Convolutional Neural Networks Model for Emotion Recognition Using EEG Signal

Isah Salim Ahmad[1], Zhang Shuai[1], Wang Lingyue[1], Sani Saminu[1], Abd El Kader Isselmou[1],
Zilian Cai1, Imran Javaid[1], Souha Kamhi[1], Ummay Kulsum[1]

[1]State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of
Technology, Tianjin, 300130, P.R. China.,

Abstract—A Brain-computer interface (BCI) using an electroencephalogram (EEG) signal has a great attraction in emotion recognition studies due to its resistance to humans' deceptive actions. This is the most significant advantage of brain signals over speech or visual signals in the emotion recognition context. A major challenge in EEG-based emotion recognition is that a lot of effort is required for manually feature extractor, EEG recordings show varying distributions for different people and the same person at different time instances. The Poor generalization ability of the network model as well as low robustness of the recognition system. Improving algorithms and machine learning technology helps researchers to recognize emotion easily. In recent years, deep learning (DL) techniques, specifically convolutional neural networks (CNNs) have made excellent progress in many applications. This study aims to reduce the manual effort on features extraction and improve the EEG signal single model's emotion recognition using convolutional neural network (CNN) architecture with residue block. The dataset is shuffle, divided into training and testing, and then fed to the model. DEAP dataset has class 1, class 2, class 3, and class 4 for both valence and arousal with an accuracy of 90.69%, 91.21%, 89.66%, 93.64% respectively, with a mean accuracy of 91.3%. 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 a mean accuracy of 94.13% on the SEED dataset. The experimental results indicated that CNN Based on residual networks can achieve an excellent result with high recognition accuracy, which is superior to most recent approaches.

I. INTRODUCTION

A brain-computer interface (BCI) is a system that translates brain activities which are directly connected to the brain of 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 is a subfield of human-computer interface (HCI) [2]HCI has become a vital and significant area of research in affective computing, it plays vital role in detecting, recognizing as well as processing and respond to user's emotion[3]. BCI assist in helping people with disabilities as well as making the user interact with devices[4].

BCI system consist of various sequential steps, which can be classified into four groups as: brain activity pattern generation, signal acquisition, feature extraction and classification [70].

Affective computing has being defined by Rosalind Picard as the computing that relates to , arises from or influence emotion[2], which include the field such as psychology, cognitive science, and artificial intelligence[3]. Emotion recognition is an important tools toward complete interaction between human and machine[5]. and had been extremely important in Biomedical engineering research.[6]-[4]. Emotion is an important tool for complete interaction between humans and machines [5].

Emotion is a psychological and physiological state accompanied by cognitive and conscious processes [7]. Emotion is a mental state experienced by a human and associated with feelings of a degree of pleasure or displeasure [3]. Emotional expression has a major role in human communication, daily life, and work. It can be defined as

positive, neutral, and negative experiences from several physiological activities [8]. There are many classes of emotions like sadness, happiness, anger, surprise, disgust, Fear among others [9]-[4] shown in figure 2.

Emotion stimuli are chosen to cover specific desire levels and valence states that are available in different structural forms such as visual, tactile, auditory, and odor. The emotional state produce by a stimulus is secure by using standard stimulus sets, like the 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.

The emotion models were divided into discrete and dimensional models - shown in figure 2. The discrete model deals with evolutionary features [10]-[2]-[10] - such as anger, sadness, fear, joy, and disgust as shown in figure 1. At the same time, 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 a strong level of emotion [11]. Valence is the positive and negative characteristics [4], Positive improves human health, work, and efficiency. While negative emotions may cause health problems [12]. Dominance is the mirror of one's status being control [13]-[14].

Alternatively, dimensional models express complex emotion in two-dimensional models as Valence-arousal (VA) or in three dimensions as : Valence, arousal and dominance (VAD)[15]. Emotion is also classified as negative, neutral and positive, the basic positive emotions are care and happiness which are necessary for life, development as well as evolution. The fundamental negative emotions include sadness, anger, disgust and fear, sometimes it operates at short period of time, while neutral emotion is based on prescriptive model[4].

With the enhancement and availability of wearable sensor technology, research on emotion recognition have become more common among researchers because it brings an important application, whether at professional, personal, or social level [16], such as in the field of Medicine [17] therapy, stock market [18], psychological counseling, computer game, education [4], driving, security, entertainment [3], workload estimation [19] among others.

There are several ways to identify an emotion, such as brain waves and facial expressions, and change in the nervous system [7]. 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. In contrast, the noninvasive method is referred to as BCI which includes Magnetoencephalogram (MEG), functional magnetic resonance imaging (fMRI), Electroencephalogram (EEG) [20].

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

Non-physiological signals are the early work, such as speech, gesture, facial expression, movement, voice intonation, text, body gesture, among others are mostly used in earlier works [7].

Recently, more research is still carried out based on physiological signals, such as Galvanic skin response(GSR), EEG [8], electrocardiogram (ECG) [21], pupillary diameter (PD), and electromyogram (EMG) [20]. EEG signals are more effective and reliable which reflects the mental state of human beings [22]. EEG records brain activity in the central nervous system via electrodes placed on the scalp and gives useful characteristics in responses to the emotional states [2].

Neurologists widely use the EEG signal for analysis and diagnosing various brain disorders such as; seizure detection, autism [14], attention deficit, and game addiction [23]. There are many EEG measurement devices available such as Neuroelectrics, Neurosky, Emotiv, and Biosemi which give different spatial and temporal resolution, resolution deals with the number of electrodes while the temporal resolution is related to the number of EEG samples computed per unit time. Usually, EEG signals have a high temporal but low spatial resolution, safety, and ease to use. EEG electrode is usually placed on the skull according to the 10-20 or 10-5 systems [24]-[25].

A strong correlation between EEG signals and emotion exists. The brain's two major regions related to emotional activities are 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 [26]-[10].

EEG feature Extractor and classifiers are two basic units in the machine learning (ML) approach: single-channel or multi-channel features [27].

Emotion recognition based on a single modality, the EEG signals have been generally used to develop efficient brain-computer interaction systems for analysis of both internal emotion and cognitive states.

Deep learning (DL) is a subfield of machine learning (ML) in artificial intelligence (AI) that deals with algorithms produced from the biological structure and brain function to help machines with intelligence[2]. DL is a specific ML algorithm in which features and classifiers are joined together, Convolutional neural network (CNN)[16], restricted Boltzmann machines (RBM), deep neural network (DNN) [14], and so on are an example of DL.

In the simplest form, the DL model is designed based on Convolutional neural network (CNN) architecture shown in figure 1; a NN is a hierarchical organization of neurons as well as connection to other neurons; these neurons send a signal to other neurons based on the input signal, neurons from the first hidden layer consume the signal, and then gives an output to the next layer and so on. Hence, resulting in the final output. Every layer may have one or many neurons, and they compute small functions that activate the function. The combination of two neurons of layers has associated with weight. The weight

determined the influence of the input to the put for the next neuron and lastly to the final output, in the CNN model, the logical building blocks such as a neuron, layer, weight, input, output, an activation function was defined inside the neuron to compute learning process.

Recently, Deep learning (DL) based on convolutional neural network (CNN) has been used for emotion recognition [19] and still at its infancy stage because parameter settings and model structure can still be adjusted and enhance [28].

It can be found that EEG-based emotional features have the problem that, the original information is easy to be lost. Furthermore, most of the network models' generalization and robustness are also poor and emotion recognition accuracy is not high [29]-[27]. CNN based residue block gives a high-speed accurate algorithms displaying excellent performance in the emotion recognition compared to the previous approaches.

The application of deep learning (DL) include: EEG signal analysis, Emotion recognition, voice recognition and classification[14], text interpretation, traffic sign recognition, lane classification, driver assistance, smartphone app, and an ATM. Recently, various biological models are inspired, such as CNN[29], DNN is giving out an excellent performance and accuracy in solving complex tasks due to developing fast and cheap hardware. 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 experts built the model. Other areas are; image and video recognition, audio processing, natural language processing, autonomous systems and robotics, medical diagnostics, computational biology, physical science, finance, economics, market analysis, cyber-security, architectural and algorithmic enhancement[2].
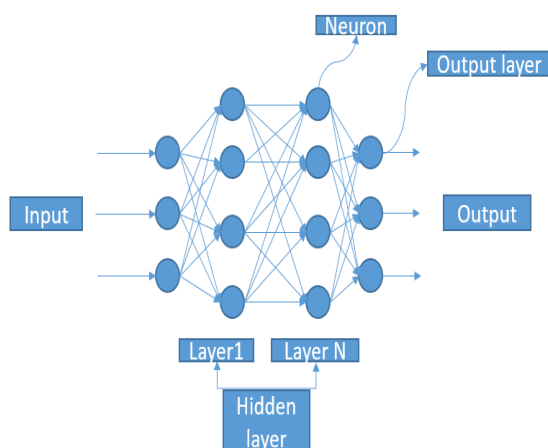

Figure 1: basic structure of CNN

This core objectives or aims of this paper is to use convolutional neural network (CNN) based on Residue block with Adam optimizer to improve emotion recognition of DEAP [30] and SEED [31] dataset.

The novelty of the proposed work is tested, which gives high-speed and produced an excellent recognition accuracy.

Contributions of this work.

The main contributions of this work to related literature in EEG-based emotion recognition can be summarized as follows:

To propose a CNN-based residual network to improve emotion recognition using DEAP and SEED datasets.

The feature extraction process is completely carried out to CNN based on the residual network model, whose extraction

Capability is highly competent in various recognition and classification tasks. This enables the model to explore useful and hidden features for classification.

The accuracy of the residual network increases as the network layers go deeper without any hyper-parameter tuning.

It also observed that for the emotion recognition problem, the accuracy of the residual network is preserved and improve by increasing depth.

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.
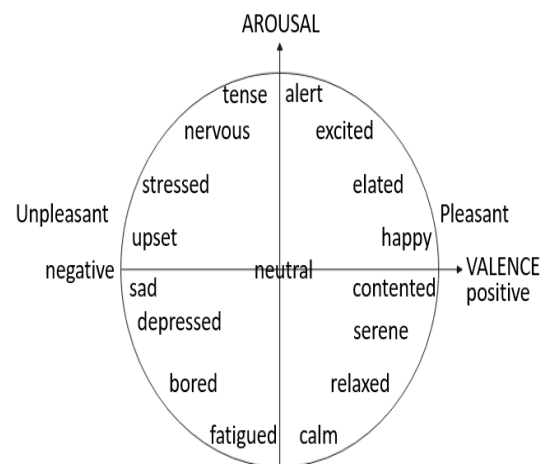

FIGURE2: Human emotions based on valence and arousal model DEAP dataset [4].

.

II. RELATED WORKED

For emotion recognition of EEG signals, several machine learning methods such as support vector machine (SVM) **[9]**, K-Nearest Neighbor(KNN) **[32]**-[6], Decision Tree(DT), Random Forest(RF) **[33]**, Linear Discriminant Analysis(LDA) **[34]**-**[35]**, Quadratic discriminant analysis (QDA) **[36]**, Discrete Wavelet Transform (DWT) **[37]**-**[38]** Multi-class Least squares support vector machine (MC-LS-SVM) [6] and logistic regression (LR) **[39]** were applied using standard

stimuli such as IAPS and IADS with remarkable results [21].

The study conducted using an Improved radial basis function neural network (I-RBF-NN) algorithm to classify and recognize emotion using EEG signal single modal and the result was compared with SVM, RBF-NN, GMM classifier, was found out the I-RBF-NN algorithm was better [29].

The correlations between different emotions and brain regions with suitable electrodes and applying modified Dampster-Shafer theory of using DEAP dataset have a mean accuracy of 91% [23].

The significance test/sequential backward selection and support vector machine (ST-SBSSVM) was applied based on high dimension feature for extracting multiple features and formation high dimensional feature on SEED and DEEP dataset. The accuracy was found to be 72% and 89% on DEAP and SEED respectively [40].

Minimum spanning tree (MST), quadratic discriminant analysis (QDA), were proposed using video as stimuli for Arousal and Valence with a mean accuracy of 88.23% and 81.255% respectively **[16]**.

Domain adaptation for emotion recognition based on latent representation similarity was proposed on the DEAP and SEED datasets, and the result outperforms the state of the arts [41].

EEG-based emotion recognition is the non-linearity and non-stationarity of EEG Signals, features set, like spectral band power of EEG channels, and it is difficult to identify features from the EEG recordings of different subjects, different sessions, and different datasets that indicate more commonality. This makes the recognition difficult with traditional methods. PS Bayes, Power different of the EEG waves were used for arousal an average accuracy of 62% and 57.6% respectively **[30]**.

A novel emotion Elicitation index using Frontal brain asymmetry for enhanced EEG-Based emotion recognition using two feature-vector extraction techniques and an SVM classifier was applied for extensive classification techniques for valence and arousal with mean accuracy up to 62.58% for user-dependent and 94.4% for user-dependent **[42]**.

Hidden Markov Models (HMM), SM, RBF, power spectral features were proposed for emotion recognition of arousal, valence, and dominance using IAPS as stimuli with the accuracy of 62% and 73% **[43]**.

Cross-subject multimodal emotion recognition based on hybrid fusion was employed on DEAP and LUMED dataset, with the mean accuracy of 53.8% and 74.2% respectively. This method helps find the correct emotional state in case of natural deceptive facial expression **[44]**.

Various machine learning approaches have been applied, such as SVM, RBF, and minimum-Redundancy-Maximum-Relevance (mRMR) were applied using video as stimuli for valence and arousal with the accuracy of 60.7% and 62.33% **[10]**.

A multimodal emotion recognition using symmetric sparse extreme learning machine-learning using Privileged Information (S-ELM-LUPI) was applied to improve the recognition effect and saving storage space with a recognition rate of over 80% at a very high speed **[45]**.

EEG signal is easy to be affected by artifacts, and drift may occur during the signal acquisition process. 1/f fluctuation theory was applied to classify EEG signal without artifact removing and drift data and achieved mean accuracy of 75% compared to KNN, SVM, for valence and arousal using music as stimuli **[46]**.

Most of the research focuses on accuracy recognition without interpretability of emotion progression. Interpretable emotion recognition using EEG signal was applied on DEAP and SEED datasets with a mean accuracy of 63.09% and 75% respectively **[33]**.

A multimodal emotion recognition framework called emotion meter was proposed, which combined brain waves and eye movements to increase feasibility and wearability. The multimodal deep neural network was significantly enhanced emotion recognition of happy, sad, fear, and neutral with a mean accuracy of 85.11% and mean accuracy of 72.39% on Emotionmeter SVM, LK, PSD, and DE were proposed for recognition of happy, sad, fear, and neutral with the accuracy of 85.11% and 72.39% **[8]**.

Integrated test/ sequential backward selection and support vector machine (ST-SBSSVM) is applied to improved emotion recognition across the subject with a mean accuracy of 72% and 89% on the DEAP and SEED datasets respectively [11].

Emotion recognition achieved good results based on traditional machine learning classifiers [21] as mentioned above but the accuracy needs to be improving.

To improve EEG emotion recognition, various methods using deep learning (DL) have been proposed such as [22]; Deep neural network(DNN) [47] and Long-Short Term Memory (LSTM), and Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) [48]. Most of these networks have shown good results as compared to the state-of-art method.

Recently, the EEG was collected, preprocessed, feature extracted, and then Improved radial basis function neural network (I-RBF-NN) is applied to study emotion recognition with a mean accuracy of 82.27% better than SVM and Gaussian mixture model (GMM) [49]. Regularized graph neural networks(RGNN) [27].

Long Short-Term Memory Recurrent Neural Network (LSTM-RNN) was developed for emotion recognition from EEG signal for valence and arousal using IAPS as a stimulus which achieves an accuracy of 81.10% for valence and 74.38% for arousal **[48]**.

Deep neural network(DNN) based Bag of in-depth features (BoDF) model was applied on SEED and DEAP datasets, the BoDF model achieved 93.8% on SEED and 77.4% on DEAP as compared with the traditional method [47].

Neuronal activities in different brain region and topology of EEG channels is studied using Regularized graph neural networks (RGNN) with SEED and SEED-IV dataset. Node-

wise domain adversarial training (NodeDAT) and emotion-aware distribution learning (Emotion DL) as regulazers were used. The result is superior to the state-of-art model [27].

Convolutional neural network (CNN) Based works.

The Deep Learning (DL) approach was proposed for emotion recognition from non-stationary EEG signals. By using a stacked Autoencoder (SAE) for feature learning to avoid overfitting caused by non-stationary EEG Signals, and then applied principal component analysis(PCA) to extract most components as well as covariant shift adaptation to minimize the non-stationary effect [50].

Hence, a multi-column CNN-based model for emotion recognition from EEG signal was proposed using the DEAP dataset for valence and arousal with a mean accuracy of 90.01% and 90.65% respectively [51].

Feedforward neural network was proposed to recognize and discriminate human emotion from EEG signals based on valence and arousal. The model has a spectral filter and an input layer, followed by shift register memory [52].

Convolutional neural network (CNN)-based emotion recognition was proposed using a DEAP dataset with two dissimilar models, a simple deep neural network and convolutional neural network(CNN) with 4.96% improvement [53].

3D Convolutional neural network (CNN) model was applied for emotion recognition from multichannel DEAP dataset using data augmentation phase to enhance the model's performance with a mean accuracy of 87.44% and 88.49% for valence and arousal respectively [54]. A hybrid neural network that contains CNN and RNN was proposed to classify human emotion by spatial-temporal representation of raw EEG signals for valence and arousal on the DEAP dataset with a mean accuracy of 90.8%, 91.03% respectively [55].

Multimodal emotion recognition was proposed to recognize eye and brain wave movement using a Deep neural network(DNN) with an emotion meter and obtained mean accuracy of 72.39% [8].

Convolutional neural network (CNN) for emotion recognition using EEG signal using brain connectivity features to account for synchronous activation of different brain regions. The approach sufficiently captures asymmetric brain activity patterns [56].

CNN for emotion recognition on EEG signal using DEAP dataset was applied to reduce the manual effort on features and improve EEG signal performance using the DEAP dataset with a mean accuracy of 77.98% and 72.98% for both valence and arousal respectively [57].

Deep learning (DL) methods such as CNN [57], DNN, long short-term memory (LSTM), a hybrid model of CNN and LSTM (CNN-LSTM) were applied on DEAP dataset, and the result indicated that CNN and CNN-LSTM models have the highest accuracy of 90.12% and 94.17% respectively [28].

The neuronal activities in the brain cause the non-stationary behavior of the EEG signals. It is challenging to extract hidden information from EEG signals as a result of poor generalizability of features across-subject. Flexible analytic wavelet transform was applied with a mean accuracy of 83.33% on SEED and 59.06% on the DEAP dataset [58].

A Dynamical Graph convolutional neural network (DGCNN) was applied with a mean accuracy of 90.4% for subject dependent to use a graph to model the multichannel and 79.95% for subject independent on SEED dataset, and DREAMER dataset means the accuracy of 86.23%, 84.54% and 85.02% for valence, arousal and dominance respectively [22].

Spatial-temporal recurrent neural network (STRNN) was used to capture spatial-current variations of human emotion with 89.5% on the SEED dataset. The result showed that STRNN is superior to state-of-the-art methods [59].

A Deep Convolutional neural network (DCNN) was applied to improve emotion recognition based on temporal features, frequency features, and their combination on the DEAP dataset for valence and arousal. The result DCNN has higher performance compared to Bagging tree (BT), SVM, LDA, and Bayesian linear discriminant analysis (BLDA) [21].

It can be found that EEG-based emotional features have the problem that, the original information is easy to be lost. Furthermore, most network models' generalization and robustness are also poor, and the accuracy of emotion recognition is not high.

An approach is necessarily needed to improve emotion recognition [6].

This article aims to improve the previous research on emotion recognition.

Contributions of this work.

The main contribution of this work to related literature is to propose a CNN-based residual block to improve emotion recognition using DEAP and SEED datasets.

III. DATA PREPARATION

This article's experiment was carried out on a dataset for emotion analysis using physiological signal (DEAP) [30]. and SJTU emotion EEG dataset (SEED) [31]. The DEEP dataset divided into 80% training and 20% for testing, while 70% for training 30% for testing. Training dataset is the sample data used to fit the model while testing dataset is sample of data used to provide an unbiased evaluation of a final model fit on the training dataset. Validation dataset is the sample of dataset held back from the training the model.

DEAP and SEED dataset are briefly introduced below.

A. DEAP dataset

DEAP dataset was produced by Katsigiannis and Ramzan and publically available at (http://www.eecs.qmul.ac.uk/mmv/datasets/deap/) it is a multimodal, consisting of electroencephalogram (EEG) signals as well as physiological signals of 32 subjects or participants such as EMG, GSR, EOG, Plethysmograph, and Respiration belt. Half of the participants were females while the remaining are male. The videos' frontal was also taken and recorded the data when the subjects watched one-minute long music. Each

subject or participant was shown a total of 40 videos. The participants rated the videos in terms of arousal, levels, valence, like/dislike, dominance, and familiarity which change between 1. The EEG signals were taken using 32 electrodes and then down-sampled to 128Hz. The EOG artifact was also removed and a band-pass filter was used to filter the signals at the frequency range of 4-4.5Hz. Then the dataset was divided or segmented into sixty seconds (60s) intervals and three-second (3s) based line data was removed [30].

B. SEED dataset

This study was carried out on the publically available emotion dataset in the 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 an 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) seconds of the hint of the movie, four (4) minutes of the clip, forty-five seconds (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 clips' contents as +1 for positive, 0 for neutral, and -1 for negative. The EEG signal was collected through 62 channels based on a 10-20 system as shown in figure 4. The signal was down-sampled to 200Hz, preprocessed, and segmented in MATLAB(.mat file), also the band-pass filter was used to filter the signals from 0-75Hz [31].
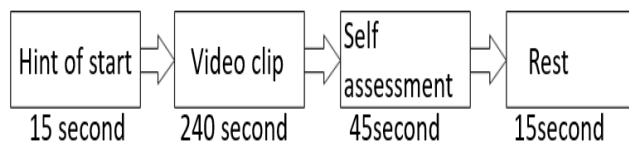

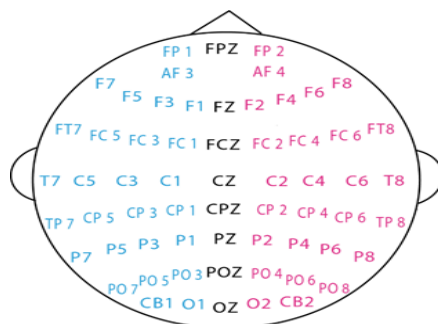FIGURE 3.   The detailed procedure of the experiment.

FIGURE 4. EEG cap of 62 channels based on 10-20 system (SEED Dataset (sjtu.edu.cn)).

IV. METHODOLOGY

A. Data processing.

The EEG signal is down-sampled to 200Hz. The sample contaminated by ECG, EMG, and EOG is removed and then filter using a Band-pass filter from the frequency 0-75Hz. The dataset is a shuffle and then divided into two, 80% training and 20% testing for DEAP and 70% training, and 30% testing for SEED. And then fed the signals into the model. The experiment was conducted in python 3.6 software, Keras in a CPU: i9-9900kf, memory 64GB Ram, and online GPU to run the algorithms: RTX 2080Ti.
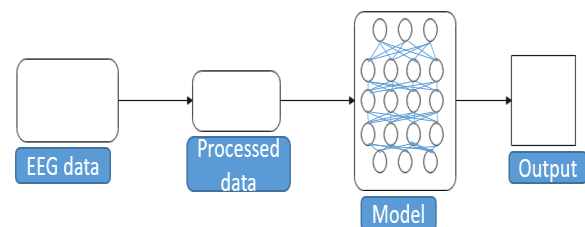

FIGURE 5. Emotion recognition procedure.

B. Emotion classification model based on CNN with residue

Figure 6. Show the network architectures for the emotion classification model based on the DEAP and SEED datasets.

Left side: a standard residual block, where the input is an EEG signal (62 * 200 for the SEED and 32*128 for DEAP). The output is the classification of three types of emotions (positive, negative, and neutral) for the SEED dataset. And two types of emotional arousal and valence for four classes on the DEAP dataset. Right side: shows more details and other variants about 2 types of residue blocks.

The shortcut connections are inserted, which turn the network into its counterpart residual version. The identity block use shortcut connections, which are the alternative path for the gradient to flow through, which solve the vanishing gradient.

Residue block with 50 convolution layers is used, which are divided into five stages. Each stage had a convolutional and an identity block, while each block had three convolution layers, with the kernel size 1x1, 3x3. The 3x3 kernels are responsible for extracting features and reducing feature size, and 1x1 kernels are responsible for reducing and then increasing or restoring dimensions. The convolution kernel is 7x7 and 3x3 shown in figure 6.

The basic structure of the network consists of 1 convolutional layer, 2 residual blocks A and B, max-pooling layers, softmax layer, Batch normalization, and dropout layer for overfitting consideration. ReLU is used as an activation function and completely connected layers.

The function of the convolutional layer is to identify the local connection feature of the existing layers. Shown in figure

6. The 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.
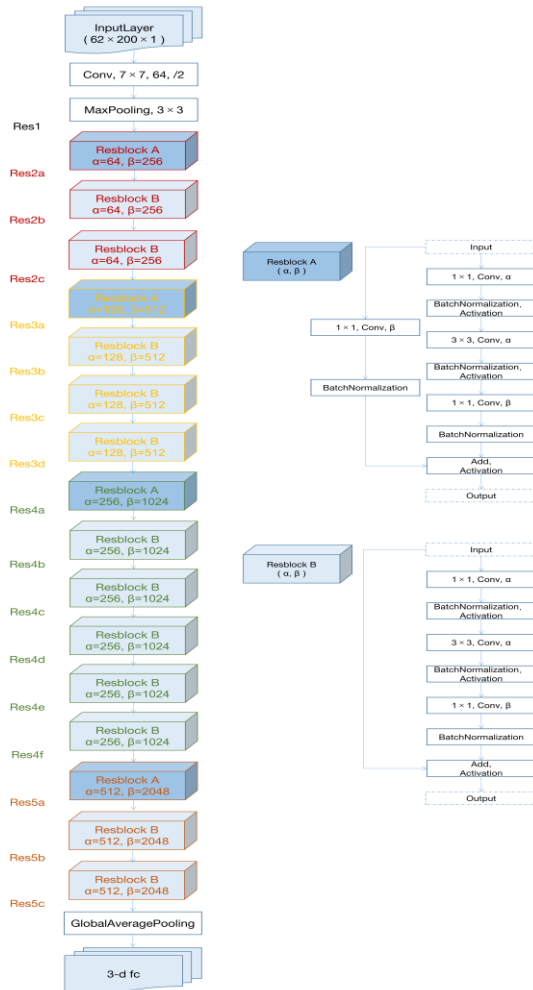

FIGURE 6. Emotion classification model based on Residue block

C. Training steps of the model.

1. Build the CNN with residue block.
2. Shuffle the dataset.
3. Divided the dataset into 80% training and 20% testing for DEAP and 70% training 30% testing for SEED.
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.
9. End.

D. Residue block

Residue block has two kinds of mapping: residual and identity mapping, identity mapping refers to the curved curves shown in Figure 7, which use shortcut connections, while residual mapping refers to the part except curved, so the final output is shown in equation 1

$$Y = F(x) + x \qquad (1)$$

Identity mapping as the name implies refers to itself, which is x in the formula, while residual mapping refers to the difference $Y - x$ then residual refers to $F(x)$

At first, the residue block performed the convolutional operation on the input, followed by a residual block, and finally performed a full connection operation to achieve classification tasks [60]. The network structure is shown in Figure 6.
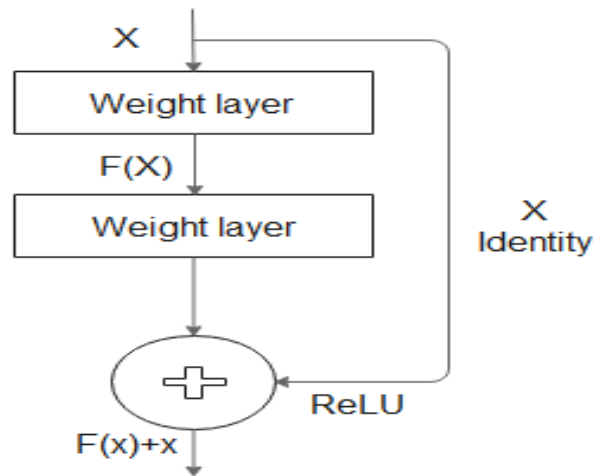

Figure 7. Residue block

The hyperparameter values for model configuration interm of accuracy, generalization speed are as follows:

- Batch size = 64;
- Optimizer =ADAM;
- Initial learning rate = 0.0001;
- Momentum optimizer = 0.9 and 0.999;
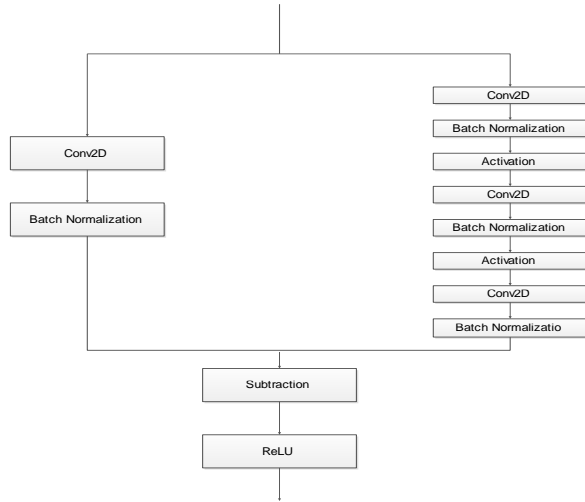- Number of epochs = 100;
- Dropout = 0.5

Figure8. Residual module

E. 2D Convolution

2D Convolution. We assumed that x is a filter inside the Signals $\mathbb{R}^{h \times w \times d_1}$, and also K be a filter in the signals of value $\mathbb{R}^{n \times n \times d_1 d_2}$, 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; $\quad$ Y= AX

Which is recognized as the forward operation of the 2D convolution of y= K ∗ x. And A has a form as given below;

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

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

F. 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.[84] The batch size used is 64 and dropout of 0.5

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

$$B = (x^i, \gamma, \beta) = \gamma \frac{\gamma(x^i - \mu)}{\sigma} + \beta \quad i = 1, 2, \ldots M \quad (3)$$

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

J. 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 $\mathbb{R}^{hwd_1} \to \mathbb{R}^{hwd_2}$ 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;

$$y = \begin{pmatrix} x_i & if \ d+1 \leq i \leq d_1, & where & d = \left\lfloor \frac{(d_2 - d_1)}{2} \right\rfloor \\ 0 & otherwise, & i = 1, 2, \ldots d_2 \end{pmatrix}$$

G. Over-fitting handling

In general, over-fitting means a situation whereby a model learns statistical regularities specific to the training set i.e ending up memorizing the irrelevant noise instead of learning signals. Memorizing the training dataset which may cause poor performance of the test dataset. As a result, the training dataset's performance is excellent while the test data set is lacking. The capacity of the network and the nature of the training can affect the model. We use the technique such as; obtaining more training dataset, dropout by using regularization and generalization [48] into the model, signal augmentation and stopping technique to avoid overfitting. It can be shown in figure 8 and 9 that, the loss drop more quickly this indicates the complexity of the model [44].

H. Activation function

In this article, Rectified linear unit (ReLU) is selected as an activation function. ReLU is a valid nonlinear function that works as an activation function. It improves the performance of the model help the number of computation to be reduced during the training phase of our model and solve the gradient dissipation[63].

ReLU has the following function;

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

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

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

Hence, α is a constant value.

I. The loss function

The loss function simply means is the matric that helps the convolutional neural network (CNN) to know if the learning process is going in the right direction. We used categorical cross-entropy in our model.

The Categorical cross-entropy: is a loss function used when categorical outcomes are non-binary that is greater than two (>2) as the possible outcomes such as class1, class2, class3, and class4 for DEAP and positive, neutral, and negative for SEED dataset.

Mathematical formula for categorical cross-entropy is

$$E(X) = -\sum_X P(X) \log P(X) \qquad (7)$$

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.

If the difference between s and t is higher, the cross-entropy increase, and vice-versa; the model is designed to learn the ground truth.

$$CE = -\sum_{i=0}^{n} y_i^{label} \log y_i^{pred} \qquad (8)$$

Hence, n stand for the number of categories, and y^{pred}, y^{label} are the probability distributions respectively.

J. Softmax function.

The softmax function is a function that turns a vector of k real values into a vector of k real values that sum to 1. The objective is to predict if the trained set of features x with its own set of weights w. thus normalizes output real values from the last fully connected to target class probability.

$$P(y = j / \Theta^{(i)}) = \frac{e^{\Theta^{(i)}}}{\sum_{j=0}^{k} e^{\Theta_k^{(i)}}} \qquad (9)$$

Where,

$$\Theta = w_0 x_0 + w_1 x_1 + \ldots \ldots w_k x_k = \sum_{i=0}^{k} w_i x_i = w^T x \qquad (10)$$

Θ is the transpose of the weights matrix w, multiplied by the feature matrix x.

K. ADAM Optimizer

Adam stands for Adaptive momentum estimation, it's the most popular and widely used optimizer in DL. We used the Adam optimizer technique to computes an adaptive learning rate for each parameter, minimize the margin loss function, accelerating and speeding up the model [64]. Adam defines the momentum and variance of the loss's gradient and leverages a combined effect to update the weight parameters. The momentum, variance together help smooth the learning curve and effectively improve the learning process.

The weights are computed and updated based on the Adam optimizer equation below;

$$w_t^i = w_{t-1}^i - \frac{\eta}{\sqrt{\hat{v}_t} + \varepsilon} . \hat{m}_t \qquad (11)$$

1st Moving momentum

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) G \qquad (12)$$

2nd Moving momentum

$$V_t = \beta_2 V_{t-1} + (1 - \beta_2)[G]^2 \qquad (13)$$

The bias correction of the first and second moment estimated as;

$$m_t = \frac{m_t}{1 - \beta_1^t} \qquad (14)$$

$$\tilde{v} = \frac{v_t}{1 - \beta_2^t} \qquad (15)$$

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

$$r_t = \sqrt{\frac{(p_t - 4)(p_t - 2) P_\infty}{(P_\infty - 4)(P_\infty - 4) P_t}} \qquad (16)$$

Hence $P_t = p_\infty - \frac{2t\beta_2^t}{1 - \beta_2^t}$, 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:

$$x_i = y \hat{w_i} + \beta \qquad (17)$$

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

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

(18)

And mini-batch mean is obtained using the following:

$$\mu_\beta = \frac{1}{m} \sum_{i=1}^{m} w_i$$

(20)

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

$$\sigma_B^2 = \frac{1}{m} \sum_{i=1}^{m} \left(w_i - u_B \right)^2$$

(21)

Use these to update the parameters just as we have seen, which yields the Adam update rule shown in equation 11. The variance of adaptive learning rate is deactivated, if not, the variance can be calculated and updating the adaptive learning rate.

In summary form, the mathematical form can be simplified as;

$$Weights = weights - \left(Momentum \quad and \quad variance \quad combined \right)$$

(22)

Where w_t represent the weights at step t, η is the learning rate, C is the cost function, the weight parameters w_t of the signal x and it is relative to a label y. β_i is the amount of information need for the next update, Hence, the values of β is 0.9, and 0.999 and γ are the parameter used for scale and shift are learning during training, m_t represent the running average of the gradients, it is called second moment or square moment. 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.

L. *Confusion matrix*

The confusion matrix has been used to ensure acceptable and reliable recognition results. The model's performance accuracy is evaluated and interpreted using the confusion matrix as shown in Figures 13(I&j) for both testing and training data. The following equations are used to obtain the accuracy and average accuracy based on a single model [27]. The comparison between the proposed CNN-based residue model and other models was evaluated based on the following values;

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

(23)

$$Sensitivity = \frac{TP}{TP + FN} \times 100\%$$

(24)

$$Specificity = \frac{TN}{TN + FN} \times 100\%$$

(25)

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}$$

(26)

Where,

TP	=	*True*	*positive*
TN	=	*True*	*Negative*
FN	=	*False*	*Negative*
FP	=	*False*	*Positive*

V. RESULTS AND DISCUSSION

In the literature review, there were various studies by researchers that used SEED or DEAP dataset for emotion recognition. Most of accuracy is not high.

An approach is necessarily needed to improve emotion recognition [6].

The method utilized in this article, which uses CNN based on residue block on SEED and DEAP dataset to recognize emotion. For the DEAP dataset, the classification is conducted on four classes. DEAP dataset labels as valence and arousal between one and nine. The classification is carried out on three classes for the SEED dataset: Positive-Neutral-Negative valence (Pos-Neu-Neg). The SEED dataset gives 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 training and feature extraction processes are based on a convolutional neural network (CNN) based residue block. We have verified the model's performance using training loss, validation loss, validation accuracy, training accuracy, and confusion matrix as discussed in Figure 9-13.

If training loss is less than validation loss, then the model is overfitting, and if the training loss is greater than validation loss then it is under-fitting. And when the training loss and validation loss are slightly higher than the training loss then it's a good fit. The result of the two datasets is compared, which is superior to most of the recent techniques as shown in Table 1. The model is sufficient to be used in real-time applications.
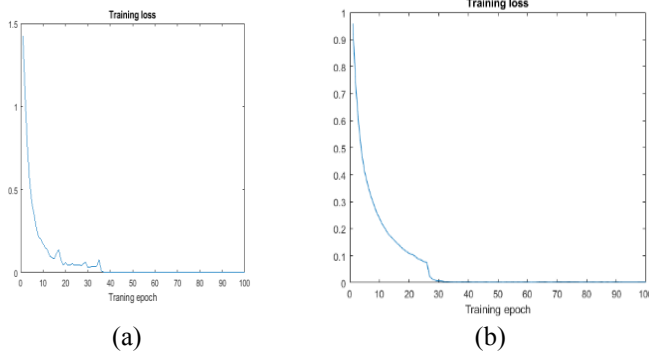
A. Training loss

(a)   (b)

Figure 9. Graph (a) represents the results of the training loss on the DEAP dataset obtained by CNN-based residue block, and graph (b) represents the result of the training loss value on the SEED dataset achieved by the CNN-based residue model.
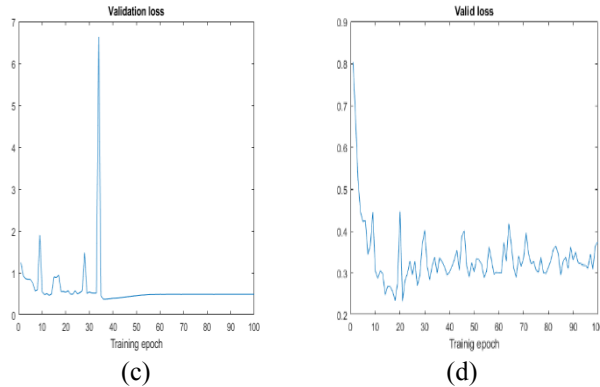
B. Validation loss

(c)   (d)

Figure 10. Graph (c) represents the results of the validation loss on the DEAP dataset obtained by the CNN-based residue block, and graph (d) represents the result of the validation loss value on the SEED dataset achieved by the CNN-based residue model.
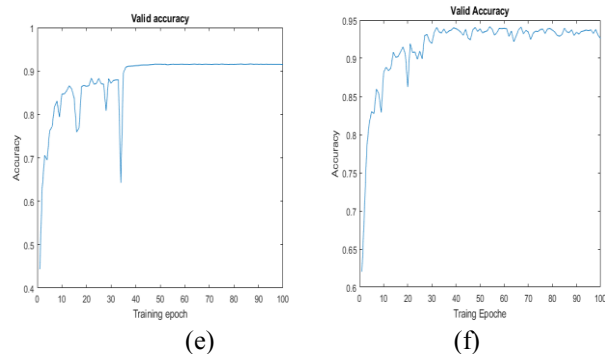
C. Validation Accuracy

(e)   (f)

Figure 11. Graph (e) represents the results of the valid accuracy on the DEAP dataset obtained by the CNN-based residue block, and graph (f) represents the result of the valid

accuracy value on the SEED dataset achieved by the CNN-based residue model.
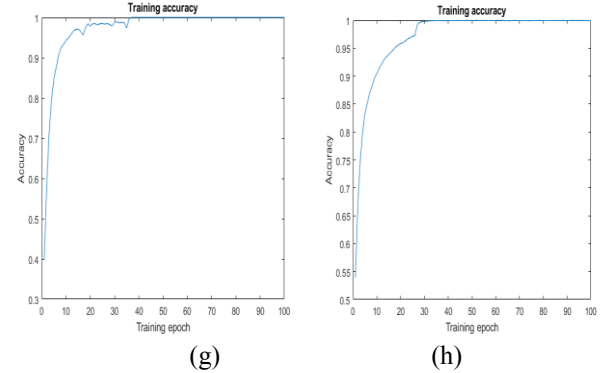
D. Training Accuracy

(g)   (h)

Figure 12. Graph (g) represents the results of the training accuracy on the DEAP dataset obtained by CNN-based residue block, and graph (h) represents the result of the training accuracy value on the SEED dataset achieved by the CNN-based residue model.

E. The confusion matrix

(i)

(j)

Figure 13. Graph (i) represents the results of the confusion matrix on the SEED dataset obtained by CNN-based residue block, and graph (j) represents the result of the confusion matrix value on the DEEP dataset achieved by the CNN-based residue model.

Table 1. Comparison of various research on DEAP and SEED datasets using different models with our proposed CNN model.

METHOD	SEED	DEAP
ST-SBSSVM [11]	89%	72%
BoDF [47]	93.8%	77.4%
LSTM-RNN [48]	*	74.38% & 81.10%
RGNN [27]	85.3%	*
CNN [44]	86.56 & 78.34%	*
ASFM [65]	80.46%	*
SAAE [66]	81.81%	*
DGCNN+DE [24]	90.4%	*
BDAE [67]	91.01%	*
DBN [31]	86.08%	*
Weight Coefficient[33]	75%	*
CNN [57]	77.98% & 72.98%	*
CNN, RNN [55]	91.03% & 90.8%	*
3DCNN [54]	88.49% & 87.44%	*
OUR PROPOSED	**94.13%**	**91.3%**

*stands for not used dataset in the worked.

VI. DISCUSSION

CNN networks have achieved great success in recent years in analyzing emotion recognition. Improving DL models required a huge datasets for emotion recognition. Our scientific team obtained huge data by DEAP and SEED dataset, which improved the proposed approach performance.

The existence of residue block solve the problems of network degradation. CNN had good ability of feature extraction and excellent in training and recognition ability. Adam optimizer was used to enhance the model performance the iteration process was carried out until iteration process was reached.

The dataset's training loss dropped very quickly while training accuracy increased as shown in figure 9. The average accuracy is approximately 0.913 for DEAP and 0.9413 for SEED after the training step. The result shows that model is suitably and conveniently fit to our training dataset. Because the training loss is greater than the validation loss of the two datasets because the Regularization applied during the training but not during the validation, training loss is measured during each epoch while validation loss is measured after each epochs and the validation set is easier than the training set, and this indicated that our model is not overfitting. It also indicated that as the number of epochs increased the loss decrease and vice versa. Both training and testing datasets showed a slightly lower loss value at 100 epochs. The training is achieved at less epoch because the curve moves straight not vibrating.

It can be seen that the validation loss dropped very quickly to a point of stability and has a small gap with the training loss. Because the architecture is a combination of CNN based on Residue block, indicate how far or close our model is generalizing the dataset record at the current epoch, validation loss is higher than the training loss of both datasets, this indicates that the model is perfectly fit. There is a convergent between training and validation loss based on Figures 10. Better insights of the model's performance. The validation loss [60] starts decreasing while validation accuracy starts increasing. A good fit id identified by a training and validation loss decreases to a point of stability with a minimal gap between the two final loss values. The loss of the model is lower on the training datasets than the validation dataset.

Validation Accuracy starts to increase rapidly from zero to its highest average value. It could reach a mean accuracy of 0.913 for DEAP and 0.94133 for SEED as the loss decrease as shown in Figure 11 which indicates how perfect the model is and able to recognize emotion with the highest level of accuracy. The validation accuracy is higher in the SEED than in the DEAP dataset.

The training accuracy based on the model is shown in Figure 12 below, it can be seen that, the recognition accuracy of the model using the different number of iterations, as the number of iterations increases, the accuracy increases and stabilizes gradually. During the training, the accuracy keeps in-creasing. This shows that the model is trained well for both datasets which leads to high accuracy.

The model's performance accuracy is evaluated, interpreted, and presented using the confusion matrix as shown in Figure 13 (i&j). Which shows the ability to recognize the emotion of valence, arousal, positive, neutral, and negative [26]. 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 stand for the actual recognition of emotion for both testing and train data. The DEAP dataset has class1, class2, class3, and class4 for both valence and arousal with the accuracy of 90.69%, 91.21%, 89.66%, and 93.64% respectively shown in Figure 10 (i). And three classes for the SEED dataset, the negative emotion has the highest accuracy followed by neutral, and positive emotion, respectively. The average accuracy was 94.13% for the SEED and 91.3% for the DEAP dataset based on a single model shown in Figure 12 (j). The result shows that our model can recognize emotion with the highest level of accuracy.

Table 1 shows a comparison between the proposed CNN-Residue and the existing models such as; [11]-[47]-[48]-[27]-[44]-[65]-[66]-[24]-[67]-[31]-[33]-[57]-[55]-[54].

In this Worked, we compared the two dataset. It can be seen from the table1 that our proposed method obtains the highest accuracy which mainly depends on the existence of residue model to solve the problems of network degradation and CNN which gives good ability

of feature extraction and excellent training ability.

The accuracy of the proposed model is much better than of the previous model. We can concluded that the proposed model offers powerful methods for accurate emotion recognition. And outperforming several recent models.

VII. CONCLUSIONS

The current study has addressed the long-term challenge of getting neural relations between human emotion, brain region, and computer-interface and improve emotion recognition. We used popular Convolutional neural network (CNN) algorithms combined with residue block, which has achieved a good effect in the single model classification task. The 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 extracts EEG signal features 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 Figure13. DEAP dataset has class1, class2, class3, and class4 for both valence and arousal with the accuracy of 90.69%, 91.21%, 89.66%, 93.64% respectively as shown in figure 12(i) with a mean accuracy of 91.3%. The negative emotion has the highest accuracy of 94.86% fellow by neutral emotion with 94.29% and positive emotion of 93.25% for the SEED dataset respectively, shown in Figure 13. We obtained the highest average recognition accuracy of 94.13% as we compare with other approaches as shown in table1, and the proposed model achieved the highest accuracy and good recognition effect in term of average accuracy.

In the future work, we will examine the overall success of our CNN based residue model and improve the network architecture which will improve the recognition and classification performance more effective than before. We will evaluate our method using Multimodal, and other classes of natural emotion in real-time will be studying.

Large data sets may improve the current performance of our proposed worked [71].

APPENDIX

BCI   Brain-computer interface.
EEG   Electroencephalogram.
ML   Machine learninig.
DL   Deep learninig.
CNN   Convolutional Neural Network.
ReLU   Rectified linear unit
Adam   Adaptive momentum.
DEAP Dataset for emotion analysis using physiological signal.
SEED Shanghai Jiao Tong University (SJTU) emotion EEG dataset.
     *   Stands for not used dataset.

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,* vol. 14, p. 43, 2020.

[2]   P. J. Bota, C. Wang, A. L. Fred, and H. P. Da Silva, "A review, current challenges, and future possibilities on emotion recognition using machine learning and physiological signals," *IEEE Access,* vol. 7, pp. 140990-141020, 2019.

[3]   J. Z. Lim, J. Mountstephens, and J. Teo, "Emotion recognition using eye-tracking: taxonomy, review and current challenges," *Sensors,* vol. 20, no. 8, p. 2384, 2020.

[4]   R. Alhalaseh and S. Alasasfeh, "Machine-Learning-Based Emotion Recognition System Using EEG Signals," *Computers,* vol. 9, no. 4, p. 95, 2020.

[5]   P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller, and S. Zafeiriou, "End-to-end multimodal emotion recognition using deep neural networks," *IEEE Journal of Selected Topics in Signal Processing,* vol. 11, no. 8, pp. 1301-1309, 2017.

[6]   S. Taran and V. Bajaj, "Emotion recognition from single-channel EEG signals using a two-stage correlation and instantaneous frequency-based filtering method," *Computer methods and programs in biomedicine,* vol. 173, pp. 157-165, 2019.

[7]   H. Chao and Y. Liu, "Emotion recognition from multi-channel EEG signals by exploiting the deep belief-conditional random field framework," *IEEE Access,* vol. 8, pp. 33002-33012, 2020.

[8]   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,* vol. 49, no. 3, pp. 1110-1122, 2018.

[9]   A. Raheel, M. Majid, M. Alnowami, and S. M. Anwar, "Physiological sensors based emotion recognition while experiencing tactile enhanced multimedia," *Sensors,* vol. 20, no. 14, p. 4037, 2020.

[10]   J. Atkinson and D. Campos, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," *Expert Systems with Applications,* vol. 47, pp. 35-41, 2016.

[11]   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, vol. 13, p. 53, 2019.

[12] L. Shu *et al.*, "A review of emotion recognition using physiological signals," *Sensors,* vol. 18, no. 7, p. 2074, 2018.

[13] 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,* vol. 130, pp. 376-382, 2018.

[14] 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.

[15] E. P. Torres P, E. A. Torres, M. Hernández-Álvarez, and S. G. Yoo, "EEG-based BCI emotion recognition: A survey," *Sensors,* vol. 20, no. 18, p. 5083, 2020.

[16] S. Farashi and R. Khosrowabadi, "EEG based emotion recognition using minimum spanning tree," *Physical and Engineering Sciences in Medicine,* vol. 43, no. 3, pp. 985-996, 2020.

[17] 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,* vol. 7, pp. 40144-40153, 2019.

[18] G. Zhang, T. Luo, W. Pedrycz, M. A. El-Meligy, M. A. F. Sharaf, and Z. Li, "Outlier processing in multimodal emotion recognition," *IEEE Access,* vol. 8, pp. 55688-55701, 2020.

[19] J. Li, Z. Zhang, and H. He, "Hierarchical convolutional neural networks for EEG-based emotion recognition," *Cognitive Computation,* vol. 10, no. 2, pp. 368-380, 2018.

[20] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing,* vol. 11, no. 3, pp. 532-541, 2018.

[21] J. Chen, P. Zhang, Z. Mao, Y. Huang, D. Jiang, and Y. Zhang, "Accurate EEG-based emotion recognition on combined features using deep convolutional neural networks," *IEEE Access,* vol. 7, pp. 44317-44328, 2019.

[22] T. Song, W. Zheng, P. Song, and Z. Cui, "EEG emotion recognition using dynamical graph convolutional neural networks," *IEEE Transactions on Affective Computing,* 2018.

[23] 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,* vol. 14, no. 1, p. 17, 2018.

[24] 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.

[25] 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,* vol. 14, no. 1, pp. 1-15, 2018.

[26] 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,* vol. 2020, 2020.

[27] P. Zhong, D. Wang, and C. Miao, "EEG-based emotion recognition using regularized graph neural networks," *IEEE Transactions on Affective Computing,* 2020.

[28] Y. Zhang *et al.*, "An Investigation of Deep Learning Models for EEG-Based Emotion Recognition," *Frontiers in Neuroscience,* vol. 14, 2020.

[29] S. Paul, A. Banerjee, and D. Tibarewala, "Emotional eye movement analysis using electrooculography signal," *International Journal of Biomedical Engineering and Technology,* vol. 23, no. 1, pp. 59-70, 2017.

[30] S. Koelstra *et al.*, "Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing,* vol. 3, no. 1, pp. 18-31, 2011.

[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,* vol. 7, no. 3, pp. 162-175, 2015.

[32] A. Heraz and C. Frasson, "Predicting the three major dimensions of the learner's emotions from brainwaves," *International Journal of Computer Science,* vol. 2, no. 3, pp. 187-193, 2007.

[33] C. Qing, R. Qiao, X. Xu, and Y. Cheng, "Interpretable emotion recognition using EEG signals," *IEEE Access,* vol. 7, pp. 94160-94170, 2019.

[34] 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.

[35] X.-W. Wang, D. Nie, and B.-L. Lu, "Emotional state classification from EEG data using machine learning approach," *Neurocomputing,* vol. 129, pp. 94-106, 2014.

[36] R. Jenke, A. Peer, and M. Buss, "Feature extraction and selection for emotion recognition from EEG," *IEEE Transactions on Affective computing,* vol. 5, no. 3, pp. 327-339, 2014.

[37] M. Li, H. Xu, X. Liu, and S. Lu, "Emotion recognition from multichannel EEG signals using K-nearest neighbor classification," *Technology and Health Care,* vol. 26, no. S1, pp. 509-519, 2018.

[38] Z. Mohammadi, J. Frounchi, and M. Amiri, "Wavelet-based emotion recognition system using

EEG signal," *Neural Computing and Applications,* vol. 28, no. 8, pp. 1985-1990, 2017.

[39] Y.-P. Lin *et al.*, "EEG-based emotion recognition in music listening," *IEEE Transactions on Biomedical Engineering,* vol. 57, no. 7, pp. 1798-1806, 2010.

[40] 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,* vol. 13, 2019.

[41] 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,* vol. 12, no. 2, pp. 344-353, 2019.

[42] 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,* vol. 15, no. 5, pp. 737-746, 2011.

[43] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE transactions on affective computing,* vol. 3, no. 1, pp. 42-55, 2011.

[44] Y. Cimtay, E. Ekmekcioglu, and S. Caglar-Ozhan, "Cross-subject multimodal emotion recognition based on hybrid fusion," *IEEE Access,* vol. 8, pp. 168865-168878, 2020.

[45] L. Yang, X. Ban, M. Mukeshimana, and Z. Chen, "Multimodal Emotion Recognition Using the Symmetric S-ELM-LUPI Paradigm," *Symmetry,* vol. 11, no. 4, p. 487, 2019.

[46] H. Li, X. Mao, and L. Chen, "An emotion classification method from electroencephalogram based on 1/f fluctuation theory," *Measurement and Control,* p. 0020294020913893, 2020.

[47] M. A. Asghar *et al.*, "EEG-based multi-modal emotion recognition using bag of deep features: An optimal feature selection approach," *Sensors,* vol. 19, no. 23, p. 5218, 2019.

[48] 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,* vol. 13, p. 37, 2019.

[49] 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,* pp. 1-12, 2020.

[50] S. Jirayucharoensak, S. Pan-Ngum, and P. Israsena, "EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation," *The Scientific World Journal,* vol. 2014, 2014.

[51] H. Yang, J. Han, and K. Min, "A multi-column CNN model for emotion recognition from EEG signals," *Sensors,* vol. 19, no. 21, p. 4736, 2019.

[52] R. Khosrowabadi, C. Quek, K. K. Ang, and A. Wahab, "ERNN: A biologically inspired feedforward neural network to discriminate emotion from EEG signal," *IEEE transactions on neural networks and learning systems,* vol. 25, no. 3, pp. 609-620, 2013.

[53] S. Tripathi, S. Acharya, R. D. Sharma, S. Mittal, and S. Bhattacharya, "Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset," in *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, 2017, pp. 4746-4752.

[54] E. S. Salama, R. A. El-Khoribi, M. E. Shoman, and M. A. W. Shalaby, "EEG-based emotion recognition using 3D convolutional neural networks," *Int. J. Adv. Comput. Sci. Appl,* vol. 9, no. 8, pp. 329-337, 2018.

[55] Y. Yang, Q. Wu, M. Qiu, Y. Wang, and X. Chen, "Emotion recognition from multi-channel EEG through parallel convolutional recurrent neural network," in *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018: IEEE, pp. 1-7.

[56] S.-E. Moon, S. Jang, and J.-S. Lee, "Convolutional neural network approach for EEG-based emotion recognition using brain connectivity and its spatial information," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018: IEEE, pp. 2556-2560.

[57] Z. Wen, R. Xu, and J. Du, "A novel convolutional neural networks for emotion recognition based on EEG signal," in *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*, 2017: IEEE, pp. 672-677.

[58] V. Gupta, M. D. Chopda, and R. B. Pachori, "Cross-subject emotion recognition using flexible analytic wavelet transform from EEG signals," *IEEE Sensors Journal,* vol. 19, no. 6, pp. 2266-2274, 2018.

[59] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, "Spatial–temporal recurrent neural network for emotion recognition," *IEEE transactions on cybernetics,* vol. 49, no. 3, pp. 839-847, 2018.

[60] A. ERRACHDI and M. BENREJEB, "Adaptive Internal Model Neural Networks Control for Nonlinear System," 2013.

[61] M. I. García Planas and S. Tarragona, "Analysis of functional output-controllability of time-invariant Composite linear systems," 2013.

[62] A. Shrestha and A. Mahmood, "Review of deep learning algorithms and architectures," *IEEE Access,* vol. 7, pp. 53040-53065, 2019.

[63] 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,* vol. 10, no. 10, p. 3501, 2020.

[64] P. Keelawat, N. Thammasan, M. Numao, and B. Kijsirikul, "Spatiotemporal emotion recognition using deep CNN based on EEG during music listening," *arXiv preprint arXiv:1910.09719,* 2019.

[65]   X. Chai *et al.*, "A fast, efficient domain adaptation technique for cross-domain electroencephalography (EEG)-based emotion recognition," *Sensors,* vol. 17, no. 5, p. 1014, 2017.

[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,* vol. 79, pp. 205-214, 2016.

[67]   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.

[68]   Ayachi Errachdi, Mohamed Benrejeb, Adaptive Internal Model Neural Networks Control for Nonlinear System pp.9-14, International Journal of Electrical Engineering and Computer Science (EEACS), Vol.2, 2020

[69]   Maria Isabel Garcıa-Planas, Sonia Tarragona, Analysis of Functional Output-controllability of Time-invariant Composite Linear Systems pp.1-8, International Journal of Electrical Engineering and Computer Science (EEACS), Vol.2, 2020

[70]   Martina Zabcikova, Zuzana Koudelkova, Roman Jasek, Investigation of Emotiv Epoc+ Headset Usability by Measuring Various Stimuli, WSEAS Transactions on Applied and Theoretical Mechanics, pp. 184-191, Volume 14, 2019.

[71]   Suprijanto, Azizah S. Noor, Ayu G. Risangtuni, Hesty Susanti, Automatic Segmentation Scheme for Effective Synchronization of EMG-EEG Quantification, International Journal of Biology and Biomedical Engineering, pp. 136-143.

[1] Isah Salim Ahmad[1]
isahsalimahmad@gmail.com
Isah Salim Ahmad was born in Kano state, Nigeria. He received his B.Eng. degree in Electrical and Electronics Engineering from Kano University of Science and Technology, Kano Nigeria. In 2016. He is currently a Masters' student at the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin, China. Biomedical engineering. His research interest includes a brain-computer interface (BCI) and Machine learning.

[2] Dr. Zhang Shuai[2]
Professor,
Vice Dean of the College of Electrical Engineering, Hebei University of Technology.
Email address: zs@hebut.edu.cn
Dr. Zhang received his Bachelor's degree from the College of Electrical Engineering at Southwest Jiaotong University (China) in 2007. Afterward, he started the joint doctoral study and received Ph.D. degrees from both the Department of Applied Physics in Ghent University (Belgium) in 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 1596 with h-index 22 and i10-index 36 by Google Scholar.

[3] Wang Lingyue
Student, Doctors Student at the Hebei University of Technology, China.
Wang Lingyue received his B.Eng. from University of Electronic Science and Technology of China, Chengdu, China in 2015. He is a currently Doctors' student majoring in Biomedical Engineering at the Hebei University of Technology, Tianjin, China. His research interest includes digital image processing and deep learning on EEG and facial images.

[4] SANI SAMINU was born in Kano State, Nigeria. He received his B.Eng. degree in Electrical and Electronics Engineering from Kano University of Science and Technology Wudil, Kano, Nigeria. He obtained M.Sc. in Electrical and Electronics Engineering from Yasar University Izmir, Turkey.
He has been working with the Biomedical Engineering Department, University of Ilorin, Nigeria. He is currently a Ph.D. candidate at the School of Electrical Engineering, Hebei University of Technology Tianjin, China.
His research interests are in signal processing, Wireless/Sensor communication, and networks and instrumentation, in particular applied to biomedical applications.

[5] ISSELMOU ABD EL KADER received the B.S. degree from the School of Science, University Technology of Malaysia, Juhar, Malaysia, in 2013 and the M.S degree in biomedical engineering from Hebei University of Technology, Tianjin, China in 2017. He is currently Ph. D candidate with the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin, China. He has published 13

papers indexing SCI and EI. His research interests include medical image processing and machine learning.

[6] **IMRAN JAVID** received his B.S from the University of Peshawar, Pakistan (1st Division). Currently, He is a master's study (Biomedical engineering) with the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin, China. He has 5 publications. His research interests include medical image processing and machine learning.

[7] **SOUHA KAMHI** received her bachelor's degree from the University of Mohamed 6th of Health Sciences, Casablanca, Morocco. Currently, she is a master's student with the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin, China. Her current research interests include biomedical signal processing and brain-computer interfacing.

[8] **UMMAY KULSUM**, Received her bachelor's degree from Gono University, Dhaka, Bangladesh. Currently, she is a master's student with the State Key Laboratory of Reliability and Intelligence of Electrical Equipment, Hebei University of Technology, Tianjin, China. Her current research interests include Electrical impedance tomography.

Author Contributions:
Isah Salim Ahmad: writing, Simulation, Methodology, validation, software and analysis.
Zhang Shuai: Supervision, Simulation, Methodology, dataset recommendation, Technical review, editing, original drafting, software validation and Conceptualization.
Lingyue Wang: Technical review, Simulation and methodology.
Sani Saminu: Technical review, editing and Methodology.
Abd El Kader Isselmou: Technical review, editing, Methodology and formatting.
Zilian Cai: Formatting, Technical Review, conceptualization, formatting and methodology.
Imran Javaid: Review and editing.
Souha Kamhi: language Review.
Ummay Kulsum: Formatting.