Electroencephalogram-Based Attention Level Classification Using Convolution Attention Memory Neural Network

CHEAN KHIM TOA[ID]**1, KOK SWEE SIM**[ID]**1, (Senior Member, IEEE), AND SHING CHIANG TAN**[ID]**2**

[1]Faculty of Engineering and Technology, Multimedia University, Melaka 75450, Malaysia
[2]Faculty of Information Science and Technology, Multimedia University, Melaka 75450, Malaysia

Corresponding author: Kok Swee Sim (kssim@mmu.edu.my)

ABSTRACT Attentive learning is an important feature of the learning process. It provides a beneficial learning experience and plays a key role in generating positive learning outcomes. Most studies widely applied electroencephalogram (EEG) to measure human attention level. Although most studies use EEG handcrafted features and statistical methods to classify attention level, a more effective feature learning technique is still needed. In this paper, we aim to analyze participants' EEG signals through a deep learning model and classify those signals as showing either attentive or inattentive behaviors. To carry out this research, we initially conducted a background study on attention and its detection in EEG. After that, we design a Troxler's fading experiment and use an EEG device to collect data on participants' attentive and inattentive behaviors during the test. The collected EEG data will be analyzed using a Convolution Attention Memory Neural Network (CAMNN) model to classify participants' attention level. The proposed CAMNN model is optimized with Vector-to-Vector (Vec2Vec) modeling, where the model can be learned through deep neural networks in an end-to-end approach. The result shows that our model can achieve 92% accuracy and 0.92 F1 score which outperforms several existing neural network models such as Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN), Deep Learning with Convolutional Neural Networks (deep ConvNets), and Compact Convolutional Network for EEG-based BCIs (EEGNet). This research can be useful for those who are interested in developing attention level monitoring or biofeedback system in areas such as educational classroom learning, medical research, and industrial operator.

INDEX TERMS Attention level, deep learning, electroencephalogram, Troxler's fading.

I. INTRODUCTION

Attention is defined as a cognitive process of focusing on key information in the environment for learning and memory. Being attentive is important because it will benefit the learning experience and can potentially contribute to positive learning outcomes. Most professions such as teacher, operator, doctor, and driver require a high degree of attentiveness to complete the work efficiently and effectively [1], [2]. Thus, developing a method that can accurately monitor the human attention level has become a central interest of research.

The associate editor coordinating the review of this manuscript and approving it for publication was Sung Chan Jun[ID].

In the past, Krüger *et al.* [3] applied facial detection and recognition to determine whether a person is attentively learning. However, this method is too subjective and could not accurately determine the attention level. Kuo *et al.* [4] proposed an attention awareness system that uses eye tracking software to detect students' attention in the classroom. Although eye movement can provide information related to attention, it can only respond to obvious visual attention shifts, which may not be enough to determine the student's attention level. Due to the reasons, most of the research studies explored the relationship between Electroencephalogram (EEG) and attention level. EEG is an electrophysiological monitoring method that measures brain activity at

milliseconds time precision. Ko *et al.* [5] investigated the use of EEG and found that there is a close relationship between the changes in EEG spectral and impaired behavior performance in a sustained attention task. List *et al.* [6] used pattern classification to conclude that the EEG signals can be used to predict attention states. Liu *et al.* [7] used a wireless mobile EEG headset to record the EEG signals of the frontal cortex and adopted a Support Vector Machine (SVM) to classify student's attention state. Similar EEG headsets, feature selection methods, and classifiers were used by Gunawan *et al.* [1] to detect the early drop of attention and Peng *et al.* [8] designed an attentiveness recognition system. Typically, most research studies rely heavily on handcrafted EEG spectral features to perform classification, and there is a lack of research on the use of deep learned features and neural networks in this field. Hence, a deep learning model is applied in the research.

In this paper, we proposed a method with the use of EEG and a deep learning model for the classification of attention level. In the proposed method, eye tracking will be used during Troxler's fading experiment to ensure the reliability of the collected EEG data. Besides, a deep learning model is used as a classification method for the following reasons. First of all, deep learning does not require performing handcrafted features from raw data, because a neural network can learn features in an end-to-end approach [9]. Secondly, deep learning uses the Multilayer Perceptron (MLP) to learn and classify the characteristic of features, which is usually more effective than statistical methods, such as k-nearest neighbors (k-NN) and SVM [10]. This study aims to introduce a more effective learning algorithm to measure the attention level. For this purpose, we propose a novel deep learning model called Convolution Attention Memory Neural Network (CAMNN). It uses Vector-to-Vector (Vec2Vec) modeling with convolution block, attention mechanism, Long Short-Term Memory, and fully connected neural network to learn high-level features of the raw EEG signal and perform the classification. Other neural network models such as Recurrent Neural Network (RNN), Long Short-Term Memory (LSTM), and Convolutional Neural Network (CNN), Deep Learning with Convolutional Neural Networks (deep ConvNets), and Compact Convolutional Network for EEG-based BCIs (EEGNet) have also been applied for comparison with the proposed model.

The contribution of this research is in three folds. First is to observe the brain activity during Troxler's fading experiment by analyzing EEG signals. Second is to propose a deep learning model with Vec2Vec modeling that can learn in an end-to-end approach by a neural network. Third is to show that our proposed model can outperform other deep learning models in terms of accuracy, precision, recall, specificity, and F1.

Section II will present the related work on the use of EEG for attention analysis and the contribution of deep learning in the field of EEG. Section III will describe the research material and proposed method. Section IV will present the result and discussion of the proposed method. Section V is the conclusion summarizing salient finding in this research.

II. RELATED WORK
A. ATTENTION ANALYSIS USING ELECTROENCEPHALOGRAM

The application of electroencephalogram (EEG) in attention analysis has attracted the attention of researchers [11], [12]. Li *et al.* [13] developed a learning system that using the combination of EEG features and the Self-Assessment Manikin (SAM) model as the input for the k-Nearest-Neighbour classifier and Naïve Bayes classifier to classify the learner's attention. The EGG features are extracted using signal processing algorithms such as Fast Independent Component Analysis (FastICA) and Approximate Entropy (ApEn). The SAM model is an assessment that measures valence and arousal associate with the learner's reaction to a variety of stimuli. Li *et al.* [13] stated that EEG signals are proven useful in analyzing the learner's attention during the learning process. Gunawan *et al.* [1] discussed the usage of EEG for detecting the early drop of attention. In this work, the Neurosky Mindwave device has been used to record the participants' brain waves when performing continuous performance tests. The collected signals are then inserted into Fourier transform and Power Spectral Density (PSD) to extract features. After that, the k-nearest neighbors (k-NN) algorithm is applied to classify the signal based on the selected features. It shows that 70% accuracy was achieved. Alirezaei and Hajipour Sardouie [14] used the Emotiv device to classify human attention in an educational environment. The recording of EEG is performed in the frontal cortex because this area can obtain significant signals of the attention state. Besides, Alirezaei and Hajipour Sardouie [14] uses the energy extraction method to compute features from raw EEG data and finds that features of the beta band can obtain important information about the attention state. Abiri *et al.* [15] developed a portable EEG-based platform to monitor the visual attention states. A wireless EEG headset called Emotiv EPOC has been used to collect the EEG signals during the visual discrimination task, where the participants were asked to respond only to the targeted subcategory while ignoring the irrelevant subcategories. An individualized SVM model is used and an average accuracy of about 77% was achieved when decoding the participant's attention state based on their brainwave signals. Peng *et al.* [8] discussed an attentiveness recognition system based on EEG. In this work, Neurosky Mindwave was used to record the brain activity of participants during spot-the-different puzzles and during the resting phase. The empirical mode decomposition (EMD) was applied to decompose the collected signal into Intrinsic Mode Functions (IMF), which was then furthered break down into several features by using Hilbert–Huang transform (HHT). Only selected features will be applied into Support Vector Machine (SVM) for the classification. It shows that an accuracy of 84.8% was achieved to distinguish between attentiveness and relaxed states. Kim *et al.* [16] studied the relationship between EEG

and eye gaze in boredom circumstances. An experiment was designed with the use of video stimulus to elicit boredom. During the experiment, the eye movement and brain activity will be recorded by eye tracker and EEG sensor. For analysis, a heat map of eye gaze data generated to confirm that the participants were bored. While for the EEG dataset, a thresholding technique and several statistical methods (standard deviation, mean, signal magnitude area) were applied to identify the significant information that could indicate the boredom. It seems that most of the related works were relied on handcrafted features, where they were manually selected and fed into the statistical methods for classification. Furthermore, there is a lack of research on the usage of neural network for the classification of attention level.

B. DEEP LEARNING IN ELECTROENCEPHALOGRAM

Deep Learning can provide a significant contribution to the classification of electroencephalogram (EEG) in different research fields. For example, An *et al.* [17] describes the advantages of using deep learning in EEG and applying Deep Belief Net (DBN) to classify EEG signals of Motor Imagery (MI) task. The results show that the DBN model can provide consistent improvements for all MI tasks. Hasib *et al.* [18] used a Hierarchical Long Short-Term Memory (H-LSTM) model to solve the non-stationarities of EEG signals and showed that the H-LSTM model could be used to perform prediction on a human decision based on EEG. Schirrmeister *et al.* [19] designed a Deep Learning with Convolutional Neural Networks (deep ConvNets) to decode movement-related information from the raw EEG signals. The deep ConvNets do not rely on the handcrafted features to decoding the movement classes. Besides, the convolution and nonlinearities of deep ConvNets can learn from non-linear features and represent as high-level features. Schirrmeister *et al.* [19] showed that the use of deep learning with an end-to-end approach can obtain satisfactory accuracy. Lawhern *et al.* [20] introduced the use of the Compact Convolutional Network for EEG-based BCIs (EEGNet) model, which is designed to classify EEG signals in different BCI paradigms with limited data. The use of depthwise and separable convolutions in the model has encapsulated the EEG feature extraction concept for BCI. Lawhern *et al.* [20] shows that the EEGNet model can provide good performance on extracting a variety of neurophysiological interpretable features over a range of BCI tasks. Kim and Choi [21] proposed an end-to-end emotional analysis model, which uses Long Short-Term Memory (LSTM) network with an attention mechanism for emotional recognition based on EEG signals. Kim and Choi [21] showed that the attention mechanism can help on improving the accuracy of emotional recognition by considering the peak-end rule.

III. MATERIALS AND METHODS
A. EXPERIMENT STIMULATOR AND PROTOCOL

In order to successfully elicit participants' attention behavior, an experiment protocol was designed as shown in Figure 1.
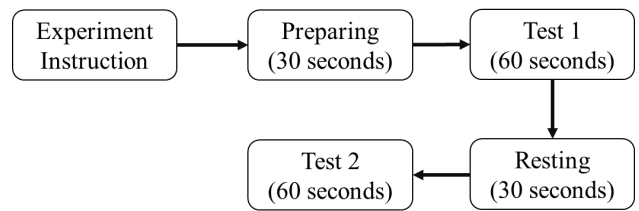

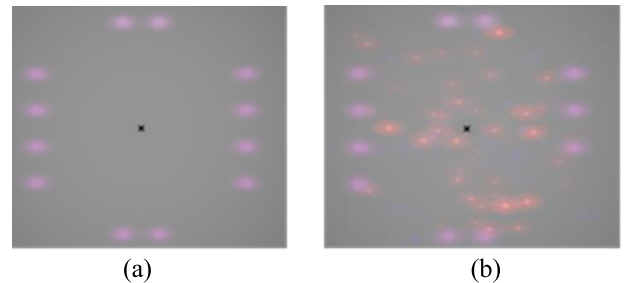**FIGURE 1. Troxler's fading experiment protocol.**

FIGURE 2. a) Control test (without distraction) and b) Experimental test (with distraction).

The duration of the experiment is 180 seconds including preparation and rest. The experiment is based on Troxler's fading, which is a visual field test that has a great influence on participants' sustained attention and selective attention [22]. Massa and O'Desky [23] and Jansiewicz *et al.* [24] showed that Troxler fading experiment can provide great help in the attention-related area. They conducted an experiment to study the impaired habituation of children and adults with attention deficit hyperactivity disorder. In the experiment, if the participants can focus or fix their attention on the targeted stimulus located in the central, the peripheral stimulus will fade from awareness. Figure 2 shows the application of Troxler's effect, where Figure 2a stimulates participants' attentiveness by focusing on the targeted stimulus (Black "X" mark) and Figure 2b stimulates participants' inattentiveness by diverting their attention from focusing on the targeted stimulus.

Before experiment, participants were briefly introduced to the nature of the experiment and obtained their consent. Next, they were asked to sit in a quiet empty room with a chair, table, and desktop computer prepared beforehand. A basic instruction sheet was provided for them to understand the experiment process. After that, the EEG device was properly attached to the surface of participants' scalp and they have a rest for 30 seconds to get used to the EEG device. Later, participants began the control test (Figure 2a) where they focused on the targeted stimulus for 60 seconds. If their gaze is properly fixed on the stimulus, the peripheral area (fuzzy purple circle) of the central dot would gradually disappear, indicating that the participant showed attentive behavior. After 60 seconds, participants were given a 30-second rest, to set the visual baseline. Then, participants proceeded to the experimental test (Figure 2b), which included red color particles as distraction factor. These particles will move randomly in the scene and cause participants to lose focus on the targeted stimulus, thus showing inattentive behavior.
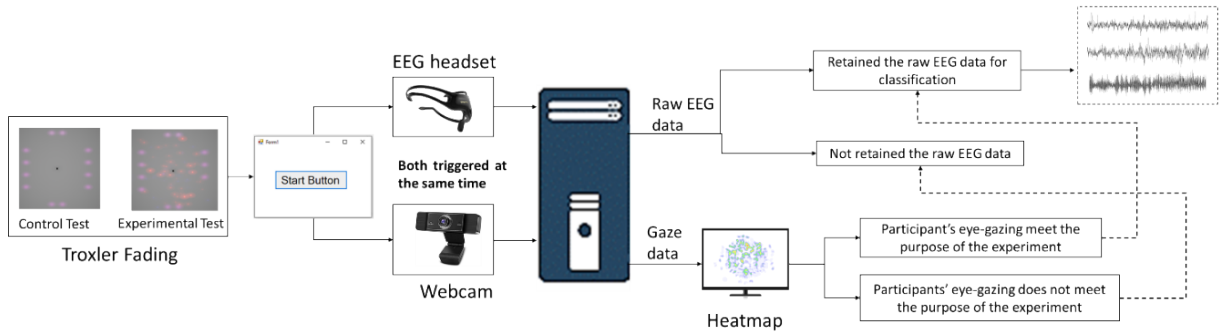
FIGURE 3. The framework of experiment process.

The experiment was monitored with a webcam to track the gaze of the participants. An eye tracking algorithm is used to obtain the gaze data of the participant when looking at the screen. Figure 3 shows the framework of the experiment process. A simple graphical user interface with 'Start Button' is created, which is used to triggered EEG headset and webcam at the same time when the participant clicks the button. The EEG headset and webcam are used for different purposes. The use of eye tracking through webcam is to ensure that the collected EEG data are useful to meet the objective of the experiment. For example, if it is found that the participant's eye is not properly fixed on the targeted stimulus during the control test, the raw EEG data of that test will not be retained. A heat map will be drawn to see where the participant eye looking at the screen. The collected raw EEG data are used as the input of Convolution Attention Memory Neural Network (CAMNN) for the binary classification. The detailed process of the CAMNN model will be described later.

B. RECORDING INSTRUMENT

There are two recording instruments used in the experiment. The first instrument is the EMOTIV Insight wireless headset, which is used to record participants' EEG signals [25]. The device has 5 Channels of electrodes which are AF3, AF4, Pz, T7, and T8. Since our study is about attention analysis, only electrodes in the frontal cortex (AF3 and AF4) and parietal cortex (Pz) are used [22]. As for the sampling frequency of EEG signals, the device can record up to 128 Hz, which means that 128 time points can be acquired in one second. Since the duration of each test sample is 60 seconds, there will be 7680 time points. Furthermore, the device has a resolution of 14 bits with one least significant bit equals to 0.51 μV.

The second instrument is webcam-based eye tracking, which is a video camera suitable for recording the eye gaze data of participants [26]. The webcam has a resolution of 1920 $\times$ 1080 (Full High Definition) with 30 frames per second. Figure 4 shows the flowchart of the eye tracking algorithm. A simple calibration will be performed for each participant eye at the start of experiment. The participant will be fixing their gaze at several points on the screen during calibration. In our study, the gaze direction is an indicator
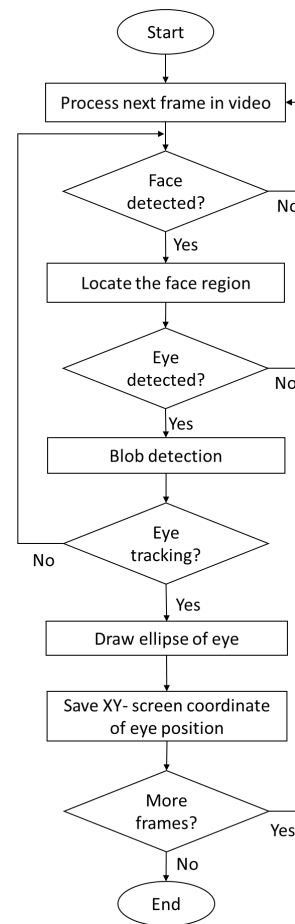

FIGURE 4. Eye tracking algorithms.

of showing spatial focus of a person's eyes, whether from left to right, or vice versa, while the gaze data are a set of XY-coordinates on the screen that show the position where the eyes are looking at. Figure 5 shows the use of eye tracking algorithm to track the eye (green circle) through a webcam.

C. PARTICIPANT DATASET

In this study, 30 healthy participants (18 males and 12 females) whose ages are between 20 and 30 years old volunteered to participate in the experiment. Written consent
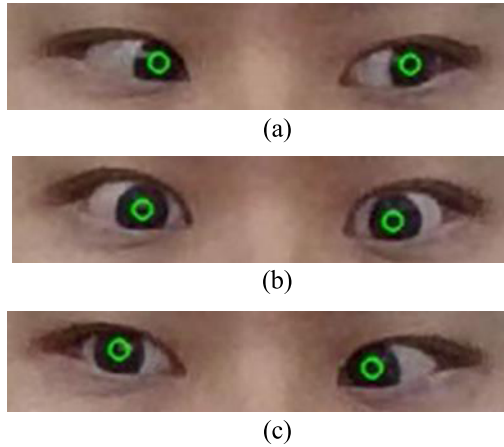
FIGURE 5. To track participant eye position using eye tracking algorithm: (a) Left, (b) Center, and (c) Right.
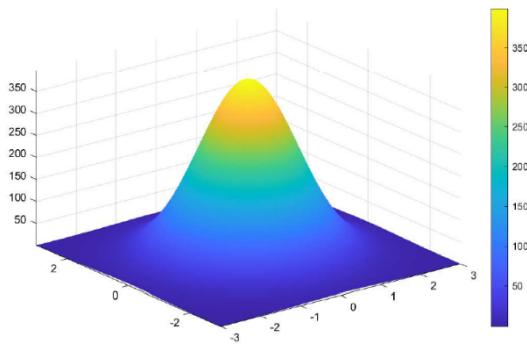

FIGURE 6. The Gaussian curve with a two-dimensional domain.

is signed by each participant prior to the experiment. All participants were informed that they could stop the experiment at any time as they felt not well. The participants were required to participate in the experiment for 4 weeks, of which they would join in the experiment 2 days per week and 5 times per day.

D. EYE GAZE DATA VISUALIZATION TECHNIQUE

To track the performance of participants during the test, a heat map is used to visualize the eye movement in the area of interest. The heat map is a useful data analytic software that uses gaze data as input and applies a hot-to-cool color spectrum to an area that receives attention [26]. In order to construct the heat map, a two-dimensional Gaussian function has been used. Figure 6 shows the curve of an elliptical Gaussian function.

The mathematical calculation of the two-dimensional Gaussian distribution is shown in Equation 1.

$$f(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x-\mu_x)^2 + (y-\mu_y)^2}{2\sigma^2}}, \qquad (1)$$

where x and y are screen coordinate of the eye position, σ is the standard deviation of the Gaussian function, and μ_x, μ_y is the mean of the x and y coordinate.
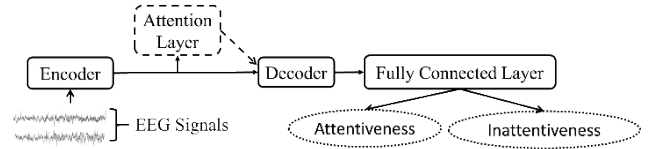

FIGURE 7. Vector-to-vector (Vec2Vec) modeling.

E. CONVOLUTION ATTENTION MEMORY NEURAL NETWORK

To effectively extract the features from EEG signals and classify these signals into either attentive or inattentive behavior, a deep learning model called Convolution Attention Memory Neural Network (CAMNN) is proposed. The model uses Vector-to-Vector (Vec2Vec) modeling as shown in Figure 7 to perform the feature extraction and classification. Vec2Vec is inspired by the work in Sequence-to-Sequence (Seq2Seq) modeling, which has shown powerful representation capabilities in applications such as language translation, voice recognition, image captioning, and text summarization [27], [28]. To the best of our knowledge, there is no research on Vec2Vec modeling applied for classifying the attention level of EEG signals. The Vec2Vec is divided into 4 parts, which are the encoder, attention layer, decoder, and fully connected layer. In the CAMNN model, the encoder is Residual Networks 34 (ResNet34), the attention layer is a soft attention mechanism, the decoder is a Long Short-Term Memory (LSTM) and the fully connected layer is a network with one hidden layer. These parts will be explained in details in the following subsections.

1) ENCODER

The encoder is used to perform the feature extraction of the two-dimensional (2-D) input EEG signals using the convolution operation of ResNet34, as shown in Figure 8. ResNet34 is a convolution neural network that has been used as a backbone for numerous computer vision tasks and image classification [29], [30]. The biggest advantage of ResNet34 is that it can increase the depth and accuracy of the network model while avoiding the negative outcome. The input EEG signals of the ResNet34 are using a single input channel (i.e. grayscale image) and are resized to a size of 700 × 700. Table 1 shows the parameters of ResNet34 such as kernel size, padding, stride, and output size. These parameters will be inserted into the operation, which consists of a convolution block, a convolution layer with an identity shortcut, and a convolution layer with a projection shortcut. Convolution block is feature learning, which is composed of convolution, batch normalization, non-linear activation function, and pooling layers as shown in Figure 9.

The two-dimensional convolution consists of 64 kernels, in which all EEG signals will be learned during the training process. These kernels will stride horizontally column by column on the input image matrix to identify 64 feature maps. Later, these feature maps will be fed into batch normalization (BN) to enable a higher learning rate and thus, speed up
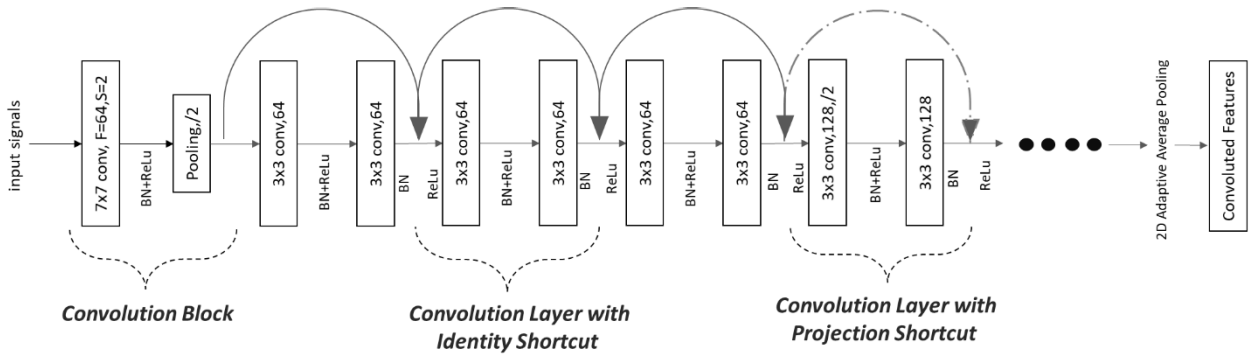
FIGURE 8. Convolution operation of ResNet34.

TABLE 1. Recall parameters of ResNET34.

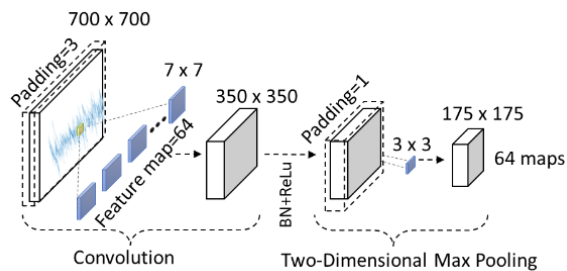| Layer name	Output size	ResNet34-layer
conv1	$350 \times 350 \times 64$	7×7, 64, stride 2, padding 3
pool1	$175 \times 175 \times 64$	3×3, max pool, stride 2, padding 1
conv2_x	$175 \times 175 \times 64$	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 3$
conv3_x	$88 \times 88 \times 128$	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 4$
conv4_x	$44 \times 44 \times 256$	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 6$
conv5_x	$22 \times 22 \times 512$	$\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 3$

FIGURE 9. Convolution block.

the training process. Equations 2-5 show the formulation for batch normalization.

$$\mu_B = \frac{1}{m} \sum_{i=1}^{m} x_i \qquad (2)$$

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^{m} (x_i - \mu_B)^2 \qquad (3)$$

$$\hat{x}_i = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \in}}, \qquad (4)$$

$$y_i = BN_{\gamma,\beta} (x_i) \qquad (5)$$

where m is the batch size, x is the activation value, μ_B is the mini-batch mean, σ_B^2 is the mini-batch variance, $\hat{x}_i$ is the normalized activation, $\in$ is a noise parameter used if the variance becomes zero and y_i is the learnable scale and shift parameter. After BN, these features will be fed into the
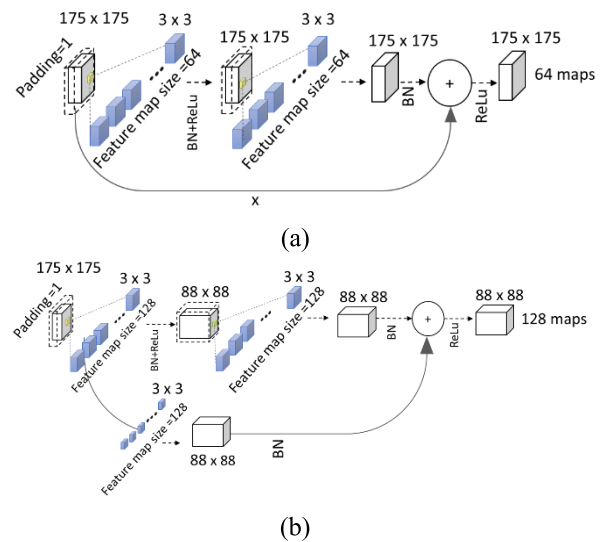

FIGURE 10. Convolution layer with a) identity shortcut and b) projection shortcut.

rectified linear unit (ReLu), which is a non-linear activation function. The purpose of ReLu is to activate only certain neurons to introduce non-linearity in the neural network [31]. These neurons are only activated if the output of the linear transform is more than or equal to 0.

Later, these feature maps will be down sampled by using the two-dimensional max pooling. The goal is to reduce computational cost so that the execution speed is faster and also to avoid model from overruns. The output feature maps (*Op*) of the convolution block are calculated using Equation 6.

$$Op = \frac{L - K + 2P}{S} + 1, \qquad (6)$$

where L is the width or height of the image, K is the kernel size, P is the number of zero paddings, and S is the stride. All the feature maps will go to the next operations which are the convolution layer with identity shortcut and the convolution layer with projection shortcut [32]. The identity shortcut as shown in Figure 10(a) is used when the output feature maps volume is the same as the input feature map volume. Thus, input feature maps volume can pass to the addition
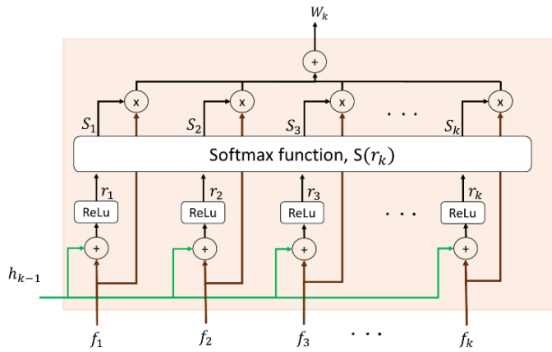
FIGURE 11. Soft attention mechanism.

FIGURE 12. Long short-term memory (LSTM) cell.
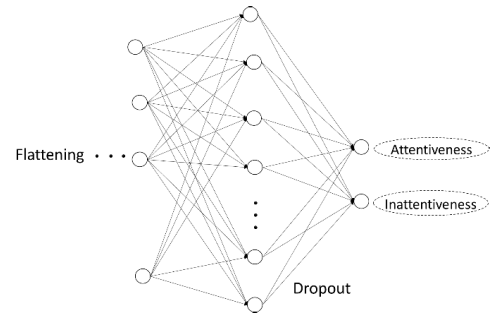
FIGURE 13. Fully connected network with one hidden layer.

operation. The formulation of the identity shortcut is shown in Equation 7.

$$Output = F(x) + x, \qquad (7)$$

where x using the previous convolution as input feature maps volume and $F(x)$ is output feature maps volume with 2 convolution layers. The projection shortcut is used when the input and output feature map volume are different from each other. So, to ensure that both can perform the additional operation, the input volume will perform the convolution operation as shown in Figure 10(b). The formulation of the projection shortcut is shown in Equation 8.

$$Output = F(x) + Wx, \qquad (8)$$

where x is input features volume, W is convolution layer, and $F(x)$ is output features volume.

2) ATTENTION LAYER

The convoluted features will then be passed to the attention layer. In this layer, a soft attention mechanism will be used to focus on the relevant part of features [33], [34]. Figure 11 shows the structure of the soft attention mechanism.

Based on Figure 11, the mechanism will use convoluted features ($f_1, f_2, \ldots, f_k$) and hidden state (h_{k-1}) as the input to calculate the weighted annotation vector which will be used for the decoder. h_{k-1} is the output of an LSTM cell that acts as a neural network memory for storing information of output data. The aggregation value of f_k and h_{k-1} with ReLu layer is first calculated, as shown in Equation 9 and Equation 10.

$$x_k = f_k + h_{k-1} \qquad (9)$$

$$r_k(x_k) = \begin{cases} 0 & for \ x_k < 0 \\ x_k & for \ x_k \geq 0, \end{cases} \qquad (10)$$

where x_k is the sum of convoluted features and hidden state at time step k and $r_k(x_k)$ is the aggregation value after the ReLu layer. Next, a softmax function as shown in Equation 11 is used to calculate the attention weight of the aggregation value, $S(r_k)$.
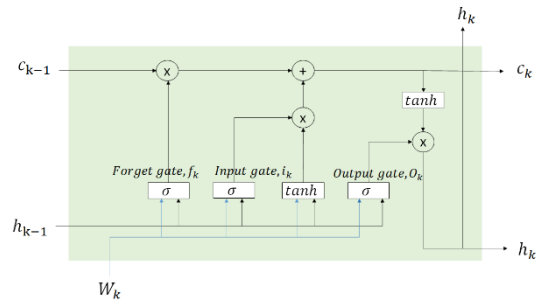
$$S(r_k) = \frac{e^{r_k}}{\sum_k e^{r_k}} \qquad (11)$$

Finally, the attention weighted annotation vector, W_k can be computed using Equation 12. W_k represents the relevant information for each feature according to the hidden state.

$$W_k = \sum_k S_k f_k \qquad (12)$$

3) DECODER LAYER

For the decoder, the Long Short-Term Memory (LSTM) cell was chosen because of its powerful characteristic, which can make the next prediction while storing the previous time-series electroencephalogram (EEG) signals [35]. The LSTM cell consists of three gates as shown in Figure 12. Each gate contains a sigmoid function, which is used to regulate the flow of information [36]. The first gate is the forget gate, which controls whether to keep or forget the input information (previous hidden state and current input) according to the output value (0 or 1). If the output value closer to 1, the input information will be kept, otherwise, it will be forgotten. The second gate is the input gate, which is used to control whether the input information needs to be updated. The gate will control the flow of input information into the cell by referring to the output value (0 or 1). If the value is 1, the input information is important, otherwise, it is not important. The third gate is the output gate, which is used to determine how much input information is needed in the cell. The output hidden state can be obtained by multiplying the new cell state with the output vector. The new cell state is calculated by making an additional between the forget vector and the input vector. Equations 13-15 show the formulation of these three gates.

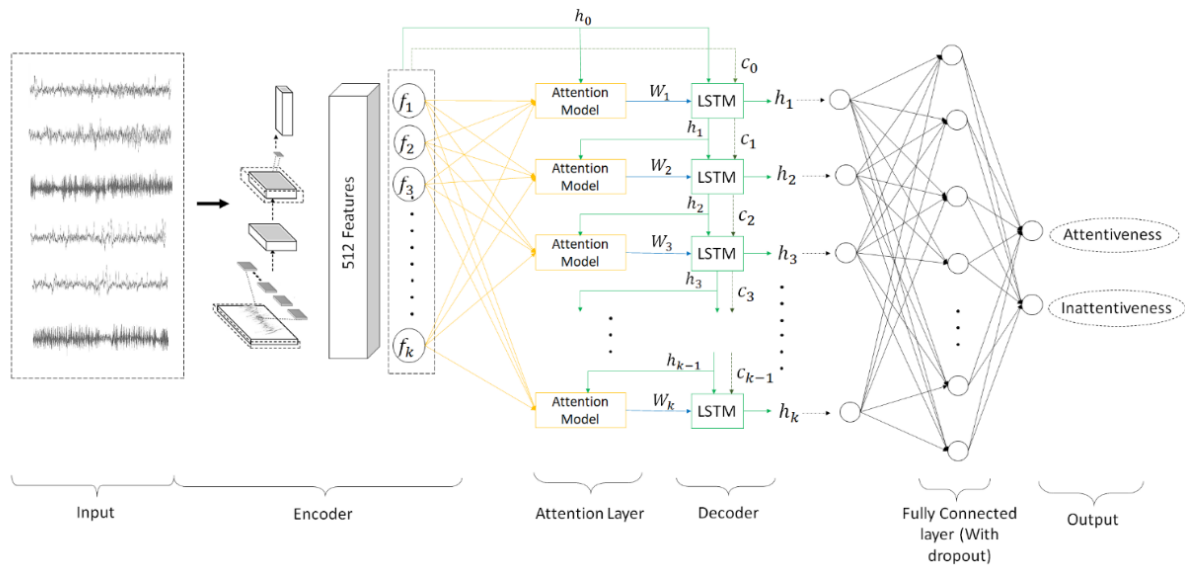$$f_k = \sigma(M_f W_k + M_f h_{k-1} + b_f) \qquad (13)$$

FIGURE 14. Complete overview of convolution attention memory neural network (CAMNN).

$$i_k = \sigma(M_i W_k + M_i h_{k-1} + b_i) \qquad (14)$$
$$o_k = \sigma(M_o W_k + M_o h_{k-1} + b_o) \qquad (15)$$

where f_k, i_k, and o_k are forget, input, and output vectors, M_f, M_i, and M_o are weight metrics, W_k is the attention weighted annotation vector, h_{k-1} is the hidden state, and b_o is the bias vector.

4) FULLY CONNECTED LAYER
The fully connected network with one hidden layer as shown in Figure 13 is used to take the output of the decoder, flattened it into a single vector of values, and calculate the probabilities for each label (Attentive or Inattentive behavior). A dropout rate of 0.5 is performed after each layer. The use of dropout is to prevent overfitting of training data [37].

5) OVERVIEW OF CAMNN
Figure 14 shows an overview of the Convolution Attention Memory Neural Network (CAMNN) model. Initially, the EEG signal will be fed into the convolution operation of ResNet34 to extract high-level convoluted features. After that, these features will be fed into the soft attention mechanism and LSTM cell to generate 512 hidden states. The purpose of the soft attention mechanism is to assist LSTM in learning important convoluted features at each step. Finally, all the hidden states will be channeled to the neural network in a fully connected layer to perform classification.

IV. RESULT AND DISCUSSION
In this section, we will present the experimental results of the participant's eye gaze and electroencephalogram (EEG) data. Since 30 participants voluntarily participated in the 4 weeks' experiment, the total number of EEG data samples obtained from each test of the experiment is 7100, of which 3550 are the control test and 3550 are the experimental test. Five-fold cross-validation is used to evaluate the CAMNN model by
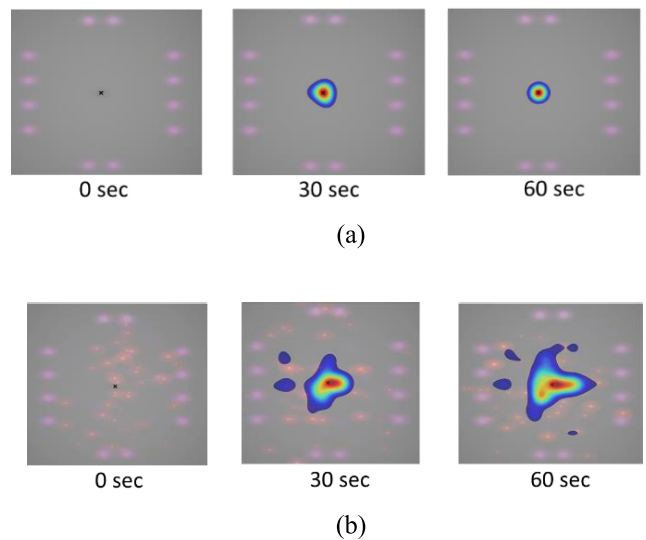

(a)

(b)

**FIGURE 15.** Heat map visualization on a) Control test and b) Experimental test.

partition the data samples into training sets and test sets. All data samples were carefully examined, and the poor EEG signal quality was eliminated.

A. EYE GAZE GRAPHICAL DATA REPRESENTATION
To visualize the gaze points of the participants in the control and experimental tests, a heat map has been created as shown in Figure 15.

Through visualization on the heat map, we can determine whether the participant's gaze is fixed on the targeted stimulus or deviated from the targeted stimulus (Black "X" mark). During the control test, if it is found that the participant's gaze is not properly fixed on the stimulus, the raw EEG data of that test will not be retained since it does not meet the objective of the experiment. Figure 15(a) shows that during
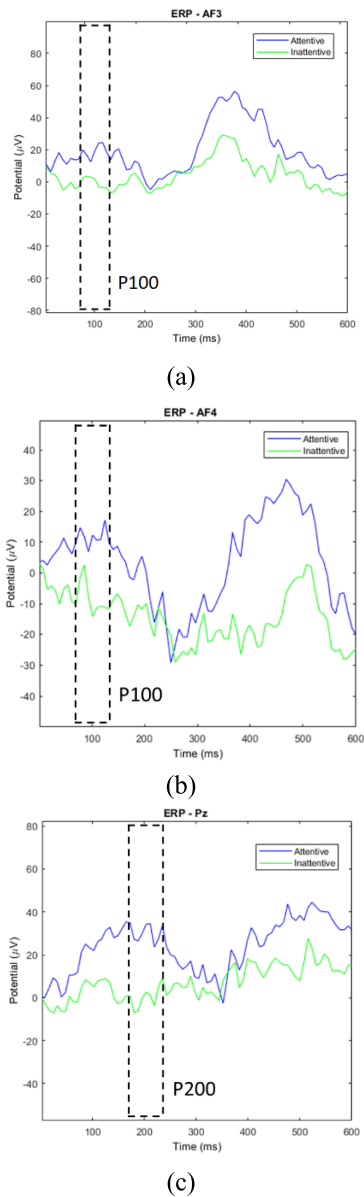
FIGURE 16. Grand-average ERP components waveforms (a) P100 for channel AF3, (b) P100 for channel AF4, and (c) P200 for channel PZ.

the control test, the color of the heat map is focused on the stimulus area. This indicate that the participant's gaze is fixed on the stimulus and thus elicited attentive behavior. As for Figure 15(b), during the experimental test, the color of the heat map is scattered over several regions, which indicates that red color particles successfully caused the participants to lost focus and deviate their gaze from the stimulus, thus eliciting inattentive behavior. The result shows that webcam-based eye tracking can provide us with participant's gaze data, visualized as a heat map, which can help to determine whether the participant's attentive and inattentive behaviors are being elicited during the control test and the experimental test, thereby ensuring the reliability of the collected EEG data.

B. GRAND-AVERAGED EVENT-RELATED POTENTIAL

Event-Related Potential (ERP) is derived from the electroencephalogram (EEG) measurement of brain response. An ERP is used to measure the voltage change results of attentive and inattentive behaviour during the experiment [38]. Figure 16 shows a plot of grand-average ERP waveform. It is constructed by first averaging all trials of EEG signals within a subject to produce the average ERP waveform. Later, the average ERP waveform across the subjects will be averaged to generate grand-average ERP waveform. By averaging, we can average out the random brain activity and remain the relevant waveform. Based on the graph, AF3 and AF4 channels having a positive peak around 100 ms, which represent the P100 of the ERP component. As for Pz channel, a positive peak is found around 200 ms, which represent the P200 of ERP component. The P100 and P200 is waveform components that are commonly related to the early processing of attention state [39]. According to the waveforms shown in the graph, the P100 mostly distributed over the frontal cortex, whereas the P200 is often distributed in the parietal cortex. Those waveforms show that when the participant elicit attentive behavior by fixing the gaze on the target stimulus, the potential of P100 and P200 will be higher. As for the inattentive behavior, since the participant's gaze is deviated from the stimulus, the P100 and P200 potential will be lower. Thus, it shows that the changes in the ERP waveform are correlated with the gaze of participants, regardless of whether their gaze is fixed on the stimulus or deviated from the stimulus.

C. EEG SIGNALS DEEP LEARNING CLASSIFICATION

Since we introduce the use of deep learning in the attention level classification, 5 existing models have been designed and compared with the proposed model. The first model is a Recurrent Neural Network (RNN), which feeds time-series EEG signals into the model and makes a prediction. The RNN model is composed of internal memory, which is used to memorize the previous output data and consider it together with the current input, thus being able to make a classification for EEG signal. The RNN model included 2 hidden layers and 28-time steps for classification. The second model is Long Short-Term Memory (LSTM), which is a special kind of RNN with the addition of three gates (input gate, forget gate, and output gate). These gates are used to update, forget or control the information in the cell so that it can make predictions while storing the previous time-series EEG signals. The LSTM model includes 3 hidden layers and 28-time steps for classification. The third model is the standard Convolutional Neural Network (CNN), which is mainly used for image classification. The CNN model includes three convolution operations for feature extraction and a fully connected layer for EEG classification. The fourth model is Deep Learning with Convolutional Neural Networks (deep ConvNets), which consists of four convolution blocks and a fully connected layer. The first convolution block of deep

TABLE 2. Evaluation metrics comparison between our camnn model and other neural network models.

Classification Models	Evaluation Metrics				
	Accuracy	Precision	Recall	Specificity	F1-Score
RNN	52%	0.60	0.46	0.64	0.52
LSTM	60%	0.54	0.89	0.53	0.67
CNN	72%	0.75	0.77	0.76	0.76
deep ConvNets	80%	0.83	0.77	0.83	0.8
EEGNet	76%	0.82	0.69	0.83	0.75
CAMNN	92%	0.93	0.92	0.90	0.92

ConvNets is different compared to the standard CNN convolution block where it will split into convolution across time and convolution across electrodes to handle a large number of input channels. The rest of the convolution blocks in ConvNets will be following the standard CNN. The fifth model is the Compact Convolutional Network for EEG-based BCIs (EEGNet), which consists of three convolution blocks and fully connected layers. Those three convolution blocks include temporal convolution to learn frequency filter, depth-wise convolution to learn frequency-specific spatial filter, and separable convolution to learn to summarize feature map and then optimally mix them. All existing models have implemented a dropout rate of 0.5 in the fully connected layer.

To show the generalization capability of the CAMNN model, five evaluation metrics which are accuracy, precision, recall, specificity, and F1-Score are used as shown in Equations 16-20. The input values of these equations are obtained from the confusion matrix, which is a tabular representation used to visualize the generalization of the model. Table 2 shows the result of the evaluation metrics of the CAMNN model and other existing models.

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \times 100\% \quad (16)$$
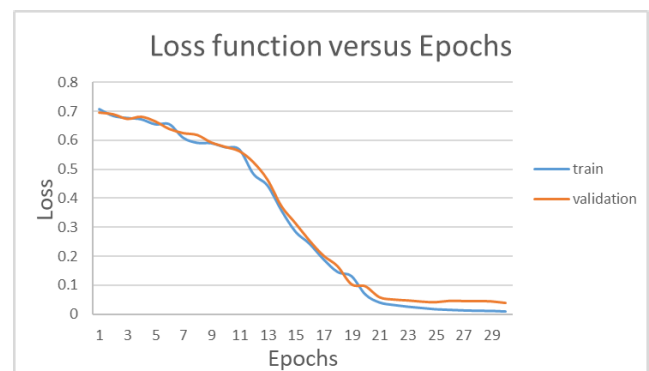
$$Precision = \frac{TP}{TP + FP} \quad (17)$$

$$Recall = \frac{TP}{TP + FN} \quad (18)$$

$$Specificity = \frac{TN}{TN + FP} \quad (19)$$

$$F1 = \frac{2 * Precision * Recall}{Precision + Recall} \quad (20)$$

For the decoding accuracy of the classification model in Table 2, it is shown that CAMNN model can achieve the highest accuracy of 92%, followed by deep ConvNets (80%), EEGNet (76%), CNN (72%), LSTM (60%), and RNN (52%). This indicates that the CAMNN model has the largest

FIGURE 17. Graph of train and validation loss function.

number of correct predictions for the input EEG samples. Figure 17 shows the training loss and validation loss of the model. The loss graph shows that the training process of the CAMNN model converges well without any signs of over-fitting, indicating that the model having good generalization capability.

For the precision in Table 2, it shows that CAMNN has the highest value of 0.93, followed by deep ConvNets (0.83), EEGNet (0.82), CNN (0.75), RNN (0.6), and LSTM (0.54). Precision is used to quantify the number of predicted classes that belong to the actual attentive classes. We found that CAMNN has the precision value close to 1.0. This indicate that the number of predicted attentive classes using the CAMNN model is close to number of actual attentive classes. For the recall in Table 2, it shows that CAMNN has the highest value of 0.92, followed by LSTM (0.89), CNN (0.77), deep ConvNets (0.77), EEGNet (0.69), and RNN (0.46). Recall is used to quantify the predicted number of attentive classes among all attentive classes in the dataset. CNN and ConvNet seem to have same recall value, which mean that both models predict the same number of attentive classes. As for the CAMNN, it has a recall value close to 1.0, which indicates that it predicted the highest number of

attentive classes. For the specificity in Table 2, it shows that CAMNN has the highest value of 0.9, followed by deep ConvNets (0.83), EEGNet (0.83), CNN (0.76), RNN (0.64), and LSTM (0.53). Specificity is used to quantify the predicted number of inattentive classes among all inattentive classes in the dataset. It shows that deep ConvNets and EEGNet have same specificity value, which mean that both models predict the same number of inattentive classes. As for the CAMNN, it has specificity value close to 1.0, which mean that it predicted the highest number of inattentive classes. For the F1-score in Table 2, it is a harmonic mean of precision and recall. Since the CAMNN has the highest precision and recall value, the F1-score undoubtedly has the highest value of 0.92, followed by deep ConvNets (0.8), CNN (0.76), EEG-Net (0.75), LSTM (0.67), and RNN (0.52). Higher F1-score indicated that the model has the higher accuracy on the dataset.

Based on the results of the evaluation metrics, the CAMNN model proved to have good generalization ability as it had the highest accuracy, precision, recall, specificity, and F1-score when classifying the input EEG signals compared to other models.

V. CONCLUSION

In this paper, comprehensive research on the use of electroencephalogram (EEG) in measuring human attention level has been conducted. The designed Troxler's fading experiment can effectively elicit the participant's attentive and inattentive behaviors during the control and experimental tests. Furthermore, the use of a heat map on visualizing the participant's eye position at the screen during the test has been successfully implemented. We also constructed an Event-Related Potential (ERP) graph to compare the participants' brain responses in attentive and inattentive behaviors. In addition, we have introduced the use of deep learning model in analyzing participants' EEG signals. The proposed model is called Convolution Attention Memory Neural Network (CAMNN) which uses Vector-to-Vector (Vec2Vec) modeling to learn in an end-to-end approach. We have shown that the proposed CAMNN model outperforms the other 5 existing models with good generalization. In CAMNN method, we use three-channels data to build the classification models and obtain 92% accuracy, 0.93 precision, 0.92 recall, 0.9 specificity, and 0.92 F1-score. Currently, we use the EEG signals of three channels located in the frontal cortex (AF3 and AF4) and the parietal cortex (Pz). We suggest using more EEG channels in frontal and parietal cortex as it may provide more features to improve the model for classification. Nevertheless, fewer EEG channels takes less time to set up and is convenient for the participants. This study shows that the proposed CAMNN model can effectively analyze the participants' EEG signals through the deep learning model and classify their attention level. Therefore, this research can be beneficial for those who are interested in developing attention level monitoring or biofeedback system in areas such as educational classroom learning, medical research, and industrial operator

REFERENCES

[1] F. E. Gunawan, K. Wanandi, B. Soewito, S. Candra, and N. Sekishita, "Detecting the early drop of attention using EEG signal," in *Proc. 4th Int. Conf. Electr. Eng., Comput. Sci. Informat. (EECSI)*, Sep. 2017, pp. 1–6, doi: 10.1109/EECSI.2017.8239175.

[2] M. Ross, C. A. Graves, J. W. Campbell, and J. H. Kim, "Using support vector machines to classify student attentiveness for the development of personalized learning systems," in *Proc. 12th Int. Conf. Mach. Learn. Appl.*, Dec. 2013, pp. 325–328, doi: 10.1109/ICMLA.2013.66.

[3] V. Krüger, U. Mahlmeister, and G. Sommer, "Attentive face detection and recognition," in *Mustererkennung 1998* (Informatik aktuell), P. Levi, M. Schanz, R. J. Ahlers, and F. May, Eds. Berlin, Germany: Springer, 1998, pp. 279–286, doi: 10.1007/978-3-642-72282-0_30.

[4] Y.-L. Kuo, J.-S. Lee, and M.-C. Hsieh, "Video-based eye tracking to detect the attention shift: A computer classroom context-aware system," *Int. J. Distance Educ. Technol.*, vol. 12, no. 4, pp. 66–81, Oct. 2014, doi: 10.4018/ijdet.2014100105.

[5] L.-W. Ko, O. Komarov, W. D. Hairston, T.-P. Jung, and C.-T. Lin, "Sustained attention in real classroom settings: An EEG study," *Frontiers Hum. Neurosci.*, vol. 11, pp. 1–10, Jul. 2017, doi: 10.3389/fnhum.2017.00388.

[6] A. List, M. D. Rosenberg, A. Sherman, and M. Esterman, "Pattern classification of EEG signals reveals perceptual and attentional states," *PLoS ONE*, vol. 12, no. 4, pp. 1–23, 2017, doi: 10.1371/journal.pone.0176349.

[7] N.-H. Liu, C.-Y. Chiang, and H.-C. Chu, "Recognizing the degree of human attention using EEG signals from mobile sensors," *Sensors*, vol. 13, no. 8, pp. 10273–10286, Aug. 2013, doi: 10.3390/s130810273.

[8] C.-J. Peng, Y.-C. Chen, C.-C. Chen, S.-J. Chen, B. Cagneau, and L. Chassagne, "An EEG-based attentiveness recognition system using Hilbert–Huang transform and support vector machine," *J. Med. Biol. Eng.*, vol. 40, no. 2, pp. 230–238, Apr. 2020, doi: 10.1007/s40846-019-00500-y.

[9] M. R. Zare, D. O. Alebiosu, and S. L. Lee, "Comparison of handcrafted features and deep learning in classification of medical X-ray images," in *Proc. 4th Int. Conf. Inf. Retr. Knowl. Manage. (CAMP)*, Mar. 2018, pp. 1–5, doi: 10.1109/INFRKM.2018.8464688.

[10] E. Raczko and B. Zagajewski, "Comparison of support vector machine, random forest and neural network classifiers for tree species classification on airborne hyperspectral APEX images," *Eur. J. Remote Sens.*, vol. 50, no. 1, pp. 144–154, Jan. 2017, doi: 10.1080/22797254.2017.1299557.

[11] S.-M. Yang, C.-M. Chen, and C.-M. Yu, "Assessing the attention levels of students by using a novel attention aware system based on brainwave signals," in *Proc. IIAI 4th Int. Congr. Adv. Appl. Informat.*, Jul. 2015, pp. 379–384, doi: 10.1109/IIAI-AAI.2015.224.

[12] A. Y. Shestyuk, K. Kasinathan, V. Karapoondinott, R. T. Knight, and R. Gurumoorthy, "Individual EEG Measures of Attention, Memory, and Motivation Predict Population Level TV Viewership and Twitter Engagement," *PLoS ONE*, vol. 14, no. 3, pp. 1–27, 2019, doi: 10.1371/journal.pone.0214507.

[13] X. Li, Q. Zhao, L. Liu, H. Peng, Y. Qi, C. Mao, Z. Fang, Q. Liu, and B. Hu, "Improve affective learning with EEG approach," *Comput. Informat.*, vol. 29, no. 4, pp. 557–570, 2010.

[14] M. Alirezaei and S. Hajipour Sardouie, "Detection of human attention using EEG signals," in *Proc. 24th Nat. 2nd Int. Iranian Conf. Biomed. Eng. (ICBME)*, Nov. 2017, pp. 1–5, doi: 10.1109/ICBME.2017.8430244.

[15] R. Abiri, S. Borhani, Y. Jiang, and X. Zhao, "Decoding attentional state to faces and scenes using EEG brainwaves," *Complexity*, vol. 2019, pp. 1–10, Feb. 2019, doi: 10.1155/2019/6862031.

[16] J. Kim, J. Seo, and T. H. Laine, "Detecting boredom from eye gaze and EEG," *Biomed. Signal Process. Control*, vol. 46, pp. 302–313, Sep. 2018, doi: 10.1016/j.bspc.2018.05.034.

[17] X. An, D. Kuang, X. Guo, Y. Zhao, and L. He, "A deep learning method for classification of EEG data based on motor imagery," in *Intelligent Computing in Bioinformatics—ICIC* (Lecture Notes in Computer Science), vol. 8590, D. S. Huang, K. Han, and M. Gromiha, Eds. Cham, Switzerland: Springer, 2014, pp. 203–210, doi: 10.1007/978-3-319-09330-7_25.

[18] M. M. Hasib, T. Nayak, and Y. Huang, "A hierarchical LSTM model with attention for modeling EEG non-stationarity for human decision prediction," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Mar. 2018, pp. 104–107.

[19] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball, "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017, doi: 10.1002/hbm.23730.

[20] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013, doi: 10.1088/1741-2552/aace8c.

[21] Y. Kim and A. Choi, "Eeg-based emotion classification using long short-term memory network with attention mechanism," *Sensors*, vol. 20, no. 23, pp. 1–22, 2020, doi: 10.3390/s20236727.

[22] M. S. Mennemeier, A. Chatterjee, R. T. Watson, E. Wertman, L. P. Carter, and K. M. Heilman, "Contributions of the parietal and frontal lobes to sustained attention and habituation," *Neuropsychologia*, vol. 32, no. 6, pp. 703–716, 1994, doi: 10.1016/0028-3932(94)90030-2.

[23] J. Massa and I. H. O'Desky, "Impaired visual habituation in adults with ADHD," *J. Attention Disorders*, vol. 16, no. 7, pp. 553–561, Oct. 2012, doi: 10.1177/1087054711423621.

[24] E. M. Jansiewicz, C. J. Newschaffer, M. B. Denckla, and S. H. Mostofsky, "Impaired habituation in children with attention deficit hyperactivity disorder," *Cognit. Behav. Neurol.*, vol. 17, no. 1, pp. 1–8, Mar. 2004, doi: 10.1097/00146965-200403000-00001.

[25] M. S. Ijjada, H. Thapliyal, A. Caban-Holt, and H. R. Arabnia, "Evaluation of wearable head set devices in older adult populations for research," in *Proc. Int. Conf. Comput. Sci. Comput. Intell. (CSCI)*, Dec. 2015, pp. 810–811, doi: 10.1109/CSCI.2015.158.

[26] K. Semmelmann and S. Weigelt, "Online webcam-based eye tracking in cognitive science: A first look," *Behav. Res. Methods*, vol. 50, no. 2, pp. 451–465, Apr. 2018, doi: 10.3758/s13428-017-0913-7.

[27] G. Szucs and D. Huszti, "Seq2seq deep learning method for summary generation by LSTM with two-way encoder and beam search decoder," in *Proc. IEEE 17th Int. Symp. Intell. Syst. Informat. (SISY)*, Sep. 2019, pp. 221–226, doi: 10.1109/SISY47553.2019.9111502.

[28] S. Hwang, G. Jeon, J. Jeong, and J. Lee, "A novel time series based Seq2Seq model for temperature prediction in firing furnace process," *Procedia Comput. Sci.*, vol. 155, pp. 19–26, Jan. 2019, doi: 10.1016/j.procs.2019.08.007.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[30] B. Li and Y. He, "An improved ResNet based on the adjustable short-cut connections," *IEEE Access*, vol. 6, pp. 18967–18974, 2018, doi: 10.1109/ACCESS.2018.2814605.

[31] H. Ide and T. Kurita, "Improvement of learning for CNN with ReLU activation by sparse regularization," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, May 2017, pp. 2684–2691, doi: 10.1109/IJCNN.2017.7966185.

[32] E. R. S. de Rezende, G. C. S. Ruppert, A. Theóphilo, E. K. Tokuda, and T. Carvalho, "Exposing computer generated images by using deep convolutional neural networks," *Signal Process., Image Commun.*, vol. 66, pp. 113–126, Aug. 2018, doi: 10.1016/j.image.2018.04.006.

[33] K. Xu, J. Ba, R. Kiros, K. Cho, A. Courville, R. Salakhudinov, R. Zemel, and Y. Bengio, "Show, attend and tell: Neural image caption generation with visual attention," in *Proc. 32nd Int. Conf. Mach. Learn., (ICML)*, vol. 3, 2015, pp. 2048–2057.

[34] Y. Chu, X. Yue, L. Yu, M. Sergei, and Z. Wang, "Automatic image captioning based on ResNet50 and LSTM with soft attention," *Wireless Commun. Mobile Comput.*, vol. 2020, pp. 1–7, Oct. 2020, doi: 10.1155/2020/8909458.

[35] P. Nagabushanam, S. Thomas George, and S. Radha, "EEG signal classification using LSTM and improved neural network algorithms," *Soft Comput.*, vol. 24, no. 13, pp. 9981–10003, Jul. 2020, doi: 10.1007/s00500-019-04515-0.

[36] N. C. Tay, C. Tee, T. S. Ong, and P. S. Teh, "Abnormal behavior recognition using CNN-LSTM with attention mechanism," in *Proc. 1st Int. Conf. Electr., Control Instrum. Eng. (ICECIE)*, Nov. 2019, pp. 1–4, doi: 10.1109/ICECIE47765.2019.8974824.

[37] P. Baldi and P. Sadowski, "The dropout learning algorithm," *Artif. Intell.*, vol. 210, pp. 78–122, May 2014, doi: 10.1016/j.artint.2014.02.004.

[38] L. Landa, Z. Krpoun, M. Kolarova, and T. Kasparek, "Event-related potentials and their applications," *ANS J. Neurocognitive Res.*, vol. 56, no. 1, pp. 17–23, 2014.

[39] E. M. Sokhadze, M. F. Casanova, E. L. Casanova, E. Lamina, D. P. Kelly, and I. Khachidze, "Event-related potentials (ERP) in cognitive neuroscience research and applications," *NeuroRegulation*, vol. 4, no. 1, pp. 14–27, Mar. 2017, doi: 10.15540/nr.4.1.14.

CHEAN KHIM TOA received the Bachelor of Engineering degree (Hons.) in electronics with a focus on robotics and automation and the Master of Engineering Science degree (by Research) from Multimedia University, Melaka, Malaysia, in 2017 and 2020, respectively, where he is currently pursuing the Doctor of Philosophy (Ph.D.) degree in engineering (by Research). His current research interests include deep learning, biomedical, image processing, virtual/augmented/mixed reality, latex glove protein estimation, and physiological measurement. He is also a member of the Center of e-Health. He received several awards, including the 1ST Runner-Up in Infineon Week 2017, the Gold Medal Award in Research Innovation Commercialization and Entrepreneurship Showcase (RICES) 2018, the International Championship in World Summit on the Information Society (WSIS) Prizes 2019, the Gold Medal Award in International Invention Innovation Competition (ICAN) 2020, and the Humanizing Innovation Award in RICES 2020.

KOK SWEE SIM (Senior Member, IEEE) is currently a Professor with Multimedia University, Melaka, Malaysia. He is also working closely with various local and overseas institutions and hospitals. He has filed 18 patents and more than 70 copyrights. He has received many international and national awards, such as the Academy of Sciences Malaysia (ASM) as the Top Research Scientists Malaysia (TRSM); the Korean innovation and special awards, in 2013, 2014, and 2015; the 2005, 2006, and 2011 World Conference in Applied Computing (USA); and the 2008 IEEE Conference at U.K. For national level achievements, he received the Gold Medal Award in the Invention, Innovative Technology Exhibition (ITEX) 2008, 2009, 2010, 2013, and 2014; the Bio Malaysia Award, in 2009 and 2010; the Malaysia Technology Expo 2011; AIK 2011; AIK 2012; and the Gold Medal Award in APICTA, in 2014 and 2015. He received the MMU Best Staff Award in 2009, 2010, and 2015. He was a recipient of the Japan Society for the Promotion of Science (JSPS) Fellowship, Japan, in 2018. He received the TM Kristal Award, in 2016, and the International Championships in World Summit on the Information Society (WSIS) Prizes, in 2016, 2017, 2018, 2019, and 2020. These awards were in the areas of biomedical Engineering, such as breast cancer detection, brain for early infarct detection, and latex glove protein detection.

SHING CHIANG TAN received the B.Tech. (Hons.) and M.Sc.(Eng.) degrees from the Universiti Sains Malaysia, Malaysia, in 1999 and 2002, respectively, and the Ph.D. degree from Multimedia University, Melaka, Malaysia, in 2008. He is currently an Associate Professor with the Faculty of Information Science and Technology, Multimedia University. His current research interests include computational intelligence (artificial neural networks, evolutionary algorithms, fuzzy logic, and decision trees), deep learning and their applications, data classification, condition monitoring, fault detection and diagnosis, stroke rehabilitation, virtual/augmented/mixed reality, and biomedical disease classification and optimization. He was a recipient of the Matsumae International Foundation Fellowship, Japan, in 2010.

●●●