Deep Learning-Based Assessment Model for Real-Time Identification of Visual Learners Using Raw EEG

Soyiba Jawed, Ibrahima Faye, *Senior Member, IEEE*, and Aamir Saeed Malik, *Senior Member, IEEE*

Abstract— **Automatic identification of visual learning style in real time using raw electroencephalogram (EEG) is challenging. In this work, inspired by the powerful abilities of deep learning techniques, deep learning-based models are proposed to learn high-level feature representation for EEG visual learning identification. Existing computer-aided systems that use electroencephalograms and machine learning can reasonably assess learning styles. Despite their potential, offline processing is often necessary to eliminate artifacts and extract features, making these methods unsuitable for real-time applications. The dataset was chosen with 34 healthy subjects to measure their EEG signals during resting states (eyes open and eyes closed) and while performing learning tasks. The subjects displayed no prior knowledge of the animated educational content presented in video format. The paper presents an analysis of EEG signals measured during a resting state with closed eyes using three deep learning techniques: Long-term, short-term memory (LSTM), Long-term, short-term memory–convolutional neural network (LSTM-CNN), and Long-term, short-term memory–Fully convolutional neural network (LSTM-FCNN). The chosen techniques were based on their suitability for real-time applications with varying data lengths and the need for less computational time. The optimization of hypertuning parameters has enabled the identification of visual learners through the implementation of three techniques. LSTM-CNN technique has the highest average accuracy of 94%, a sensitivity of 80%, a specificity of 92%, and an F1 score of 94% when identifying the visual learning style of the student out of all three techniques. This research has shown that the most effective method is the deep learning-based LSTM-CNN technique, which accurately identifies a student's visual learning style.**

Index Terms— **Raw-electroencephalogram, deep learning, machine learning, visual learner, classification, learning styles.**

I. INTRODUCTION

LEARNING style plays a vital role in acquiring learning skills; hence, the part of learning style cannot be ignored in overall learning [1]. Learning style depends on one's personality, preference for learning in a group or individual, and level of intelligence [2]. For achieving the ultimate learning goals in a classroom setting, it is vital to use a combination of methods to enhance Learning by addressing all the learning styles. The learning styles are categorized into four types that correspond to the four learning modalities: (1) visual, (2) kinesthetic, (3) audio, and (4) tactile [3]. According to statistics, 70% of students are visual learners, 20 % are auditory learners, and 10% fall into the remaining two categories. Hence, the focus of this work is on the identification of the visual learning style, as most people are visual learners in real time.

A real-time learning system is important as assessing learning styles is helpful for teachers and students to improve their grades. Therefore, obtaining this assessment during learning is essential to enhance learning. That's why it is crucial to have a real-time system. Real-time assessment tailored for visual learners has numerous applications across education. Some examples of applications include 1) Interactive Quizzes, 2) Digital whiteboards, 3) Simulations, 4) Data visualization, and 5) Video feedback. These applications cater to the visual learning style, enhancing engagement and comprehension through real-time visual assessment methods. However, these are all subjective measures and suggesting someone's learning style without looking into their brain patterns can only increase the cognitive load. Therefore, it is important to consider objective measures by looking into brain patterns using neuroimaging techniques such as EEG. In current practices, objective measures are used for feature extraction using handcrafted feature extraction techniques. That requires preprocessing of EEG data, which cannot be done in real-time, resulting in delayed feedback; however, to enhance learning outcomes, it is required to provide feedback in real-time just after the learning is performed.

Visual learners are those who learn best through visuals, and these visuals include pictures, videos, presentations, etc.

[3]. In other words, visual learners are the ones who learn by visual descriptors. The subjective measures to identify visual learning style are based on self-assessment; hence, they can be dubious, that is, one identifies oneself as a visual learner. But, in reality, this is just a learned behaviour, and the brain patterns suggest something different [3]. As Learning and, correspondingly, the learning style is related to brain neuronal dynamics, it is therefore important to directly investigate brain patterns to assess the learning style. Thus, brain modalities such as electroencephalogram (EEG), functional near-infrared spectroscopy (fNIRS), magnetoencephalography (MEG), and functional magnetic resonance imaging (fMRI) can be used to compute objective measures for tracing brain patterns during visual learning [4], [5], [6], [7]. However, EEG is the only feasible modality for real-time processing in an outside environment (like a classroom) because of its mobility, accessibility, and ease of use. Hence, this paper focuses on the use of EEG to study brain patterns for real-time identification of visual learning styles in people.

In neural engineering, neuroscience, and biomedical research, electroencephalography (EEG) is very popular. It is commonly used for many medical and non-medical applications, such as the diagnosis of depression, assessment of mental workload, and seizure detection [8]. Due to its noninvasive nature, resolution, high mobility, high temporal, and low cost, it is commonly used. The challenge lies in real-time automatic Classification of the EEG data, which could be more beneficial for practical applications and more attractive for professionals due to lower training complexity. The brain generates EEG signals that are recorded from the scalp. These signals can be recorded during rest or while performing tasks. Next, the EEG dataset stores voltage and channel information in a matrix. Due to its time-series nature, EEG data is beneficial for machine learning tasks.

Since this work focuses on visual learning styles, reviewing some of the most recent works to identify visual learning styles using EEG is important. Visual learning style is essential as most learners are visual learners [9], [10], [11]. The category encompasses various items, including visuals like pictures, graphs, videos, images, and written materials such as words. Here, the overview of some selected works is presented that used EEG to classify visual learning styles. In [1], the visual and reading part of the Visual, Audio, Reading, Kinesthetic (VARK) test was used along with P200 amplitude to measure event-related potential at the occipital site on the scalp. These students were divided into two groups under visual learning modality groups; visual learners (who learn by looking at pictures) and verbal learners (who learn by reading (books, articles, and flashcards). The study concluded that there was a psychological difference between students who were visual learners and those who were readers. The visual Learner's students had a higher peak of P200 compared to the readers [1]. This study sees the difference between the two categories, that is, between visual and reading learners. Thus, additional work is required to classify all the categories of visual learning styles, such as videos, presentations, multimedia, and animations.

A study examined the brain reactions of individuals presented with a learning style that doesn't match their brain pattern [12]. The study compared visual and verbal learners using 135 healthy, right-handed students. The indexed learning style test [12] determined their learning style, and their EEG was recorded while they viewed pictures of animals. The data categorized students as visual or verbal learners. It showed that visual learners' theta waves decreased and beta waves increased during a verbal task, while verbal learners had higher theta waves and lower beta waves. The conclusion was that visual learners' concentration decreases during a verbal task [12]. The drawback was it does not cover all the aspects of visual modalities, such as Learning using texts, videos, and PowerPoint presentations. Thus, further analysis is required.

Another research [13] identified students' visual learning styles who used electronic learning mediums for Learning. The number of participants in this study was one hundred and eighteen. The idea was to use a bio-inspired chatbot using VARK modified test to identify the visual learning style of students who are either introverts or extroverts. After two minutes of learning content, the VARK test was recorded to identify the visual learning style. It was observed that while watching the learning content, the beta waves recorded from the whole brain of visual learners were higher. This EEG-recorded data was then classified using Naïve Bayes and Clustering algorithms. The classification accuracy was 93 %. The result showed that the bio-inspired chatbot method took less time to classify learners. The limitation of this work [13] is that it could only be used to identify the learning style of specific personality types, such as introverts and extroverts, for E-learning systems, not covering all personality types and not all aspects of visual learning modality.

Another work [14] was done to identify the visual learning style of the students using the VARK test and physiological signals such as heart rate and blood pressure. The study group included thirty university students and forty primary school students. The students were asked first to see the video, and later, they were supposed to fill out the VARK Performa. The physiological signals were recorded while watching the videos. The increased blood pressure and heart rate were observed in students who were not visual learners. Later the data were classified using a decision tree classifier. For university students, 90 % classification accuracy and 85 % accuracy were achieved for primary school students. The results also concluded that most of the participants were visual learners. Although the author claimed that physiological signals could be used for the Classification, they did not consider factors such as; (1) the nervousness of students and (2) introverted personality type. If any of the above-mentioned factors are true, the results are jeopardized, and the model's classification accuracy is no longer reliable. Thus, a more reliable approach is directly looking at the brain patterns addressed in the current study.

Although many existing studies work on identifying the visual learning style of students [15], [16], [17], [18], [12], [14], and [13], they have certain limitations that are summarized below.

Researchers use learning models [19], [20], [21] to identify students' learning styles, but these models are based on self-assessment methods that can be biased. Existing studies [15], [16], [17], [18] use resting state EEG and synthetic EEG to correlate with self-assessment-based learning styles [19], [20], but synthetic EEG is not accurate and may not depict real scenarios. ML models have task dependencies [19], [20],

meaning one model may not fit all scenarios and datasets. Some models may work well for certain datasets but fail for others [15], [16], [17], [18].

The biggest limitation of existing studies is that their methodology cannot be implemented on raw EEG for real-time applications, as the existing studies used machine learning techniques with handcrafted features. The accuracy of handcrafted features from EEG data is typically low. This is especially true when dealing with a small dataset, which can result in bias. Handcrafted feature extraction is also expensive and requires professional knowledge [22], [23]. It often overlooks high-level features derived from lower-level features. The solution is hierarchical Learning, also known as deep Learning, which creates high-level model abstractions of the data. This paper proposes using deep Learning for the real-time classification of visual learners using raw EEG.

II. DEEP LEARNING FOR REAL-TIME CLASSIFICATION OF VISUAL LEARNING STYLE USING EEG

Deep Learning was not considered the best option for classifying EEG signals in the past due to its long computational time and vanishing gradient problems [24]. Recent developments, such as large datasets and GPUs, have given deep learning researchers an inexpensive solution to their hardware bottleneck. Investigating deep learning architectures with many hidden layers is now easier. These technological discoveries make people more interested in deep-learning applications. Deep networks optimize parameters automatically, making them useful for medical research with massive datasets that are difficult to interpret. Even experts face difficulty in interpreting medical data [25]. Another limitation of dealing with a large dataset is the need for preprocessing, which is tedious and time-consuming. In this scenario, deep learning-based models can simplify the processing by doing end-to-end preprocessing, feature extraction, and Classification while performing well.

Deep Learning has been used with EEG data [26], [27]. It can improve existing EEG models in various ways. For example, it can automatically extract features from raw EEG, eliminating the need for data preprocessing. These features are more expressive and capable of high-level performance. Deep Learning can also help develop EEG-based generative modeling, improving performance and representation for different tasks and subjects [26], [27].

EEG datasets are often analyzed using a Convolutional neural network (CNN), Recurrent neural network (RNN), and Long-term short-term memory (LSTM). The most commonly used activation function is Rectified Linear (ReLU), followed by Sigmoid (S-Shape curve) and SoftMax (logistic function). A literature review found that the highest accuracy of 87% [28] was achieved with two LSTM layers and one dense layer, while the next best accuracy of 82% was achieved with CNN alone. This suggests that combining multiple deep learning architectures with dense layers can improve classification accuracy [28].

Researchers have not extensively explored the use of deep learning in learning tasks related to mental workload [29]. Further research is necessary to determine the best combinations and arrangements of layers and to compare and interpret raw

versus clean EEG. This study uses deep learning to identify visual learning styles, which is a stable way to access learning styles without self-assessment. The study does not use any tools that require self-assessment. The main contribution of this work is identifying visual learning styles using deep Learning, which is summarized as follows:

1) A reliable and unbiased method for identifying students' learning styles is proposed. EEG is recorded during Learning without any self-assessment tool, making it unbiased. The use of raw EEG data makes it more reliable. Using the EEG data, real-time results are achieved by identifying the student's learning style.

2) To improve accuracy, the proposed framework can automatically learn delicate and thorough features from raw EEG. This is achieved through the use of trainable LSTM, LSTM-CNN, and LSTM-FCNN architectures. Thus, this new assessment model is introduced to identify visual learning styles using the F-measure. This model represents state-of-the-art modality-based assessment. The proposed method combines temporal modelling, feature extraction, and summary generation into a single end-to-end architecture.

Deep learning models have shortcomings. Existing models have single structures, and there aren't enough samples to train them. To address this, batch normalization and dropout layers are added to LSTM layers to learn features effectively. Raw EEG segments are divided into non-overlapping chunks to increase available data for training and testing. This proposed scheme is beneficial for real-time identification of visual learning styles using raw EEG signals.

Inspired by the lack of identifying visual learning style in real-time in global features. We proposed a deep learning-based model for identifying visual learning styles in real time using raw EEG signals considering global features.

In this work, the main contribution can be summarized as follows:

1) An unbiased and reliable method is proposed to identify student's learning styles. The EEG is recorded, and no self-assessment tool is involved in the data collection process, making it unbiased. Secondly, the EEG data is not synthetic, which makes it more reliable. For the first time, we have used deep learning on raw EEG to identify the student's learning style to achieve better results.

2) Delicate and thorough features of raw EEG are automatically obtained in the proposed framework using trainable LSTM, LSTM-CNN and LSTM-FCNN architecture. These features are especially learned from the raw EEG to improve the accuracy of the proposed method. We set a new state-of-art modality-based assessment model for the identification of learning styles measured by F-measure. Our proposed method combines temporal modeling, feature extraction, and summary generation into an end-to-end architecture.

3) The following shortcomings of deep learning models are also addressed by our proposed model. First, the existing deep learning models consist of comparatively single model structures. Secondly, the small number of available samples which are not enough to train a deep model. To solve the above-mentioned concerns, first, the

Q. The damage of epithelial wall activates the platelets to form_____.
1. Fibrin thread
2. Good cholesterol
3. HDL
4. Clot
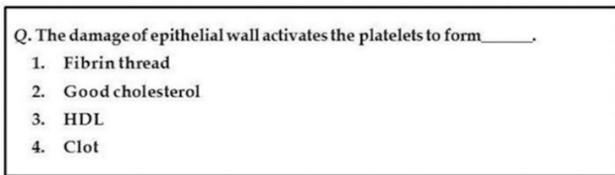
Fig. 1. Example of multiple choice question [30].

batch normalization layer and dropout layer are added into the traditional LSTM layers for learning features, that can help in identifying the learning style effectively. Second, to increase the number of samples for training and testing the segments of raw EEG were divided into non-overlapping chunks, which can benefit the network by increasing the available data for fully training a deep model. The research findings of this study have shown that the proposed scheme is beneficial in identifying the student's learning style using raw EEG signals.

III. MATERIAL AND METHODS

This section discusses the research work's data collection and explains the overall process comprehensively.

A. Participants

Thirty-four (34) university participants, aged 18 to 30 years with an average of 23.17±3.04, were recruited for the experiment. They had normal or "corrected to normal" vision, no neurological disorders or hearing impairments, and were not taking any medications. Before the trials, participants signed an informed consent document, and the Ethics Coordination Committee approved the research study [30].

B. Tasks

The test had two tasks: Task 1 was for learning, and Task 2 was for memory retrieval. Task 1 used 8-10 minutes of animated human anatomy materials to teach participants without prior knowledge. It was suitable for assessing Learning and skills. Task 2 consisted of 20 multiple-choice questions related to the animated content. Participants had 30 seconds to answer each question, and each question had four options with only one correct answer. Figure 1 displays an example question from the test.

C. Procedure

The participants followed a particular order to perform the test. First, the memory test was conducted. Then, the participants were divided into two groups based on the memory test score. The details are explained in the experiment and the result section of this paper. The process was followed by the eyes open/eyes closed EEG recording. The third step of the process was to show learning tasks to participants, followed by the retrieval (recall) sessions. After the learning tasks, the participants were given a 30-minute break before starting the retrieval task (recall session 1). Then, the participants performed the same retrieval task after 2 months

(recall session 2). Next, the EEG was recorded for eyes open resting state; eyes closed resting state, learning sessions, recall session 1, and recall session 2. We have used resting state data for this analysis to investigate the Classification of the visual learning style.

The experiment has 2D visuals without the speech. The participants learned in the experiment using visual pathways only. The EEG is recorded from the whole brain. However, the frontal, parietal, and occipital EEG signals were considered for analysis as these regions are activated during visual information processing and learning.

The first three sessions were Learning sessions representing the encoding of new information because, in these sessions, new information was given to the participants. On the other hand, the results of the recall sessions explain how well the students retain new information. Both the learning and recall sessions provided the ground truth data of whether the participant was a visual learner. Research suggests that individuals who identify as visual learners often exhibit better comprehension and retention of newly acquired information, leading to improved performance on assessments.

However, resting-state data, where a subject is not engaged in a specific task, provides valuable insights into inherent brain connectivity patterns. During the resting state, the brain's spontaneous activity reflects its baseline functioning. Analysing this intrinsic activity can reveal distinct patterns associated with various cognitive traits, including visual learning styles. Resting-state data allows researchers to identify consistent neural networks related to visual processing, even without a specific learning task. Patterns detected during rest can indicate how an individual processes visual information, aiding in the identification of visual learners. Machine learning techniques analyse these resting-state patterns, discerning unique brain signatures linked to visual learning [36]. In rest state Eyes-closed resting state data is often selected in studies because it provides a stable and consistent baseline for measuring brain activity. When participants close their eyes, it reduces visual input and minimizes external stimuli, allowing researchers to focus on intrinsic brain activity. Their controlled environment enhances the reliability of data analysis and interpretation. Therefore, analyzing the results of resting state sessions allows the development of a model for real-time identification of a visual learner that is highly reliable and can also be implemented in real-life scenarios.

A pre-test was conducted to ensure that the participants had no background knowledge. Participants were asked to read and answer 10 questions in the pre-test based on learning the animated content. Based on the results of this pre-test, exclusion criteria were created. It separates and excludes participants who managed to answer more than 10 %. The learning task was presented to the participants using a 42-inch TV screen positioned 1.5 meters away from them. The tasks were executed using E-Prime Professional 2.0 (Psychology Software Tools, Inc., Sharpsburg, PA) [34].

D. Electroencephalogram (EEG) Recording

The EGI EEG device with 128 channels HydroCel Geodesic Sensor Net was used to continuously record the subject

responses. The raw signals were amplified using the EGI NetAmps 300 amplifier's band-pass filter (0.1–100 Hz) and sampled at a rate of 250 Hz with an impedance below 50 kΩ. Cz served as a reference electrode in the standard net configuration.

E. Behavioral Data Analysis

The behavioural data was analyzed to evaluate the accuracy of the information retrieved by participants during visual learning and memory recall. The participants were given a set of twenty questions, each with a duration of one minute. The time window's total length was calculated to be 1200 seconds, which is equivalent to 20 minutes multiplied by 60 seconds. The assessment of learning and memory performance relied on the accuracy of responses and the time taken by each participant to answer the questions. The reaction time is an indicator of information processing speed related to cognitive ability. The measurement of learning performance was based on the percentage of accurate responses. The total number of trials available per subject was calculated by multiplying the number of subjects (34) by the number of multiple-choice questions (20), resulting in a total of 680 trials. The assessment of learning ability was conducted using the RAPM (Raven Advanced Progressive Matrices). It is a test of working memory. Here, a memory test is used for categorizing Visual learners and non-visual learners as memory plays an important part in learning as the learned information is stored in memory. Thus, learning and memory are related. The control variable that is used for categorizing groups is fluid intelligence, which is measured using Raven's advanced progressive matrix (RAPM). It is a non-verbal test that commonly and directly measures two components of a fluid's cognitive ability, defined as: (i) "the ability to draw meaning out of confusion" and (ii) "the ability to recall and reproduce information that has been made explicit and communicated from one to another". The subjects are divided into two equal groups using the median score of the RAMP test [27]. Based on the median score, the subjects who scored equal or above the median are considered visual learners, and those who scored less than the median are considered non-visual learners.

IV. METHODOLOGY

Deep learning models, a subset of machine learning, are designed to automatically learn and improve functions by analyzing algorithms. Deep learning models learn from data to perform specific tasks. Deep learning models learn patterns and extract global features from data. It has the ability to handle unstructured data. In our case, raw EEG data, which is unstructured, was given to the deep learning algorithm, which finds patterns from noise in the data automatically. At the same time, learning and adapting to new patterns make it highly efficient for the task at hand.

The combination of Long Short-term Memory (LSTM) and CNN forms a powerful model for handling time series data, such as EEG signals.

Long Short-Term Memory (LSTM) Networks:

Sequential Learning: LSTMs are excellent at capturing sequential patterns and long-term dependencies in time series data.

Memory Cells: LSTMs use memory cells to store and access information over long periods, making them ideal for modelling temporal sequences.

Convolutional Neural Networks (CNN):

1D Convolutions: In time series data, especially EEG signals, 1D convolutions are applied across the temporal dimension to capture local patterns and dependencies in the data.

Hybrid Model:

Combining Strengths: By combining CNNs for feature extraction and LSTMs for sequential pattern recognition, the LSTM-CNN hybrid model can capture both local patterns and long-term dependencies in the time series data.

Advantages:

1) The synergy of CNNs and LSTMs enhances the accuracy of predictions on time series data.

2). This architecture is versatile and can be adapted for different types of time series data beyond EEG signals.

Challenges:

1) Building and training hybrid models require expertise due to the complexity of integrating different neural network architectures.

2) Computational Resources: Training such models demands substantial computational resources, especially for large-scale datasets.

Based on the above information it can be concluded that Deep learning can particularly benefit from unfiltered data where a high SNR ratio facilitates the deep learning-based models to learn to discriminate features for robust classification. Therefore, this study used the raw EEG data, which has high SNR, and showed that such data, when used to train a deep learning model, improves performance. The use of raw data is also shown to work well for hyperdimensional data. This proves right here as well, and higher classification accuracy is achieved. Some other advantage of using deep learning is: 1) There is no need to clean the data. 2) It is capable of automatically detecting the most discriminating hidden features that are overlooked by handcrafted methods. 3) The performance accuracy is way higher than machine learning classifiers for all the sessions.

Based on that the problem is formulated as following:

1) Dataset: Let D be the dataset, consists of N number of subjects i.e such visual learner and non visual learner.

$$D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\}$$

x_i represents the features of the *ith subject*, and y_i is the corresponding label or class.

2) Input Features: Let X represent the matrix input of features

$$X = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}$$

The *i-th* row of $X(x_i)$ correspondiong to the feature of i-th subject.

3) Target Labels: Let Y represents the vectors of labels

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

y_i is the label of the *i-th* subject.
4) Model parameters: Let θ represent the parameter of the classification model.
5) Prediction function:
The prediction function f_0 maps input

$$\hat{Y} = f_0(X)$$

$\hat{Y}$ is the vector of the predicted label.
6) Loss function:
The objective function, $J(\theta)$, measures the difference between true labels. It quantifies the cost of the model

$$J(\theta) = Loss(Y, \hat{Y})$$

7) Optimization:
The goal is to find the parameter θ that minimizes the objective function:

$$\theta^* = arg min J(\theta)$$

8) Decision Boundary for Binary Classification:
In binary classification, the decision boundary separates the instances of one class from the other. For a linear model, it can be represented as:

$$\theta_0 + \theta_1 x_1 + \theta_2 x_2 \ldots + \theta_N x_N = 0$$

Here, θ_0 is the basic term, $\theta_1, \theta_2, \ldots, \theta_N$ are the weights. For the problem of classification of visual learners, a hybrid LSTM-CNN model is used. The model is formulated mathematically as follows:
1) LSTM network:
- Input: $x(t)$ represents the input of raw EEG data with time t
- LSTM operation: $h(t-1) and c(t-1)$ represent the cell state and the cell state from the previous time step, respectively.
- LSTM operation: LSTM $x(t), h(t-1), c(t-1)$ denotes the including forget, input, cell state, and output gates
- The complete LSTM operation is donated as follows: $h(t), c(t-1) = $ LSTM $(x(t), h(t-1), c(t-1))$
2) CNN network:
- Input $w(t)$
- Convolution operation $*$ signifies the convolution operator applied to the input sequence.
- Activation function: $\sigma()$ = denotes the activation function RELU. This will applied after the convolution operator.
- The CNN operation can be represented as:

$$CNN(x(t)) = \sigma(x(t) * filters)$$

3) Hybrid LSTM-CNN model:
- Combining LSTM and CNN: The CNN output and LSTM hidden states are combined to produce a hybrid LSTM-CNN
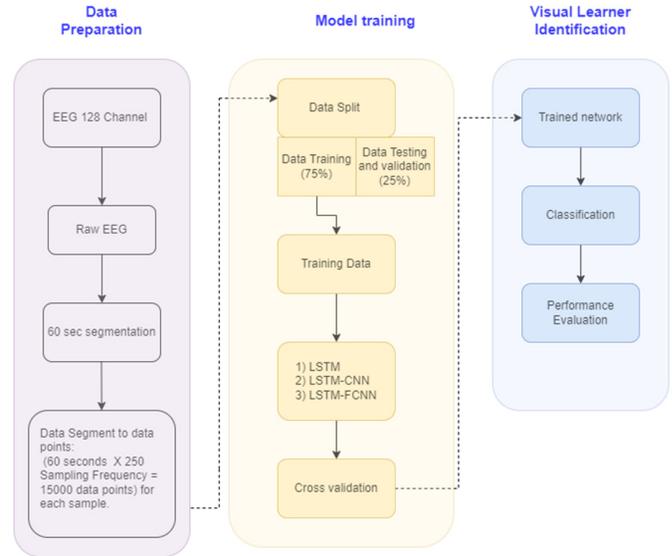
Fig. 2. The proposed assessment model uses deep learning classifiers to recognize visual learning styles. The input to the pipeline is 1D raw EEG signals and class labels at the training stage. The proposed model is evaluated using a confusion matrix and predicted class labels.

- Fully connected layer: A fully connected layer is applied for classification.

$$Hybrid\ output = Fully connected(CNN(x(t), h(t))$$

The above-mentioned notation integrates pattern recognition from CNN and sequential pattern learning from LSTMs for robust modelling of complex data, such as time series data like EEG signals.

In short, the methodology proposed involves the use of three deep learning techniques: Long-term, short-term memory (LSTM), Long-term, short-term memory – Convolutional neural network (LSTM-CNN), and Long-term short-term memory – Fully convolutional neural network (LSTM-FCNN). The classification models of deep Learning, including LSTM, LSTM-CNN, and LSTM-FCNN, have been optimized for raw EEG by hypertuning parameters that are outlined in the following subsections. After completing the classification process, the model's performance is evaluated using metrics. The models are compared to determine which one has the highest level of accuracy. The sub-sections provide a detailed explanation of every step in the deep learning pipeline. Figure 2 presents the pipeline proposed for deep Learning.

A brain-learning model has been developed for classifying visual learners using deep learning techniques like LSTM, LSTM-CNN, and LSTM-FCNN. This is necessary as conventional machine learning techniques cannot process the raw EEG data. LSTM, LSTM-CNN, and LSTM-FCNN are implemented for classification to identify the optimum combination that produces the best results.

A. Data Preparation

The dataset comprised an eyes-closed session. This session consists of 34 subjects, with each subject's EEG recording lasting between 2 to 8 minutes. Each minute of these signals is then segmented into samples of 15001. The segments extracted

from various signals are pooled as a large dataset regardless of origin. The large dataset was divided into training, validation, and testing sets. There was no overlap between these sets, as the signals for training and testing sets are divided and separated before feeding to the classifier; therefore, there is no mixing of testing and training datasets. The results presented are based on completely unseen data and not on the same signals.

Further, The EEG was recorded while answering the MCQ's questions. The EEG data is then divided into training and testing from where 75% of the data goes to training and 15% of the data goes to testing where 10 % of the data is used for validation which is the standard practice in the Machine learning community. The instances are the EEG data collected during resting state eyes open eyes close and during performing learning tasks from the participants. After the model development, the data is randomized into three parts here 3-fold cross-validation is used for training, testing, and validation. The training set is used to train the model; however, testing is performed to evaluate the overall ability of the dataset's training part. Since there were 34 subjects, we separated 26 subjects for training and 8 for testing. The 8-minute data for each subject were segmented into one minute. Each data segment was then multiplied by a sampling frequency of 250 Hz; hence, 15000 data points (60 seconds and 250 samples per second) for each sample were obtained. The data is arranged in this way to facilitate the implementation process. Out of 128 electrodes, 102 were selected that covered all the brain regions. The excluded 16 electrodes cover the areas of the neck and face, which may be contaminated with artifacts; also, the data captured from these electrodes are not required to analyze this problem. The input (15000,102) consists of 15000 samples and 102 electrodes. The implementation of this format aims to increase the number of samples. The segmented data undergoes the application of classification models. The participants from class 1 and class 2 are randomly assigned to the groups. This randomization helps in distributing both known and unknown confounding factors evenly across groups, reducing the impact of artifacts on one specific class.

The network specifications during the training phase are summarized as follows.

- The categorical cross-entropy loss function is commonly utilized in binary classification problems.
- The Adam optimizer was utilized with an initial learning rate of 0.001. The first moment has an exponential decay rate of 0.9, and the second moment has an exponential decay rate of 0.999 through the heuristic approach. The default value for the constant stability epsilon is 1e-7, and the default value for the ASMGrad variant is (False).
- Through experimentation, the values of 16 for batch size, 15 for a number of epochs, and 400 for hidden size were chosen for LSTM. The specifications for CNN include a filter size of 256, a kernel size of 3, and the use of the SoftMax activation function. The subsequent model utilized is LSTM-FCNN. The chosen parameters for LSTM are a batch size of 16, 15 epochs, and a hidden size of 400. CNN uses a filter size of (128, 256, 128), a kernel size of 3, and the activation function SoftMax.

- The real-time implementation utilizes a 3-fold validation process.
- The early stopping policy with a patience of 10 is used during training to prevent overfitting or underfitting and optimize resources.
- The initial parameters for the models are selected arbitrarily, which is one of the standard practices in deep learning community researchers. Later, these parameters are fine-tuned to improve the models' accuracy, sensitivity, and specificity.

B. Long Short-Term Memory

This section offers a detailed description of the technical concept of LSTM and its practical application as a classifier for visual learners. The LSTM operates as a classifier, similar to a conventional neural network, taking input from a particular time step and the hidden state of a prior time step. LSTM is employed for classification tasks, like other forms of neural networks. The choice to use LSTM for this research was based on its demonstrated ability to effectively manage time series data in comparison to other deep learning techniques [31]. Human cognition is characterized by persistent focus and high levels of attention. Conventional Recurrent neural networks lack this persistence attribute and address this issue [32] by including loops in their approach. One limitation of RNN is its inability to retain long-term dependencies. The Long-Short Term Memory (LSTM) network is a type of RNN that can effectively learn long-term dependencies. It possesses the capacity to retain information for an extended period. The structure of LSTM is comparable to RNN, featuring a chain-like formation and consisting of four layers. Gate structures regulate the cell state by adding or removing information, which helps to retain long-term information. A deep learning model based on LSTM has been suggested for the classification of visual and non-visual learner states using brain waves. The chosen technique for implementation is one-hot encoding, utilizing the tensor flow platform.

The process of classifying raw EEG using LSTM involves the following steps.

- The embedded layer is the initial layer that utilizes the unprocessed (raw) EEG data.
- The next layer of the LSTM has 400 memory units. The term "cells" is often used to refer to these memory units. The cell and hidden states are first passed on to the succeeding cell. The cell state is integral to the data flow. The cellular mechanism involves the transmission of data through linear transformations without altering it. Sigmoid gates in the cell state allow for the addition or removal of data. A gate can be described as a set of individual weights used in layering or matrix operations. During the first layer of the LSTM network, input information is evaluated, and if deemed unnecessary, it is excluded from the cell in this stage. The decision to include or exclude data is made by the sigmoid function, which utilizes the output of the previous LSTM unit $(h_t - 1)$ at time t − 1 and the current input (X_t) at time t. The sigmoid function plays a role in determining which portion of the previous output should be discarded. This
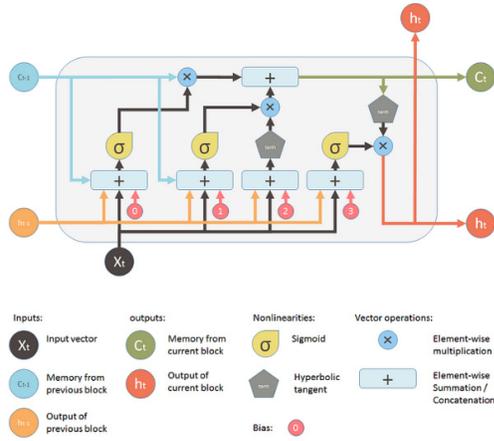
Fig. 3. LSTM architecture [34].

Fig. 4. 1D CNN structure.

gate is referred to as the forget gate (or f_t); and it utilizes a vector f_t with values between 0 and 1 that correspond to the numbers in the cell state, $C_t - 1$.

$$f_t = \sigma(W_f[h_t - 1, X_t] + b_f) \tag{1}$$

The sigmoid function is represented by σ, and W_f and b_f are the weight matrices and bias of the forget gate are represented by and, respectively. The next step involves storing and decoding information from the new input (X_t) in the cell state, as well as updating the cell state. This step has two parts: the sigmoid layer determines whether the new information should be updated or ignored (0 or 1), and the tanh function assigns weight to the passed values, which determines their level of importance (-1 to 1). The two values are multiplied to update the new cell state, which is then added to the old memory $C_t - 1$, resulting in C_t.

$$i_t = \sigma(W_i[h_t - 1, X_t] + b_i) \tag{2}$$
$$N_t = tanh(W_n[h_{t-1} - 1, X_t] + b_n), \tag{3}$$
$$C_t = C_t - 1 f_t + N_t i_t \tag{4}$$

The cell state at time $t - 1$ and t are represented by $C_t - 1$ and C_t, respectively. Weight matrices and bias are denoted by W and b. The output values (h_t) are based on the output cell state (O_t) but are a filtered version. A sigmoid layer decides which parts of the cell state make it to the output. The output of the sigmoid gate (O_t) is multiplied by the new values created by the tanh layer from the cell state (C_t). The value ranges between -1 and 1.

$$O_t = \sigma(W_0[ht - 1, X_t] + b_0), \tag{5}$$
$$h_t = O_t tanh(C_t) \tag{6}$$

The weight matrices and bias for the output gate are represented by W_0 and b_0. The architecture of LSTM is illustrated in Figure 3.
- The LSTM layer is followed by a dense output layer that utilizes a SoftMax activation function to predict either 0 or 1 for the two classes, namely L - Visual Learner and NL - Non-Visual Learner. The selection of hidden layers for LSTM is typically determined by the specific
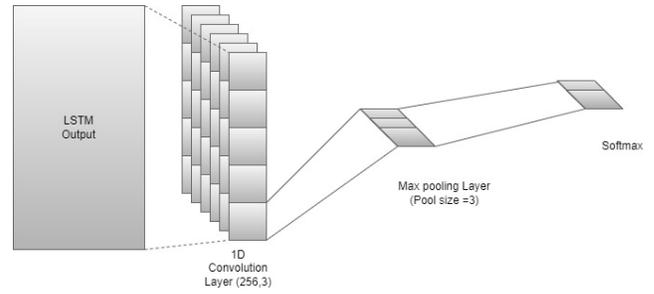
problem being addressed. Thus, The arbitrary parameter setting is utilized for selecting hidden layers [33].

C. Long Short-Term Memory Convolutional Neural Network

Temporal information is considered crucial in the classification of Learning based on EEG signals. The CNN-LSTM architecture is being studied to determine its effectiveness in classifying visual learning styles. The LSTM-CNN model incorporates LSTM for sequence analysis and feature extraction, which CNN subsequently interprets. The architecture utilized for this study includes a 1D CNN structure, which is illustrated in Figure 4.

This model's hidden layer works for a 1D sequence. To avoid overfitting, the dataset is shuffled at each epoch. Training involves 15 epochs and 16 batch sizes. A dense, fully connected layer is connected to the convolutional and pooling layers, explaining the extracted features. A dropout layer randomly turns off 20% of neurons at each iteration to ensure generalization. A flattened layer reduces feature maps to a single one-dimensional vector. Standard practices are employed: SoftMax as an activation function along with Adam optimizer for Classification. The loss function used is cross-entropy, suitable for Classification. Max-pooling after convolution retains essential features while reducing the dimension of the feature matrix. Feature reduction methods are employed for better resource optimization. The dataset is divided into training, validation, and testing data to avoid underfitting and overfitting. Table I shows the optimized parameters obtained through accuracy optimization and model architecture details.

The implementation of the LSTM-CNN model involves the utilization of the Keras library in Python. The model follows the protocol illustrated in Figure 5. The EEG data is fed into the pipeline as input.
- The next step adds the LSTM layer as a front-end layer.
- The subsequent layers include the CNN, dense, and output layers.

D. Long Short-Term Memory Fully Convolutional Neural Network (LSTM-FCNN)

This combination is commonly utilized when the data dimension is uneven, and it is also efficient for segmented data. The LSTM-FCNN technique is used to differentiate between visual learners and non-visual learners. Research suggests that temporal convolution is a successful model for classifying time series data [32], [33]. FCNN utilizes temporal convolutions to

TABLE I
DEEP LEARNING ARCHITECTURE

(a)LSTM ARCHITECTURE

S.No.	Layer name	Layer parameter
1	LSTM	(15001,102)
2	Dropout	10% dropout
3	Dense	
4	loss	Categorical cross entropy
5	Optimizer	Adam
6	Activation function	SoftMax

(b) CNN ARCHITECTURE

Sr.No.	Layer name	Layer parameter
1	Conv1D	(256,3)
2	Dropout	10% dropout
3	MaxPooling1D	Pool size = 3
4	Dense	
5	Flatten	
5	Activation function	SoftMax

(C)LSTM-CNN ARCHITECTURE

No	Layer name	Layer parameter
1	LSTM	(15000,102))
2	Conv1D	(256,3)
3	Dropout	20 % dropout
4	MaxPooling1D	Pool size = 3
5	Dense	
6	Flatten	
7	Activation function	SoftMax

Fig. 5. (a) LSTM-CNN model (b) LSTM-CNN classifier architecture. Input is EEG data, and the LSTM layer is added as a front-end layer followed by CNN and a dense layer. The output is classified as Learner and non Learner.

extract features and reduces the number of parameters through global average pooling before classification [33]. The FCNN block is expanded with the addition of the LSTM block in this pipeline. A dropout layer is added to the architecture [34], as depicted in Figure 6. The FCNN consists of three CNN blocks featuring filters of sizes 128, 256, and 128. The CNN architecture shows similarities to the proposed architecture by Graves et al. [32]. The CNN block comprises a temporal convolution layer and batch normalization. The activation function used here is ReLu, which helps prevent the vanishing gradient problem. The global average pooling is applied prior to the final block.

The EEG data obtained from the dimension shuffle is transferred to the LSTM block. The LSTM block consists of an LSTM layer and a dropout layer. The global pooling layer and LSTM block are combined and fed into a Soft-Max classification layer. The fully convolutional block and LSTM block interpret the EEG data input from two distinct perspectives. The fully convolutional block considers EEG data containing multiple time steps. The FCNN will gather information throughout all time steps in order to compute the ultimate outcome.

In contrast, the proposed architecture utilizes the LSTM block to process input data, a multivariate time series with a single time step. The temporal dimension of the data is transposed through the use of the dimension shuffle layer. After transforming, univariate EEG data with a length of N will be perceived as a multivariate time series. It will have a single-time step and N variables. This approach is key to the enhanced performance of the proposed architecture [36].

V. RESULTS

This section discusses the analysis of deep learning techniques for the Classification of visual learners from non-visual learners. The LSTM, LSTM-CNN, and LSTM-FCNN-based
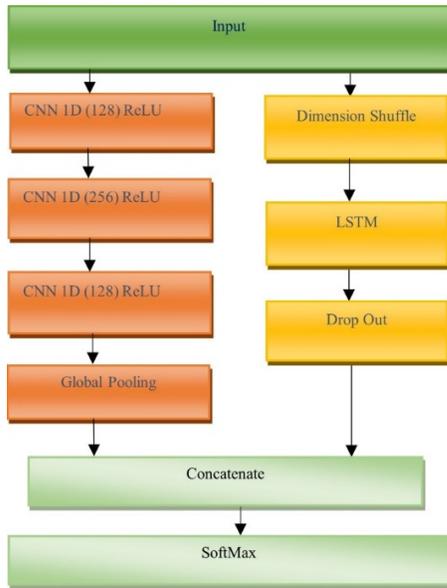
Fig. 6. LSTM-FCNN architecture. The input is EEG data to FCNN and LSTM. FCNN consists of three CNN layers, each having ReLU as an activation function. LSTM is applied to the same data then both FCNN and LSTM model layer has concatenated. The output of this is given to the softmax layer for the classification of individuals as visual or non-visual learners.

TABLE II
MODEL EVALUATION

(a) LSTM Eyes close condition

		Specificity	Sensitivity	f1-score	accuracy	predicted	support
LSTM	L	0.80	0.92	0.96	0.94	12	13
	NL	1.00	1.00	0.89		4	4

(b) LSTM-CNN Eyes close condition

		Specificity	Sensitivity	f1-score	accuracy	predicted	support
LSTM-CNN	L	0.80	0.92	0.96	0.94	12	13
	NL	1.00	1.00	0.89		4	4

(c) LSTM-FCNN Eyes close condition

		Specificity	Sensitivity	f1-score	accuracy	predicted	support
LSTM-FCNN	L	0.75	0.92	0.92	0.89	12	13
	NL	0.92	0.80	0.80		3	4

deep learning models have been used for the Classification of visual learners (L) and non-visual learners (NL) with raw EEG as input data. One hot encoding technique is used for all the models. The model implementations are done using Keras. The results of the classification models are discussed in detail in the following sections.

A. Performance Analysis of LSTM, LSTM-CNN, and LSTN-FCNN Classifier Using Raw EEG Eyes Closed Condition

The features were visualized using t-SNE method as shown in Figure 7. The t-SNE shows the best clustering distribution and is the most decentralized for all data. It clearly separates the Visual learner features (clustered together in purple) from the Non visual learner features (clustered together in orange), which has a positive influence on the classification of the low-dimensional data representation, as it increases the separation between the natural clusters in the data.

The model performance is evaluated by computing the accuracy, precision, and recall parameters and obtaining the receiver operating characteristic (ROC) curve. The computation of accuracy, precision, and recall involves the use of the confusion matrix. The recall parameter is determined by dividing the predicted number of learners (L) by the total number of learners as calculated by the model. A confusion matrix provides a means of evaluating the accuracy of a classification model by comparing its predictions to a set of known true values.

Table II(a) presents the accuracy, precision, recall, and confusion matrix for the raw EEG data obtained from various iterations. According to Table II(a), the LSTM model predicts 12 Ls and 4 NLs. Table II(b) shows that actual learners and non-learners are 13 Ls and 4 NLs, as indicated in the support

column. The components of the confusion matrix include true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The TPs indicate when individuals predicted to be visual learners are indeed visual learners. The TNs represent cases, where individuals predicted to be non-visual learners are, in fact, non-visual learners. The term "FPs" refers to the prediction of visual non-learners as visual learners. The FN (False Negative) occurs when individuals who are visual learners are incorrectly identified as non-visual learners. Table II(a) also displays the confusion matrix, which computes the accuracy, sensitivity, and specificity.

The ROC can be generated by plotting the FP rate along the x-axis and the TP along the y-axis. The ROC curve for the LSTM classifier is shown in Figure 7(a), where the backline represents visual learners and the green line represents non-visual learners. The AUC values for both classes are equal to 1. The system's accuracy is 96%, and it has an F1 score of 0.96 and 0.89. The ROC analysis serves several purposes, including assessing the ability of continuous predictors to correctly classify two groups, determining the optimal cut-off point to avoid misclassification, and demonstrating the effectiveness of the predictor.

The LSTM-CNN model was used in the subsequent step. According to Table II(b), there are 13 Ls and 4 NLs, as indicated in the support column. The model predicted 12 Ls and 4 NLs, as Table II(b) shows. The ROC can be generated by plotting the FP rate along the x-axis and the TP along the y-axis. The ROC curve of the LSTM-CNN classifier is depicted in Figure 7(b) with the visual Learner indicated by the backline and the non-visual Learner indicated by the green line. The AUC values for both classes were 0.95 and 0.96. The system's overall accuracy is 94%, with an F-score of 0.96 and 0.89.

According to Table II(c), the classifier indicated 5 Ls and 11 NLs. However, there are actual 6 Ls and 11 NLs. The ROC can be generated by plotting the FP rate along the x-axis and the TP along the y-axis. The ROC curve of the LSTM-FCNN classifier is presented in Figure 7(c), with the backline representing the visual Learner and the green line representing the non-visual Learner. The AUC values for both classes are 0.89 and 0.90, respectively. The system has an accuracy rate of 89%, with an F-score of 0.92 and 0.80.
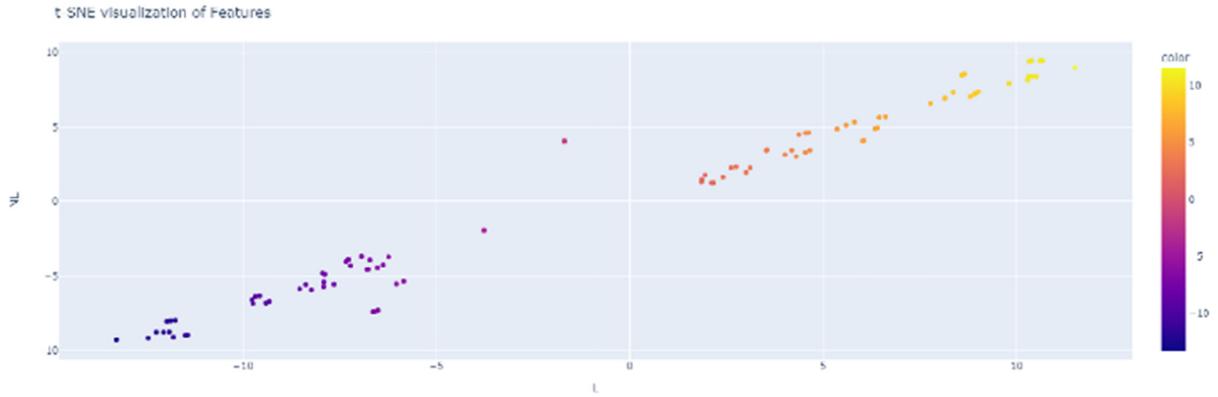
Fig. 7.   Visualization of learned features using t-SNE.

TABLE III
COMPARISON OF EXISTING FEATURES AND PROPOSED FEATURES

ML methods	Eyes closed condition Hand Crafted Feature extraction					
	alpha			gamma		
	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy
SCF	83	87	0.85	72	75	0.74
Amplitude ratio	76	77	0.77	72	75	0.74
PSD	94	99	0.97	94	94	0.94
DL methods	Automatic Feature extraction					
	Sensitivity		Specificity		Accuracy	
LSTM	80		92		94	
LSTM-CNN	80		92		94	
LSTM-FCNN	75		92		94	

TABLE IV
COMPARISON OF COMPUTATION TIME (INTEL(R) CORE (TM) I-5
10400 CPU @2.90GHz, 16.0 GB (15.8 GB USABLE, 64-BIT
OPERATING SYSTEM, C64-BASED PROCESSOR)

Machine Learning based Classification (averged time)		Deep learning-based Classification	
k-NN		LSTM	
Preprocessing time	10 minutes	Preprocessing	Not required
Feature extraction	30 minutes	Feature extraction	Automatic
Features selection	10 minutes	Training time	30 minutes
Training + Testing	10 minutes	Testing time	1 minute

The classification accuracy of LSTM and LSTM-CNN classifiers is similar, as evidenced by the results. However, the classification accuracy of LSTM-FCNN is on the lower side. According to our results, the LSTM-CNN model is recommended over LSTM for real-time applications due to its lower computation time.

B. Performance Analysis of Handcrafted Existing Features Extraction Methods to Identify Learning Style With the Proposed Deep Learning Method

For comparison, the existing feature extraction techniques to identify the learning style are selected and implemented

TABLE V
SUMMARY OF THE EXISTING WORK TO IDENTIFY THE LEARNING
STYLE OF THE STUDENT

Ref	The technique used for classification and validation	No of subjects	Advantages	Disadvantages	Classification accuracy
12]	Wilcoxon signed-rank test on the average relative power Mapping results of relative power by each task type	135	Accurately classify visual learner and verbal learner based on beta and theta band.	Small data set Does not fully represent all the aspects of learning modality.	Not-applicable
16]	KNN K-fold cross-validation	68	SCF feature has reduced feature complexity Use of EEG removing the drawback of subjective tests	KNN is using Euclidian distance thus misclassification of data is possible Use only male students Use synthetic EEG	Training accuracy is 100% Testing accuracy is 97%
17]	KNN K-fold cross-validation	8	New feature Spectral Centroid Features are introduced Successfully characterized resting brainwave at different development stages of human learning.	Since the study only focuses on the resting brain, the scope will be limited to the alpha and theta bands	Training accuracy 100 Testing accuracy 91.5%

on current study data. Those methods are selected from the more related literature and have shown high accuracy. These feature extraction methods include Spectral centroid frequencies (SCF) and amplitude ratio. These methods are used with the k-NN classifier and Euclidian distance, as published in the literature [18]. These methods are implemented for comparison with the proposed methodology. Results show that the Power Spectrum Density (PSD) features outperform the SCF and amplitude ratio features to identify the visual learning style of
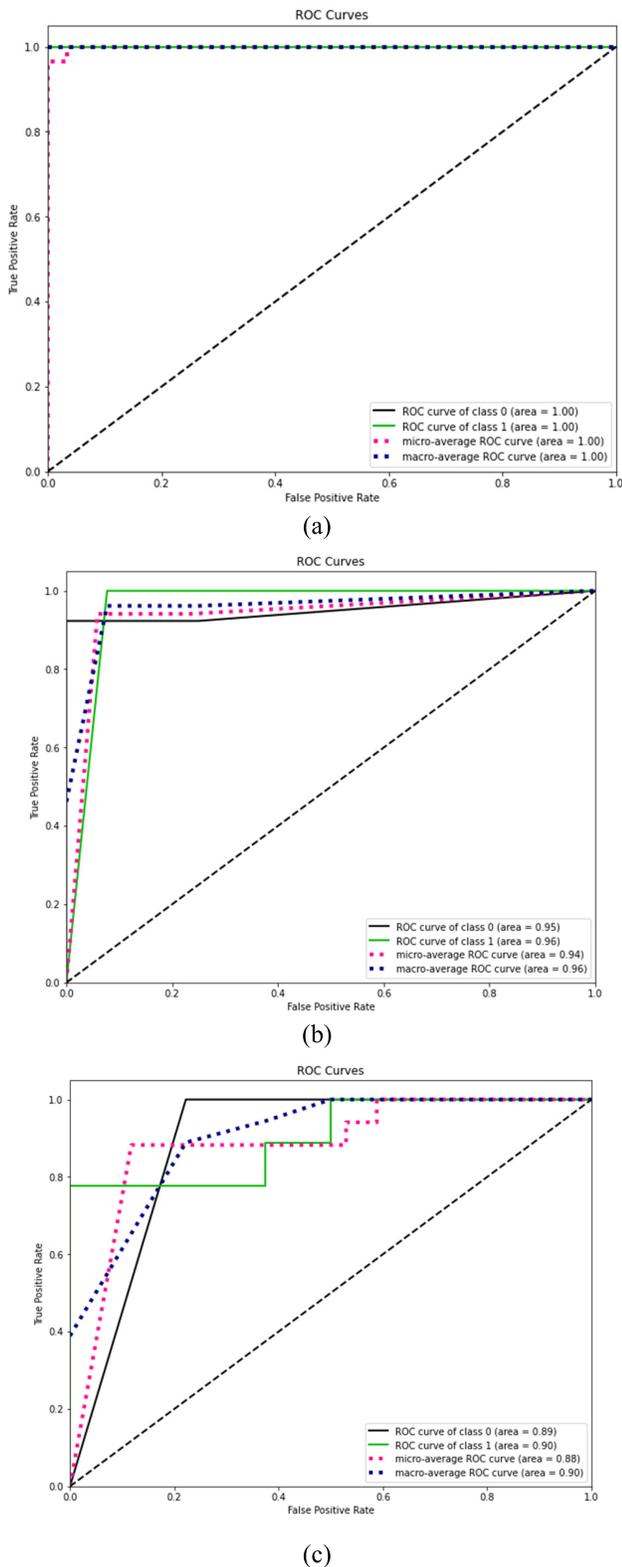
Fig. 8. The ROC curve of LSTM, LSTM-CNN, and LSTM-FCNN classifier for the Eyes close condition is presented here.

the students. Overall, the results show that the deep learning approach outperforms the machine learning-based methods. Table IV. shows the comparison of handcrafted features with automatic features. From Table III, it is observed that handcrafted feature PSD has better sensitivity and specificity

as compared to automatically extracted features. However, PSD cannot be employed for real-time processing. Hence, we choose automatic feature extraction methods because of their suitability for real-time applications, such as 1) no preprocessing requirements and 2) low computational complexity (having testing time as low as 1 minute), as shown in Table IV.

VI. DISCUSSION & CONCLUSION

Using the EEG dataset, the idea is to develop an assessment model to distinguish visual learners from non-visual learners. The learning style of the student can be identified in two ways. The first is based on Learning modalities and the other is based

on learning models. The advantage of using learning modalities over learning models is their stability as learning models lack a common framework. This study uses visual modality and the raw EEG data with deep learning techniques for classification. Using raw EEG with deep learning techniques can identify learning styles in real-time settings as it does not require offline processing and feature extraction. Therefore, this model is beneficial for educators and students, as it can help them identify their preferred learning style in real-time settings.

In the literature, it has been posited that the features based on EEG data can be utilized for identifying students' learning style. Conventional methods to identify students' learning style are mainly subjective and based on learning models. These models assess the learning style. However, they are based on self-assessment and may not manifest the learning style. For example, there might be a chance that a visual learner may mistakenly assess as an audio learner. This can be improved by using the objective evidence provided by neuroimaging modalities. Hence, to identify the learning style, objective measures based on neuroimaging modalities are required. The neuroimaging modalities provide additional evidence to perform a more accurate and valid assessment of learning style

In this context, the features extracted from various EEG frequency bands can be used to gather enough evidence to assess the learning style. More recently, the features extracted from the EEG are used as inputs to the ML methods. On the other hand, studies that involve the identification of learning styles utilize features extracted from both EEG and synthetic data. The synthetic EEG based features include EEG Spectral Centroid Features and asymmetry and ML techniques are used for the classification. The studies discussed have shown promising research results. However, due to certain limitations, these methods do not give the full picture and are not suitable. Some of the limitations include the use of synthetic EEG. Also, the studies have utilized small sample sizes. Moreover, the studies did not present their results in standard metrics, such as sensitivity, and specificities. Hence, the studies cannot be compared with each other and pose a hindrance in the generalization of the finding. Table II.8 shows a summary of the existing work to identify the learning style of the students.

There are some existing works that recognize the learning style of the students using EEG. These works mostly use local features, having an accuracy of 74% [18]. In other work, 1D- CNN is used. They classify the learning style with 71.2% accuracy [35]. In this work, local features are extracted and

then fed to a deep learning model for identification of learning style.

In our work, raw EEG is used to eliminate the preprocessing steps and use global features, increasing the model's accuracy by 94 %. This work can be used in real-time scenarios.

The high accuracy of this model makes it highly reliable. This designed model is not computationally expensive as it requires as low as a few milliseconds to complete the classification task. Because of this, it can train, test, and validate large datasets.

Thus, the presented paper outlines a model for identifying visual and non-visual learners through the utilization of deep learning techniques, including LSTM, LSTM-CNN, and LSTM-FCNN. The study shows deep learning techniques are more effective than conventional machine learning techniques in identifying visual learning styles. The proposed approach effectively identifies the visual learning style of students and demonstrates better performance in scenarios where machine-learning approaches fall short. The proposed method's performance was evaluated based on the quality of features used to identify learning styles from raw EEG data. The proposed deep learning method has achieved a higher level of accuracy, making it a reliable model that can be utilized in classroom settings to identify the visual learning style of the student.

REFERENCES

[1] S. Thepsatitporn and C. Pichitpornchai, "Visual event-related potential studies supporting the validity of VARK learning styles' visual and read/write learners," *Adv. Physiol. Educ.*, vol. 40, no. 2, pp. 206–212, Jun. 2016.

[2] S. K. M. Sreenidhi and M. T. C. Helena, "Styles of learning based on the research of Fernald, Keller, Orton, Gillingham, Stillman, Montessori and Neil D fleming," *Int. J. Innov. Res. Multidisciplinary Field*, vol. 3, no. 4, pp. 17–25, 2017.

[3] A. Pourhosein Gilakjani, "Visual, auditory, kinaesthetic learning styles and their impacts on English language teaching," *J. Stud. Educ.*, vol. 2, no. 1, pp. 104–113, Dec. 2011.

[4] M. Chen, J. Han, X. Hu, X. Jiang, L. Guo, and T. Liu, "Survey of encoding and decoding of visual stimulus via FMRI: An image analysis perspective," *Brain Imag. Behav.*, vol. 8, no. 1, pp. 7–23, Mar. 2014.

[5] N. M. van Atteveldt, M. van Kesteren, B. Braams, and L. Krabbendam, "Neuroimaging of learning and development: Improving ecological validity," *Frontline Learn. Res.*, vol. 6, no. 3, pp. 186–203, Dec. 2018.

[6] C. Cabral, M. Silveira, and P. Figueiredo, "Decoding visual brain states from fMRI using an ensemble of classifiers," *Pattern Recognit.*, vol. 45, no. 6, pp. 2064–2074, Jun. 2012.

[7] I. Dahlstrom-Hakki, J. Asbell-Clarke, and E. Rowe, "Showing is knowing: The potential and challenges of using neurocognitive measures of implicit learning in the classroom," *Mind, Brain, Educ.*, vol. 13, no. 1, pp. 30–40, Feb. 2019.

[8] G. Murugesan et al., "Assessment of mental workload by visual motor activity among control group and patient suffering from depressive disorder," *Comput. Intell. Neurosci.*, vol. 2022, pp. 1–10, Mar. 2022.

[9] *Advantages and Disadvantages of Visual Learning*. Accessed: Mar. 3, 2023. [Online]. Available: https://www.reference.com/education/advantages-disadvantages-visual-learning-5ede107612d4e80

[10] *Visual Learning Style*. Accessed: Mar. 3, 2023. [Online]. Available: https://www.mindtools.com/

[11] P. Honey, *The Learning Styles Helper's Guide*, P. Honey and A. Mumford, Eds. Maidenhead, U.K.: Peter Honey, 2006.

[12] E. Kim, I. Yang, and S. Lim, "An EEG study on visual learners' performance a scientific classifying task composed of pictures and words," *Adv. Sci. Technol. Lett.*, vol. 92, pp. 86–91, Jan. 2015, doi: 10.14257/astl.2015.92.18.

[13] R. Rajkumar and V. Ganapathy, "Bio-inspiring learning style chatbot inventory using brain computing interface to increase the efficiency of E-learning," *IEEE Access*, vol. 8, pp. 67377–67395, 2020.

[14] L. I. F. Dutsinma and P. Temdee, "VARK learning style classification using decision tree with physiological signals," *Wireless Pers. Commun.*, vol. 115, no. 4, pp. 2875–2896, Dec. 2020.

[15] M. M. Anoor, A. H. Jahidin, H. Arof, and M. S. A. M. Ali, "EEG-based intelligent system for cognitive behavior classification," *J. Intell. Fuzzy Syst.*, vol. 39, no. 1, pp. 177–194, Jul. 2020.

[16] M. S. A. Megat Ali, A. H. Jahidin, N. Md Tahir, and M. N. Taib, "Learning style classification via EEG sub-band spectral centroid frequency features," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 4, no. 6, pp. 931–938, Dec. 2014.

[17] M. S. M. Ali, M. N. Taib, N. Tahir, and I. M. Yassin, "Characterization of resting brainwave at different developmental stages of human learning," *Int. J. Control Theory Appl.*, vol. 9, no. 6, pp. 2985–2991, 2016.

[18] M. S. M. Ali, A. H. Jahidin, M. N. Taib, N. M. Tahir, and I. M. Yassin, "Classification of Kolb's learning styles using EEG sub-band spectral centroid frequencies and artificial neural network," *Asian J. Sci. Res.*, vol. 9, no. 5, pp. 234–241, Aug. 2016.

[19] C. Kelly, "David Kolb, the theory of experiential learning and ESL," *Internet TESL J.*, vol. 3, no. 9, pp. 1–5, 1997.

[20] R. Dunn, "Learning style: State of the science," *Theory Pract.*, vol. 23, no. 1, pp. 10–19, Jan. 1984.

[21] R. M. Felder and L. K. Silverman, "Learning and teaching styles in engineering education," *Eng. Educ.*, vol. 78, no. 7, pp. 674–681, 1988.

[22] C.-J. Lin and M.-H. Hsieh, "Classification of mental task from EEG data using neural networks based on particle swarm optimization," *Neurocomputing*, vol. 72, nos. 4–6, pp. 1121–1130, Jan. 2009, doi: 10.1016/j.neucom.2008.02.017.

[23] S. Alhagry, A. Aly, and R. A. El-Khoribi, "Emotion recognition based on EEG using LSTM recurrent neural network," *Int. J. Adv. Comput. Sci. Appl.*, vol. 8, no. 10, pp. 355–358, 2017.

[24] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," *J. Neural Eng.*, vol. 16, no. 5, Oct. 2019, Art. no. 051001.

[25] W. Zhao et al., "A novel deep neural network for robust detection of seizures using EEG signals," *Comput. Math. Methods Med.*, vol. 2020, pp. 1–9, Apr. 2020, doi: 10.1155/2020/9689821.

[26] M. Clerc, L. Bougrain, and F. Lotte, *Brain-Computer Interfaces 1: Methods and Perspectives*. Hoboken, NJ, USA: Wiley, 2016.

[27] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[28] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *J. Neural Eng.*, vol. 16, no. 3, Jun. 2019, Art. no. 031001.

[29] S. Yang, Z. Yin, Y. Wang, W. Zhang, Y. Wang, and J. Zhang, "Assessing cognitive mental workload via EEG signals and an ensemble deep learning classifier based on denoising autoencoders," *Comput. Biol. Med.*, vol. 109, pp. 159–170, Jun. 2019.

[30] H. U. Amin et al., "Feature extraction and classification for EEG signals using wavelet transform and machine learning techniques," *Australas. Phys. Eng. Sci. Med.*, vol. 38, no. 1, pp. 139–149, Mar. 2015.

[31] S. Jawed, H. U. Amin, A. S. Malik, and I. Faye, "EEG visual and non-visual learner classification using LSTM recurrent neural networks," in *Proc. IEEE-EMBS Conf. Biomed. Eng. Sci. (IECBES)*, Dec. 2018, pp. 467–471.

[32] A. Graves, "Supervised sequence labelling," in *Supervised Sequence Labelling With Recurrent Neural Networks*. Berlin, Germany: Springer, 2012, pp. 5–13.

[33] D. Stathakis, "How many hidden layers and nodes?" *Int. J. Remote Sens.*, vol. 30, no. 8, pp. 2133–2147, Apr. 2009.

[34] S. Yan. *Understanding LSTM and Its Diagrams*. Accessed: Feb. 15, 2023. [Online]. Available: https://medium.com/mlreview/understanding-lstm-and-its-diagrams-37e2f46f1714

[35] A. Miljevic, N. W. Bailey, F. Vila-Rodriguez, S. E. Herring, and P. B. Fitzgerald, "Electroencephalographic connectivity: A fundamental guide and checklist for optimal study design and evaluation," *Biol. Psychiatry, Cognit. Neurosci. Neuroimaging*, vol. 7, no. 6, pp. 546–554, Jun. 2022.

[36] Z. Jiang et al., "Decreased default mode network functional connectivity with visual processing regions as potential biomarkers for delayed neurocognitive recovery: A resting-state fMRI study and machine-learning analysis," *Frontiers Aging Neurosci.*, vol. 14, Jan. 2023, Art. no. 1109485.