This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3070373, IEEE Sensors Journal

# Exploring Deep Learning Features For Automatic Classification Of Human Emotion Using EEG Rhythms

Fatih Demir, Nebras Sobahi, Siuly Siuly, Abdulkadir Sengur

***Abstract*— Emotion recognition (ER) from Electroencephalogram (EEG) signals is a challenging task due to the non-linearity and non-stationarity nature of EEG signals. Existing feature extraction methods cannot extract the deep concealed characteristics of EEG signals from different layers for efficient classification scheme and also hard to select appropriate and effective feature extraction methods for different types of EEG data. Hence this study intends to develop an efficient deep feature extraction based method to automatically classify emotion status of people. In order to discover reliable deep features, five deep convolutional neural networks (CNN) models are considered: AlexNet, VGG16, ResNet50, SqueezeNet and MobilNetv2. Pre-processing, Wavelet Transform (WT), and Continuous Wavelet Transform (CWT) are employed to convert the EEG signals into EEG rhythm images then five well-known pretrained CNN models are employed for feature extraction. Finally, the proposed method puts the obtained features as input to the support vector machine (SVM) method for classifying them into binary emotion classes: valence and arousal classes. The DEAP dataset was used in experimental works. The experimental results demonstrate that the AlexNet features with Alpha rhythm produces better accuracy scores (91.07% in channel Oz) than the other deep features for the valence discrimination, and the MobilNetv2 features yields the highest accuracy score (98.93% in Delta rhythm ( with channel C3) for arousal discrimination.**

***Index Terms*— EEG based emotion classification, EEG rhythms, CWT, Deep features, Pretrained CNN models.**

## I. Introduction

Emotion itself is constituted of neuro-physiological variations associated with thoughts, behavioral responses and a degree of pleasure or displeasure [1], [2]. Emotions, which is understood from facial expressions, have important role in human-to-human communication. Sounds and body gestures are the other important emotion indicators. The major part of the emotions in human's daily life is the management of attention [3], [4] and decision-making [5], [6]. Emotion detection can improve various artificial intelligence based, applications namely, patient monitoring, criminal detection disabled assistance, security services, robotics, and communication [7], [8]. There have been several methods for classification of emotions such as bio signals (e.g. EEG, electrocardiogram (ECG) and electromyography (EMG) etc.), imaging techniques and videos [9], [8]. The EEG data is defined as non-stationary time series, which refers to the recording of the brain's spontaneous electrical activity [10], [11]. Many research studies have been undertaken to investigate the effects of emotion on EEG signals and ER from EEG signals. Issa et al. [12] used CWT and broad learning system for EEG based emotion recognition (ER). After CWT, the gray scale image feature extraction was employed and three classifiers were employed for classification into four emotions. Atkinson et al. [13] used fourteen EEG electrodes for a novel feature-based approach for EEG based ER. Authors used various statistical parameters such as standard deviation, median and kurtosis. Beside statistical features, Hjorth parameters, band power and the fractal dimension were used for feature extraction. A feature selection mechanism was employed after feature extraction and the SVM classifier was preferred in classification stage of the proposed approach. Kumar et al. [14] used only 2 EEG electrodes for achieving the EEG based ER. The higher order spectrum analysis was used to obtain the bispectrum features on the EEG signals. Extracted features were classified by using the SVM classifier. Gupta et al. [15] used 6 electrodes and flexible analytic wavelet transform (FAWT) to perform EEG based ER. Authors used various information based features and SVM and random forest (RF) classifiers in their proposed work. Zhang et al. [16] used ontology for EEG based ER. The proposed method was depended on two steps. The users' contexts, EEG and the environmental parameters were considered in the first step and in the second one, modeling of the reasons on users' emotions were extracted. Zhang et al. [17] proposed an methodology for EEG based ER. The empirical mode decomposition (EMD) and sample entropy were used for feature extraction and SVM classifier was used for classification. Candra et al. [18] used the WT for EEG based ER. The Shannon entropy and SVM classifier was used for construction of the proposed method. Rozgiç et al. [19] used an approach for EEG based ER which was composed of three steps. An overlapping window was used for feature extraction in the first step. In the second step, feature transformation was carried out by using the non-parametric nearest-neighbor model. Finally, the obtained features were classified. Al-Nafjan et al. [20] proposed a methodology for EEG based ER. Power spectral density and frontal asymmetry

Manuscript received January 15, 2021. (Corresponding author: Abdulkadir Sengur).
F. Demir is with Firat University, Technology Faculty, Electrical and Electronics Engineering Department, Elazig, Turkey (E-mail: fatihdemir@firat.edu.tr)
N. Sobahi is from King Abdulaziz University, Department of Electrical and computer Eng., Jeddah, Saudi Arabia. (E-mail: Nsobahi@kau.edu.sa)
S. Siuly is from Institute for Sustainable Industries & Liveable Cities, Victoria University, Australia (E-mail: Siuly.Siuly@vu.edu.au).
A. Sengur is from Firat University, Technology Faculty, Electrical and Electronics Engineering Department, Elazig, Turkey (E-mail: ksengur@firat.edu.tr).

features and deep neural networks were considered in the proposed approach. Chen et al. [21] proposed convolutional neural networks (CNN) for EEG based ER. A CNN model was proposed and trained on the EEG signals. Zhuang et al. [22] proposed an approach where EMD was used EEG based ER. Energy and phase information were used in classification where SVM classifier was considered. Li et al. [23] used ensemble of SVM classifiers for EEG based ER. A scheme, which was composed of weighted fusion, was considered.

In this paper, a hybrid approach is proposed. The proposed approach employsEEG rhythms for ER. The EEG signals were initially low-pass filtered for denoising. Then, the WT is employed for rhythm extraction. The extracted EEG rhythm signals are then converted to the EEG rhythm images by using the CWT method. A series of pre-trained CNN models consisting of AlexNet, VGG16, ResNet50, SqueezeNet and MobilNetv2 are used to extract deep features [18]–[22]. The obtained deep features are used in SVM classifier for the classification process. The well-known DEAP dataset was considered for performance evaluation of proposed method [16].

The contributions of the proposed method are summarized as follows:

(1) An effective automated ER framework is presented which combines EEG rhythm, deep CNN features and SVM to define valence and arousal emotions from EEG signals;

(2) Investigating the effectiveness of deep features through five deep learning CNN models: AlexNet, VGG16, ResNet50, SqueezeNet and MobilNetv2;

(3) Discovering appropriate deep features with appropriate EEG band and channel through statistical analysis for the SVM classifier;

The reminder of this paper is as following. Next section introduces the proposed methodology. The proposed method and the related theories were briefly introduced in the proposed methods section. Section 3 gives the dataset, experimental works and the results. The discussions on experimental works were given in Section 3. The conclusions were given in Section 4.

## II. Proposed Methodology

Our proposed method is designed in an emotional classification model for recognizing valence and arousal. The valence is considered in 2 classes, low valence (LV) and high valence (HV) and arousal is considered as low arousal (LA) and high arousal (HA). Fig. 1 illustrates the framework of the proposed methodology. As seen in Fig. 1, the EEG signals are initially denoised by using a low-pass filtering. And, WT is employed for decomposition of the EEG signals into its rhythms. These rhythm signals are converted into rhythm images by employing the CWT. The constructed rhythm images are conveyed to the pretrained CNN models namely; AlexNet, VGG16, ResNet50, SqueezeNet and MobilNetv2 for deep feature extraction. Finally, the classification of the rhythm images into emotion categories is carried out by using the SVM classifier.

### A. *Pre-processing: Noise removing by low-pass filtering*

In general, the EEG data is very noisy and often affected by artifacts that can bias the analysis of the data and lead to incorrect conclusions. That's why, before going the analysis the EEG data, we first removed the noises and artifacts by applying of the low-pass filtering. The reason for employing low-pass filtering is that the low-pass filter helps to keep each EEG channel in its horizontal zone, eliminating large up or down shifts in the space of other channels. Using low-pass filtering we denoised the raw EEG signals and then used for the following step.

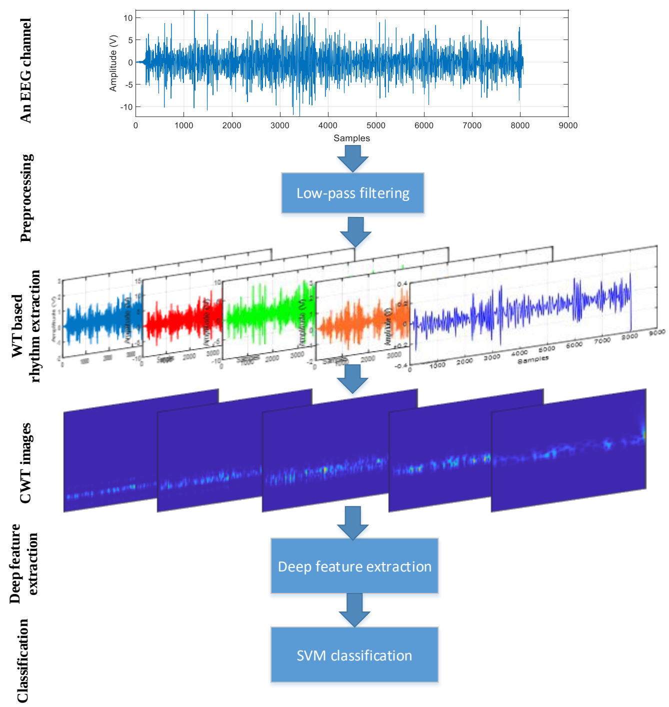



Fig. 1 Illustration of the proposed method

It is worth to mentioning that each EEG channel was investigated independently in order to determine the efficient EEG channel and EEG rhythm on ER. The pseudo-code of the proposed study is given in Table 1.

TABLE 1 THE PSEUDO-CODE OF THE PROPOSED EEG BASED EMOTION DETECTION METHOD

```
Input: Raw EEG signals
Output: ER
For all channels
  For subjects
      Denoising by using low-pass filtering
      Wavelet Decomposition based rhythm extraction
      Convert rhythm signals into rhythm images by using CWT

      For rhythms
          Extract deep features by using pretrained CNN models
      end
      Apply SVM classification
  end
end
```

### B. *EEG rhythms extraction*

As EEG signals convey different emotional information in different frequency band, thus we divided EEG signals into five bands in the frequency scale in this study. We employed the WT to extract the rhythm of the full-band EEG signal. We performed a four level decomposition by using the Daubechies $4^{th}$ order wavelet (db4) function [24]. In the first level decomposition, the detail coefficients indicate the Gamma rhythm (30-60 Hz). In the second level decomposition, the detail coefficients indicate the Beta rhythm (15-30 Hz). In the third level decomposition, the detail coefficients indicate the Alpha rhythm (8-15 Hz). In the last decomposition level, while the detail coefficients show the Theta rhythm (4-7 Hz), the approximation coefficients show the Delta rhythm (0-4 Hz). We extracted below five types of EEG rhythms by the WT from original EEG signals:

- Delta rhythm covers the frequency band less than 4 Hz (associated with Deep Dreamless Sleep, Loss of Body Awareness, quietness, lethargy, fatigue)
- Theta rhythm covers the frequency band greater than or equal to 4 and less than 8 Hz (Associated with Dreams, Deep Meditation, REM sleep, Creativity , awake state, or the emotion gradually becomes calmer),
- Alpha rhythm covers the frequency band greater than or equal to 8 and less than or equal to 14 (associated with Relaxation (while awake), Pre-sleep Drowsiness),
- Beta rhythm covers the frequency band greater than 14 and less than 40 Hz (associated with Active Thinking, Concentration, and Cognition)
- Gamma rhythm covers the frequency band greater than or equal to 40 Hz, (associated with Higher Mental Activity, Consciousness, Perception, multisensory information integration).

### C. Converting EEG rhythms to EEG images

The CWT, which is a TF method, is employed for transforming the EEG rhythms into EEG images. And, it is defined as [25];

$$F_{\omega}(a,b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{\infty} f(t)\bar{\psi}\left(\frac{t-a}{b}\right) dt \qquad (1)$$

where, ψ(t) is called as the mother wavelet. The mother wavelet is a wavelet function that is represented in both time and frequency domains. The over line symbolizes the complex conjugate operation. The Morse wavelet function is used in CWT.

### D. Extraction of Deep Features

This study considered five pre-trained CNN models: AlexNet, VGG16, ResNet50, SqueezeNet and MobilNetv2 that were trained on 25 million of images. An investigation was performed about these five models, which one is better to produce higher performance in the ER. Table 2 shows some important properties of these CNN models. AlexNet is known as the first deep CNN model, which has gained much attention in artificial intelligence community. Its depth is 8 and the image input size is 227×227 [26]. The 'fc6' layer was used for feature extraction and the obtained deep feature vectors are 4096 dimensional. VGG16 was designed as a deeper CNN model compared to the AlexNet [27].

TABLE 2 SOME PROPERTIES OF THE PRETRAINED CNN MODELS

| Pretrained CNN model | Depth | Image input size | Feature length |
|---|---|---|---|
| AlexNet | 8 | 227×227 | 4096 |
| VGG16 | 16 | 224×224 | 4096 |
| ResNet50 | 50 | 224×224 | 1000 |
| SqueezeNet | 18 | 227×227 | 1000 |
| MobileNetv2 | 53 | 224×224 | 1000 |

The depth of the VGG16 is 16 and the input size is 224×224. The 'fc6' layer of the VGG16 was used for feature extraction and the obtained deep feature vectors are 4096 dimensional. The ResNet50 is a deeper model compared to AlexNet and VGG16 [28]. It was composed of 50 layers. Its image input size is 224×224. The 'fc1000' layer was used for feature extraction and the extracted feature vectors are 1000 dimensional. SqueezeNet was developed as a fast and effective CNN model [29]. As its depth is 18 but it achieves AlexNet-level accuracy with 50x fewer parameters. The image input size of SqueezeNet is 227×227. The 'pool10' layer is used for 1000 dimensional feature vector extraction. MobileNetv2, of which depth is 53, produces 1000 dimensional feature vector by using the 'logits' layer [30]. The input size of MobileNetv2 model is 224×224.

### E. Classification

The SVM classifier, which is a well-known machine learning approach, is considered in this study. The SVM classifier was preferred due to its high performance in wide range of its applications [31].

Let's assume a two class data denoted by $(x_1,y_1),(x_2,y_2)\dots,(x_i,y_i)$ and $i=1,2,\dots,M$. $x_i$ shows the input feature vector and $y$ shows the class labels. Thus, Eq. (2) can be used for obtaining of the hyper-plane that used to separate the two classes.

$$f(x) = w^T x + b = 0 \qquad (2)$$

Where, $w$ and $b$ indicates the weight vector and the bias value used to determine the position of the hyper-plane, respectively. SVM was initially developed for the datasets that can be separated linearly. If the dataset cannot be separated linearly, a kernel function is employed to turn dataset into another hyper-plane. To this end, different kernel functions such as polynomial and radial basis function can be used. SVM employs an optimization procedure to detect the optimum separating hyper-plane, which is given in Eq. (3).

$$\begin{cases} \min \dfrac{\|w\|^2}{2} \\ y_i(w^T x_i + b) \geq 1 \quad i = 1,2,\dots,M \end{cases} \qquad (3)$$

Eq. (3) can be revised by adding $C$ regularization parameter and a positive artificial $\xi_i$ variable as follows:

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3070373, IEEE Sensors Journal

$$min\, f(w,\xi) = \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{M}\xi_i \tag{4}$$

## III. EXPERIMENTAL WORKS

### A. DEAP dataset

The DEAP data set, which contains 32-electrode EEG signals from 32 subjects, was constituted while each subject was watching forty one-minute music videos. The subjects rated each music video in terms of valence, arousal, like / dislike, dominance, and familiarity levels. In addition to the EEG signals, a facial video was recorded for 22 of the 32 subjects while watching each music video. Ratings of subjects during the experiment were carefully analyzed. Correlations among EEG signal frequencies and subjects' evaluations were also analyzed. EEG signals were recorded with a sampling frequency of 512 Hz and then down-sampled up with a sampling frequency of 128 Hz. Thus, a size of 40 × 32 × 8064 data was constituted for each subject.

### B. Results and Discussions

For experiments, a computer, which has Intel (R) CPU (3.20 GHz) with 32 GB of memory, was utilized. The extracted features were divided into training and test sets in the ratio of 0.75 and 0.25, respectively. This process was carried out in a random fashion. The SVM classifier was preferred due to its high performance in wide range of applications [31]. Also, it determines the maximum-margin hyperplane to reduce the prediction error. The hyperparameter C was searched in the range of [$10^{-4}$-$10^{3}$]. The radial basis function kernel was used in SVM and its parameters were tuned by using the grid search optimization algorithm. Double tagging was considered during the labeling of emotions. Valence ratings below five were assumed to contain negative emotions, and valence scores higher than five were considered to have positive feelings. In addition, arousal rating scales range from passive to active. Arousal ratings less than five were considered passive, and other rating scales higher than five were considered active. Therefore, the class labels obtained were high arousal (HA) with low valence (LV) and high valence (HV) and low arousal (LA), respectively.

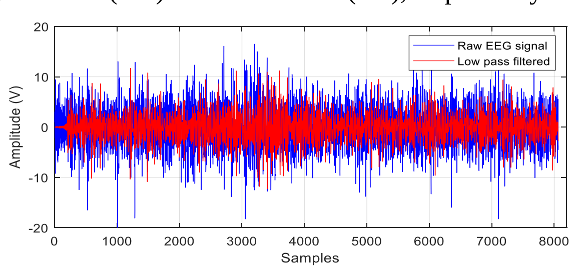



Fig. 2 The raw and low-pass filtered of an EEG signal

Fig. 2 illustrates a raw signal and low-pass filtered signal of an EEG signal and Fig. 3 shows the gamma, beta, alpha, theta and delta rhythms of an EEG signal, respectively.

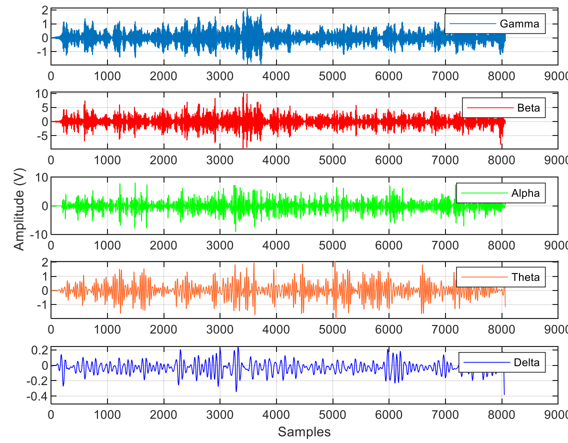



Fig. 3 The rhythms of an EEG signal

Fig. 4 shows the EEG rhythm images that were obtained by using the CWT. The x-axis shows the time and the y-axis shows the frequency. As seen in Fig. 4, each rhythm is covering a frequency band.

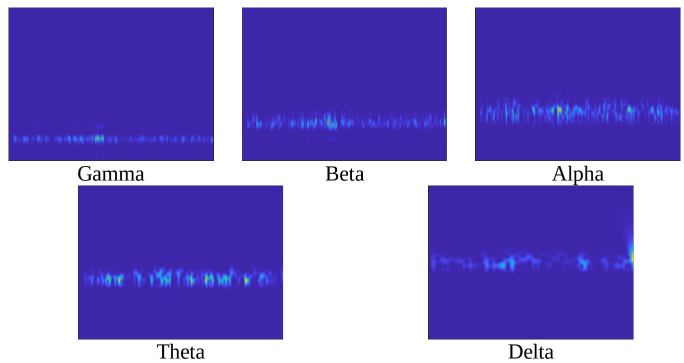
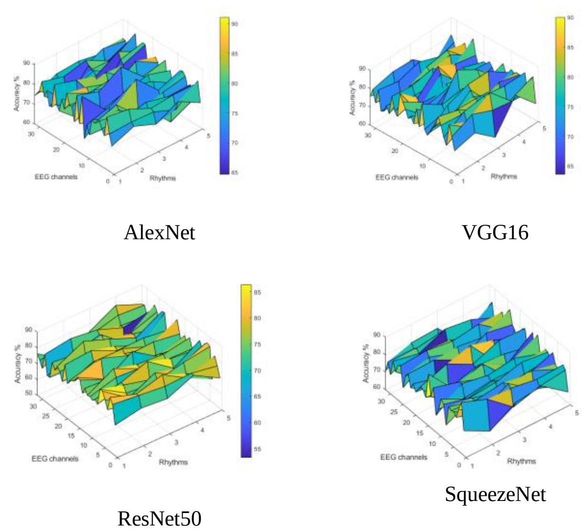
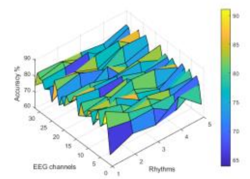



Fig. 4 The rhythm images of an EEG signal that are obtained via CWT

The initial experiments were conducted on HV vs LV classification and the obtained all results were illustrated in Fig. 4. In Fig. 4 each cell shows the achievement of a pretrained CNN model.



This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3070373, IEEE Sensors Journal



MobilNetv2

Fig. 5 The performance of deep feature + SVM on LV and HV classification

The x-axis, y-axis and z-axis of the given plots show the rhythms, EEG channels and accuracy scores, respectively. For the x-axis, alpha, beta, delta, gamma and theta rhythms are denoted by the numbers 1, 2, 3, 4 and 5, respectively. Moreover, Tables 3-7 show the achievements of the five EEG channels on all EEG rhythms. These channels were selected due to the highest average accuracy scores of the rhythms achievements. In Table 3, the performances of AlexNet features are provided based on different channels and bands. As observed from Table 3, Oz channel and Delta rhythm produced a 91.07% accuracy score, which was the highest accuracy score that AlexNet features have produced.

TABLE 3 CLASSIFICATION ACCURACY FOR ALEXNET MODEL IN HV VS LV FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| **EEG channel** | **Alpha** | **Beta** | **Delta** | **Gamma** | **Theta** | **Average** |
|---|---|---|---|---|---|---|
| Oz | 8.21% | 84.29% | **91.07%** | 80.36% | 82.86% | 85.34% |
| F3 | 86.79% | 85.00% | 79.64% | 80.71% | 80.36% | 82.50% |
| P4 | 83.93% | 80.71% | 83.21% | 77.14% | 82.14% | 81.43% |
| PO3 | 89.29% | 77.14% | 80.36% | 82.86% | 79.29% | 81.79% |
| CP5 | 80.36% | 76.07% | 89.64% | 78.21% | 82.50% | 81.36% |
| **Average** | 85.72% | 80.64% | 84.78% | 79.86% | 81.43% | |

The second highest accuracy score 89.64% was produced by the CP5 and Delta rhythm and the third highest accuracy score 89.29% was obtained by the PO3 and Alpha rhythm, respectively. From Table 3, it can be inferred that Alpha rhythm and Oz channel were produced better achievements with AlexNet and SVM classifier.

TABLE 4 CLASSIFICATION ACCURACY FOR VGG16 MODEL IN HV VS LV FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| **EEG channel** | **Alpha** | **Beta** | **Delta** | **Gamma** | **Theta** | **Average** |
|---|---|---|---|---|---|---|
| FC5 | 76.07% | 82.86% | 84.29% | 82.50% | 82.50% | 81.64% |
| PO3 | 87.50% | 80.00% | 81.43% | 76.79% | 80.71% | 81.29% |
| Cz | 82.50% | 81.07% | 83.21% | 74.64% | 85.00% | 81.28% |
| F3 | 84.64% | 76.07% | 87.86% | 76.07% | 80.71% | 81.07% |
| C3 | 85.71% | 82.14% | 75.00% | 82.86% | 79.64% | 81.07% |
| **Average** | 83.28% | 80.43% | 82.36% | 78.57% | 81.71% | |

Table 4 gives the achievements of the VGG16 features and SVM classifier on HV vs LV discrimination. As seen in Table 4, the Alpha rhythm and the FC5 channel produced the highest average accuracy scores. With VGG16 features, the best achievement 87.86% accuracy score was produced by F3 channel and Delta rhythm.

TABLE 5 CLASSIFICATION ACCURACY FOR RESNET50 MODEL IN HV VS LV FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| **EEG channel** | **Alpha** | **Beta** | **Delta** | **Gamma** | **Theta** | **Average** |
|---|---|---|---|---|---|---|
| Cz | 78.21% | 86.43% | 80.36% | 73.57% | 82.86% | 80.29% |
| F3 | 84.29% | 78.93% | 83.57% | 73.93% | 80.36% | 80.22% |
| Fp2 | 77.50% | 84.29% | 78.57% | 79.64% | 77.14% | 79.43% |
| FC5 | 80.36% | 73.57% | 77.50% | 81.43% | 83.93% | 79.36% |
| FC1 | 84.64% | 72.86% | 81.07% | 72.50% | 85.00% | 79.21% |
| **Average** | 81.00% | 79.22% | 80.21% | 76.21% | 81.86% | |

In Table 5, the accuracy scores, which were obtained by ResNet50 features, were given for each rhythm and EEG channel, respectively. According to the average accuracy scores, Theta rhythm and Cz channel produced the highest average accuracy scores. While Cz produced an average 80.29% accuracy score, Theta rhythm obtained an average 81.86% accuracy score. The best achievement 86.43% was produced by the Cz channel and Beta rhythm.

TABLE 6 CLASSIFICATION ACCURACY FOR SQUEEZENET MODEL IN HV VS LV FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| **EEG channel** | **Alpha** | **Beta** | **Delta** | **Gamma** | **Theta** | **Average** |
|---|---|---|---|---|---|---|
| AF4 | 78.93% | 84.29% | 76.43% | 82.50% | 81.79% | 80.79% |
| Oz | 89.64% | 73.57% | 80.00% | 72.14% | 83.21% | 79.71% |
| AF3 | 89.29% | 83.57% | 71.43% | 75.71% | 77.50% | 79.50% |
| T8 | 81.07% | 82.14% | 80.00% | 82.14% | 71.07% | 79.28% |
| P4 | 80.00% | 76.79% | 83.21% | 81.07% | 74.64% | 79.14% |
| **Average** | 83.79% | 80.07% | 78.21% | 78.71% | 77.64% | |

In Table 6, the achievements of EEG channels against rhythms with SqueezeNet features on HV vs LV classification were given. As seen in Table 6, the best average rhythm achievement 83.79% was obtained by the Alpha rhythm and the highest average accuracy score 80.79% was produced by the AF4 channel, respectively. The best accuracy score 89.64% was obtained by the Alpha rhythm and Oz channel. The mobilNetv2 features achievements were represented in Table 7. As observed in Table 7, alpha rhythm and Cz channel produced a 85.57% and a 83.00% average accuracy scores respectively. The best achievement with MobilNetv2 features was a 91.07%, and it was obtained by F3 channel and Alpha rhythm.

TABLE 7 CLASSIFICATION ACCURACY FOR MOBILNETV2 MODEL IN HV VS LV FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| **EEG channel** | **Alpha** | **Beta** | **Delta** | **Gamma** | **Theta** | **Average** |
|---|---|---|---|---|---|---|
| Cz | 86.43% | 83.57% | 76.43% | 84.64% | 83.93% | 83.00% |
| Oz | 86.79% | 83.93% | 83.93% | 78.93% | 80.36% | 82.79% |
| F3 | 91.07% | 82.50% | 79.29% | 78.57% | 82.50% | 82.79% |
| P3 | 78.57% | 81.79% | 86.43% | 85.00% | 78.57% | 82.07% |
| PO3 | 85.00% | 81.43% | 84.64% | 73.57% | 82.14% | 81.36% |
| Average | 85.57% | 82.64% | 82.14% | 80.14% | 81.50% | |

Experiments were also conducted on HA vs LA classification and the obtained all results with deep features were illustrated in Fig. 5 similar to the Fig. 4. Tables 8-12 show the achievements of the five EEG channels that produced the highest average accuracy scores of the rhythms achievements. Table 8 shows the achievements of the AlexNet features and SVM. As observed from Table 8, Gamma rhythm and Fp2 channel produced a 94.29% and 90.21% accuracy scores, respectively, that were the highest average accuracy

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JSEN.2021.3070373, IEEE Sensors Journal

scores. The best accuracy score 96.43% was also produced by the Gamma rhythm and T7 channel.

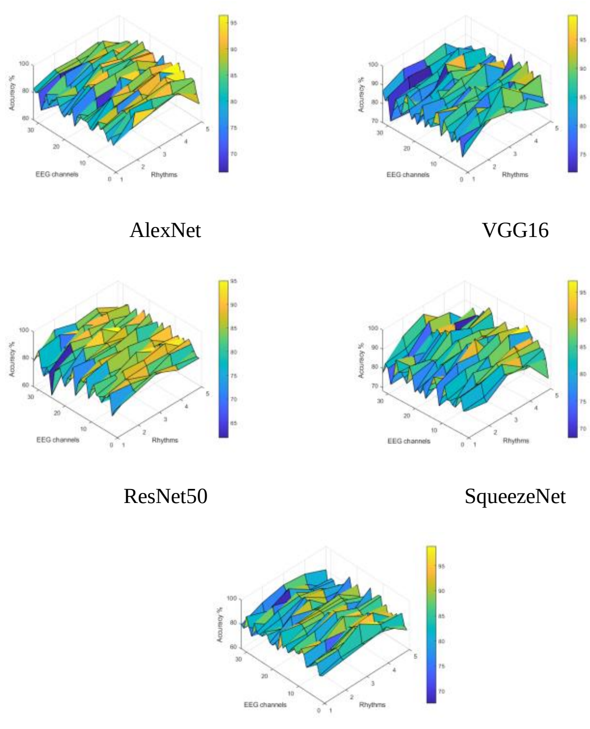
TABLE 8 CLASSIFICATION ACCURACY FOR ALEXNET MODEL IN HV VS LV FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| EEG channel | Alpha | Beta | Delta | Gamma | Theta | Average |
|---|---|---|---|---|---|---|
| Fp2 | 86.43% | 93.93% | 93.57% | 93.93% | 83.21% | 90.21% |
| T7 | 84.64% | 90.00% | 92.86% | 96.43% | 85.36% | 89.86% |
| P7 | 83.57% | 86.79% | 91.79% | 95.00% | 89.29% | 89.29% |
| FC5 | 86.43% | 90.00% | 90.36% | 92.14% | 85.36% | 88.86% |
| F7 | 83.57% | 88.57% | 88.21% | 93.93% | 87.50% | 88.36% |
| Average | 84.93% | 89.86% | 91.36% | 94.29% | 86.14% | |

Table 9 represents the obtained accuracy scores for VGG16 features. From Table 9, it is seen that Beta and Theta rhythms obtained average accuracy scores above 90.00% and the CP5 channel also produced a 90.29% average accuracy score, respectively. A 94.64% accuracy score, which the highest among all results, was produced by the CP5 channel and Theta rhythm.

TABLE 9 CLASSIFICATION ACCURACY FOR VGG16 MODEL IN HA VS LA FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| EEG channel | Alpha | Beta | Delta | Gamma | Theta | Average |
|---|---|---|---|---|---|---|
| CP5 | 90.36% | 92.86% | 88.21% | 85.36% | 94.64% | 90.29% |
| P7 | 91.79% | 88.57% | 87.14% | 89.29% | 91.79% | 89.72% |
| F7 | 82.50% | 91.07% | 90.71% | 90.00% | 91.43% | 89.14% |
| FC5 | 87.86% | 93.57% | 86.43% | 88.57% | 90.00% | 89.29% |
| Fp2 | 91.43% | 88.57% | 88.57% | 93.21% | 83.57% | 89.07% |
| Average | 88.79% | 90.93% | 88.21% | 89.29% | 90.29% | |



AlexNet VGG16

ResNet50 SqueezeNet

MobilNetv2

Fig. 6 The performance of deep feature + SVM on LA and HA classification.

Table 10 shows the achievements of the ResNet50 features and SVM on HA vs LA classification. As seen in Table 10, Delta rhythm and P7 channel produced a 92.36% and 91.50% accuracy scores, respectively, that were the highest average accuracy scores. The best accuracy score 93.93% was also produced by the Delta rhythm and P7 channel and Gamma rhythm and FC1 channel, respectively.

TABLE 10 CLASSIFICATION ACCURACY FOR RESNET50 MODEL IN HA VS LA FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| EEG channel | Alpha | Beta | Delta | Gamma | Theta | Average |
|---|---|---|---|---|---|---|
| P7 | 89.29% | 92.86% | 93.93% | 90.71% | 90.71% | 91.50% |
| Fp2 | 91.43% | 88.93% | 88.57% | 91.79% | 87.50% | 89.64% |
| FC6 | 86.07% | 84.64% | 95.00% | 82.86% | 90.36% | 87.79% |
| CP5 | 88.21% | 86.07% | 93.21% | 84.64% | 86.79% | 87.78% |
| FC1 | 78.57% | 92.50% | 91.07% | 93.93% | 81.79% | 87.57% |
| **Average** | 86.71% | 89.00% | 92.36% | 88.79% | 87.43% | |

Table 11 shows the performance of the SqueezeNet features on HA vs LA classification. As seen in Table 11, Theta rhythm and P7 channel produced a 90.86% and 90.64% accuracy scores, respectively, that were the highest average accuracy scores. The best accuracy score 96.43%was also produced by Gamma rhythm and T7 channel, respectively.

TABLE 11 CLASSIFICATION ACCURACY FOR SQUEEZENET MODEL IN HA VS LA FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| EEG channel | Alpha | Beta | Delta | Gamma | Theta | Average |
|---|---|---|---|---|---|---|
| P7 | 86.43% | 91.43% | 90.00% | 94.29% | 91.07% | 90.64% |
| FC6 | 88.57% | 85.71% | 94.29% | 89.29% | 90.71% | 89.71% |
| T7 | 81.79% | 84.29% | 91.79% | 96.43% | 91.79% | 89.22% |
| CP5 | 88.21% | 90.71% | 90.00% | 82.14% | 94.29% | 89.07% |
| F3 | 87.14% | 88.21% | 91.43% | 83.57% | 86.43% | 87.36% |
| **Average** | 86.43% | 88.07% | 91.50% | 89.14% | 90.86% | |

Table 12 shows the performance of the MobilNetv2 features on HA vs LA classification. As seen in Table 12, Delta rhythm and P7 channel produced a 94.93% and 90.57% accuracy scores, respectively, that were the highest average accuracy scores. The best accuracy score 98.93% was also produced by Delta rhythm and F3 channel, respectively.

TABLE 12 CLASSIFICATION ACCURACY FOR MOBILNETV2 MODEL IN HA VS LA FOR DIFFERENT RHYTHMS WITH RESPECT TO EACH EEG CHANNELS

| EEG channel | Alpha | Beta | Delta | Gamma | Theta | Average |
|---|---|---|---|---|---|---|
| P7 | 89.29% | 91.43% | 96.07% | 88.21% | 87.86% | 90.57% |
| F3 | 88.57% | 88.57% | 98.93% | 87.50% | 85.71% | 89.86% |
| CP5 | 92.14% | 85.36% | 96.79% | 85.71% | 84.29% | 88.86% |
| Fp2 | 94.64% | 90.36% | 88.57% | 87.50% | 80.00% | 88.21% |
| FC5 | 85.00% | 86.43% | 94.29% | 82.86% | 82.50% | 86.22% |
| **Average** | 89.93% | 88.43% | 94.93% | 86.36% | 84.07% | |

In Table13, the proposed method was compared with the existing methods that use the same dataset. As shown in Table 13, the proposed method outperformed the existing other methods on HV vs LV categorization. Alzarzi et al. produced a second-best accuracy score where the accuracy was 85.8% [32]. Tripathi et al. [33] and Li et al. [23] reported 81.4% and 80.7% accuracy scores, respectively. Rozgić et al. [19], Zhang et al. [17], and Atkinson et al. [13] reported accuracy scores among 76.9%, 75.2% and 73.1% and Zhuang et al. [22], Huang et al. [34], and Candra et al. [18] reported performance of their method among 60.0% and 70.0%. The worst result reported by Koelstra et al. [35] was 57.6%.

TABLE 13 PERFORMANCE COMPARISON OF THE PROPOSED METHOD WITH SOME OF THE STATE-OF-ART-RESULTS.

| **Method** | **Accuracy (%)** | |
|---|---|---|
| | **HV vs LV** | **HA vs LA** |
| Koelstra et al. [35] | 57.6 | 62.0 |
| Alazrai et al. [32] | 85.8 | 86.6 |
| Huang et al. [34] | 66.1 | 82.5 |
| Candra et al. [18] | 65.1 | 65.3 |
| Rozgic et al.[[19] | 76.9 | 69.1 |
| Zhang et al. [17] | 75.2 | 81.7 |
| Atkinson et al. [13] | 73.1 | 73.0 |
| Tripathi et al. [33] | 81.4 | 73.3 |
| Zhuang et al.[22] | 69.1 | 71.9 |
| Li et al. [23] | 80.7 | 83.7 |
| **Proposed study** | **91.1** | **98.9** |

The achievement comparisons on HA and LA categorization were also shown in Table 13. As shown in Table 13, the proposed method obtained the best classification score. The second-best result was produced by Alazrai et al. [32]. Huang et al. [34], Zhang et al. [17], and Li et al. [23] reported accuracy results between 80.0% and 85.0%, respectively. Atkinson et al. [13], Tripathi et al. [33], and Zhuang et al. [22] reported accuracy results between 70.0% and 75.0% respectively. Also, Koelstra et al. reported the worst accuracy score [35].

In order to further evaluate the performance of the proposed method, we performed a multiclass classification between four different classes of emotions: HVHA, LVHA, LVLA, and HVLA in the same DEAP database. The labelling process is given in Table 14. These results were obtained for Oz channel, AlexNet features and Alpha rhythm. The SVM was also tuned as mentioned previously.

TABLE 14 2D EMOTION CLASSIFICATION LABELS

| **Emotion Label** | **Scores** |
|---|---|
| HVHA | $V \geq 5,\ A \geq 5$ |
| LVHA | $V < 5,\ A \geq 5$ |
| LVLA | $V < 5,\ A < 5$ |
| HVLA | $V \geq 5,\ A < 5$ |

The best accuracy was 90.6% for HVHA emotion state, 89.23% for LVHA emotion state, 83.1% for LVLA emotion state, and 83.4% for HVLA emotion state.

## IV. CONCLUSION

This study intended to develop an emotion classification framework that can use EEG signal data for recognizing valance and arousal from humans. The proposed framework consists of several steps such as noise removing by low-pass filtering; EEG rhythms extraction by the WT, the rhythms conversation to EEG images by the CWT, deep feature extraction and investigation by the five CNN models and the SVM classification of the extracted deep features. The followings are important conclusions of the proposed method.

1) In the HV vs LV discrimination, the AlexNet features with Alpha rhythm produced better average accuracy scores than the other deep features. Besides, the Oz channel with AlexNet features produced the best average classification score. The MobilNetv2 features achieved the second best performance for HV vs LV classification. It is surprising that an initial model (AlexNet) produced better results than the recent deep models.

2) In the HV vs LV classification, generally Cz, F3, PO3 and Oz channels produced considerable results than the other channels.

3) In the HA vs LA discrimination, the AlexNet features with Gamma rhythm produced better average accuracy scores than the other deep features. Besides, the P7 channel with ResNet50 features produced the best average classification score. The MobilNetv2 features achieved the second best performance for HV vs LV classification. The highest accuracy score 98.93% was obtained by the F3 channel, Delta rhythm and MobilNetv2 features.

4) The proposed method produced better results in HA vs LA classification than HV vs LV classification.

## REFERENCES

[1] H. Chao, L. Dong, Y. Liu, and B. Lu, "Emotion recognition from multiband eeg signals using capsnet," *Sensors (Switzerland)*, vol. 19, no. 9, p. 2212, 2019, doi: 10.3390/s19092212.

[2] D. Şengür and S. Siuly, "Efficient approach for EEG-based emotion recognition," *Electron. Lett.*, vol. 56, no. 25, pp. 1361–1364, 2020, doi: 10.1049/el.2020.2685.

[3] R. J. Dolan, "Emotion, cognition, and behavior," *Science (80-. ).*, vol. 298, no. 5596, pp. 1191–1194, 2002.

[4] D. Sengür and M. Turhan, "Prediction Of The Action Identification Levels Of Teachers Based On Organizational Commitment And Job Satisfaction By Using K-Nearest Neighbors Method," *Fırat Univ. Turkish J. Sci. Technol.*, vol. 13, no. 2, pp. 61–68, 2018.

[5] J. S. Lerner, Y. Li, P. Valdesolo, and K. S. Kassam, "Emotion and decision making," *Annu. Rev. Psychol.*, vol. 66, 2015.

[6] D. Şengür, "Investigation of the relationships of the students' academic level and gender with Covid-19 based anxiety and protective behaviors: A data mining approach," *Turkish J. Sci. Technol.*, vol. 15, no. 2, pp. 93–99, 2020.

[7] Y. Luo, G. Wu, S. Qiu, S. Yang, W. Li, and Y. Bi, "EEG-based Emotion Classification Using Deep Neural Network and Sparse Autoencoder," *Front. Syst. Neurosci.*, vol. 14, p. 43, 2020.

[8] H. Akbari, M. T. Sadiq, and A. U. Rehman, "Classification of normal and depressed EEG signals based on centered correntropy of rhythms in empirical wavelet transform domain," *Heal. Inf. Sci. Syst.*, vol. 9, no. 1, pp. 1–15, 2021.

[9] K. Giannakaki, G. Giannakakis, C. Farmaki, and V. Sakkalis, "Emotional state recognition using advanced machine learning techniques on EEG data," in *2017 IEEE 30th International Symposium on Computer-Based Medical Systems (CBMS)*, 2017, pp. 337–342.

[10] H. R. Al Ghayab, Y. Li, S. Siuly, and S. Abdulla, "A feature extraction technique based on tunable Q-factor wavelet transform for brain signal classification," *J. Neurosci. Methods*, vol. 312, pp. 43–52, 2019.

[11] J. Yin, J. Cao, S. Siuly, and H. Wang, "An integrated spectral-temporal analysis based framework for MCI detection using resting-state EEG signals," *Int. J. Autom. Comput.*, vol. 16, no. 3, pp. 1–14, 2019.

[12] S. Issa, Q. Peng, and X. You, "Emotion Classification Using EEG Brain Signals and the Broad Learning System," *IEEE Trans. Syst. Man, Cybern. Syst.*, 2020.

[13] J. Atkinson and D. Campos, "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers," *Expert Syst. Appl.*, vol. 47, pp. 35–41, 2016.

[14] N. Kumar, K. Khaund, and S. M. Hazarika, "Bispectral analysis of EEG for emotion recognition," *Procedia Comput. Sci.*, vol. 84, pp. 31–35, 2016.

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

[16] X. Zhang, B. Hu, J. Chen, and P. Moore, "Ontology-based context modeling for emotion recognition in an intelligent web," *World Wide Web*, vol. 16, no. 4, pp. 497–513, 2013.

[17] Y. Zhang, X. Ji, and S. Zhang, "An approach to EEG-based emotion recognition using combined feature extraction method," *Neurosci. Lett.*, vol. 633, pp. 152–157, 2016.

[18] H. Candra *et al.*, "Investigation of window size in classification of EEG-emotion signal with wavelet entropy and support vector machine," in *2015 37th Annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 2015, pp. 7250–7253.

[19] V. Rozgić, S. N. Vitaladevuni, and R. Prasad, "Robust EEG emotion classification using segment level decision fusion," in *2013 IEEE international conference on acoustics, speech and signal processing*, 2013, pp. 1286–1290.

[20] A. Al-Nafjan, M. Hosny, A. Al-Wabil, and Y. Al-Ohali, "Classification of human emotions from electroencephalogram (EEG) signal using deep neural network," *Int. J. Adv. Comput. Sci. Appl*, vol. 8, no. 9, pp. 419–425, 2017.

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

[22] N. Zhuang, Y. Zeng, L. Tong, C. Zhang, H. Zhang, and B. Yan, "Emotion recognition from EEG signals using multidimensional information in EMD domain," *Biomed Res. Int.*, vol. 2017, 2017.

[23] X. Li, J.-Z. Yan, and J.-H. Chen, "Channel Division Based Multiple Classifiers Fusion for Emotion Recognition Using EEG signals," in *ITM Web of Conferences*, 2017, vol. 11, p. 7006.

[24] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 9, pp. 1545–1551, 2007.

[25] R. Polikar, "The wavelet tutorial. Glassboro," *Rowan Univ., NJ, USA*, 1999.

[26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[27] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv Prepr. arXiv1409.1556*, 2014.

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[29] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," *arXiv Prepr. arXiv1602.07360*, 2016.

[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 4510–4520.

[31] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Min. Knowl. Discov.*, vol. 2, no. 2, pp. 121–167, 1998.

[32] R. Alazrai, R. Homoud, H. Alwanni, and M. I. Daoud, "EEG-based emotion recognition using quadratic time-frequency distribution," *Sensors*, vol. 18, no. 8, p. 2739, 2018.

[33] 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.

[34] D. Huang, C. Guan, K. K. Ang, H. Zhang, and Y. Pan, "Asymmetric spatial pattern for EEG-based emotion detection," in *The 2012 International Joint Conference on Neural Networks (IJCNN)*, 2012, pp. 1–7.

[35] S. Koelstra *et al.*, "Deap: A database for emotion analysis; using physiological signals," *IEEE Trans. Affect. Comput.*, vol. 3, no. 1, pp. 18–31, 2011.