

Deep Learning Based on CNN for Emotion Recognition Using EEG Signal

ISAH SALIM AHMAD, SHUAI ZHANG, SANI SAMINU, LINGYUE WANG, ABD EL KADER ISSELMOU,
ZILIANG CAI, IMRAN JAVAID, SOUHA KAMHI, UMMAY KULSUM

State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of
Technology, Tianjin, 300130, P.R. CHINA.

Abstract— Emotion recognition based on brain-computer interface (BCI) has attracted important research attention despite its difficulty. It plays a vital role in human cognition and helps in making the decision. Many researchers use electroencephalograms (EEG) signals to study emotion because of its easy and convenient. Deep learning has been employed for the emotion recognition system. It recognizes emotion into single or multi-models, with visual or music stimuli shown on a screen. In this article, the convolutional neural network (CNN) model is introduced to simultaneously learn the feature and recognize the emotion of positive, neutral, and negative states of pure EEG signals single model based on the SJTU emotion EEG dataset (SEED) with ResNet50 and Adam optimizer. The dataset is shuffle, divided into training and testing, and then fed to the CNN model. The negative emotion has the highest accuracy of 94.86% fellow by neutral emotion with 94.29% and positive emotion with 93.25% respectively. With average accuracy of 94.13%. The results showed excellent classification ability of the model and can improve emotion recognition.

Keywords— BCI, EEG, Deep learning (DL), CNN, ResNet50, Emotion recognition, Adam optimizer, SEED.

Received: February 2, 2021. Revised: April 7, 2021. Accepted: April 8, 2021. Published: April 14, 2021.

1. Introduction

A brain-computer interface (BCI) is a system that translates brain activities which are directly connected to the brain of a living organisms such as human and animal. BCI acts as an interface between the human brain and external devices for various applications, such as emotion recognition[1]. BCI had been extremely important in Biomedical engineering research.

Emotional expression has a major role in human communication, daily life[2], and work. It can be defined as positive, neutral, and negative experiences from several physiological activities, and there are many classes of emotions like sadness, happiness, anger, surprise, disgust, fear[3]-[4], and so on.

With the enhancement in the development and availability of wearable sensor technology, research on emotion recognition have become more common among researchers because emotion recognition can bring the important application, whether at professional, personal or social level[5], such as in the field of Medicine[6] education, psychology, computer game, driving, security, entertainment[7], workload estimation[8] among others.

There are several ways to identify emotion such as brain wave and facial expression. Brain signals are a possible method to obtained human emotion, it can be an invasive or noninvasive method. In an invasive BCI, electrodes are inserted on the exposed surface of a brain using surgical operation. This method is not allowed for the human being, while the noninvasive method is referred to BCI, it gives an easy, convenient, and favorable method to collect the brain signal, which includes Magnetoencephalogram (MEG), functional magnetic resonance imaging (fMRI), Electroencephalogram (EEG).

In the practical application of emotion recognition, many signals had been approved, adopted, and approximately classified into non-physiological and physiological signals. A

non-physiological signals are the early work, such as speech, gesture, facial expression, movement, voice intonation, text among others is mostly used in earlier work.

Recently, more research is still carried out based on physiological signals, such as EEG, electrocardiogram (ECG)[9]-[10], pupillary diameter (PD), and electromyogram (EMG)[11]-[12] which are more effective and reliable. EEG record brain activity [13] in the central nervous system via electrodes placed on the scalp, and gives useful characteristics in responses to the emotional states [14].

The EEG Signal is widely used by neurologists for analysis and diagnosing various brain disorders such as; seizure detection[15] autism, attention deficit, and game addiction[16] Emotion is an idea, feeling, or conscious experiences in which people are faced with internal or external stimuli, Emotion has an important role in natural communication among human being and others living.

Recently, the research in human-computer interaction (HCI) enhances the system's capacity to detect, process, and respond to user emotional states[17]. Emotional intelligence (EI) is an essential aspect of human intelligence. EI simply refers to the capacity to recognized and perceive emotional intelligence. It is also called affective computing in computer technology[18]. Affect is generally defined as spontaneous mental feeling or state which has a major role in our daily life. In the case of human psychology, affect sensing was classified as; physiological recordings, self-reports, and behavioral observation. Physiological recording refers to an implicit method to recognize emotional reactions by identifying the user's physiological changes by using biosensors. While self-reporting refers to an explicit method to acquired information concerning a person's emotions or feelings using questionnaires as well as interviews to get one's state. Behavioral observations refer to the method employ to identify one's emotional state by observing their externalized reactions like facial expression and speech[19].

The eye movement analysis of children with autism has a major role in emotion recognition, electrical potential is induced or generated as a result of the shifting the eyeballs in the surrounding region of the eye[13]. Research on psychology related to emotions and the brain reveals a strong influence of cognitive processes in emotions.

As a result of that, the EEG signals have important information about one's emotions[20]. It recorded using active electrodes placed according to the international 10-20 system. Voltage fluctuations from the cortex in the brain, and can be measured or detected by EEG signal, can display vital information about the human emotional state[21] and well known for measuring brain activity[15].

The recent development of the EEG signal for emotion recognitions gives strong interest to researchers[22]. EEG signal gives the cheap, portable, easy solution in classification and identifying emotions. The strong correlation between EEG signals and emotion exists. The two major regions of the brain that is related to emotional activities is the amygdala and frontal lobe. Research indicated that the frontal scalp store more emotional activation as compared to other areas of the brain like the temporal, and parietal occipital[23]. A common

physical signal that is extremely selected for emotion analysis is the EEG[24]. EEG feature Extractor and classifiers are two basic units in the ML approach, it can be a single-channel or multi-channel features[25].

Emotion stimuli are chosen to cover specific desire levels as well as valence states that are available in the different structural forms such as visual, tactile, auditory, and odor. Emotional state produce by a stimulus is secure by using standard stimulus sets, like international affective picture system (IAPs) and international affective digitized sound system (IADs). IAPs give a list of standard information for emotional stimuli to produce attention levels and emotional changes, while IADs express acoustic stimuli to produces emotions, sometimes, IADs work together with IAPs[19]. The discrete model and dimensional model are used as a reference to help researchers to build an emotion recognition system, the discrete model deal with evolutionary features[24] or discrete emotions such as anger, sadness, fear, joy, and disgust. While the dimensional model conveyed the emotion of provoking people, and has two main types of dimensions arousal and valence or arousal, pleasure, and dominance. Arousal shows the strong level of emotion[20]. Valence is the positive and negative characteristics. While dominance is the mirror of one's status being control or in control[26]. Emotion recognition has an essential effect in both engineering and scientific research[27]. Emotion recognition based on single modality, the EEG signals have been generally used to develop efficient brain computer interaction system for analysis of both internal emotion and cognitive states[28] Recently, DL had been used for emotion recognition[8] and still at its infancy stage, because parameter settings and model structure can still be adjusted[29]. CNN and DNN[30]-[31]. Single modality achieved a remarkable results.[25] The objectives of this article are to recognize positive, neutral and negative emotion from SJTU emotion EEG dataset (SEED) dataset based on deep convolutional neural network (CNN) using residual neural network (ResNet50) and Adam optimizer. This paper is organized as follows, section I introduction the background about Emotion recognition with the cognitive system. Section II Brief review of related work Section III Experiments. Section IV Methodology, Section V Result and Discussion, and Section VI Conclusion and future work.

2. Related Work

This section briefly described the recent work in the field of emotion recognition, and several approaches had been employed for emotion recognition to identified and classified from EEG signals if appropriate stimuli were applied, automatic recognition is sometimes limited to some emotions classes as a result of signal's feature, EEG constraint, noise, as well as the subject dependent problems[24]. The classifier and feature extractors are the two basic component in ML. The EEG feature can be either single-channel or multi-channel feature, while EEG classifier can be either topology invariant or topology- aware[25]-[32]

Two classifiers were trained with a multimodal Emotional dataset called SEED-V for EEG, Eye, and feature level fusion (FLF) with average accuracies of 70.8%, 59.87%, and 75.13%

respectively [8]. Several feature extractors such as power spectral density (PSD), Differential Caudality (DCAU), asymmetry (ASM), rational asymmetry (RASM), DE, differential asymmetry (DASM), and linear dynamic system (LDS) are used as feature smoothing, while k nearest neighbors (KNN), logistic regression (LR), SVM, discriminative Graph regularized extreme learning machine (GELM) as classifiers were proposed for emotion classification on DEAP and SEED dataset. GELM with DE features achieves the best accuracy of 69.675% and 91.07% on DEAP and SEED datasets respectively, neural patterns were relatively stable among the sessions[33].

Multilayer weight sharing multimodal residual network for emotion recognition (sadness, happiness, fear, and neutral) under sleep deprivation was proposed[13], the method combined EEG single-channel, DE features, EEG functional strength features with topological correlation connectivity, and eye movement features with accuracies of 86.86% and 82.03% in subject dependent and cross-subject emotion recognition task for 30 hours sleep deprivation respectively[34]. Feature level fusion strategy and decision level fusion level strategy was proposed for emotion recognition with 15 emotional film clip of three categories (positive, neutral, and negative) of EEG signals and pupillary response collected from the eye tracker, the accuracy based on EEG and eye-tracking data are 71.77% and 73.98% respectively, fusion strategy and decision level strategy achieved an accuracy of 73.59% and 72.98% respectively[35].

Independent component analysis (ICA) algorithm was proposed to remove artifact and extract independent components, the channels were then selected based on threshold average activity value for obtaining components, and then applied modified Dempster-Shafer theory of evidence with the accuracy of 91%[36]. SVM, KNN, and RF classifier were proposed to classified the emotion of 360° videos presented using virtual reality (VR) headset as stimuli. SVM has the best performance with an accuracy of 57.05% for identifying a particular class KNN was found to be best with the accuracy of 70.83%[15].

Improved radial basis function neural network (I-RBF-NN) algorithm was used to classify and recognize the emotional signal and the result was compared with SVM, RBF-NN, GMM classifier, it was found out that I-RBF-NN algorithm for emotion recognition of EEG was better[3]. Relative power values and Bayesian network was proposed for emotion recognition with EEG signals, PSD divide EEG signal into five frequency ranges, artifact were removed by removing low frequency from 0-4Hz and the remaining range frequency was based on percentage selected range, and then compared to Bayesian network, the result indicated that emotion was human face avatar[37].

Multidimensional information in empirical mode decomposition(EMD) was proposed for emotion recognition from EEG signals, intrinsic mode functions(IMFs) were used for emotion signals decomposition automatically, these features were effective for emotion recognition[12]. Multimodal approximation for valence classification was used for body and brain emotional responses, EEG signal, ECG

signals, and skin temperature of 24 subjects were used, emotions classification was possible at central as well as peripheral levels[16]. Mel Frequency Cepstral coefficient (MFCC) was proposed with 13 acceleration and velocity component as feature extractor, CNN as well as LSTM method was used for emotion classification and found the accuracy of 80%[19]. DE base emotion recognition was proposed and compared with other symmetrical electrodes (RASM, DASM) DE was found to be more suited with the accuracy of 84.22%[38]. The significance test/sequential backward selection and support vector machine (ST-SBSSVM) was proposed based on high dimension feature as a result of extracting multiple features as well as formation high dimensional feature on SEED and DEEP dataset, the accuracy was found to be 72% and 89% on DEAP and SEED respectively[39].

Dynamical Graph convolutional neural network (DGCNN) was proposed for emotion recognition on SEED and DREAMER dataset, the classification accuracy on SEED is 90.4% and 79.95% for subject dependent and independent respectively, while on DREAMER, The accuracies were 86.23%, 84.54%, and 85.02% respectively[40]. DNN based Bag of deep features (BoDF) Model was proposed on SEED and DEAP datasets, the BoDF model performance was 93.8% on SEED and 77.4% on DEAP as compared with the traditional method for human emotion[41]. MST, QDA, SVM, ST were proposed using video as stimuli for Arousal and Valence with 88.23% and 81.255% respectively[5]. PS Bayes, Power different of the EEG waves were used for arousal, valence and obtained the accuracy of 62% and 57.6% [42]. SVM, K-means extract, FFT Spectra, and Power Spectrum analysis were proposed with IAPS as stimuli for arousal, valence with accuracy 70%, and 71.85%[43]. KNN, DT, Bagging predictor, Correlation, and regression Brain wave were proposed using video as stimuli for pleasure, arousal, and dominance and obtained an accuracy of 58.54% to 75.16% [44]. KNN, LR, and GELM were proposed using video as stimuli for recognition of positive, negative, and neutral with the accuracy of 69.67% and 91.07% [45]. KNN and LDA was proposed for Disgust, Happy, Surprise, Fear, and Neutral using IAPS as stimuli with the accuracy of 83.26% and 75.21%[46] respectively. SVM, MLP, STFT were used for emotion recognition of joy, anger, sadness, and pleasure using IAPS as stimuli with the accuracy of $82.29\% \pm 3.06\%$ [47].

QDC, SVM KNN, and statistical analysis were used for the emotion of positive, negative, and neutral with IAPS as stimuli obtained an accuracy of 82%[48]. SVM method with IAPS as a stimulus was used for emotion recognition of valence and arousal with the accuracy of 62.58% and 94.40% [49]. HMMs, SM, RBF, power spectral features were proposed for emotion recognition of arousal, valence, dominance with IAPS as stimuli with the accuracy of 62% and 73% [50]. SVM, LDS, PSW, and dynamical analysis were proposed for the emotion of positive and negative using IAPS as a stimulus with the accuracy of 91.77% [51]. QDA, ML, ERP, Time-domain, HOC, HOS, and HHS were proposed using IAPS as stimuli for valence, arousal, and dominance with the accuracy of 25% to 47.5%[32]. Binary fisher linear classifier, KNN, LDA, SVM,

ANN was used for negative and positive emotion with an accuracy of 95.5781% [52].

CNN, SVM, IC, Pre-trained CNN models were used for neutral, negative, and positive with music and video as a stimulus with an accuracy of 86.56% and 78.34% for SEED, 72.81% for DEAP, and 26 81.8%, for LUMED respectively [53]. LSTM-RNN, SAE, PSD, Band Power Feature for valence and arousal using IAPS as a stimulus with the accuracy of 81.10% for valence and 74.38% for arousal [54]. KNN, DWT was proposed for valence, arousal, and dominance using video as stimuli with the accuracy of 89.54%, 92.28%, and 93.72%, respectively [55]. KNN, SVM, and DWT were used for valence, arousal, and dominance with the accuracy of 86.75% and 84.05 % [56]. SVM, RBF, and mRMR were proposed to use video as stimuli for valence and arousal with the accuracy of 60.7% and 62.33%[24]. S-ELM-LUPI, EOH, LDN, and LBP were used for valence and arousal with an accuracy of 80%[57]. WT, FT based, HHT PSD, and wavelet were used for valence and arousal[58]. 1/f fluctuation theory, KNN, SVM, was proposed for pleasure, arousal, and dominance using music as stimuli with an accuracy of 75%[59].

FIOA, SVM, KNN, DT, RF, Hjorth, FIOA, PSO binary, and FA binary were used for positive and negative using music as stimuli[60]. PNN, PSO, GA, ACO, SA, and Time-frequency were proposed using video and music as stimuli for arousal and valence with an accuracy of $67.5 \pm 3.4\%$ for DEAP[17]. KNN, DWT was proposed for the emotion of neutral, happy, sad, anger using music as stimuli with an accuracy of 72.05%, 66.05%, and 71.25%[61]. KNN, SVM, BN, ANN, Fourier transform were proposed for valence, arousal, and dominance with an accuracy of 77.78% [62]. NN, HRV, GSR were used for valence and arousal with video as stimuli with the accuracy of 80.2%[63]. KNN, GSR, and PPG were proposed for the emotion of Happy, Angry, Sad, and Relaxed using video as stimuli with the accuracy of 79.76% [4].

KNN, SVM DBN, LR, and DE were proposed for the emotion of positive, neutral, and negative with the accuracy of 86.08%, 83.99%, 82.70%, and 72.60% for DBN, SVM, LR, and KNN respectively [31]. DT, KNN, and RF feature tensor, Autoencoder were proposed for the arousal, valence, and dominance with an accuracy of 63.09% for DEAP and 75% for SEED [64]. SVM, LK, PSD, and DE were proposed for recognition of happy, sad, fear, and neutral with the accuracy of 85.11% and 72.39% [2]. An approach is necessarily needed to improve emotion recognition [21]. This article aims to improve the previous research on emotion recognition.

3. Experiments

3.1 SEED Dataset

This study was carried out on the publically available emotion dataset in BCMI laboratory (<http://bcmi.sjtu.edu.cn/~seed/>) at Shanghai Jiao Tong University (SJTU). Fifteen (15) video clips (positive, neutral, and negative emotions) were selected and used as stimuli in the experiment. Each subject participating in SEED underwent experiments for three

sessions with a one-week interval. The EEG signal of fifteen (15) Chinese participants (seven 7 males and eight 8 females, with the average age of 23.27 and standard deviation of 2.37) was recorded using an EEG cap based on a 10-20 international system with 62 channels. Each stimuli session is composed of five (5) second of the hint of the movie, four (4) minutes of the clip, forty-five second (45s) of self-assessment as well as fifteen-second (15s) of rest as shown in figure1. All the participants had three (3) sessions on different days, which is forty-five (45) sessions of recording the EEG signal. The labels are assigned according to the contents of the clips as +1 for positive, 0 for neutral, and -1 negative. The EEG signal was collected through 62 channels based on 10-20 system as shown in figure 2, the signal down sampled, preprocessed and segmented version of the EEG data in MATLAB (.mat file) the signal down-sample to 200Hz, band-pass filter was used to filter the signals from 0-75Hz. [38]- [31]

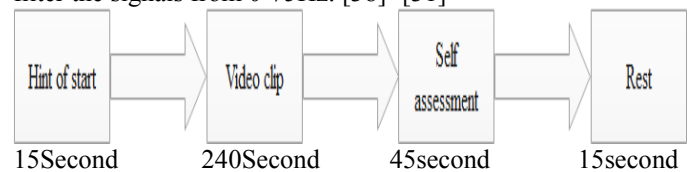


Figure: 1 the detailed procedure of the experiment

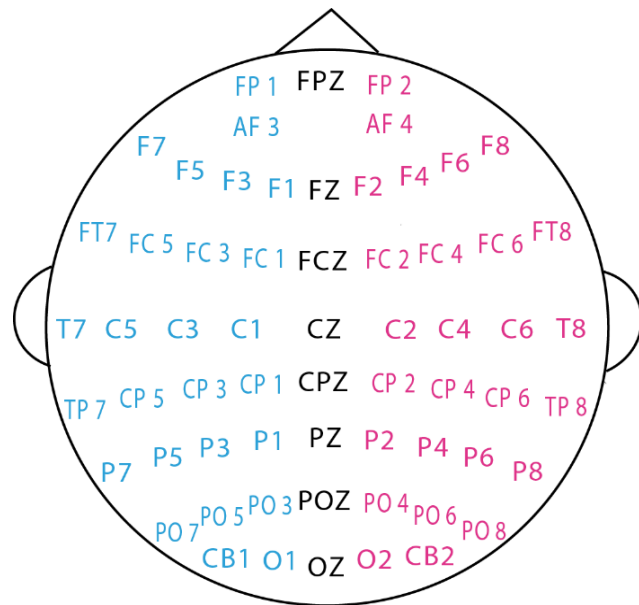


Figure 2: EEG cap of 62 channels based on 10-20 system (SEED Dataset (sjtu.edu.cn))

4. Methodology

4.1 Data processing

The EEG signal is down-sampled to 200Hz. The sample contaminated by ECG[65] EMG and EOG is removed[66] and then filter using a Band-pass filter[67] from the frequency 0-75Hz. The dataset is a shuffle, and then divided into two, 70% for training and 30% for testing. And then fed the signals into the model shown in figure3. The experiment was conducted in

python 3.6 software, Keras in a CPU: i9-9900kf, memory 64GB Ram, and with online GPU to run the algorithms: RTX 2080Ti.

4.2 Emotion classification model based on ResNet50

Figure 3. Example network architectures for emotion classification model based on resnet50. Left: a standard residual network 50, where the input is an EEG signal (62×200) and the output is the classification of three types of emotions (positive, negative and neutral). Right: shows more details and other variants about 2 types of resblocks, there are two parameters (α and β) that determine the type of convolution structure.

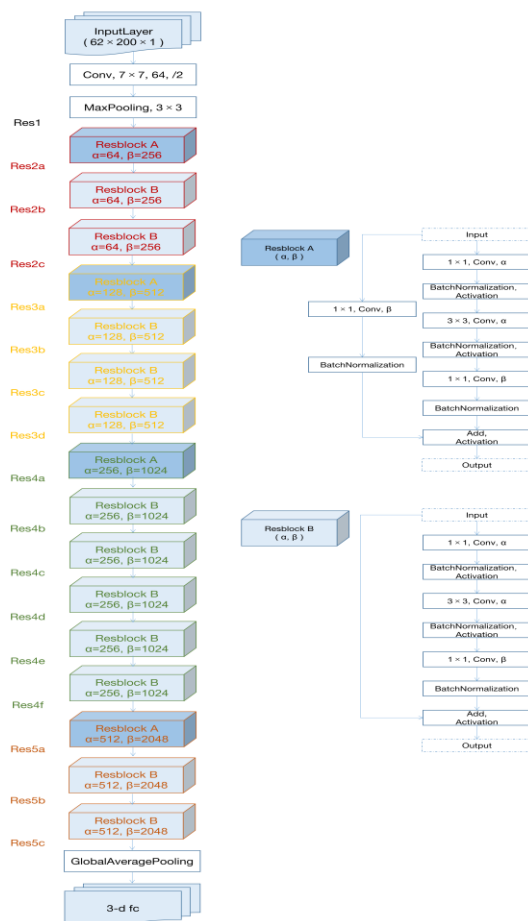


Figure3: Emotion classification model based on ResNet50

4.3 Training steps of the model.

1. Build the CNN model with ResNet50.
2. Shuffle the dataset.
3. Divided the dataset into 70% training and 30% testing.
4. Build the model by the use of tensor flow and Set the model key parameters.
5. Fed the signal into the model and used the tensor board to make sure our model is not overfitted.

6. Calculate the corrective parameter of Adam optimization algorithms to search for the optimal solutions.
7. Train the model and make predictions.
8. Test, and evaluate the model.

4.4 Deep learning (DL).

Deep learning (DL) is a branch of machine learning (ML) in artificial intelligence (AI) that deal with algorithms to imitates the human brain[68], the example DL is CNN, DNN, and RNN[69]. The term deep simply means the number of network layers, the higher the layers the deeper the network. There are 2 to 3 layers with the traditional network and more than a hundred layers with the deep networks. We propose the CNN model to learn the feature, recognize and classify the emotions of pure EEG Signals. Deep learning differs from ordinary methods as we use the CNN model on the pure EEG signal without explicit hand-crafted of feature extraction[70].

4.5 Application of deep learning

EEG signal analysis, Emotion recognition, voice recognition and classification, text interpretation, traffic sign recognition, lane classification, driver assistance, smartphone app, and an ATM. Recently, a various biological model is inspired, such as CNN[67], DNN is giving out a very good performance and accuracy in solving complex tasks as a result of developing fast and cheap hardware.[71] Deep learning has a degree of accuracy compared to other models due to the following properties: mass data can be access easily, a large amount of data can be accessed within a short period, and the model was built by experts. Other areas are; image and video recognition, audio processing and natural language processing, autonomous systems and robotics, medical diagnostics, computational biology, physical science, finance, economics, market analysis, cyber-security, architectural and algorithmic enhancement[72].

4.6 CNN

Convolutional neural network (CNN) is one of the DL algorithms use for learning feature[73]-[64] in emotion recognition[67], it functions well in recognizing emotion, the CNN model is built, simplified in Keras, and it makes CNN layers available. Weight sharing, local sensing field and down sampling are three important characteristics of CNN. The model has 18 layers of convolution, with nonlinear activation function like rectified linear unit (ReLU) or sigmoid function for obtaining recognition accuracy. The CNN nodes of the input layer are usually connected with the node of the hidden layer and also the hidden layer nodes are also completely connected with the nodes of the output layer. In CNN, convolution is usually applied to the input layer to obtain the output. For the sampling of the input from the different filters, the pooling layer is usually used for that. It also reduces the size of the input, CNN has a stack of layers like convolution, activation, and pooling, for converting input dataset to output through special function. ResNet50 is used to evaluate the effectiveness of the CNN model[74].

G. Max pooling layer structure is employed to down sample the feature obtained from convolutional layers, which minimize over-fitting as well as computational process, thus increase the performance of the model.

4.7 ResNet50

ResNet50 is an abbreviation that stands for Residual Network, 50 stands for the number of layers and is a CNN with 50 layers. ResNet50 is just a residue learning model, is a popular deep network model which can gives high accurate result on signal processing. Other ResNet modules include; Resnet18, ResNet101, and ResNet152, it is easy and simple for residue learning to understand. it uses all the features that joint from the boundary point of the nth layer with the $(n+x)$ layer shown in figure 3 [75]. ResNet50 model was build based on the CNN architectures and gives the best accuracy and convergence behaviors to our model. It is designed by many stacked residual units as well as the dissimilar number of layers such as 1202, 152, 101, 50, 34, and 18. Different architecture varies with the operation. ResNet50 is a deeper model and has 49 convolutional layers completely- connected at the end of the network layer. It saves computing resources and training time.[76] ResNet models are in various depths, the accuracy of the models increases with the depth of the network. Moreover, the signal tends to change the weights with increasing depth, this happens at the end of the network by examining the ground-truth and prediction.[77] CNN layers learn high- or low-level features during training, while ResNet learns residuals instead of features.[78] The residual module is shown in figure4.

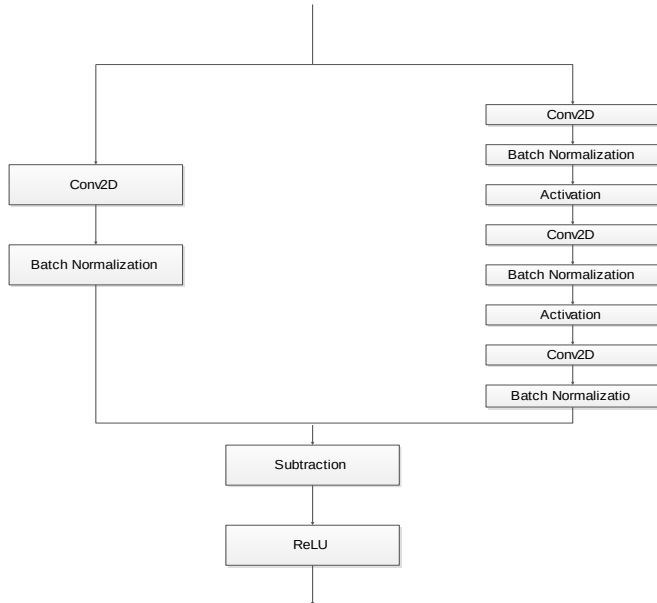


Figure4: Residual module

4.8 2D Convolution

2D Convolution. We assumed that x is a feature inside $\mathbb{R}^{h \times w \times d_1}$, and also K be a filter in $\mathbb{R}^{n \times n \times d_1 d_2}$, and $Y = K * X$

Is a feature in $\mathbb{R}^{h \times w \times d_2}$, with given that $X = Vec(x)$ and $Y = Vec(y)$, and the linear system can be derived as; which is recognized as the forward operation of the 2D convolution of $y = K * x$, And A has a form as given below;

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,d_1} \\ A_{2,1} & A_{2,2} & \cdots & A_{2,d_1} \\ \vdots & \vdots & \ddots & \vdots \\ A_{d_2,1} & A_{d_2,2} & \cdots & A_{d_2,d_1} \end{pmatrix} \quad (1)$$

Given that $A_{i,j} \in \mathbb{R}^{h \times w \times h \times w}$ is the block-wise circulant matrix related to Sub filter $K_{j,i}$ (where $i = 1, 2, \dots, d_1$ and $\eta \sim N(0, \sigma^2)$ is the probability distributions, CX with structure noise (i.e., $X + \epsilon^{X_0}$ with χ_0) as the predicted probability distributions, $j = 1, 2, \dots, d_2$) and the equation $= A_{Is=a^x}$ donates that the stride in the convolution $Y = K * X$ is a.

4.9 Batch Normalization

The weights and parameters varied when the network had been trained, the layer of CNN also change while applying the exact inputs dataset. It may be too difficult to train the model as a result of data in the layer is too small or large, Normalization is used to improve the accuracy as well as training time.[79] The batch size used is 64 and dropout of 0.5

Batch Normalization assumes that $\beta = \{x^{(1)}, x^{(2)}, \dots, x^{(m)}\}$ be a batch of features. Batch normalization β is given as;

$$B = (x^i, \gamma, \beta) = \gamma \frac{\gamma(x^i - \mu)}{\sigma} + \beta \quad i = 1, 2, \dots, M \quad (2)$$

Where $\mu = \sum_{i=1}^m x^i / m$ and $\sigma^2 = \sum_{i=1}^m (x^{(i)} - \mu)^2 / m$

4.10 Padding

Padding as we let x to be a feature in $\mathbb{R}^{h \times w \times d_1}$. Padding operates with parameter $d_2 > d_1$ denoted by E $\mathbb{R}^{h \times w \times d_1} \rightarrow \mathbb{R}^{h \times w \times d_2}$ as defined as $E(Vec(x); d_2) := vec(y)$

Given that y is the feature in $\mathbb{R}^{h \times w \times d_2}$ and we can define each channel as $y_1 \in \mathbb{R}^{h \times w}$ of Y as defined $\mathbb{R}^{h \times w \times d_2}$ of y as defined as;

$$y = \begin{cases} x_i & \text{if } d+1 \leq i \leq d_1, \text{ where } d = \left\lceil \frac{(d_2 - d_1)}{2} \right\rceil \\ 0 & \text{otherwise, } i = 1, 2, \dots, d_2 \end{cases}$$

4.11 ReLU

ReLU simply means rectified linear unit was used as an activation function[73] due to its simplicity of carry out. It helps the number of calculations to be reduced while training process, convergent and is an operation which is applied component-wise to any multi-dimensional feature X.

ReLU has the following function as;

$$F(x) = \max(0, z); \quad (4)$$

$$F(z) = z; \text{ When } z > 0 \quad (5)$$

$$F(z) = \alpha z; \text{ When } z < 0 \quad (6)$$

Hence, α is a constant value,

We used the ReLU activation function as a view by applying a step in the dynamical system which defines the forward propagation. The function only needs a nonlinear activation function so that the model will learn properly. It solves gradient disappearance as well as faster convergent speed its drawback is that it affects neurons not been active due to weakening and dying of gradient during the training of the model. ReLU has the following significant benefit in CNN over conventional, it minimizes the disappearing problem in the CNN model, resolves cancellation effect as well as the implementation of mathematical operations... Assume that:

$I_{\mathbb{R}_+^d}$ represent the indicator function of the set $\mathbb{R}_+^d$ defined as

$$I_{\mathbb{R}_+^d}(x) = \begin{cases} 0, & \text{if } x \in \mathbb{R}_+^d \\ \infty, & \text{if } x \notin \mathbb{R}_+^d \end{cases} \quad (7)$$

The close operator related with $I_{\mathbb{R}_+^d}$ is in fact ReLU that is

$$\text{proximal}_{y/\mathbb{R}_+^d}(x) = \arg \min y I_{\mathbb{R}_+^d}(x) + \frac{1}{2} \|x - y\|_{\ell^2}^2 \quad (8)$$

$$= \text{Proj}_{\mathbb{R}_+^d}(x) = \sigma(x), \quad (9)$$

And it is independent of it $y > 0$.

M. Cross-Entropy (CE) is a loss function that measures the randomness or degree of disorder of a signal[56]. The cross-entropy is used for the analysis and computation of various loss. Category cross-entropy is used. It is also used to measure classification loss as applied it in the method. We assumed that x represents a random variable, $p(x)$ represents the probability function, the entropy x is given by;

$$E(X) = -\sum_X P(X) \log P(X) \quad (10)$$

Hence, t is the ground truth and s assumed to be estimated class distribution, then t_i and s_i is the real and estimated probability at class i .

$$\text{The cross entropy (CE)} = -\sum t_i \log s_i \quad (11)$$

If the difference between s and t is higher, the cross-entropy increase and vice-versa, as a result of that, the model is designed to learn the ground truth.

$$CE = -\sum_{i=0}^n y_i^{\text{label}} \log y_i^{\text{pred}} \quad (12)$$

Where n is the category number and $y^{\text{pred}}, y^{\text{label}}$ are the probability distributions respectively.

4.12 ADAM Optimizer

The most important part during training the model is the optimizer for accelerating or speeding up[67] the model, Adam optimizer is use. Adam simply means Adaptive moment optimization is an algorithm that brings and combines the useful performance of Nesterov momentum, RMS Prop, and AdaGrad algorithms[80]. It adapts a single learning rate and weight remain unchanged during training. The weights are computed and updated based on Adam optimizer equation below at learning rate of 0.001 with 100 epoch and dropout of 0.5.

$$w_t^j = w_{t-1}^j - \frac{\eta}{\sqrt{\hat{v}_t + \epsilon}} \cdot \hat{m}_t \quad (13)$$

1st Moving momentum

$$m_t = (1 - \beta_1) \sum_{i=0}^t \beta_1^{t-i} g_i \quad (14)$$

2nd Moving momentum

$$v_t = (1 - \beta_2) \sum_{i=0}^t \beta_2^{t-i} g_i^2 \quad (15)$$

The bias correction of the momentums

$$m_t = \frac{m_t}{1 - \beta_1^t} \quad (16)$$

$$\tilde{v} = \frac{v_t}{1 - \beta_2^t} \quad (17)$$

By combining the rectification term in equation 13, then rectified Adam can be written as;

$$w_t = w_{t-1} - \eta r_t \frac{\hat{m}_t}{\sqrt{\tilde{v}_t}} \quad (18)$$

The step size η is a variable hyper-parameter, and the rectification rate is obtained using the following as:

$$r_t = \sqrt{\frac{(p_t - 4)(p_t - 2)P_\infty}{(P_\infty - 4)(P_\infty - 2)P_t}} \quad (19)$$

$$\text{Hence } P_t = p_\infty - \frac{2t\beta_2^t}{1-\beta_2^t}, \text{ and also } P_\infty = \frac{2}{1-\beta_2^t} - 1.$$

If the approximated length ≤ 4 , then the variance of adaptive learning rate is deactivated, if not the variance rectification can be calculated and updating the adaptive learning rate.

We use batch normalization to the activation function result of CNN layers and getting normalize as:

$$\hat{x}_i = y w_i + \beta \quad (20)$$

Hence β and γ are the parameter used for scale and shift are learn during training. For mini-match normalization weight is as follow:

$$\hat{w}_t = \frac{w_t - \mu_\beta}{\sqrt{\sigma_\beta^2 - \varepsilon}} \quad (21)$$

And mini-batch mean is obtained using the following:

$$\mu_\beta = \frac{1}{m} \sum_{i=1}^m w_i \quad (22)$$

The variance of the mini-batch can be obtained as:

$$\sigma_\beta^2 = \frac{1}{m} \sum_{i=1}^m (w_i - \mu_\beta)^2 \quad (23)$$

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) G \quad (24)$$

$$V_t = \beta_2 V_{t-1} + (1 - \beta_2) [G]^2 \quad (25)$$

$$G = \nabla_w C(W_t) \quad (26)$$

In summary form, the mathematical form can be simplified as:
Weights = weights - (Momentum and variance combined) (27)

The sigmoid activation function is used for binary classification between 0 and 1, it improves the learning process, is similar to the fundamental assumption of lower input and higher output, it available in python, obtained using the mathematical model bellow with nonlinear s shape output.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (28)$$

Hence, w_t represent the weights at step t , η is the learning rate, C is the cost function, $\nabla_w C(w_t)$ represents the gradient of the weight parameters w_t of the signal x and it is relative to a label y . β_t is the amount of information need for the next update, m_t represent the running average of the gradients, it is called second moment or square moment as $\beta_t \in [0, 1]$. When the first and second moment initialized at zero, they are biased toward it, zero biased problems were

solved by bias-corrected i.e., by dividing them by their respectively β .

4.13 Over-fitting handling

In general, over-fitting means memorizing the training dataset which may cause poor performance of the test dataset. As a result of that, the performance of the training dataset is excellent while that of the test data set is poor. The capacity of the network as well as the nature of the training can affect the model. We use the technique such as; dropout by using regularization and generalization[54] into the model, signal augmentation as well as stopping technique to avoid over fitting[80]. It can be shown in figure 4 and 5 that, the loss drop more quickly this indicates the complexity of the model.

4.14 Confusion matrix

The confusion matrix has been used to ensure acceptable and reliable recognition results, the performance accuracy of the model is evaluated and interpret using the confusion matrix as shown in figure 8, for both testing and training data. The following equations are used obtained the accuracy and average accuracy based on a single model[25].

$$\text{Accuracy} = \left(\frac{\text{number of correct trials}}{\text{total number of trials}} \right) \times 100\% \quad (29)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100\% \quad (30)$$

$$\text{Specificity} = \frac{TN}{TN + FN} \times 100\% \quad (31)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + FN + TN + FP} \quad (32)$$

Where,

$TP = \text{True positive}$

$TN = \text{True Negative}$

$FN = \text{False Negative}$

$FP = \text{False Positive}$

5. Results and Discussion

The SEED dataset gives the labels as negative, neutral, and positive. The most important factor that determines the success and accuracy of the process is the data itself, the quality and the reliability labels of the data. The model was successfully applied to the SEED dataset. The result is superior to most of the recent techniques shown in table1. The model is sufficient to be used in real-time applications.

5.1 Training loss

The average accuracy is approximately 0.9413 after the training step[81]. Based on figure 4 below, it shows the control number of complete pass through the training, the vertical number from 0-1 indicate skill of the model while horizontal

line indicates the training epoch. The result shows the model is suitably and conveniently fit to our training dataset. It also indicated that from figure 4, as the number of epochs increased the loss decrease and vice versa. Both training and testing datasets showed a slightly lower loss value based on 100 epochs. the training is achieved at less epoch because the curve moves straight not vibrating[82]. As shown the loss dropped quickly to zero, which increase the training accuracy and rapid drop of the loss.

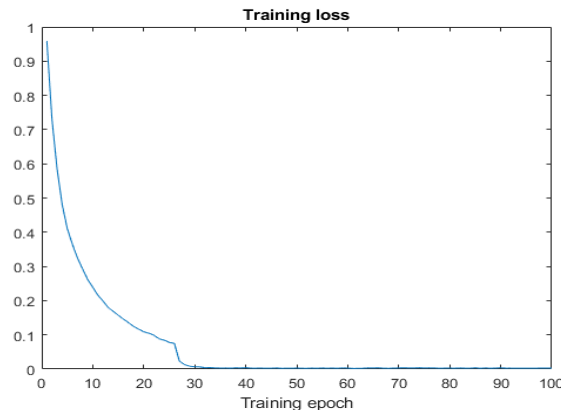


Figure: 4 Training loss for SEED dataset

5.2 Validation loss

Validation loss indicate how far or close of our model is generalizing unseen dataset record at the current epoch, validation loss is higher than the training loss, this indicates that the model is perfectly fit. There is a convergent between training and validation loss based on figure 5 better insights of the model's performance. The validation loss[83] starts decreasing and validation accuracy starts increasing and the build model leaning and works perfectly.

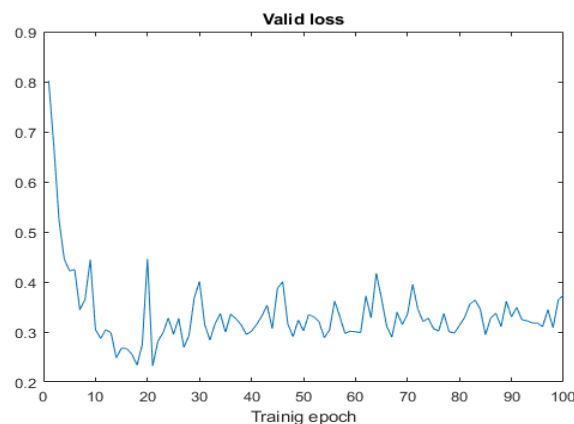


Figure: 5 valid loss for SEED dataset

5.3 Validation Accuracy

Validation Accuracy starts to increase rapidly from zero to its highest average value and could reach[1] 0.94133 as the loss decrease as shown in figure 6, which indicates how perfect is model and able to recognize emotion with the highest level of accuracy.

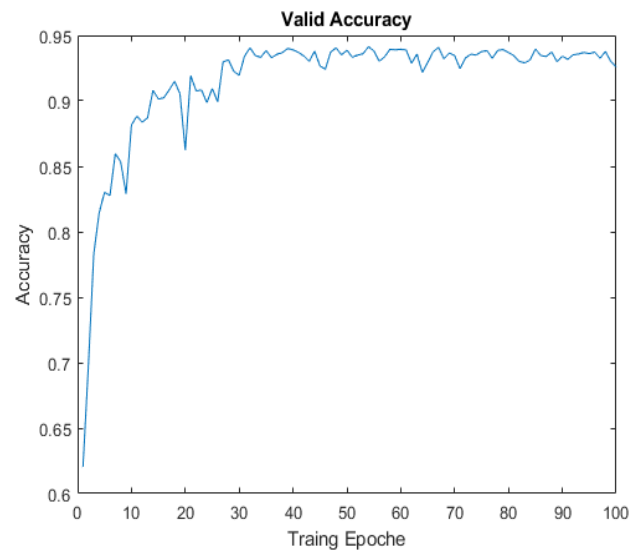


Figure: 6 valid accuracies for SEED dataset

5.4 Training Accuracy

The training accuracy based on the model is shown in figure 7 below, during the training the accuracy keeps increasing. This shows that the model was train well and working perfectly. Which causes high accuracy.

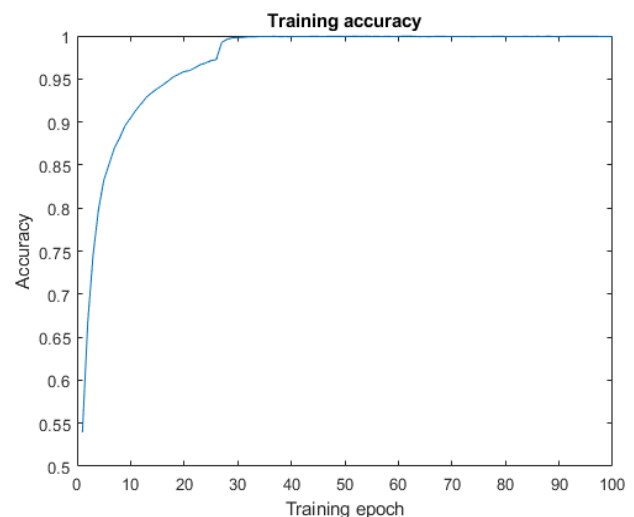


Figure: 7 Training accuracy for SEED dataset

5.6 The confusion matrix

The performance accuracy of the model is evaluated and interpreted using the confusion matrix as shown in figure 8, which shows the ability to recognize the emotion of positive, neutral, and negative[25]. The confusion matrix shows the strength of the model, the row represents the target class while the column represents the predicted class, and the diagonal elements of the confusion matrix stands for the actual recognition of emotion. For both testing and train data. The negative emotion has the highest accuracy fellow by neutral,

and positive emotion respectively. The average accuracy was found to be 94.13% for the SEED dataset based on a single model. Based on the result, it shows that our model can recognize emotion with the highest level of accuracy.

positive	0.9325	0.0356	0.0242
neutral	0.0362	0.9429	0.0272
negative	0.0313	0.0215	0.9486
	positive	neutral	negative

Figure 8: Confusion matrix for SEED dataset

Table1: Comparison of the various research on SEED dataset using deep learning.

METHOD	SEED
RGNN[25]	85.3%
CNN[53]	86.56% ,78.34%
ASFM[9]	80.46%
SAAE[66]	81.81%,
DGCNN+DE [84]	90.4%
BDAE[85]	91.01%
BDGLS+DE[86]	93.7%
DBN[31]	86.08%,
OUR PROPOSED	94.13%

6. Conclusions and Future Work

The current study has address the long-term challenge of getting neural relations between human emotion, brain region,

and computer interface. Convolutional neural network (CNN) with the ResNet50 model is used to learn, recognize and classify emotion of pure EEG signals. The model differs from the traditional or ordinary approaches because it automatically extract EEG signal feature without hand-craft feature extraction. The validity of the proposed approach is confirmed and verified. The single model classification accuracy is shown using the confusion matrix in figure8. The negative emotion has the highest accuracy of 94.86% fellow by neutral emotion with 94.29% and positive emotion of 93.25% respectively. We obtained the highest average recognition accuracy of 94.13% as we compare with other approaches shown in table 1, and the proposed model achieved the highest accuracy.

In the future work, we will evaluate our method with pure EEG signal using Multimodal, and other classes of natural emotion in real-time will be studying.

Conflicts of Interest: The authors declare no conflict of interest.

References

- [1] Y. Luo, G. Wu, S. Qiu, S. Yang, W. Li, and Y. Bi, "EEG-based Emotion Classification Using Deep Neural Network and Sparse Autoencoder," *Frontiers in Systems Neuroscience*, 2020, vol. 14, p. 43.
- [2] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "Emotionmeter: A multimodal framework for recognizing human emotions," *IEEE transactions on cybernetics*, 2018, vol. 49, no. 3, pp. 1110-1122.
- [3] S. Paul, A. Banerjee, and D. Tibarewala, "Emotional eye movement analysis using electrooculography signal," *International Journal of Biomedical Engineering and Technology*, 2017, vol. 23, no. 1, pp. 59-70.
- [4] A. Raheel, M. Majid, M. Alnowami, and S. M. Anwar, "Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia," *Sensors*, 2020, vol. 20, no. 14, p. 4037.
- [5] S. Farashi and R. Khosrowabadi, "EEG based emotion recognition using minimum spanning tree," *Physical and Engineering Sciences in Medicine*, 2020, vol. 43, no. 3, pp. 985-996.
- [6] H. Ullah, M. Uzair, A. Mahmood, M. Ullah, S. D. Khan, and F. A. Cheikh, "Internal emotion classification using EEG signal with sparse discriminative ensemble," *IEEE Access*, 2019, vol. 7, pp. 40144-40153.
- [7] J. Z. Lim, J. Mountstephens, and J. Teo, "Emotion Recognition Using Eye-Tracking: Taxonomy, Review and Current Challenges," *Sensors*, 2020, vol. 20, no. 8, p. 2384.
- [8] J. Li, Z. Zhang, and H. He, "Hierarchical convolutional neural networks for EEG-based emotion recognition," *Cognitive Computation*, 2018, vol. 10, no. 2, pp. 368-380.
- [9] X. Chai *et al.*, "A fast, efficient domain adaptation technique for cross-domain electroencephalography (EEG)-based emotion recognition," *Sensors*, 2017, vol. 17, no. 5, p. 1014.
- [10] W.-L. Zheng, B.-N. Dong, and B.-L. Lu, "Multimodal emotion recognition using EEG and eye tracking data," in *2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2014: IEEE, pp. 5040-5043.
- [11] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, 2018, vol. 11, no. 3, pp. 532-541.
- [12] L. J. Zheng, J. Mountstephens, and J. Teo, "Four-class emotion classification in virtual reality using pupillometry," *Journal of Big Data*, 2020, vol. 7, no. 1, pp. 1-9.
- [13] E. Kanjo, L. Al-Husain, and A. Chamberlain, "Emotions in context: examining pervasive affective sensing systems, applications, and analyses," *Personal and Ubiquitous Computing*, 2015, vol. 19, no. 7, pp. 1197-1212.

- [14] J. Sorinas, J. M. Ferrández, and E. Fernandez, "Brain and Body Emotional Responses: Multimodal Approximation for Valence Classification," *Sensors*, 2020, vol. 20, no. 1, p. 313.
- [15] K.-E. Ko, H.-C. Yang, and K.-B. Sim, "Emotion recognition using EEG signals with relative power values and Bayesian network," *International Journal of Control, Automation and Systems*, 2009, vol. 7, no. 5, p. 865.
- [16] M. Z. Soroush, K. Maghooli, S. K. Setarehdan, and A. M. Nasrabadi, "A novel approach to emotion recognition using local subset feature selection and modified Dempster-Shafer theory," *Behavioral and Brain Functions*, 2018, vol. 14, no. 1, p. 17.
- [17] B. Nakisa, M. N. Rastgoo, D. Tjondronegoro, and V. Chandran, "Evolutionary computation algorithms for feature selection of EEG-based emotion recognition using mobile sensors," *Expert Systems with Applications*, 2018, vol. 93, pp. 143-155.
- [18] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from EEG signals using multidimensional information in EMD domain," *BioMed research international*, 2017, vol. 2017.
- [19] M.-K. Kim, M. Kim, E. Oh, and S.-P. Kim, "A review on the computational methods for emotional state estimation from the human EEG," *Computational and mathematical methods in medicine*, 2013, vol. 2013.
- [20] F. Yang, X. Zhao, W. Jiang, P. Gao, and G. Liu, "Cross-subject emotion recognition using multi-method fusion from high-dimensional features," *Frontiers in Computational Neuroscience*, 2019, vol. 13, p. 53.
- [21] T.-H. Li, W. Liu, W.-L. Zheng, and B.-L. Lu, "Classification of five emotions from EEG and eye movement signals: Discrimination ability and stability over time," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2019: IEEE, pp. 607-610.
- [22] M. Z. Soroush, K. Maghooli, S. K. Setarehdan, and A. M. Nasrabadi, "A novel approach to emotion recognition using local subset feature selection and modified Dempster-Shafer theory," *Behavioral and Brain Functions*, 2018, vol. 14, no. 1, pp. 1-15.
- [23] N. S. Suhaimi, J. Mountstephens, and J. Teo, "EEG-Based Emotion Recognition: A State-of-the-Art Review of Current Trends and Opportunities," *Computational Intelligence and Neuroscience*, 2020, vol. 2020.
- [24] J. Atkinson and D. Campos, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," *Expert Systems with Applications*, 2016, vol. 47, pp. 35-41.
- [25] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Transactions on Affective Computing*, 2020.
- [26] T. Xu, Y. Zhou, Z. Wang, and Y. Peng, "Learning emotions EEG-based recognition and brain activity: A survey study on BCI for intelligent tutoring system," *Procedia computer science*, 2018, vol. 130, pp. 376-382.
- [27] J. Li, S. Qiu, C. Du, Y. Wang, and H. He, "Domain Adaptation for EEG Emotion Recognition Based on Latent Representation Similarity," *IEEE Transactions on Cognitive and Developmental Systems*, 2019.
- [28] L.-Y. Tao and B.-L. Lu, "Emotion Recognition under Sleep Deprivation Using a Multimodal Residual LSTM Network," in *2020 International Joint Conference on Neural Networks (IJCNN)*, 2020: IEEE, pp. 1-8.
- [29] Y. Zhang *et al.*, "An Investigation of Deep Learning Models for EEG-Based Emotion Recognition," *Frontiers in Neuroscience*, 2020, vol. 14.
- [30] Y. Wang *et al.*, "Silent speech decoding using spectrogram features based on neuromuscular activities," *Brain Sciences*, 2020, vol. 10, no. 7, p. 442.
- [31] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Transactions on Autonomous Mental Development*, 2015, vol. 7, no. 3, pp. 162-175.
- [32] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Transactions on Affective computing*, 2014, vol. 5, no. 3, pp. 327-339.
- [33] J. Zhang, Y. Zhou, and Y. Liu, "EEG-based emotion recognition using an improved radial basis function neural network," *JOURNAL OF AMBIENT INTELLIGENCE AND HUMANIZED COMPUTING*, 2020.
- [34] C. Aracena, S. Basterrech, V. Snáel, and J. Velásquez, "Neural networks for emotion recognition based on eye tracking data," in *2015 IEEE International Conference on Systems, Man, and Cybernetics*, 2015: IEEE, pp. 2632-2637.
- [35] S. Basu, J. Chakraborty, and M. Aftabuddin, "Emotion recognition from speech using convolutional neural network with recurrent neural network architecture," in *2017 2nd International Conference on Communication and Electronics Systems (ICCES)*, 2017: IEEE, pp. 333-336.
- [36] S. Hickson, N. Dufour, A. Sud, V. Kwatra, and I. Essa, "Eyemotion: Classifying facial expressions in VR using eye-tracking cameras," in *2019 IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2019: IEEE, pp. 1626-1635.
- [37] V. Raudonis, G. Dervinis, A. Vilkauskas, A. Paulauskaite, and G. Kersulyte, "Evaluation of human emotion from eye motions," *Evaluation*, 2013, vol. 4, no. 8.
- [38] R.-N. Duan, J.-Y. Zhu, and B.-L. Lu, "Differential entropy feature for EEG-based emotion classification," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2013: IEEE, pp. 81-84.
- [39] F. Yang, X. Zhao, W. Jiang, P. Gao, and G. Liu, "Multi-method fusion of cross-subject emotion recognition based on high-dimensional EEG features," *Frontiers in computational neuroscience*, 2019, vol. 13.
- [40] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing*, 2018.
- [41] M. A. Asghar *et al.*, "EEG-based multi-modal emotion recognition using bag of deep features: An optimal feature selection approach," *Sensors*, 2019, vol. 19, no. 23, p. 5218.
- [42] S. Koelstra *et al.*, "Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing*, 2011, vol. 3, no. 1, pp. 18-31.
- [43] H. Lee, D. Shin, and D. Shin, "A Study on the Emotion Classification as well as the Algorithm of the Classification Applying EEG-Data," in *Future Information Technology*: Springer, 2014, pp. 515-521.
- [44] A. Heraz and C. Frasson, "Predicting the three major dimensions of the learner's emotions from brainwaves," *International Journal of Computer Science*, 2007, vol. 2, no. 3, pp. 187-193.
- [45] W.-L. Zheng, J.-Y. Zhu, and B.-L. Lu, "Identifying stable patterns over time for emotion recognition from EEG," *IEEE Transactions on Affective Computing*, 2017.
- [46] M. Murugappan, N. Ramachandran, and Y. Sazali, "Classification of human emotion from EEG using discrete wavelet transform," *Journal of biomedical science and engineering*, 2010, vol. 3, no. 04, p. 390.
- [47] Y.-P. Lin *et al.*, "EEG-based emotion recognition in music listening," *IEEE Transactions on Biomedical Engineering*, 2010, vol. 57, no. 7, pp. 1798-1806.
- [48] L. Brown, B. Grundlehner, and J. Penders, "Towards wireless emotional valence detection from EEG," in *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2011: IEEE, pp. 2188-2191.
- [49] P. C. Petrantonakis and L. J. Hadjileontiadis, "A novel emotion elicitation index using frontal brain asymmetry for enhanced EEG-based emotion recognition," *IEEE Transactions on information technology in biomedicine*, 2011, vol. 15, no. 5, pp. 737-746.
- [50] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE transactions on affective computing*, 2011, vol. 3, no. 1, pp. 42-55.
- [51] X.-W. Wang, D. Nie, and B.-L. Lu, "Emotional state classification from EEG data using machine learning approach," *Neurocomputing*, 2014, vol. 129, pp. 94-106.

- [52] T. S. Rached and A. Perkusich, "Emotion recognition based on brain-computer interface systems," *Brain-computer interface systems-Recent progress and future prospects*, pp. 253-270, 2013.
- [53] Y. Cimtay, E. Ekmekcioglu, and S. Caglar-Ozhan, "Cross-subject multimodal emotion recognition based on hybrid fusion," *IEEE Access*, 2020, vol. 8, pp. 168865-168878.
- [54] X. Xing, Z. Li, T. Xu, L. Shu, B. Hu, and X. Xu, "SAE+ LSTM: A New framework for emotion recognition from multi-channel EEG," *Frontiers in neurorobotics*, 2019, vol. 13, p. 37.
- [55] M. Li, H. Xu, X. Liu, and S. Lu, "Emotion recognition from multichannel EEG signals using K-nearest neighbor classification," *Technology and Health Care*, 2018, vol. 26, no. S1, pp. 509-519.
- [56] Z. Mohammadi, J. Frounchi, and M. Amiri, "Wavelet-based emotion recognition system using EEG signal," *Neural Computing and Applications*, 2017, vol. 28, no. 8, pp. 1985-1990.
- [57] L. Yang, X. Ban, M. Mukeshimana, and Z. Chen, "Multimodal Emotion Recognition Using the Symmetric S-ELM-LUPI Paradigm," *Symmetry*, 2019, vol. 11, no. 4, p. 487.
- [58] S. G. Aydin, T. Kaya, and H. Guler, "Wavelet-based study of valence-arousal model of emotions on EEG signals with LabVIEW," *Brain informatics*, 2016, vol. 3, no. 2, pp. 109-117.
- [59] H. Li, X. Mao, and L. Chen, "An emotion classification method from electroencephalogram based on 1/f fluctuation theory," *Measurement and Control*, 2020, p. 0020294020913893.
- [60] H. He, Y. Tan, J. Ying, and W. Zhang, "Strengthen EEG-based emotion recognition using firefly integrated optimization algorithm," *Applied Soft Computing*, 2020, p. 106426.
- [61] F. Shah, "Discrete wavelet transforms and artificial neural networks for speech emotion recognition," *International Journal of Computer Theory and Engineering*, 2010, vol. 2, no. 3, p. 319.
- [62] A. T. Sohaib, S. Qureshi, J. Hagelbäck, O. Hilborn, and P. Jerčić, "Evaluating classifiers for emotion recognition using EEG," in *International conference on augmented cognition*, 2013: Springer, pp. 492-501.
- [63] S. K. Yoo, C. K. Lee, Y. J. Park, N. H. Kim, B. C. Lee, and K. S. Jeong, "Neural network based emotion estimation using heart rate variability and skin resistance," in *International conference on natural computation*, 2005: Springer, pp. 818-824.
- [64] C. Qing, R. Qiao, X. Xu, and Y. Cheng, "Interpretable emotion recognition using EEG signals," *IEEE Access*, 2019, vol. 7, pp. 94160-94170.
- [65] H. He, Y. Tan, J. Ying, and W. Zhang, "Strengthen EEG-based emotion recognition using firefly integrated optimization algorithm," *Applied Soft Computing*, 2020, vol. 94, p. 106426.
- [66] X. Chai, Q. Wang, Y. Zhao, X. Liu, O. Bai, and Y. Li, "Unsupervised domain adaptation techniques based on auto-encoder for non-stationary EEG-based emotion recognition," *Computers in biology and medicine*, 2016, vol. 79, pp. 205-214.
- [67] P. Keelawat, N. Thammasan, M. Numao, and B. Kijirikul, "Spatiotemporal emotion recognition using deep CNN based on EEG during music listening," *arXiv preprint arXiv:1910.09719*, 2019.
- [68] J.-M. López-Gil *et al.*, "Method for improving EEG based emotion recognition by combining it with synchronized biometric and eye tracking technologies in a non-invasive and low cost way," *Frontiers in computational neuroscience*, 2016, vol. 10, p. 85.
- [69] C. Lin, Q. Chang, and X. Li, "A deep learning approach for MIMO-NOMA downlink signal detection," *Sensors*, 2019, vol. 19, no. 11, p. 2526.
- [70] Y. Gao, H. J. Lee, and R. M. Mehmood, "Deep learning of EEG signals for emotion recognition," in *2015 IEEE International Conference on Multimedia & Expo Workshops (ICMEW)*, 2015: IEEE, pp. 1-5.
- [71] R. A. Naqvi, M. Arsalan, A. Rehman, A. U. Rehman, W.-K. Loh, and A. Paul, "Deep Learning-Based Drivers Emotion Classification System in Time Series Data for Remote Applications," *Remote Sensing*, 2020, vol. 12, no. 3, p. 587.
- [72] W. G. Hatcher and W. Yu, "A survey of deep learning: Platforms, applications and emerging research trends," *IEEE Access*, 2018, vol. 6, pp. 24411-24432.
- [73] M. Lee, Y. K. Lee, M.-T. Lim, and T.-K. Kang, "Emotion recognition using convolutional neural network with selected statistical photoplethysmogram features," *Applied Sciences*, 2020, vol. 10, no. 10, p. 3501.
- [74] P. Kim, "Deep learning," in *MATLAB Deep Learning*: Springer, 2017, pp. 103-120.
- [75] Y. Zhou, F. Ren, S. Nishide, and X. Kang, "Facial Sentiment Classification Based on Resnet-18 Model," in *2019 International Conference on Electronic Engineering and Informatics (EEI)*, 2019: IEEE, pp. 463-466.
- [76] V. Maeda-Gutiérrez *et al.*, "Comparison of Convolutional Neural Network Architectures for Classification of Tomato Plant Diseases," *Applied Sciences*, 2020, vol. 10, no. 4, p. 1245.
- [77] L. V. Fulton, D. Dolezel, J. Harrop, Y. Yan, and C. P. Fulton, "Classification of Alzheimer's disease with and without imagery using gradient boosted machines and ResNet-50," *Brain Sciences*, 2019, vol. 9, no. 9, p. 212.
- [78] T. Rahman *et al.*, "Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection using Chest X-ray," *Applied Sciences*, 2020, vol. 10, no. 9, p. 3233.
- [79] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access*, 2019, vol. 7, pp. 53040-53065.
- [80] I. Kandel, M. Castelli, and A. Popović, "Comparative Study of First Order Optimizers for Image Classification Using Convolutional Neural Networks on Histopathology Images," *Journal of Imaging*, 2020, vol. 6, no. 9, p. 92.
- [81] Y. Shi, Y. Wang, L. Zhao, and Z. Fan, "An event recognition method for Φ -OTDR sensing system based on deep learning," *Sensors*, 2019, vol. 19, no. 15, p. 3421.
- [82] J.-H. Lee, J. Kang, W. Shim, H.-S. Chung, and T.-E. Sung, "Pattern Detection Model Using a Deep Learning Algorithm for Power Data Analysis in Abnormal Conditions," *Electronics*, 2020, vol. 9, no. 7, p. 1140.
- [83] J. Cho and H. Hwang, "Spatio-Temporal Representation of an Electroencephalogram for Emotion Recognition Using a Three-Dimensional Convolutional Neural Network," *Sensors*, 2020, vol. 20, no. 12, p. 3491.
- [84] X. Wu, W.-L. Zheng, and B.-L. Lu, "Investigating EEG-Based Functional Connectivity Patterns for Multimodal Emotion Recognition," *arXiv preprint arXiv:2004.01973*, 2020.
- [85] W. Liu, W.-L. Zheng, and B.-L. Lu, "Emotion recognition using multimodal deep learning," in *International conference on neural information processing*, 2016: Springer, pp. 521-529.
- [86] X.-h. Wang, T. Zhang, X.-m. Xu, L. Chen, X.-f. Xing, and C. P. Chen, "EEG emotion recognition using dynamical graph convolutional neural networks and broad learning system," in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018: IEEE, pp. 1240-1244.

Author Contributions:

Isah Salim Ahmad: writing, Simulation, Methodology, and analysis.

Shuai Zhang: Supervision, Simulation, Methodology, dataset recommendation, Technical review, editing, original drafting. And conceptualization.

Sani Saminu: Simulation, Technical review, editing and Methodology

Lingyue Wang: Technical review and methodology

Abd El Kader Isselmou: Review, editing, and methodology.

Zilian Cai: Technical Review and methodology

Imran Javaid: Review and editing

Souha Kamhi: language Review.

Ummay Kulsum: Formatting.



[1] Isah Salim Ahmad¹

Student, Masters Student at the Hebei University of Technology, China.

Email: isahsalimahmad@gmail.com

Isah Salim Ahmad¹ received his B.Eng. from Kano University of Science and Technology (kust), Wudil, Nigeria in 2016. He is a Masters' student at Hebei University of technology Tianjin China. His research interest includes a brain-computer interface (BCI) and Machine learning.



[2] Dr. Zhang Shuai²

Professor,

Vice Dean of the College of Electrical Engineering, Hebei University of Technology.

Email address: zs@hebut.edu.cn

Dr. Zhang received his Bachelor degree from the College of Electrical Engineering at Southwest Jiaotong University (China) at 2007. Afterwards, he started joint doctoral study and received PhD degrees from both the Department of Applied

Physics in Ghent University (Belgium) at 2012, and the College of Electrical & Electronic Engineering in Huazhong University of Science & Technology (China) at 2013, in the field of plasma discharges.

In 2013 he joined the Chongqing University (China) as a researcher supported by the Hundred Talents Program. He is also a research member in the State Key Laboratory of Power Transmission Equipment & System Security and New Technology, Chongqing University. From November 2013 to July 2014, he was a visiting scholar in the Cold Plasma Diagnostics and Application Lab. at the Department of Mechanical Engineering, University of Minnesota (Twin Cities Campus), USA. His research interests include atmospheric pressure plasma discharges, advanced optical diagnostics, plasma interactions with liquids and solids, and novel applications of plasma discharges in energy and material science. He has been co-authored more than 60 publications, and since 2015 the total citations of his work have achieved to 1596 with h-index 22 and i10-index 36 by Google Scholar.

**Creative Commons Attribution License 4.0
(Attribution 4.0 International, CC BY 4.0)**

This article is published under the terms of the Creative Commons Attribution License 4.0

https://creativecommons.org/licenses/by/4.0/deed.en_US