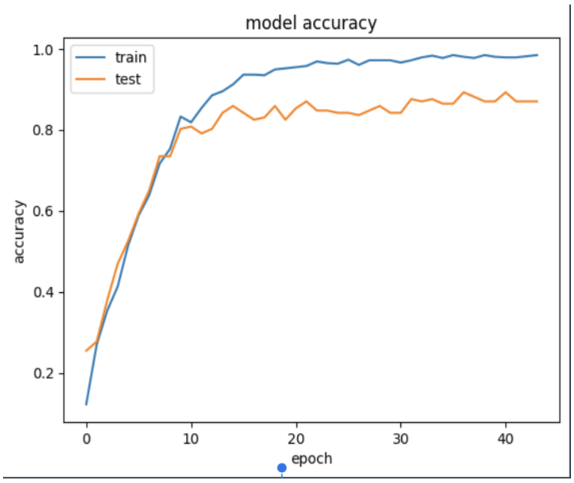

FIGURE 6.2: Accuracy curve for CNN+LSTM training with all 14 channel on BED dataset

The notable improvement in our dataset as seen in table I and table III highlights the effectiveness of our ensemble learning strategy and shows how it may be used to better utilise the distinct information contained in particular EEG channels for biometric identification. These findings support the possibility of using ensemble learning techniques to improve biometric identification systems' performance.

Comparing our findings with previously published works, our results demonstrate a significant improvement in accuracy, as evident from Table II and Table IV. This enhancement underscores the efficacy of our ensemble learning approach in harnessing the unique information embedded in specific EEG channels for improved biometric identification accuracy. These results affirm the potential of ensemble learning methodologies in enhancing the performance of biometric identification systems, particularly when dealing with complex datasets such as the BED.

Table 6.3: Metrics comparison between model with all channels taken at once vs ensemble of channel independent models (BED dataset)

Metrics	All channels	Ind.ensemble 14	Ind.ensemble 4
Accuracy	0.89	0.990	0.960
Precision	0.896	0.993	0.970
Recall	0.892	0.996	0.970
F1 score	0.879	0.992	0.965

Table 6.4: Accuracy comparison between pre-existing models vs our proposed model for (BED dataset)

Models	Accuracy
CNN [7]	83.51%
EEGNET [7]	86.74%
CNN + deep FC [8]	98.2%
CNN+LSTM	89.6%
CNN+LSTM with attention	89.3%
CNN+LSTM with Ensemble learning (4 channels)	96%
CNN+LSTM with Ensemble learning (14 channels)	99%

Chapter 7

Conclusion and Future Scope

In conclusion, this research presents a novel methodology for EEG-based biometric identification, leveraging both the Auditory Evoked Potential EEG-Biometric dataset (AEP dataset) and the Biometric EEG dataset (BED). Through the proposed approach, we have demonstrated significant advancements in the field, achieving high accuracy rates and addressing key challenges associated with variable electrode configurations and limited recording times.

Our methodology, which incorporates Short-Time Fourier Transform (STFT) for feature extraction, channel-wise processing, and ensemble learning with a hard voting classifier, has proven to be highly effective. By analyzing EEG signals obtained from auditory and visual stimuli, we have achieved notable accuracy rates, surpassing existing approaches in biometric identification. In our experiments, we achieved an impressive accuracy of 98% using the AEP dataset and 95% using the BED dataset. These results signify the robustness and reliability of our proposed methodology in accurately identifying individuals based on EEG signals.

Although the achieved results seem promising, there are many issues need to be addressed in future work. The performance needs to be evaluated over larger population database and over multiple session recordings to assure the uniqueness and the stability of these signals over time, i.e., time-permanence. Also explore scalability, real-time implementation, and addressing potential challenges in practical deployment. Looking ahead, there are several avenues for future research and improvement:

Scalability and Real-world Deployment: While our methodology shows promise in controlled experimental settings, its scalability and feasibility in real-world deployment need further exploration. Future research should focus on optimizing the system for real-time processing and deployment in practical scenarios, such as authentication systems in security applications or healthcare settings.

Longitudinal Studies and Time-Permanence: Assessing the long-term stability and reliability of EEG-based biometric identification systems is crucial. Future work should involve longitudinal studies to evaluate the consistency of EEG signals over time and their resilience to environmental factors and subject-specific changes. Understanding the time-permanence of EEG biometrics will enhance the system's reliability and applicability in long-term authentication scenarios.

Enhanced Model Architectures: Continual advancements in deep learning and neural network architectures offer opportunities for further improving biometric identification systems. Future research could explore more sophisticated model architectures, including attention mechanisms, graph-based networks, or hybrid models combining CNNs and recurrent networks, to capture complex temporal and spatial dependencies in EEG data more effectively.

Robustness and Generalization: Investigating the robustness and generalization capabilities of the proposed methodology across diverse demographic groups, such as age, gender, and cultural backgrounds, is essential. Future studies should involve larger and more diverse datasets to ensure the system's effectiveness across various population segments and mitigate potential biases or inaccuracies!

Security and Privacy Considerations: As with any biometric authentication system, ensuring user privacy and security is paramount. Future research should address privacy concerns related to EEG data collection, storage, and transmission, implementing robust encryption and anonymization techniques to protect user information from unauthorized access or misuse.

In summary, this research represents a significant step forward in EEG-based biometric identification, offering a comprehensive methodology and promising results. By addressing current limitations and laying the groundwork for future advancements, we aim to contribute to the development of reliable, efficient, and secure biometric authentication systems with broad applications in diverse domains.

Bibliography

[1] KavithaP.Thomas,CuntaiGuan,C.T.Lau,Vinod.A.PandK.K.Ang, "Adaptive Tracking of Discriminative Frequency Components in EEG for a Robust Brain-Computer Interface," Journal of Neural Eng., vol. 8, no: 3, pp. 1-15, 2011.

[2] Värbu K, Muhammad N, Muhammad Y. Past, Present, and Future of EEG-Based BCI Applications. Sensors (Basel). 2022 Apr

[3] E. Niedermeyer, F.H. Lopes da Silva (Eds.), "Electroencephalography: Basic Principles, Clinical Applications, and Related Fields," Williams & Wilkins, Baltimore, 1993.

[4] DelPozo-Banos M., Travieso C. M., Weidemann C. T. and Alonso J. B., "EEG biometric identification: a thorough exploration of the time- frequency domain," Journal of neural engineering, Sep 23;12(5):056019, 2015.

[5] Picton, T.W.: ' Human auditory evoked potentials' (Plural Publishing, San Diego, CA, USA, 2010)

[6] Hegerl, U., Priebe, S., Wildgrube, C., et al.,: 'Expressed emotion and auditory evoked potentials', Psychiatry, 1990, 53, (2), pp. 108–114

[7] Bakırcıoglu, K., Tanimu, M.B., Özkurt, N., Seçmen, M., Güzeliş, C., Yıldız, O. (2022). Multi Channel EEG Based Biometric System with a Custom Designed Convolutional Neural Network. In: Seyman, M.N. (eds) Electrical and Computer Engineering. ICECENG 2022.

[8] Hendrawan, Muhammad Afif & Ratsanjani, M & Noprianto, Noprianto & Ed Dien, Habibie. (2024). A Short-Length Single Channel EEG Based Personal Identification System. 10.1109/COMNETSAT59769.2023.10420768.

[9] Effect of Auditory Stimuli on Electroencephalography-based Authentication Alzahab, Nibras Abo;Iorio, Angelo Di;Baldi, Marco;Scalise, Lorenzo 2022-01-01

[10] Abo Alzahab, N., Di Iorio, A., Apollonio, L., Alshalak, M., Gravina, A., Antognoli, L., Baldi, M., Scalise, L., & Alchalabi, B. (2021). Auditory evoked potential EEG-Biometric dataset (version 1.0.0).

[11] P. Arnau-González, S. Katsigiannis, M. Arevalillo-Herráez and N. Ramzan, "BED: A New Data Set for EEG-Based Biometrics," in IEEE Internet of Things Journal, vol. 8, no. 15, pp. 12219-12230, 1 Aug.1, 2021

[12] M. Asif, S. Mishra, M. T. Vinodbhai and U. S. Tiwary, "Emotion Recognition Using Temporally Localized Emotional Events in EEG With Naturalistic Context: DENS# Dataset," in IEEE Access, vol. 11, pp. 39913-39925, 2023, doi: 10.1109/ACCESS.2023.3266804.

[13] Altahat SH. Robust EEG Channel Set for Biometric Application (Doctoral dissertation, University of Canberra).

[14] Ravi KV, Palaniappan R. A minimal channel set for individual identification with EEG biometric using genetic algorithm. InInternational Conference on Computational Intelligence and Multimedia Applications (ICCIMA 2007) 2007 Dec 13 (Vol. 2, pp. 328-332). IEEE.

[15] Marcel S, Millán JD. Person authentication using brainwaves (EEG) and maximum a posteriori model adaptation. IEEE transactions on pattern analysis and machine intelligence. 2007 Feb 20;29(4):743-52.

[16] Lecture on RNN, Deep Learning course 2023, CVBL Labs, IIITA