IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 7 April 2022, accepted 24 May 2022, date of publication 29 June 2022, date of current version 14 July 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3181887

Wavelet ELM-AE Based Data Augmentation and Deep Learning for Efficient Emotion Recognition Using EEG Recordings

BERNA ARI[1], KAMRAN SIDDIQUE[2], (Senior Member, IEEE), ÖMER FARUK ALÇIN[3], MUZAFFER ASLAN[4], ABDULKADIR ŞENGÜR[1], AND RAJA MAJID MEHMOOD[2], (Senior Member, IEEE)

[1]Department of Electrical and Electronics Engineering, Firat University, 23000 Elazığ, Turkey
[2]Information and Communication Technology Department, School of Computing and Data Science, Xiamen University Malaysia, Sepang 43900, Malaysia
[3]Department of Electrical and Electronics Engineering, Malatya Turgut Özal University, 44000 Malatya, Turkey
[4]Department of Electrical and Electronics Engineering, Bingol University, 12000 Bingol, Turkey

Corresponding author: Raja Majid Mehmood (rmeex07@ieee.org)

This work was supported in part by the Xiamen University Malaysia Research Fund (XMUMRF) under Grant XMUMRF/2022-C9/IECE/0035.

ABSTRACT Emotion perception is critical for behavior prediction. There are many ways to capture emotional states by observing the body and copying actions. Physiological markers such as electroencephalography (EEG) have gained popularity, as facial emotions may not always adequately convey true emotion. This study has two main aims. The first is to measure four emotion categories using deep learning architectures and EEG data. The second purpose is to increase the number of samples in the dataset. To this end, a novel data augmentation approach namely the Extreme Learning Machine Wavelet Auto Encoder (ELM-W-AE) is proposed for data augmentation. The proposed data augmentation approach is both simple and faster than the other synthetic data augmentation approaches. For deep architectures, large datasets are important for performance. For this reason, data multiplexing approaches with classical and synthetic methods have become popular recently. The proposed synthetic data augmentation is the ELM-W-AE because of its efficiency and detail reproduction. The ELM-AE structure uses wavelet activation functions such as Gaussian, groove gap waveguide (GGW), Mexican, Meyer, Morlet, and Shannon. Deep convolutional architectures classify EEG signals as images. EEG waves are scalograms using Continuous Wavelet Transform (CWT). The ResNet18 architecture recognizes emotions. The proposed technique uses GAMEEMO data collected during gameplay. Each of these states is represented in the GAMEEMO data collection. The visual data set created from the signal was divided into two groups 70% training and 30% testing. ResNet18 has been fine-tuned with augmented photos, training images only. It achieved 99.6% classification accuracy in tests. The proposed method is compared with the other approaches on the same dataset, and an approximately 22% performance improvement is achieved.

INDEX TERMS Auto-encoder, data augmentation, deep learning, emotion recognition, emotion measurement.

I. INTRODUCTION

Emotion is an abstract term that is described as a reaction to external and internal stimuli and expresses human consciousness. In recent years, emotion detection has been a popular research area. However, the concept of abstraction contained within it has made it difficult to comprehend people's emotional states [1]. When the studies are examined, there is a tendency for two headings in emotion recognition. The first is a discrete model that distinguishes between positive and negative emotions. Joy, anger, anticipation, surprise, disgust, sadness, confidence, and fear are the eight primary emotions in the discrete models [2]. The second one is the dimensional model. The dimensional model has been

The associate editor coordinating the review of this manuscript and approving it for publication was Siddhartha Bhattacharyya.

looked at under Russell and Plutchick model. The first is Russell's arousal-valence coordinate system, which divides emotions into four sections on the coordinate plane, with the left side representing negative emotions and the right side representing positive emotions. On the other side, the arousal axis shows the progression of emotions from sedentary to active [3]. The wheel of emotion described by Plutchick is another dimensional paradigm [4]. Emotions are ranked according to their intensity on the emotion wheel. In addition, in this concept, emotions combine to form complex emotions. However, the difficulty of some languages to convey every feeling generated by this wheel calls into question the model's universality.

While several studies have been conducted on emotion detection based on facial expressions, its reliability has been questioned. The fact that facial expressions can be imitated, meaning that the real emotion felt on the inside is conveyed or silenced differently with gestures and facial expressions, has harmed the studies' accuracy [5]. This has conducted emotion detection toward physiological signals like EEG and electrocardiogram (ECG) [6]–[8]. EEG records the electrical activity via electrodes on the scalp in the brain. The monitoring of conscious brain activities obtained from EEG signals and the identification of human emotions has been facilitated using Brain-Computer Interface (BCI) technology [9].

While the researchers have focused on deep learning in emotion recognition, the data sets' scarcity has become a significant limitation. Today, studies in many disciplines are carried out on data sets obtained from research areas. In some cases, data sets are sufficient for modeling the situation, while sometimes they are insufficient in terms of the number of samples. The number of samples in the data set is significant for many classification methods [10]. This number has a significant impact on classifier generalization capacity and accuracy. The term ''data augmentation'' is come up to describe methods for creating iterative optimization or sampling algorithms using unobserved data or hidden variables as a result of this need [11]. Various methods for increasing data are available in the literature.

The amount of data available, particularly in medical fields, is limited, and obtaining it is expensive. Furthermore, the amount of data obtained is inadequate and inconsistent for most scientific fields due to factors such as the inability to access previously generated data and the time-consuming data collection process. In addition to these limitations, the amount of data is of great importance in achieving the desired goal in deep architectures, which have increased their popularity day by day. There are several different ways to acquire new data pieces using conventional data augmentation methodologies, such as taking the symmetry of the image based on various axes, cutting out a random sample piece from the image, changing the axes' location, changing the color ratios in the image, noise addition and so on [12], [13]. The model's propensity to memorize has been avoided thanks to this multiplication/increment. Although these methods can be used to perform operations like object detection, they don't always produce beneficial results in photos. Recently, some methods such as Variational Auto-Encoder (VAE), Generative Adversarial Nets (GAN), and Extreme Learning Machine Auto Encoder (ELM-AE) have attracted attention in terms of generating real data [11], [14], [15].

The key objective of this study is to present a reliable method for EEG-based emotion recognition. The proposed method is based on the conversion of signals to images, data augmentation, and deep learning. Additionally, the approach is predicated on two major scenarios. In the first scenario, signal-to-image conversion and deep learning steps are used to transform the original data. In the alternative scenario, images obtained following the signal-to-image conversion step are synthetically enhanced and then subjected to deep learning steps. Separate scenarios of retesting the data following synthetic enhancement were examined. The CWT method is used to convert the signal to the image. The synthetic data were obtained using the wavelet-based ELM-AE structure, which will be referred to as ELM-W-AE throughout the remainder of the article. In the experiments, the proposed approach is validated against a variety of wavelet functions, including Gaussian, GGW, Mexican, Meyer, Morlet, and Shannon. Six different wavelet functions were used to generate synthetic data, and the effect of wavelet structure on performance was investigated. For classification, deep learning is preferred. The advantage of synthetic data acquisition is that it enables learning and training for the desired deep architectural structure. The proposed method has been validated against the GAMEEMO dataset. In the GAMEEMO dataset, EEG signals were collected while subjects participated in emotional evaluation games. The GAMEEMO dataset is used to detect emotions using the ResNet18 architecture. The ResNet18 model is used in a fine-tuned transfer learning format. The final three layers of the pre-trained ResNet18 model are altered to address our issue.

So, in this study, a different data augmentation approach namely ELM-W-AE is presented for data augmentation. The proposed data augmentation method is both uncomplicated and faster compared to previous synthetic data augmentation methods. Furthermore, the proposed method produces a higher accuracy score than the other approaches due to the proposed data augmentation methodology.

The rest of the paper is organized as follows. Section II includes the literature review on emotion recognition, data augmentation and the most commonly used datasets in the literature. Section III consists of a detailed explanation of the proposed method. Section IV includes the dataset used in the study experimental studies, their results, and a comparison of its with other studies, Section V provides results for the conclusion of the study.

II. RELATED WORKS

The related works are arranged in threes sub-section. The related works about emotion recognition are detailed in the first sub-section. The second sub-section listed AE works that

used data augmentation. Emotion recognition data sets are briefly described in the final sub-section.

A. LITERATURE REVIEW ON EMOTION RECOGNITION

The authors of [16] wanted to see if brain signals might be exploited to discern emotions. 28 people were asked to play five-minute games with an EMOTIV EPOC + wearable 14-channel EEG interface. These games monitored participants' EEG data for 20 minutes and contained four emotions: boredom, calm, horror, and fear. The discrete wavelet transform (DWT) was used to analyze the signals' time frequency. A Second-order Daubechies filter was used to recover D1-D4 data and A4 approximation coefficients. Feature extraction included detrended fluctuation analysis, Shannon entropy, standard deviation, variance, and zero transitions. The EEG channels were categorized using Support Vector Machine (SVM), K-Nearest Neighbors (k-NN), and multilayer perceptron neural network (MLPNN) classifiers based on positive-negative emotion prediction and arousal-valence dimension conditions. The best classification accuracy rates were 75.0 for k-NN, 72.2 for SVM, and 82.2 for MLPNN. In [5], the researchers used a 32-channel EEG equipment to capture signals from 44 subjects to establish a new dataset. The EEG signals were recorded throughout 12 films, three of which were happy, scary, sad, or neutral. The signals are then adjusted to zero mean and unit variance using a DWT. The retrieved parameters such as average amplitude change, absolute square root sum, and root mean square were used to classify the four emotions. The FP1-F7 channel's gamma sub-band had the best ELM performance, with 94.7% accuracy. [17] used an online semi-supervised learning method to recognize emotions from EEG signals. 14-channel EEG data from 28 subjects playing four distinct video games were analyzed using the Fourier spectrum. The collected features showed the Evolving Gaussian Fuzzy Classification (eGFC)'s efficiency in real-time learning of EEG data, with 72.2 percent performance for four category classifications using the arousal-valence method. According to [9], BCI technologies are employed as an interface between sensors and the brain. A Spiking Neural Network (SNN) was used to analyze the Database for Emotion Analysis using Physiological Signals (DEAP) data [18] and 60 EEG samples. The method is recommended because the SNN neuron structure is more realistic than the Artificial Neural Network (ANN). On average, 84.6% of the valence mood level was correctly identified using the SNN architecture NeuCube. [19] created an EEG-based emotion identification system using fractal pattern feature extraction. The 14-channel GAMEEMO data collection was decomposed using a fractal design and Tunable Q-factor Wavelet Transform (TQWT). An Iterative Chi-square Selector (IChi2) was utilized for feature extraction. The model was tested using 10-fold cross-validation using linear discriminant analysis (LDA), k-NN, and SVM. The SVM classifier had the greatest accuracy of 99.8%. [20] proposed an arousal-valence-based real-time emotion classification system for four emotional classes. DEAP's 10-channel EEG recordings were acquired by first dividing them into overlapping intervals of 2-4 seconds duration. With k-NN (k = 3), arousal classification accuracy is 86.8% and valence is 84.1. The study also claims it is more accurate than higher frequency bands, especially the gamma band. The authors used EEG signals from their own GAMEEMO data set in [21]. The data were utilized to distinguish positive and negative emotions. The investigation began with determining the signals' spectral entropy. Then the classifier gets the values. The deep bidirectional long-short term memory (BiLSTM) architecture was employed as a classifier. The approach yielded 76.91 percent accuracy and a 90% Receiver operating characteristic (ROC) score.

B. LITERATURE REVIEW ON DATA AUGMENTATION

In [22], three alternative methods were used to attempt to tackle the problem of insufficient data in EEG emotion recognition: VAE, GAN, and classical. According to the authors, they achieved the best results with GAN in trials using the DEAP [18] dataset and Shanghai Jiao Tong University Emotion EEG Dataset (SEED) [23]. They improved performance by over 10% during the test phase of their networks trained with SVM and deep architectures. In [24], they aimed to generate more data to increase the performance of the classifier. They proposed a model that duplicates data from ELM-AE images in their paper. They stated that they chose the auto-encoder approach over other data augmentation strategies because it was simpler and more efficient. They tested their methods on The Japanese Female Facial Expression (JAFFE) dataset [25], which contains Japanese female facial expressions, and looked at the impact of data enhancement on results, using k-NN, SVM, and Convolutional Neural Networks (CNN). They stressed that their approach was a viable alternative for data enhancement tasks and that it produces better results in most cases than other common strategies, according to the findings. In [26], a data enhancement and feature extraction method using a VAE for acoustic modeling is described. The authors declared that the VAE was a helpful model based on variational Bayesian learning and a deep learning framework. A VAE can generate new information by extracting hidden values of input variables. VAE was a popular method for building images and sentences. A VAE was used to improve speech structure data for acoustic modeling and feature vector extraction from their research. The size of a speech ensemble was doubled by using a VAE system to encode the hidden variables extracted from the original expressions. Latent variables inferred from speech waveforms are said to have concealed ''meanings'' of the waveforms, allowing them to be used as acoustic properties for automatic speech recognition (ASR). They used a VAE system to show the efficacy of data augmentation and that latent variable-based features can be used in ASR in their research. In [27], a two-stage model is proposed to improve the recognition rate by examining a data set of documents containing Japanese letters. The model's first step was to figure out how to distribute data and fit the shape vectors of

the characters on the page, while the second is to generate new examples. The study's VAE structure was created to divide data diversity into regions, create simple examples of in-class multi-modality, and avoid mode reduction. They accomplished this by organizing the VAE model and proposing a gradual and unregulated feature extraction structure for the VAE model. The CNN-based classification network achieved 94.02% efficiency for the non-multiplexed data set, while this rate could be improved to 95.56% for the multiplexed data set. It is viewed as the study's focal point, where the recognition rate was increased by using an enriched data collection.

C. DATASETS ABOUT EMOTION RECOGNITION

With three primary stimuli aspects, it is possible to manage the emotion recognition system, whose research and the application field are expanding by the day. Audio, visual, and audio-visual are the three types. These categories were used to build datasets, and the studies' accuracy was comparable.

- The Belfast facial expression dataset was developed to explore gender, cultural, and individual variations in emotion interpretation from TV shows and interview recordings [28].
- The Human-Machine Interaction Network on Emotion (HUMAINE) dataset has been developed, including various scenarios in terms of emotion recognition, audio-visual recordings, and an expanded version of Belfast [29].
- The MAHNOB physiological dataset was developed, including sound signals, mimics, and EEG signals [30].
- A 32-channel ECG and a MAHNOB HCI dataset containing EEG signals, which evaluate the participants' feelings after the movie according to the valence-arousal scale, were created [31].
- The Interactive Emotional Dyadic Motion Capture (IEMOCAP) audio-visual dataset was developed, which communicates emotional states such as happiness, frustration, sadness, disappointment, and neutrality and was collected from participants due to a double session [32].
- The VAM (Vera Am Mittag) dataset was generated using audio-visual recordings, including the participants' natural responses during a TV show [33].
- The DEAP audio-visual dataset was created based on 32-channel valence-arousal scale using music clips in the environment [22].
- The eNTERFACE dataset has been developed, which contains audio-visual recordings from various countries and includes tags for enjoyment, rage, sadness, surprise, disgust, and panic [34].

III. PROPOSED METHOD

In this paper, a novel approach is proposed for EEG-based emotion recognition. Fig. 1 shows the arousal-valence emotion model that is considered. As seen in Fig.1, excited, happy, and pleased emotions are in the region that is indicated by high-arousal-positive-valence (HAPV). Annoying, angry, and nervous emotions are in the region high-arousal-negative-valence (HANV) and similarly, sad, bored and sleepy emotions are located in the region low-arousal-negative-valence

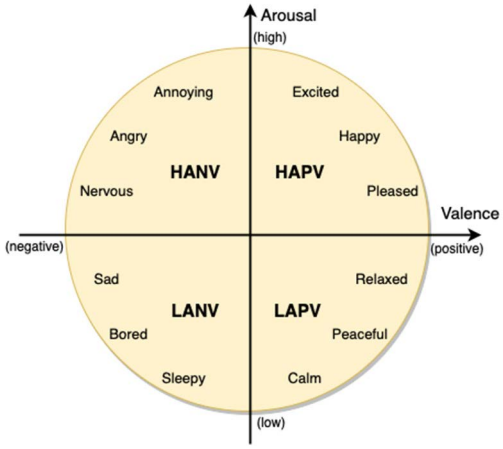


FIGURE 1. Arousal-valence emotion model.

(LANV). Lastly, calm, peaceful, and relaxed emotions are in the area low-arousal-positive-valence (LAPV). Through the arousal axis, the feeling of emotions changes from high to low. Likewise, in the valence axis, the emotions are changed trough negative to positive. In addition, the four games labeled boring (B), calm (C), horror (H), and funny (F) in the dataset used correspond to LANV, LAPV, HANV, and HAPV, respectively.

In Fig.2, the illustration of the ELM-W-AE method is given. As seen in Fig. 2, the input EEG signals are initially

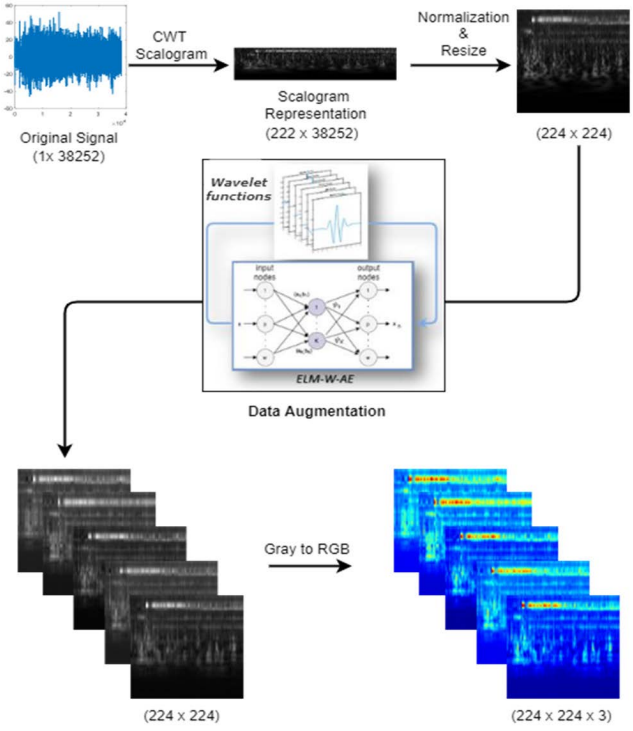


FIGURE 2. An example of EEG to CWT scalogram images and ELM-W-AE data augmentation.

converted to CWT scalogram images. The following formula is used to transform a function $x\,(t)$ given a mother wavelet $\psi\,(t)$ through CWT:

$$X\,(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} \psi\left(\frac{t-b}{a}\right) x\,(t)\,dt \tag{1}$$

where a denotes the scale or dilation parameter related with frequency, and b is the shifting parameter that denotes the time information in the transform [35].

The wavelet used for CWT is the analytic Morse wavelet as it has better time-frequency localization. For Morse wavelet symmetry parameter (gamma) and time-bandwidth product were kept at 3 and 60 respectively. Voices per octave were kept 10 [36].

The analytic Morse wavelet is utilized for CWT because it has greater time-frequency localization [37]. The symmetry parameter (gamma) and time-bandwidth product for Morse wavelets were preserved at 3 and 60, respectively. The number of voices per octave was chosen as 18 which is best in case of our EEG emotion recognition experiments. Thus, a scalogram image 222 × 38252 sized was constructed from an EEG signal.

Then, the scalogram images are normalized and resized to 224×224 for being appropriate for the input of the next building block. Initially, the scalogram images are in grayscale and are then converted to the color images by assigning the grayscale image in the color channels of a new image. After the previous procedure, the dataset is constructed. A data augmentation procedure comes after the data construction. To this end, the ELM-W-AE is considered. Various wavelet kernels are used in ELM-AE architecture. The wavelet functions are briefly explained in Section III.A.

After data augmentation, deep transfer learning is used in the classification stage of the proposed method. The pre-trained ResNet18 model, which has 18 layers, is further trained (fine-tuned) in the classification procedure. The flow of the process is given in Fig. 3.

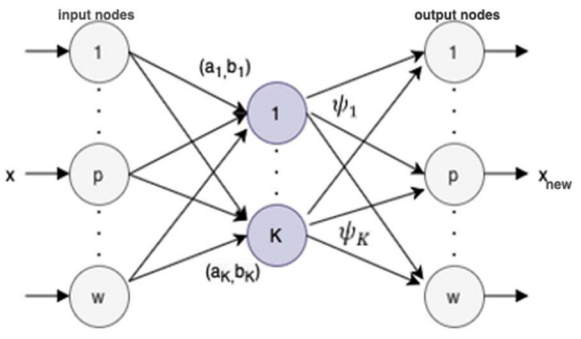


FIGURE 3. ELM-AE architecture representation.

A. WAVELET BASED EXTREME LEARNING MACHINE AUTO-ENCODER

The AE is an unsupervised learning system in which the input data are often used as output data [40]. It is made up of two parts: encoder and decoder. The input data are projected to the hidden layer in the encoder section, and an estimate of the input data is obtained in the decoder part. In the AE, the input $X_{N\times w}$ is initially encoded to a higher-level space and then an approximation of the input $X'_{N\times w}$ is obtained by using the encoded input X. Figure 4 shows the architecture of the ELM-AE structure.

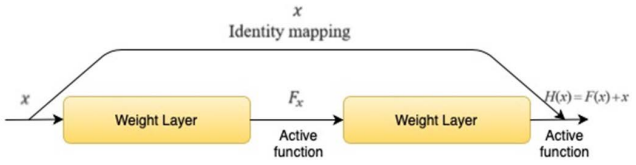


FIGURE 4. The residual unit.

By using the wavelet kernel ψ, the hidden layer output matrix H_ψ of size N × K is re-defined as:

$$H_\psi = \begin{bmatrix} \psi\left(a_1^T x_1 + b_1\right) & \dots & \psi(a_K^T x_1 + b_K) \\ \vdots & \ddots & \vdots \\ \psi(a_1^T x_N + b_1) & \dots & \psi(a_K^T x_N + b_K) \end{bmatrix}_{N\times K} \tag{2}$$

where x is inputs of ELM-AE, a and b denote the input weights and biases of hidden neurons, respectively. N denotes the training data number and K symbolizes the number of the hidden neurons. The Morlet, Gaussian, Mexican, Shannon, Meyer and GGW wavelet activation functions are defined in Table 1 [41], [42].

TABLE 1. Wavelet kernel types and their functions.

Wavelet Kernel Type	Function
Morlet	$\psi(t) = \cos(1.75t)exp(-\frac{t^2}{2})$
Gaussian	$\psi(t) = \frac{1}{\sqrt{2\pi}}exp(-\frac{t^2}{2})$
Mexican	$\psi(t) = c(1-t^2)exp\left(\frac{t^2}{2}\right), c = \frac{2}{\sqrt{3}}\pi^{(-\frac{1}{4})}$
Shannon	$\psi(t) = \frac{sin\pi(t-1/2)-sin2\pi(t-1/2)}{\pi(t-1/2)}$
Meyer	$\psi(t) = 35t^4 - 84t^5 + 70t^6 - 20t^7$
GGW	$\psi(t) = \sin(3t) + \sin(0.3t) + \sin(0.03t)$

Similar to the ELM algorithm, the output weights β of the ELM-W-AE are calculated by using the Moore-Penrose inverse [37], [42]:

$$\beta = H'X \tag{3}$$

where H' is the Moore-Penrose inverse of H and X is given input data.

B. ResNet18 ARCHITECTURE

In image classification, ResNet18, which was introduced for the problem of performance degradation with increasing depth, is frequently preferred [43]. ResNet18 could develop deep network architectures with this property by using residual units depicted in Figure 5.

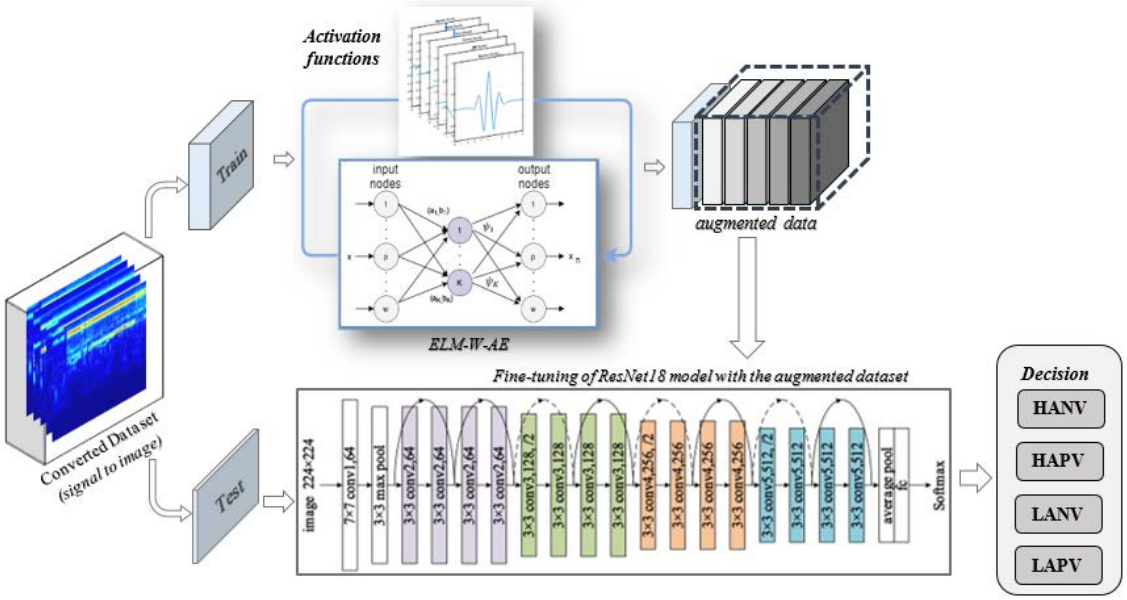


FIGURE 5. The illustration of the proposed method.

Assuming that the neural network's input parameter is x and the target output is h(x), the target output is likely to be extremely complex. In this case, residual F(x) = H(x)-x changes the goal output to F(x) + x to avoid the performance deterioration problem caused by too many convolution layers. This is referred to as a short link. These mentioned linkages, which can perform identity matching by avoiding two or more levels, are defined as [44];

$$x_{l+1} = f[x_l + F(x_l, k_l)] \tag{4}$$

In this equation, x_l and x_{l+1} represent the input and output of the first residual unit, respectively, the active function f, the residual function F, and the convolution kernel k.

IV. EXPERIMENTAL WORKS AND RESULTS

A. GAMEEMO DATASET

The GAMEEMO dataset was developed by Alakus *et al.* using EEG signals collected from participants while playing video games using a wearable-portable system with 14 EEG channels [16]. The GAMEEMO dataset will be used to test the methods used in the analysis to achieve the emotion recognition objective. EEG signals were collected with the EMOTIV EPOC + Mobile EEG device from 28 students between the ages of 20-27 in the Software Engineering Department of Firat University Faculty of Technology. On the scalp, EEG electrodes were placed in 16 different regions as AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, AF4, P3, and P4. Since P3 and P4 are reference electrodes, a 14-channel EEG device is used. The device's sampling rate was set to 128 Hz, and the signal bandwidth was set to 0.16 - 43 Hz. The participants played four different computer games categorized as boring, calm, horror, and funny for five minutes each, yielding a total of 20 minutes of EEG data from each participant.

The EEG signal is containing 38,252 samples within the period. In addition, the four games labeled boring (B), calm (C), horror (H) and funny (F) in the dataset used correspond to LANV, LAPV, HANV and HAPV, respectively.

B. RESULTS

The experimental works were carried out on MATLAB. The 14-channel EMOTIV EPOC+, a wearable and compact EEG unit, was used to capture EEG signals from 28 different subjects. Subjects played four separate video games for five minutes each to catch emotions (boring, quiet, horror, and funny), with a total of 20 minutes of EEG data available for each subject. The participants scored every video game on a scale of arousal and valence. The participants used the Self-Assessment of Manikin (SAM) type to score each video game on a scale of arousal and valence [16]. During the scalogram image construction, the scale parameter of the CWT is chosen empirically as twelve. The obtained dataset was randomly divided into two parts, where 70% of it was used for training, and the rest 30%, was used for testing the proposed approach and stayed constant in the process of the experiments. The training images only were augmented with ELM-W-AE by using Gauss, GGW, Mexican, Meyer, Morlet, and Shannon wavelet functions.

The training images increased by six times, which is including the originals. The ResNet18 was fine-tuned by using the stochastic gradient descent optimizer method (SGDM), where the input batch size, number of epochs, and the initial learning rate parameters were set to 32, 30, and 0.05, respectively. The achievement of the proposed

method was evaluated based on the various evaluation metrics, such as accuracy, sensitivity, specificity, precision, F1-Score, Mathew correlation coefficient (MCC), and Kappa [45], [46].

Also, ROC curve, which is defined as a plot of test sensitivity or true positive rate (TPR) as the y-axis versus its 1-specificity or false positive rate (FPR) as the x-axis, is a useful tool for assessing the quality or performance of medical diagnostic tests. It's widely used in radiology to assess the performance of many classifiers tests [47].

Area Under the ROC Curve (AUC) is a measure of a diagnostic test's overall performance, defined as the average value of sensitivity for all conceivable values of specificity. AUC can have any value between 0 and 1, with a higher value indicating greater overall diagnostic test performance [48].

The obtained results were represented in Table 2. The columns of Table 2 show the performance metrics, and the rows show the methods that were used. In the first row of Table 2, the results, which were obtained without data augmentation, were given. As seen in Table 2, without data augmentation, 77.66% accuracy, 77.66% sensitivity, 92.55% specificity, 77.96% precision, 0.78 F1-score, 0.70 MCC and 0.40 Kappa values were obtained. In the other rows of Table 2, the results with data augmentation were given. As seen in the other rows, various activation functions were examined in the data augmentation case.

TABLE 2. The classification results for original and augmented images.

Method	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1.	MCC	Kap.
Original	77.66	77.66	92.55	77.96	0.78	0.70	0.40
Gauss	98.72	98.72	99.57	98.71	0.99	0.98	0.97
GGW	99.57	99.62	99.86	99.53	0.99	0.99	0.99
Mexican	98.51	98.52	99.51	98.45	0.98	0.98	0.96
Meyer	99.36	99.44	99.78	99.42	0.99	0.99	0.98
Morlet	98.51	98.69	99.52	98.37	0.98	0.98	0.96
Shannon	98.30	98.37	99.4	98.28	0.98	0.98	0.96

Acc.= Accuracy, Sens.= Sensitivity, Spec.= Specificity, Prec.= Precision, F1= F1-Score, Kap.=Kappa

When one examined the ELM-W-AE based data augmentation of results, obvious performance improvements were noticeable. The improvement ratio was almost over 20% for all activation functions. The best evaluation scores, 99.574 % accuracy, 99.621% sensitivity, 99.862% specificity, 99.53% precision, and 0.996 F1- score, and 0.994 MCC and 0.989 Kappa values obtained for the GGW activation function. Besides, the Meyer activation function produced the second-best evolution scores, where 99.362% accuracy, 99.444% sensitivity, 99.780% specificity, 99.421% precision, and a 0.994 F1-score 0.992 MCC and 0.983 Kappa values were obtained. It is worth mentioning that the Shannon kernel produced the worst evaluation metrics.

In Fig. 6 and 7, the ROC curve presentation and the confusion matrix were given for the original dataset (without data augmentation). In Fig. 6, while the x-axis shows the false-positive rates, the y-axis shows the true-positive rates. Each ROC curve shows an emotion where different colors were used.

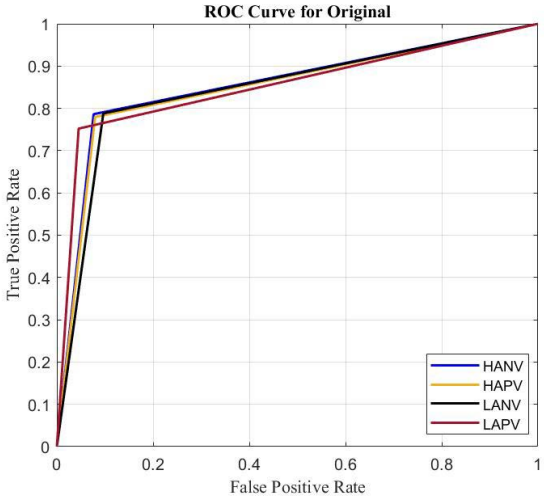


FIGURE 6. The ROC curves for original images.

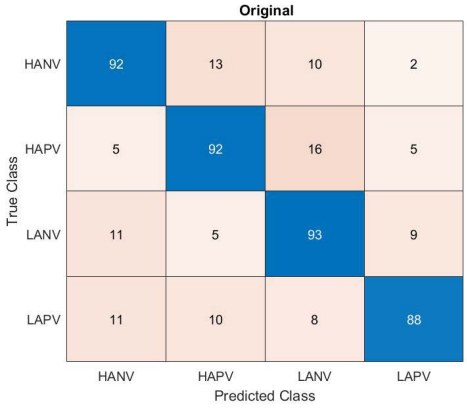


FIGURE 7. The confusion matrix for original images.

As seen in Fig. 6, all ROC curves raised through to the 0.8 true positive rate value in the 0-0.1 range of the false positive rate. Then, they went to the one true positive rate value when the false positive rate was one.

As mentioned earlier, Fig. 7 shows the confusion matrix for the original dataset. For the confusion matrix, the rows show the true class, and the columns show the predicted class. As seen in Figure 7; 92, 92, 93, and 88, test samples from the HANV, HAPV, LANV, and LAPV classes were correctly predicted, respectively. In addition, 27, 28, 34, and 16 test samples were wrongly predicted for the HANV, HAPV, LANV, and LAPV classes, respectively.

ROC curves for the HANV, HAPV, LANV, and LAPV classes are given in Fig. 8. The GGW activation function was used in ELM-W-AE. Besides, the confusion matrix for GGW activation function is given in Fig. 9. The illustrations for GGW activation function are given as that activation function

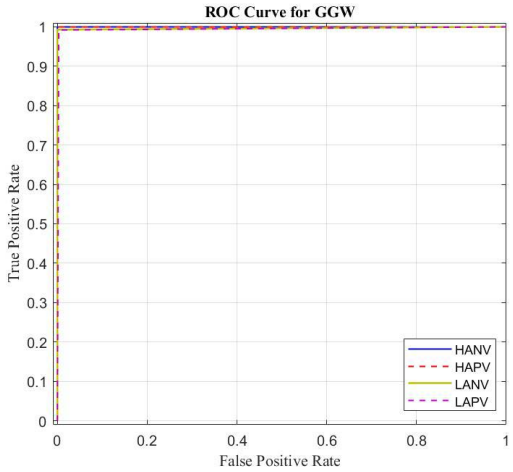


FIGURE 8. The ROC curves for data augmentation with GGW activation function.

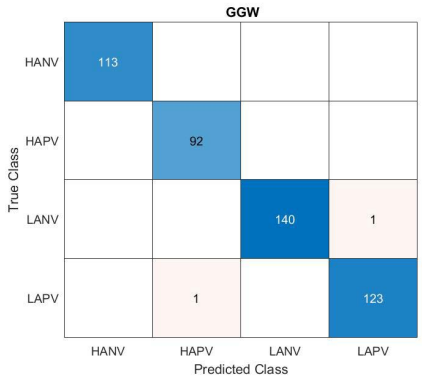


FIGURE 9. The confusion matrix for data augmentation with GGW activation function.

TABLE 3. Individual class results for original images.

Original	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1.	MCC	Kappa
HANV	78.63	78.63	92.35	77.31	0.78	0.71	0.53
HAPV	77.97	77.97	92.05	76.67	0.77	0.70	0.53
LANV	78.81	78.81	90.34	73.23	0.76	0.68	0.52
LAPV	75.21	75.21	95.47	84.62	0.79	0.74	0.56

yielded the best performance among the other activation functions. When Fig. 8 is examined, it is seen the false positive rate is around zero, while the curves increased by one true positive rate for all four classes. This shows an huge area under ROC, which means high performance. It was observed that the ROC curves immediately rose along the true positive velocity axis for all classes, and almost all classes were classified with 100% accuracy, creating a nearly 90-degree gap with the false-positive velocity axis.

When the confusion matrix that is given in Fig. 9 is analyzed, it is seen that HANV and HAPV classes were classified with 100% accuracy scores. Besides, for LANV and LAPV classes one sample in each was wrongly classified.

TABLE 4. Individual class results for ELM-W(Gauss)-AE images.

Gauss	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1.	MCC	Kappa
HANV	100.0	100.0	99.44	98.26	0.99	0.99	0.52
HAPV	97.83	97.83	99.74	98.90	0.98	0.98	0.61
LANV	97.87	97.87	99.69	99.28	0.99	0.98	0.41
LAPV	99.19	99.19	99.42	98.40	0.99	0.98	0.47

TABLE 5. Individual class results for ELM-W(GGW)-AE images.

GGW	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1.	MCC	Kappa
HANV	100.0	100.0	100.0	100.0	1.00	1.00	0.52
HAPV	100.0	100.0	99.74	98.93	0.99	0.99	0.61
LANV	99.29	99.29	100.0	100.0	0.99	0.99	0.40
LAPV	99.19	99.19	99.71	99.19	0.99	0.99	0.47

TABLE 6. Individual class results for ELM-W(Mexican)-AE images.

Mexican	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1.	MCC	Kappa
HANV	100.0	100.0	99.16	97.41	0.99	0.98	0.52
HAPV	97.83	97.83	99.47	97.83	0.98	0.97	0.61
LANV	97.87	97.87	99.39	98.57	0.98	0.98	0.41
LAPV	98.39	98.39	100.0	100.0	0.99	0.99	0.48

TABLE 7. Individual class results for ELM-W (Meyer)-AE images.

Meyer	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1.	MCC	Kappa
HANV	100.0	100.0	100.0	100.0	1.00	1.00	0.52
HAPV	100.0	100.0	100.0	100.0	1.0	1.0	0.61
LANV	98.58	98.58	99.69	99.29	0.98	0.98	0.41
LAPV	99.19	99.19	99.42	98.40	0.98	0.98	0.47

TABLE 8. Individual class results for ELM-W (Morlet)-AE images.

Morlet	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1.	MCC	Kappa
HANV	99.12	99.12	99.44	98.25	0.98	0.98	0.52
HAPV	100.0	100.0	99.27	96.84	0.98	0.98	0.60
LANV	96.45	96.45	100.0	100.0	0.98	0.98	0.42
LAPV	99.19	99.19	99.42	98.40	0.98	0.98	0.47

In Tables 3-9, the evaluation metrics for each class for original and the examined activation functions are given. While Table 3 shows the results for the original dataset, Tables 4-9 show the performance evaluation metrics for an individual class for Gauss, GGW, Mexican, Meyer, Morlet, and Shannon activation functions, respectively. In Table 3, the results for the original dataset are given. As seen in Table 3, 78.63%, 77.97%, 78.81%, and 75.21% correct classification

TABLE 9. Individual class results for ELM-W (Shannon)-ae images.

Shannon	Acc. (%)	Sens. (%)	Spec. (%)	Prec. (%)	F1.	MCC	Kappa
HANV	99.12	99.12	98.88	96.55	0.98	0.97	0.52
HAPV	98.91	98.91	99.74	98.91	0.98	0.98	0.61
LANV	97.87	97.87	99.69	99.28	0.98	0.98	0.41
LAPV	97.58	97.58	99.42	98.37	0.98	0.97	0.48

rates are obtained for HANV, HAPV, LANV, and LAPV classes, respectively. Table 3 also gives the other evaluation metrics for the mentioned individual classes.

Table 4 indicated the evaluation scores for each class for the Gauss activation function. As seen in Table 4, data augmentation highly improved the obtained results, where 100%, 97.83%, 97.87% and 99.19% correct classification rates were obtained for HANV, HAPV, LANV and LAPV classes, respectively. The other evaluation metrics were also increased when the ELM-W-AE based data augmentation was considered.

Table 5 indicated the evaluation scores for each class for the GGW activation function. As seen in Table 5, using the GGW activation function, the results of the original dataset were highly improved. Besides, the GGW activation function produced better results than the Gauss activation function. 100%, 100%, 99.29%, and 99.19% correct classification rates were obtained for HANV, HAPV, LANV, and LAPV classes, respectively.

Table 6 indicated the evaluation scores for each class for the Mexican activation function. Using Mexican activation function, 100%, 97.82%, 97.87%, and 98.39% correct classification rates were obtained for HANV, HAPV, LANV, and LAPV classes, respectively. The results obtained by the Mexican activation function were also worse than the achievements of the GGW activation function.

Table 7 shows the evaluation scores for each class for the Meyer activation function. The Meyer activation function produced similar scores to the GGW activation function where 100%, 100%, 98.58%, and 99.19% correct classification rates were obtained for HANV, HAPV, LANV, and LAPV classes, respectively.

Table 8 shows the evaluation scores for each class for the Morlet activation function. With the Morlet activation function 99.12%, 100%, 96.45%, and 99.19%, correct classification rates were obtained for HANV, HAPV, LANV, and LAPV classes.

Lastly, Table 9 shows the evaluation scores for each class for the Shannon activation function. By using the Shannon activation function, 99.12%, 98.91%, 97.8% and 97.58% correct classification rates were obtained for HANV, HAPV, LANV, and LAPV classes, respectively.

C. COMPARISON WITH OTHER METHODS

Existing studies using the same dataset have been compared to our suggested technique in this section. Alakus *et al.* [16] proposed the GAMEEMO dataset used in this study. We compared three research publications utilizing the data set with our technique [16], [19], [21]. Table 10 shows performance comparison results.

TABLE 10. Studies with the same dataset.

References	Method and classifiers	Accuracy
Alakus et al. [16]	DWT, SVM, k-NN, MLPNN	75.0% (k-NN), 72.2% (SVM) and 82.2% (MLPNN)
Tuncer et al. [19]	FFP, TQWT, k-NN, LDA, and SVM	98.3% (k-NN), 87.2% (LDA), and 98.9% (SVM)
Alakus et al. [21]	Spectral entropy values, BiLSTM	76.9%
Proposed Method	**Wavelet ELM-AE with GGW kernel and ResNet18**	**99.6%**

Alakus *et al.* [16] was investigated classification accuracy at the channel level and obtained that the average accuracy rates for k-NN, SVM, and MLPNN were 75.0%, 72.2%, and 82.2%, respectively. Tuncer *et al.* [19] reported an average accuracy of 98.3% for k-NN, 87.2% for LDA, and 98.9% for SVM in their channel-based research. Alakus *et al.* [21] obtained 76.9% success with the Bidirectional long short-term memory (BiLSTM) approach in their subsequent analysis with the same dataset. Our method fared better classification accuracy when compared to the average accuracy rates of the three current approaches. Our research with the ELM-AE structure, combined with wavelet function types individually, yielded a classification success rate of 99.6% using the ResNet18 design.

V. CONCLUSION

In this paper, a novel approach was proposed for data augmentation. The proposed method was based on the ELM-W-AE. Various activation functions abilities were examined for the EEG-based emotion classification. The following conclusions are inferred from the experimental works.

1-) Initially, it was observed that the data augmentation was quite effective in EEG-based emotion classification. With data augmentation, almost 20% improvement in accuracy score was observed.

2-) When the achievements of the various activation functions were examined in the ELM-W-AE structure, it was seen that GGW activation function produced the best evaluation scores.

3-) Except Morlet and Shannon activation functions, for all other activation functions, the HANV class was classified with a 100% accuracy score. As it is seen that the best emotion prediction is for the HANV class, which includes annoying, angry, and nervous emotions. It is seen that the next best class prediction is in the HAPV class, which includes excited, happy, and pleased emotions.

REFERENCES

[1] K. Michalopoulos and N. Bourbakis, "Application of multiscale entropy on EEG signals for emotion detection," in *Proc. IEEE EMBS Int. Conf. Biomed. Health Informat. (BHI)*, Feb. 2017, pp. 341–344.

[2] P. Ekman, "An argument for basic emotions," *Cognition Emotion*, vol. 6, nos. 3-4, pp. 169–200, 1992.

[3] J. A. Russell, "Culture and the categorization of emotions," *Psychol. Bull.*, vol. 110, no. 3, pp. 426–450, 1991.

[4] R. Plutchik, "The nature of emotions: Human emotions have deep evolutionary roots, a fact that may explain their complexity and provide tools for clinical practice," *Amer. Scientist*, vol. 89, no. 4, pp. 344–350, 2001.

[5] A. Seal, P. P. N. Reddy, P. Chaithanya, A. Meghana, K. Jahnavi, O. Krejcar, and R. Hudak, "An EEG database and its initial benchmark emotion classification performance," *Comput. Math. Methods Med.*, vol. 2020, Aug. 2020, Art. no. 8303465.

[6] S. Katsigiannis and N. Ramzan, "DREAMER: A database for emotion recognition through EEG and ECG signals from wireless low-cost off-the-shelf devices," *IEEE J. Biomed. Health Inform.*, vol. 22, no. 1, pp. 98–107, Jan. 2018.

[7] R. M. Mehmood, H. J. Yang, and S. H. Kim, "Children emotion regulation: Development of neural marker by investigating human brain signals," *IEEE Trans. Instrum. Meas.*, vol. 70, pp. 1–11, 2020.

[8] S. H. Kim, H. J. Yang, N. A. T. Nguyen, R. M. Mehmood, and S. W. Lee, "Parameter estimation using unscented Kalman filter on the gray-box model for dynamic EEG system modeling," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 9, pp. 6175–6185 Jan. 2020.

[9] A. Al-Nafjan, K. Alharthi, and H. Kurdi, "Lightweight building of an electroencephalogram-based emotion detection system," *Brain Sci.*, vol. 10, no. 11, p. 781, 2020.

[10] M. M. Najafabadi, F. Villanustre, T. M. Khoshgoftaar, N. Seliya, R. Wald, and E. Muharemagic, "Deep learning applications and challenges in big data analytics," *J. Big Data*, vol. 2, no. 1, pp. 1–21, 2015.

[11] L. L. C. Kasun, H. M. Zhou, G. B. Huang, and C. M. Vong, "Representational learning with ELMs for big data," *IEEE Intell. Syst.*, vol. 28, no. 6, pp. 31–34, Nov. 2013.

[12] F. Wang, S. H. Zhong, J. Peng, J. Jiang, and Y. Liu, "Data augmentation for EEG-based emotion recognition with deep convolutional neural networks," in *Proc. Int. Conf. Multimedia Model.*, Feb. 2018, pp. 82–93.

[13] S. Hussein, R. Gillies, K. Cao, Q. Song, and U. Bagci, "TumorNet: Lung nodule characterization using multi-view convolutional neural network with Gaussian process," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 1007–1010.

[14] W.-N. Hsu, Y. Zhang, and J. Glass, "Unsupervised domain adaptation for robust speech recognition via variational autoencoder-based data augmentation," in *Proc. IEEE Autom. Speech Recognit. Understand. Workshop (ASRU)*, Dec. 2017, pp. 16–23.

[15] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, "Synthetic data augmentation using GAN for improved liver lesion classification," in *Proc. IEEE 15th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2018, pp. 289–293.

[16] T. B. Alakus, M. Gonen, and I. Turkoglu, "Database for an emotion recognition system based on EEG signals and various computer games—GAMEEMO," *Biomed. Signal Process. Control*, vol. 60, Jul. 2020, Art. no. 101951.

[17] D. Leite, V. Frigeri, Jr., and R. Medeiros, "Adaptive Gaussian fuzzy classifier for real-time emotion recognition in computer games," 2021, *arXiv:2103.03488*.

[18] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "DEAP: A database for emotion analysis; Using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, Jun. 2012.

[19] T. Tuncer, S. Dogan, and A. Subasi, "A new fractal pattern feature generation function based emotion recognition method using EEG," *Chaos, Solitons Fractals*, vol. 144, Mar. 2021, Art. no. 110671.

[20] Z. Mohammadi, J. Frounchi, and M. Amiri, "Wavelet-based emotion recognition system using EEG signal," *Neural Comput. Appl.*, vol. 28, no. 8, pp. 1985–1990, Aug. 2017.

[21] T. B. Alakus and I. Turkoglu, "Emotion recognition with deep learning using GAMEEMO data set," *Electron. Lett.*, vol. 56, no. 25, pp. 1364–1367, 2020.

[22] Y. Luo, L.-Z. Zhu, Z.-Y. Wan, and B.-L. Lu, "Data augmentation for enhancing EEG-based emotion recognition with deep generative models," *J. Neural Eng.*, vol. 17, no. 5, Oct. 2020, Art. no. 056021.

[23] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "EmotionMeter: A multimodal framework for recognizing human emotions," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1110–1122, Mar. 2019.

[24] J. Ferreira, M. Ferro, B. Fernandes, M. Valenca, C. Bastos-Filho, and P. Barros, "Extreme learning machine autoencoder for data augmentation," in *Proc. IEEE Latin Amer. Conf. Comput. Intell. (LA-CCI)*, Nov. 2017, pp. 1–6.

[25] M. J. Lyons, S. Akamatsu, M. Kamachi, J. Gyoba, and J. Budynek, "The Japanese female facial expression (JAFFE) database," in *Proc. 3rd Int. Conf. Autom. Face Gesture Recognit.*, Apr. 1998, pp. 14–16.

[26] H. Nishizaki, "Data augmentation and feature extraction using variational autoencoder for acoustic modeling," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Dec. 2017, pp. 1222–1227.

[27] G. Cao and S.-I. Kamata, "Data augmentation for historical documents via cascade variational auto-encoder," in *Proc. IEEE Int. Conf. Signal Image Process. Appl. (ICSIPA)*, Sep. 2019, pp. 340–345.

[28] E. Douglas-Cowie, N. Campbell, R. Cowie, and P. Roach, "Emotional speech: Towards a new generation of databases," *Speech Commun.*, vol. 40, nos. 1–2, pp. 33–60, 2003.

[29] P. Petta, C. Pelachaud, and R. Cowie, *Emotion-Oriented Systems: The HUMAINE Handbook*. Cham, Switzerland: Springer, 2011.

[30] S. Petridis, B. Martinez, and M. Pantic, "The MANHOB laughter database," *Image Vis. Comput.*, vol. 31, no. 2, pp. 186–202, 2013.

[31] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 1–14, Jan. 2012.

[32] C. Busso, M. Bulut, C.-C. Lee, A. Kazemzadeh, E. Mower, S. Kim, J. N. Chang, S. Lee, and S. S. Narayanan, "IEMOCAP: Interactive emotional dyadic motion capture database," *Lang. Resour. Eval.*, vol. 42, no. 4, pp. 335–359, Dec. 2008.

[33] M. Grimm, K. Kroschel, and S. Narayanan, "The Vera am Mittag German audio-visual emotional speech database," in *Proc. IEEE Int. Conf. Multimedia Exp.*, Jun. 2008, pp. 865–868.

[34] O. Martin, I. Kotsia, B. Macq, and I. Pitas, "The eNTERFACE'05 audio-visual emotion database," in *Proc. 22nd Int. Conf. Data Eng. Workshops (ICDEW)*, Atlanta, GA, USA, 2006, pp. 1–8.

[35] R. Mahamune and S. H. Laskar, "Classification of the four-class motor imagery signals using continuous wavelet transform filter bank-based two-dimensional images," *Int. J. Imag. Syst. Technol.*, vol. 31, no. 4, pp. 2237–2248, Dec. 2021.

[36] S. Chaudhary, S. Taran, V. Bajaj, and A. Sengur, "Convolutional neural network based approach towards motor imagery tasks EEG signals classification," *IEEE Sensors J.*, vol. 19, no. 12, pp. 4494–4500, Jun. 2019.

[37] G. B. Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: A new learning scheme of feedforward neural networks," in *Proc. IEEE Int. Joint Conf. Neural Netw.*, vol. 2, Jul. 2004, pp. 985–990.

[38] A. Güner, Ö. F. Alçin, and A. Sengür, "Automatic digital modulation classification using extreme learning machine with local binary pattern histogram features," *Measurement*, vol. 145, pp. 214–225, Oct. 2019.

[39] B. Ari, A. Ari, A. Sengur, and S. A. Tuncer, "Classification of apricot leaves with extreme learning machines using deep features," in *Proc. 1st Int. Informat. Softw. Eng. Conf. (UBMYK)*, Nov. 2019, pp. 1–5.

[40] G. Altan and Y. Kutlu, "Hessenberg elm autoencoder kernel for deep learning," *J. Eng. Technol. Appl. Sci.*, vol. 3, no. 2, pp. 141–151, Aug. 2018.

[41] G. Wang, L. Guo, and H. Duan, "Wavelet neural network using multiple wavelet functions in target threat assessment," *Sci. World J.*, vol. 2013, pp. 1–7, Feb. 2013.

[42] D. Şengür, "EEG, EMG and ECG based determination of psychosocial risk levels in teachers based on wavelet extreme learning machine autoencoders," *Politeknik Dergisi*, early access, p. 1, 2021.

[43] P. Korfiatis, T. L. Kline, D. H. Lachance, I. F. Parney, J. C. Buckner, and B. J. Erickson, "Residual deep convolutional neural network predicts MGMT methylation status," *J. Digit. Imag.*, vol. 30, no. 5, pp. 622–628, Oct. 2017.

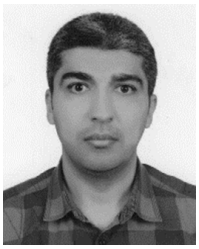
[44] D. Pakhomov, V. Premachandran, M. Allan, M. Azizian, and N. Navab, "Deep residual learning for instrument segmentation in robotic surgery," in *Proc. Int. Workshop Mach. Learn. Med. Imag.* Cham: Springer, Oct. 2019, pp. 566–573.

[45] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, pp. 1–13, 2020.

[46] O. Allouche, A. Tsoar, and R. Kadmon, "Assessing the accuracy of species distribution models: Prevalence, Kappa and the true skill statistic (TSS)," *J. Appl. Ecol.*, vol. 43, no. 6, pp. 1223–1232, Sep. 2006.
[47] K. Hajian-Tilaki, "Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation," *Caspian J. Internal Med.* vol. 4, no. 2, pp. 627–635, 2013.
[48] R. Kumar and A. Indrayan, "Receiver operating characteristic (ROC) curve for medical researchers," *Indian Pediatrics*, vol. 48, no. 4, pp. 277–287, 2011.

BERNA ARI received the B.Sc. degree in computer technology and the M.Sc. degree in electrical and electronics engineering from Firat University, Turkey, in 2013 and 2017, respectively, where she is currently pursuing the Ph.D. degree in electrical and electronics engineering. Her research interests include data augmentation, machine learning, and artificial intelligence.

KAMRAN SIDDIQUE (Senior Member, IEEE) received the Ph.D. degree in computer engineering from Dongguk University, South Korea. He is currently an Associate Professor with Xiamen University Malaysia. His research interests include big data, cybersecurity, machine learning, the IoT, and cloud computing.



ÖMER FARUK ALÇIN received the bachelor's degree in electronics and computers education, the master's degree in electronics education, and the Ph.D. degree in electrical electronics engineering from Firat University, Turkey, in 2007, 2011, and 2015, respectively. He became a Research Assistant with the Technical Education Faculty, Firat University, in February 2009. He is currently an Associate Professor with the Electrical Engineering Department, Malatya Turgut Özal University. His research interests include signal processing, machine learning, medical image processing, and deep learning.

MUZAFFER ASLAN received the B.Sc. degree in electrical and electronics engineering from Gazi University, Turkey, in 1993, and the M.Sc. degree in electronics education and the Ph.D. degree in electrical and electronics engineering from Firat University, Turkey, in 2004 and 2016, respectively. From 1993 to 2018, he worked as a Teacher with the Department of Electrical and Electronics, Technical Anatolian High School, Ministry of National Education. He is currently an Assistant Professor with the Engineering and Architecture Faculty, Bingol University. His research interests include pattern recognition, medical image processing, and computer vision.

ABDULKADIR ŞENGÜR received the B.Sc. degree in electronics and computers education, the M.Sc. degree in electronics education, and the Ph.D. degree in electrical and electronics engineering from Firat University, Turkey, in 1999, 2003, and 2006, respectively. He became a Research Assistant with the Technical Education Faculty, Firat University, in February 2001. He is currently a Professor with the Technology Faculty, Firat University. His research interests include signal processing, image segmentation, pattern recognition, medical image processing, and computer vision.

RAJA MAJID MEHMOOD (Senior Member, IEEE) received the B.S. degree in computer science from Gomal University, Pakistan, the M.S. degree in software technology from Linnaeus University, Sweden, and the Ph.D. degree in computer engineering from the Division of Computer Science and Engineering, Chonbuk National University, South Korea. He is currently an Assistant Professor with the Information and Communication Technology Department, School of Computing and Data Science, Xiamen University Malaysia, Sepang, Malaysia. He worked as a Research Professor with the Department of Brain and Cognitive Engineering, Korea University, Seoul, South Korea. He was a Lecturer with the Software Engineering Department, King Saud University, Saudi Arabia. He has authored more than 50 research papers and is supervising or co-supervising several graduate (M.S./Ph.D.) students. His research interests include affective computing, brain–computer interfaces, information visualization, image processing, pattern recognition, multi-task scheduling, and software engineering. He has served as a Reviewer of various international journals, including *IEEE Communications Magazine*, IEEE TRANSACTIONS ON AFFECTIVE COMPUTING, *International Journal of Information Technology and Decision Making*, IEEE ACCESS, *Multimedia Tools and Applications*, *ACM Transactions on Multimedia Computing Communications and Applications*, *Journal of King Saud University—Computer and Information Sciences*, *Biomedical Engineering* (BME), IEEE INTERNET OF THINGS JOURNAL, and IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS.