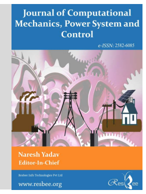
Vol.4 No.2 April 2021

Journal of Computational Mechanics, Power System and Control

Received 17 December, Revised 26 February, Accepted 28 March



Deep Convolutional Neural Network for Emotion Recognition via EEG Signal

Meherkandukuri
University of Sheffield, London
Kandukurimeher@gmail.com

Abstract: In current days, human emotional state recognition through Electroencephalogram (EEG) is considered as the up-and-coming topics which grip the concentration of researchers. For numerous real-time applications, generally, this EEG-based recognition is an effective technique, particularly for disabled persons. Regarding precise emotion recognition, numerous researchers are in advancement to create the recognition technique effectively. Nevertheless, it is not fulfilled in the accurate development, therefore, this work tries in the human emotion recognition stated or it affects via EEG signal by exploiting the classifier techniques as well as developed features. Initially, this work uses the Wavelet Transformation as well as 2501 (EMCD) in the recognition process to indicate the EEG signal in minimum dimension and expressive. The redundancy of EEG is removed using the EMCD, as well as the important information can be extracted. By exploiting the extracted features, the classification procedures are performed with the help of a classifier called Convolutional Neural Network (CNN). The developed method performance is evaluated regarding the metrics such as positive and negative measures and the results also exhibit the dominance of the developed model in emotion recognition in an accurate manner.

Keywords: Classifier; EEG; Emotion Recognition;EMCD;Metrics

Nomenclature

Abbreviations	Descriptions
DSP	Digital Signal Processing
BCI	Brain Computer Interface
HCI	Human Computer Interface
NN	Neural Network
DWT	Discrete Wavelet Transform
ML	Machine Learning
K-NN	K-Nearest Neighbor
STFT	Short Time Fourier Transform
MLP	Multi-Layer Perceptron
WT	Wavelet Transform
FAWT	Flexible Analytic Wavelet Transform
SVM	Support Vector Machine
NLP	Natural language processing
ANN	Artificial Neural Network
ML	Machine Learning
ICA	Independent Component Analysis
AFBD	Average Frequency Band Division
IMFs	Intrinsic Mode Functions
TF	Time–Frequency
PSD	Power-Spectral-Density
EMD	Empirical Mode Decomposition

1. Introduction

In day today's life of human beings, emotion plays an important role in decision making, communication activities, and multimedia [1]. For instance, automatic human emotion recognition received huge attention among researchers in the framework of multimedia applications. By means of living standards enhancements in people, multimedia demand has grown highly. Conventionally, audio and video content

https://doi.org/10.46253/jcmps.v4i2.a3

has comprised of multimedia systems that mainly motivate hearing and vision. In multimedia experiences, to improve the real sensation, the smell is integrated with the multimedia systems since it directly motivates memory as well as prompts strong emotions [2].

A non-invasive recording of the electric potentials produced is known as the EEG signal and it is performed using the activity of the neurons in the brain. To extract the information, the multi-channel EEG signals are evaluated regarding the several statuses of the brain involving diagnosis as well as BCI. For diagnosing epilepsy and sleep disorder the initially favored technique is the EEG analyzing model when exploiting EEG the BCI applications are extensively examined to discover the concealed information in the brain to control any peripheral. Additionally, by exploiting EEG signals the emotion state recognition is a current and well-known to discover how signal alters are pretentious by emotional condition [3].

In a real-world application, with the development of technology numerous studies have been conducted and the exploit the DSP and ML techniques for HCI. For the recognition of human emotion, several models have been proposed. Nevertheless, still there are numerous conflicts are presented in the reliability of these models. The most important problems of these models are the non-linear EEG signal characteristics. Naturally, the EEG signals are extremely nonlinear which consequent in sudden alterations in the features, wherein, the emotions of humans change slowly from one condition to the other.

In numerous state-of-the-art works, several feature extractions, as well as EEG signal processing methods, are carried out for the recognition of emotion. As EEG features, the power difference of electrode-pair has been extracted by the STFT that are used in non-hierarchal as well as hierarchal classification models for the recognition of emotions. From DWT, the energy-based features and statistical analysis are calculated, which provides the emotions classifications and sub-bands by exploiting the K-NN classifier [10]. By exploiting the MLP classifier, the frequency, time as well as wavelet-based features from EEG signals are extracted. For classification of emotion, the WT, needy EEG features are exploited, which input to the NN [7].

Using WT, the surface laplacian filter processed EEG signal is examined for the statistical analysis as well as rhythms separation by exploiting the KNN classifier [8]. For emotion recognition, the FAWT based features were developed. On SVM and MLP, the EEG signals which extract the emotion-specific features are classified for the recognition of emotion. The SVM classifier is exploited to identify the emotions by the frequency domain features. For the real-time emotion classification applications, the stability of the non-statistical and statistical features is examined. For feature extraction, the EEG magnitude squared coherence is contemplated subsequently the boundaries amid the separable areas of features are identified using the self-organizing map that can be trained to KNN classifier for the recognition of emotion [9].

The main contribution of this research is to work on the human emotion recognition model or that affects via EEG signals. Here, a developed features and classifier techniques are exploited. In the initial phase of the recognition process, the DWT is exploited to extract the features, and also this work uses the EMCD procedure to minimize their dimensions. Here, a renowned classifier is exploited a deep learning technique called the CNN approach.

2. Literature Review

In 2020, Hui-Rang Hou et al [1], worked on the integration of SVM as well as the AFBD technique. Here, the extraction of PSD features was performed by exploiting the AFBD technique from the EEG signals, which persuades by smelling diverse odors. The aforesaid AFBD technique calculates each PSD feature on the basis of equivalent frequency bandwidth compared with the conventional EEG rhythm-based bandwidth. Here, 13 odors were exploited to persuade the olfactory EEGs and their equivalent emotions.

In 2019, SachinTaran, Varun Bajaj et al [2], worked on the EEG signals and it possesses a two phase filtering technique that was proposed for emotion recognition. Initially, to remove the noisy IMFs, a correlation criterion was recommended and it was performed by using the EMD on the raw EEG signal. In the next phase, for the denoised EEG signal reconstruction, the VMD was applied and therefore the filtered modes were maintained. Subsequent to the 2-phase filtering, to classify the emotion recognition, as the input features the non-linear measures of filtered EEG signals were exploited to MC-LS-SVM.

In 2018, Ahmet Mert and Aydin Akan [3], examined the feasibility by exploiting the TF indication of EEG signals for recognition of the emotional state. Here, a feature extraction technique named an up-to-date and developed TF analyzing technique, MSST was exploited because of multi-channel signal processing and compact module localization abilities. Initially, exploiting MSST, 32participants' EEG recordings from DEAP emotional EEG database were analyzed to expose oscillations. Next, feature selection, as well as ICA, was exploited to minimize the high dimensional 2D TF distribution without bringing up the rear distinctive module information in the 2D image.

In 2021, Ante Topic and Mladen Russo [4], developed a novel technique for emotion recognition which was on the basis of the generation of the feature maps on the basis of the holographic and topographic indication of EEG signal characteristics. As a feature extraction, the deep learning model was exploited on feature maps and then the extracted features were combined jointly for the classification procedure for the recognition of diverse types of emotions. In 2020, Rab Nawaz et al [5], exploited three-dimensional models such as valence, arousal as well as dominance technique of emotion for the emotion recognition that was induced by music videos. Here, a video was watched by the participants when their EEG was recorded. The most important contribution of this research was to recognize the features which were optimal to discriminate the emotions. Entropy, power, statistical features, fractal dimension, and wavelet energy were extracted from the EEG signals. Moreover, the feature effects were analyzed and the optimal features were recognized.

3. Proposed Emotion Recognition approach via EEG Signal

Fig 1 demonstrates the developed model of EEG-based emotion recognition technique. This technique comprises three stages such as feature extraction, classification, and decomposition. Moreover, EEG signal $s(x)$ represents the input of the corresponding video which is gathered from the 32 channel enobio device specifically, using two manners the emotions are gathered such as a) In subjects the different videos are shown, b)VR device or VR videos, as well as the equivalent emotions are automatically identified by calculating the signal vibrant. Initially, by exploiting the DWT the features are extracted, which ensues in the signal form. Subsequently, by exploiting the EMCD procedure, the ensuing signal will be decomposed. For the classification model, the decomposed features are subjected as the input. Here, the DBN classifier is exploited and outcomes in the classified result as emotions.

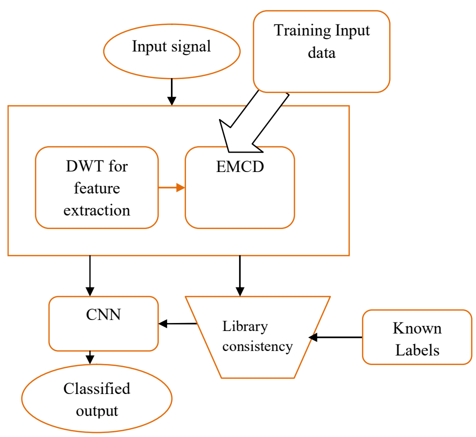


Fig. 1. *Schematic model of developedemotion recognition via EEG signal*

4. Feature Extraction Component

4.1. DWT

The initial phase is the feature extraction that divides the process into two sub-phases such as initial phase uses the DWT, which extracts the features from the input $s(x)$. The DWT procedure is stated as follows: Eq. (1) states the wavelet and it represents $\psi(x)$function with zero average.

$$\int_{-\infty}^{\infty} \psi(x)dx = 0 \tag{1}$$

Fromthe mother wavelet $\psi(x)$, the daughter wavelets are derived that is the scale family and the shifter versions and it is exhibited in eq. (2), whereas $i \in \Re$ and $g \in \Re^{+}$ indicates shifting or translation and the scaling or dilation, correspondingly.

$$\psi_{g,i}(x) = \frac{1}{\sqrt{g}}\psi\left(\frac{x-i}{g}\right) \tag{2}$$

As stated in eq. (3), the input $s(x)$ continuous wavelet at i location and g scale, wherein WC signifies the wavelet coefficients, $*$ signifies the complex conjugation, $\langle .,. \rangle$ signifies the inner product. As the shifting parameter i and scaling parameter g deviated wavelet coefficients $WC(g,i)$ can be obtained.

$$WC(g,i;s(x),\psi(x)) = \int_{-\infty}^{\infty} s(x)\psi^{*}_{g,i}(x)dx = \langle s(x),\psi_{g,i}(x)\rangle \tag{3}$$

By discretizing both g and i DWT is obtained. Generally, by DWT, a dyadic sampling is used with i and g that is based on the power to $i = r2^{p}$ and $g = 2^{p}$ with $p,r \in Z$. Whilein eq. (2), the discretized parameter is substituted, and in eq. (4), a dyadic wavelet is attained. Therefore, on the basis of eq. (5), the DWT of $s(x)$ is described, wherein $d_{p,r}$ indicates the wavelet coefficient at p level and r position.

$$\psi_{p,r}(x) = \frac{1}{\sqrt{2^{p}}}\psi\left(\frac{x}{\sqrt{2^{p}}} - r\right) \tag{4}$$

$$d_{p,r} = \int_{-\infty}^{\infty} s(x)\psi^{*}_{p,r}(x)dx = \langle s(x),\psi_{p,r}(x)\rangle \tag{5}$$

For the input $s(x)$, the multi-resolution decomposition at LE level is determined in Eq. (6), where $\psi(x)$ denotes the wavelet function whereas $\varphi(x)$ refers to the companion function. Hence DWT results with the features (signal form) of $s(x)$, which is representedas $s_{d}(x)$.

$$\begin{aligned} s(x) = & \sum_{r=-\infty}^{\infty} g_{LE,r}\, 2^{-LE/2}\varphi\left(2^{-LE}x - r\right) + \\ & \sum_{p=-\infty}^{LE}\sum_{r=-\infty}^{\infty} d_{p,r}\, 2^{-p/2}\psi\left(2^{-p}x - r\right) \\ = & AP_{LE}(x) + \sum_{p=-\infty}^{LE} SD_{p}(x) \end{aligned} \tag{6}$$

The main advantage of the wavelet transform is it shows a concurrent localization in frequency as well as time domain. Then, one of the most important benefits is it has the ability to separate the fine details in a signal. Additionally, to isolate the fine details in a signal, small wavelets are used, wherein large wavelets can recognize the coarse details. To decompose a signal into module wavelets, a wavelet transform is used.

4.2EMCD

A well-known decomposition model called EMCD is used to decompose the attained signal from the DWT model $s_{d}(x)$.

$s_{d}(x)$ represents the input signals, wherein $x = 1,......,X;\ X\ samples$. The maxima series of $s_{d}(x)$ represents $\{(r,s[r]), r = 1,.....,X_{r}\}$, wherein, X_{r} signifies the number of maxima and r signifies the maxima time indices. The minima series of $s_{d}(x)$ indicates $\{(t,s[t]), t = 1,....,X_{t}\}$, wherein X_{t} signifies the number of minima and t signifies the minima time indices. Subsequently, from the equivalent minima as well as maxima, both inferior, as well as superior envelopes, are described.

i. Superior Envelope: Superior envelope $s^{\sup}[x]$ of input signal represents the upper trend curve that passes through each and every maxima. Generally, eq. (7) represents the B-spline BS interpolation which is exploited to interpolate the maxima.

$$s^{\sup}[x] = BS\{(r,s[r]), s_{d}(x)\}, x = 1,....X \tag{7}$$

ii. Inferior Envelope: Inferior envelope $s^{\inf}[x]$ of the input signal is the lower trend curve that passes through all minima. As stated in eq. (8), B-spline BS is exploited to interpolate the minima.

$$s^{\inf}[x] = BS\{(t,s[t]), s_{d}(x)\}, \quad x = 1,.....,X \tag{8}$$

iii. Mean Curve: Eq. (9) indicates the mean curve $s^{mean}[x]$ indicates the average of both superior and inferior envelopes.

$$s^{mean}[x] = \left(s^{\sup}[x] + s^{\inf}[x]\right)/2, \quad x = 1,......,X \tag{9}$$

iv. Mode: Eq. (10) represents the Mode MO of the input signal as the mean of a number of minima X_t and number of maxima X_r.

$$MO(s_d(x)) = (X_r + X_t)/\mathbf{2} \tag{10}$$

v. Empirical Waveform (EWF): The simplified EWF is stated in eq. (11), during the determination of the mean curve, a new signal is produced by extrema, and the series of alternating maxima and minima are indicated as EWF. The simplified EWF is defined in Eq. (11).

$$EWF(s_d(x)) = \{(r, s[r]), (t, s[t])\} \tag{11}$$

By exploiting the EWF, the mean curve is indicated, wherein the mode of this characterizes the EWF. Furthermore, either in one mode, the full sine wave presents such as maximum else minimum. Hence, the EWF mode ascertains in a conventional Fourier analysis. As stated in eq. (12), the Empirical period P_E of EWF is derived. In eq. (13), the empirical frequency E_F is explained.

$$'P_E = X / MO(s_d(x)) \tag{12}$$

$$E_F = MO(s_d(x))X \tag{13}$$

E_F and P_E are estimated over the entire signal in impermanent, due to this is established as highly effectual than considering model parameters. Therefore, this recreation extremely improves the signal ability which is molded from the oscillatory sources. In a while, the important purpose of the EMCD approach is to decompose the signal at diverse scale levels which result in the elucidated EWF. Furthermore, EWFsare indicated as the essential model of the EMCD approach that does not necessitate previous knowledge of the signal model. D^f indicates the decomposed output (feature).

4.3. Convolutional Neural Network

A CNN is a part of the ANN that exploits the ML approach, a perceptron, for supervised learning to examine a huge number of data.

The proposed CNN is exploited in numerous applications and NLP is also exploited for numerous cognitive tasks.

This approach possesses an output, input layer as well as numerous amounts of hidden layers. By exploiting the mathematical formulations, some of these layers are convolved to perform the outcomes to the following layers.

The raw pixel values of the color channel are possessed by the input image. The subsequent layer is known as the conv layer. It is used to estimate the output of the neuron which is associated with the local areas to the input layer. A point product is computed by each amid their weights and a few areas linked to the input layer.

The Rectified linear unit (relu) layer is known as the third layer that is used as a unit-wise activation function. It creates the volume size unaffected. The subsequent layer is known as the pooling layer that carries out the downsampling operation beside the spatial dimensions (height and width) ensuing in size. The final layer is known as the fully connected layer and it will compute the score ensuing in size volume. Similar to the conventional NN, in this layer, each neuron is linked to all the weights in the preceding set.

From an input image, the layer extracts the features. Using the learning image feature, convolution conserves the link pixels using small squares of input image data. For this mathematical formulation, 2 inputs namely filter and image matrix is exploited. By means of several filters, convolution of an image data performs operations such as bluer, edge detection as well as sharpening of images.

The number of parameters is reduced by the pooling layer while many images are subjected as input. By taking the highest element from the modified feature map the max-pooling is performed. The main contribution of the max-pooling is to downsample the input image, minimizing its dimensions, and so on.

$$f_{X,Y}(S) = \max_{a,b=\mathbf{0}}^{\mathbf{1}} S_{\mathbf{2}X+A,\mathbf{2}Y+b} \tag{14}$$

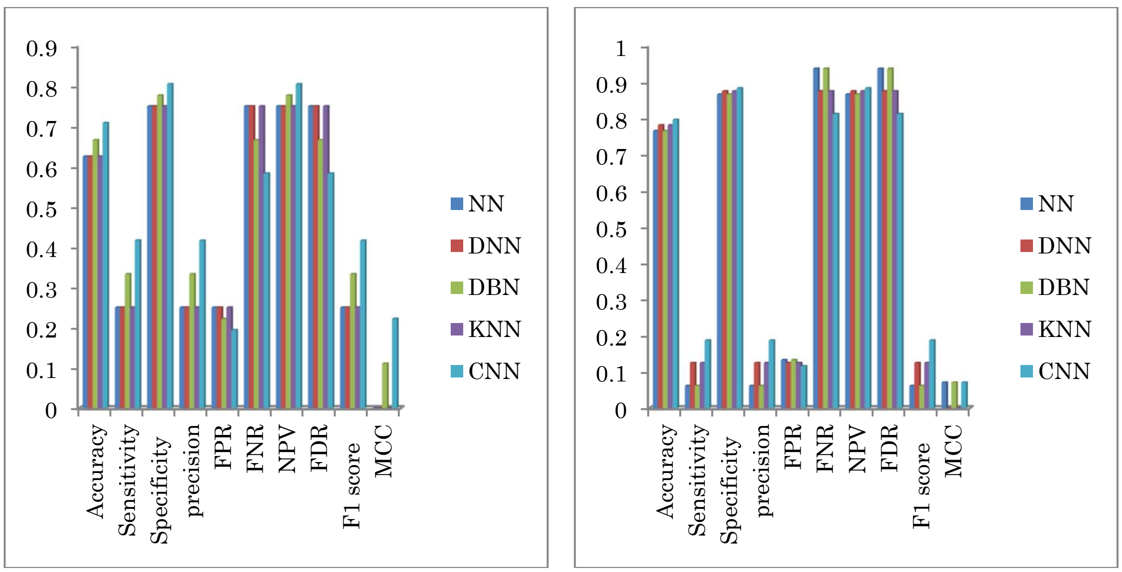
To convert the 2-dimensional array set into a single, the flattening process is exploited, a high continuous linear vector. From the convolutional layer, it obtains the output, which flattens its model to make a single, high feature to exploit the subsequent layer for the ultimate classification.

The fully connected layer is represented as the hidden layer which is present in the convolutional layer. This is a particular kind of hidden layer that should be exploited within the CNN. Moreover, it is exploited to integrate the feature into more attributes that predict the outputs in a precise way.

5. Experimentation Procedure

In this section, the performance analysis of the proposed and conventional model has been experimented with regarding the recognition of emotion. Here, the emotion recognition based on EEG signals was analyzed from ten diverse persons by linking 32electrodes,as well as emotions were stored in 2 databases. In the first database, the videos are present, which captures by exploiting the normal cameras. Here, it comprises 4 emotions. The next database comprises videos which capture by exploiting VR videos. It contains 8 emotions. The proposed model was compared with the conventional models such as NN, Deep Neural Network, Deep Belief Network, and K-Nearest Neighbour. The analysis was performed for both the positive as well as negative measures.

Fig 2 demonstrates the proposed and conventional models for two video datasets such as normal as well as VR videos. Fig 2 (a) exhibits the analysis of the proposed and existing techniques for normal video. The analysis of the proposed and existing techniques for VR video for various classifiers exhibited in Fig 2(b). It is seen that the accuracy of the proposed method is better than the conventional methods.



(a) (b)

Fig. 2. *Analysis of the proposed and existing models features for (a) normal video (b) VR video*

6. Conclusion

In multimedia, emotion recognition plays an important role to improve the real sensation. The EEG evaluates the brain's neurophysiology for recognition of diverse emotional conditions.This work has presented a new EEG-based emotion recognition technique. Here, two phases were used in this paper. In the initial phase, the feature extraction was performed by exploiting the EMCD as well as DWT. Then, the classification was performed by exploiting the classifier named the CNN model. Finally, the developed model performance was evaluated with the existing models. From the results, it was evident that the betterment of the proposed model over the conventional models in terms of the positive and negative metrics for emotion recognition.

Compliance with Ethical Standards

Conflicts of interest: Authors declared that they have no conflict of interest.

Human participants: The conducted research follows the ethical standards and the authors ensured that they have not conducted any studies with human participants or animals.

References

[1] Hui-Rang HouXiao-Nei ZhangQing-Hao Meng,"Odor-induced emotion recognition based on average frequency band division of EEG signals", Journal of Neuroscience Methods,vol.334, 21 January 2020.

[2] Sachin TaranVarun Bajaj,"Emotion recognition from single-channel EEG signals using a two-stage correlation and instantaneous frequency-based filtering method", Computer Methods and Programs in Biomedicine,vol. 173, pp. 157-165, 22 March 2019.

[3] Ahmet Mert,Aydin Akan,"Emotion recognition based on time–frequency distribution of EEG signals using multivariate synchrosqueezing transform", Digital Signal Processing, vol. 81, pp.106-115, 24 July 2018.

[4] Ante Topic,Mladen Russo,"Emotion recognition based on EEG feature maps through deep learning network", Engineering Science and Technology, an International JournalAvailable online ,16 April 2021.

[5] Rab Nawaz,Kit Hwa, CheahVooi, Voon Yap,"Comparison of different feature extraction methods for EEG-based emotion recognition", Biocybernetics and Biomedical Engineering, vol. 40, no. 3, pp. 910-926, 11 May 2020.

[6] Yumiao Wang,Zhou Zhang,Qingyun Du,"A new attention-based CNN approach for crop mapping using time series Sentinel-2 images", Computers and Electronics in Agriculture, vol. 184, 21 March 2021.

[7] Dr.Sivaram Rajeyyagari,"Automatic Speaker Diarization using Deep LSTM in Audio Lecturing of e-Khool Platform",Journal of Networking and Communication Systems, vol. 3, no. 4, October 2020.

[8] Bhagyalakshmi V,Dr.Ramchandra and Dr.Geeta D,"Arrhythmia Classification Using Cat Swarm Optimization Based Support Vector Neural Network",Journal of Networking and Communication Systems, vol 1, no. 1, October 2018.

[9] Quazi M. H and Dr. S. G. Kahaleka," Adaptive filtering in EEG Signal for Artifacts Removal using Learning Algorithm",Journal of Networking and Communication Systems, vol 2, no. 2, April 2019.

[10] Ashu Gupta,"Vertical Handover Using Mutated-Salp Swarm Optimization Algorithm Based on Radial Basis Function Neural Network in Heterogeneous Networks",Journal of Networking and Communication Systems, vol 3, no. 3, July 2020.