

Human emotion recognition from EEG signals: model evaluation in DEAP and SEED datasets

Mohit Kumar^{1,*}, Marta Molinas¹

¹Norwegian University of Science and Technology, Trondheim

Abstract

The automatic distinction of human emotional states can provide the technological basis for applications in healthcare, education, marketing, and manufacturing sectors that rely on human-machine interfaces. Emotion recognition from Electroencephalography (EEG) signals is a challenging task entailing the development of classification models that should accurately distinguish among diverse human emotions. In this work, two publicly available EEG emotion datasets, SEED, and DEAP, are used to develop automatic emotion detection models and to evaluate their performance for emotion recognition. Models are built based on a two-dimensional (valence or pleasantness and arousal or intensity) and a positive/neutral/negative emotion models using multilayer perceptron (MLP) and convolution neural network (CNN) classification algorithms. First, the preprocessed EEG signals of these datasets are decomposed into five rhythms, namely, delta, theta, alpha, beta, and gamma. The differential entropy (DE) is computed from the rhythms of the EEG signals and used as a feature for the classification algorithms. Epochs of 1-sec. duration are considered for the computation of DE features. The CNN-based method achieves a better F1-score (93.7%) for the SEED dataset as compared to the MLP-based method. However, for DEAP dataset, 94.5% and 94%, F1-scores are achieved for high vs. low arousal and high vs. low valence classes respectively, with no significant difference between the performance of the two methods.

Keywords

Emotion detection, EEG, Brain rhythms, Features, Arousal-Valence

1. Introduction

Emotion is considered an important factor in human life as it affects the human working ability, communication, personal, social life, mental state, and physical health. Human brain is responsible for the generation and regulation of different emotions. Humans have a unique ability to adjust their behavior while interacting with each other. The human can do this because they are able to decipher the emotional states of others. Human-machine interaction can also be improved if the machines are able to infer human emotional states. Hence, automatic emotion recognition can be very useful to improve human-machine interaction. Moreover, the study of human emotion encompasses research in various fields such as cognitive science, computer science, neuroscience, and psychology [1]. Electroencephalogram (EEG) based emotion recognition has opened a vast space for exploration and innovation in these disciplines [2]. A correlation between human emotions and EEG signal is observed in [3]. EEG-based methods are

Italian Workshop on Artificial Intelligence for Human-Machine Interaction (AIXHMI 2022), December 02, 2022, Udine, Italy

*Corresponding author.

✉ mohit.kumar@ntnu.no (M. Kumar); marta.molinas@ntnu.no (M. Molinas)

© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0)

CEUR Workshop Proceedings (CEUR-WS.org)

considered to be more reliable than facial expression and gesture [4] as they are less susceptible to counterfeit.

Automatic detection approaches traditionally utilize hand-crafted features from EEG signals and implement a classification model for EEG emotion recognition [5]. Various methodologies such as Fourier transform, power spectral density, and wavelet transform are widely used to analyse the EEG signal for emotion recognition [2]. The support vector machine (SVM) based methodology is used to classify joy, sadness, anger, and pleasure feelings [6]. The method proposed in [7] is based on frequency band searching to find an optimal band for emotion recognition to then apply the SVM method for emotion classification. It is observed that the gamma band is suitable for EEG-based emotion classification. In [8], differential entropy (DE) is extracted as a feature and found to be effective in representing the emotional states in EEG signals. A transfer recursive feature elimination based approach is proposed in [9], which is found to be useful for selecting the optimal feature subset.

Recently, various studies have focused on deep learning based methods for the classification of EEG signals [10]. A classification method based on a deep belief network (DBN) is introduced for the classification of EEG signals related to the different types of emotion [11, 1]. The performance of the DBN is found to be better than that of SVM and K-nearest neighbour classifiers. A regularized graph neural network is proposed to capture both local and global relations among different EEG channels for emotion recognition [12]. A method based on long short-term memory (LSTM) classifier is introduced to extract temporal information to discriminate the different emotions using EEG signals [13]. In [14, 15], a model based on LSTM and convolution neural network (CNN) is proposed which incorporated both temporal and spatial information for emotion identification.

In this work, our aim is to develop an automated method that can accurately detect human emotions using EEG signals. The current state of the art methods [14, 15] for detecting human emotions are generally based on rather complex network architectures. Hence, in this work, we focused on less complex classification models and for that, we have chosen two classification approaches, one based on multilayer perceptron (MLP) and another based on CNN. The rationale for choosing the CNN model is to utilise the spatial information provided by the location of the different EEG channels on the scalp. For comparison purposes and in order to understand the value of spatial information, the simpler MLP architecture is chosen since its input is one dimensional and does not take into account the spatial information. We then compare the performance of the CNN and MLP models on the SEED and DEAP datasets.

The remaining paper is arranged as follows: the two datasets used to evaluate the proposed method are described in section 2. Section 3 presents the methodology which includes the feature extraction and classification methods. Results and conclusions are provided in section 4 and section 5, respectively.

2. Dataset description

In the present work, the following two datasets are analysed:

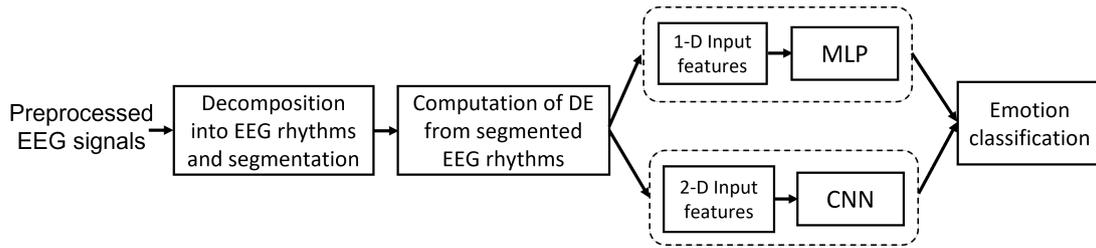


Figure 1: The steps involved in the proposed method applied to the SEED and DEAP datasets.

2.1. DEAP dataset

The DEAP dataset contains the EEG signals of 32 subjects [16, 17]. The recording was performed when the subjects were watching 40 pieces of music videos. The duration of the music video was 1 minute. The signals were collected using 32 EEG channels. The EEG signals were recorded at a sampling frequency of 512 Hz and further down-sampled to 128 Hz. The EEG signals were passed through a bandpass filter of 4-45 Hz and EOG artifacts were also removed. The pre-processed EEG signals in each trial consisted of 3 sec. baseline data and 60 sec. trial data. During the recordings of EEG signals, the participants were asked to rate the levels of arousal, valence, liking, and dominance for each video from 1 to 9 using the self-assessment manikin. Further details of the experiment are available in [16]

2.2. SEED dataset

The SEED dataset [1, 18] contains the EEG signals of 15 subjects obtained from an emotion elicitation paradigm. It contains the EEG signals of 3 sessions for each subject and 15 trials for each session. The recording was performed when the subjects were watching Chinese film clips of 4 minutes duration. The signals were collected using a 62 channels EEG system with a sampling frequency of 1000 Hz that was further down-sampled to 200 Hz. The EEG signals were passed through a bandpass filter of 0-75 Hz [18]. During the recordings of EEG signals, the participants were asked to rate each film clip with three types of emotions, namely positive, neutral, and negative. The class labels -1, 0, and 1 are used to represent negative, neutral, and positive classes. Further details of the experiment are available in [1].

3. Methodology

The overall methodology is summarised in Figure 1. The steps involved in the method are described as follows:

3.1. Decomposition and segmentation

In the present work, the Butterworth filter of order 3 is used to decompose the EEG signals into the EEG rhythms. The EEG signals of SEED dataset are decomposed into five EEG rhythms

named, delta (1- 4Hz), theta (4-8Hz), alpha (8-14Hz), beta (14-31Hz), and gamma (31- 51Hz). The EEG signals of the DEAP dataset are decomposed into four rhythms namely, theta (4-8Hz), alpha (8-14Hz), beta (14-31Hz), and gamma (31-45Hz), as the DEAP dataset is preprocessed using 4 Hz to 45 Hz band pass filter. In [19], the 1 sec. segment length was found to be most suitable for emotion recognition. Hence, the EEG rhythms are segmented into epochs of 1 sec. duration. After segmentation, the total number of epochs for each subject are 3394 and 2400 for SEED and DEAP datasets, respectively.

3.2. Feature extraction

In this work, we have used DE as a feature for the classification of different types of emotions. DE is known to provide a measure of the complexity of a signal and has been applied with success to non-stationary and non-linear signals such as EEG. DE can differentiate between low and high frequency energy [1]. It is defined as:

$$h(X) = - \int_X g(x) \log(g(x)) dx \quad (1)$$

where x is a random variable and in this work, x represents the segmented EEG rhythms. The $g(x)$ is the probability density function. For a Gaussian random variable, DE can be estimated as follows:

$$h(X) = \frac{1}{2} \log(2\pi e\sigma^2) \quad (2)$$

where e and σ^2 are the Euler's constant and the variance of a time series, respectively. In [20], it is shown that the sub-band EEG signals can meet the Gaussian distribution criterion. Hence, the DE features are computed from the EEG rhythms.

For the SEED dataset, the DE feature is extracted from the five EEG rhythms. The obtained feature matrix has the dimension of (3394, 62, 5) for each subject and each session, where, 3394 represents the total number of epochs collected from the 15 trials, 62 represents the number of channels, and 5 is corresponding to the five different EEG rhythms.

For the DEAP dataset, the DE feature is extracted from the four EEG rhythms. The feature matrix has the dimension of (2400, 32, 4) for each subject, where, 2400 represents the total epochs collected from all 40 trials for each subject, 32 represents the number of channels, and 4 is corresponding to the four subbands. The feature matrix is termed as DE_{main} . the DE feature is also extracted from the 3 secs. baseline provided with the DEAP dataset. First, we have divided the 3 secs. baseline into 3 segments of 1 sec. duration. Then, we computed the DE from all 3 segments and marked it as DE_{base} . Finally, we consider the average value of the three DE_{base} as the baseline DE features. The baseline DE feature matrix has a dimension of (40, 4) for each subject, where, 40 represents the number of trials and 4 features correspond to the four subbands. The baseline DE feature is subtracted from the main DE feature before applying it to the input of the classifiers as suggested in [21]. However, in [15], the baseline is segmented into six parts. Each one of them is of 0.5 sec. duration and the average baseline DE features are computed over six segments.

3.3. Classification methods

In this work, we have applied the following two classification approaches:

3.3.1. MLP based classification

MLP is a widely used tool in various classification problems. It is also used for emotion classification in [21]. It handles large amounts of input data well and can make quick predictions after training. A two-layer neural network has an input layer, an output layer, and one hidden layer between the input and output layers [22]. However, the MLP may have more than one hidden layers. In this work, the employed architecture of the MLP consists of 4 hidden layers. The hidden layers consist of 256, 256, 128, and 64 units corresponding to hidden layer 1 to hidden layer 4. The number of units and layers are selected via the trial and experimentation method so that the model can perform well on both SEED and DEAP datasets. For the SEED dataset, the feature matrix is reshaped into the size of (3394,310) for each subject and each session to feed to the MLP classifier, and for the DEAP dataset is reshaped into the size of (2400,128).

3.3.2. CNN based classification

CNN is a special kind of neural network which has shown tremendous success in practical applications [23]. In [21, 15], it is utilized for emotion classification using EEG signals. For the CNN classifier, the DE feature matrix is arranged into the 2D features space according to the location of electrodes as suggested in [15]. This mapping can preserve the spatial structure information of the electrode location. The input data size to the CNN classifier is (3394, 8, 9, 5) for SEED dataset and (2400, 8, 9, 4) for DEAP dataset. The convolution layer (CLr) is an essential part of CNN architecture, it applies a set of filters to the input data that provides activation of the features from the input data. The output can be referred to as the feature map [23]. The details of the employed CNN architecture are as follows: It has two convolution blocks CB-1 and CB-2. CB-1 consists of 2 CLrs and one maxpooling layer. Each CLr of CB-1 has 256 units with a filter size of 5X5. CB-2 contains one CLr having 128 units with a filter size of 4x4 and one maxpooling layer. The filter size of maxpooling layer for CB-1 and CB-2 is 2x2 with a stride of 2. The output of CB-2 is flattened and connected to the dense layer having 64 units. Then, it is finally connected to the classification layer having a softmax activation function. The number of units, convolution blocks, and CLrs are selected via the trial and experimentation method so that the model can perform well on both SEED and DEAP datasets.

Generally, a CLr is followed by a pooling layer [14]. It is used to downsample the output of the CLr layer. In case of a small translation in the input, it helps to keep the representation to be nearly invariant [23]. The studies on the DEAP dataset [14, 21] suggested that the pooling layer is not necessary as the size of the feature matrix is smaller than the size of the feature matrix in the computer vision field. However, in [15], one pooling layer is added after the last CLr. Hence, in our proposed CNN architecture, each CLr is not followed by a pooling layer. Each convolution block is followed by a pooling layer. All the CLrs of CB-1 and CB-2 are used with the same padding to preserve the size of the feature matrix within each convolution block as the size of the feature matrix is not large.

4. Results

In this work, a 5-fold cross-validation approach is applied to test the performance of the model. In this approach, the dataset is divided into five equal sets. The dataset is iterated over five times. At a time, four sets are used for the training, and the remaining one set is used for testing. In each iteration, a different set is used for testing. Finally, the average classification accuracy over five iterations is reported. The results are obtained for each subject separately. The maximum number of iterations and batch size is kept 100 and 128 for both MLP and CNN classifiers as suggested in [15]. Both of the classification models are implemented using the Keras library [24]. The Adam optimizer [25] is used with the default learning rate provided in the Keras library. The RELU activation function is used in each layer of both classifiers except for the final classification layer. It incorporates the nonlinearity in the neural network [23]. In the final layer, the softmax activation function is used.

4.1. Results on SEED dataset

The performance of the MLP model for SEED dataset for all 15 subjects in terms of accuracy is shown in Figure 2. The highest accuracy of 98.4% is observed for the first session of the 15th subject and the lowest accuracy of 69.6% is obtained for the third session of 4th subject. From Figure 2, it can be noted that the accuracy is found to be greater than 90% for subjects 1 (session 1), 5 (session 2 and 3), 6, 7 (session 2), 8 (session 2 and 3), 13 (session 1), and 15. The average performance of MLP is summarised in Figure 4. The average accuracy and F1-score over 15 subjects and three sessions are 86.8% and 86.5%, respectively. The performance of the CNN

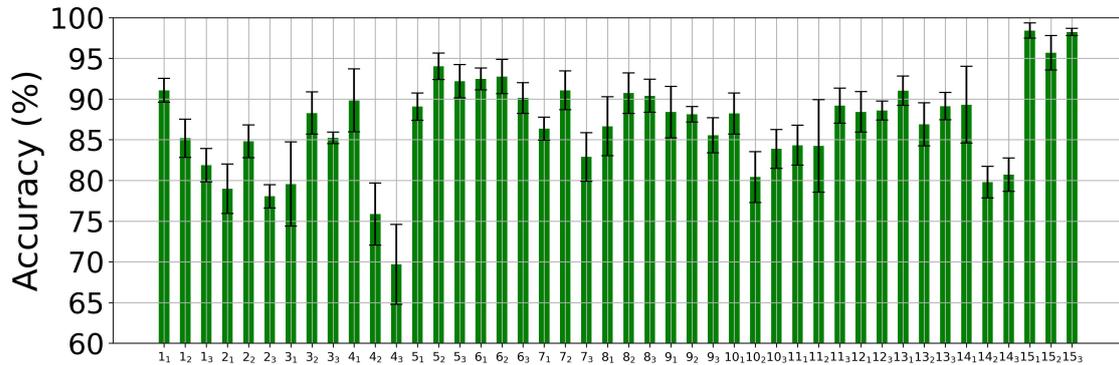


Figure 2: Accuracy (average over 5-folds) for all 15 subjects for SEED dataset using MLP classifier. The X-axis represents subjects and the subscript represents the session.

model for SEED dataset for all 15 subjects in terms of accuracy is shown in Figure 3. The highest accuracy of 99.6% is observed for the third session of the 15th subject and the lowest accuracy of 84.6% is obtained for the third session of the 4th subject. From Figure 3, it can be noted that the accuracy is found to be greater than 95% for subjects 5 (session 2 and 3), 6 (session 1 and 2), 8 (session 3), 10 (session 1), 12 (session 1), 13 (session 3), 14 (session 1), and 15. The average accuracy and F1-score over 15 subjects and three sessions are 93.8% and 93.7%, respectively, which are shown in Figure 4. The accuracy is significantly improved with the CNN classifier as

compared to that of the MLP classifier (p -value <0.001). The improvement of accuracy with the CNN classifier can be explained by the fact that the CNN takes tensor (2D) as input so it can understand spatial relation between EEG channels better and this can explain why the CNN performs better than MLP, since MLP takes vector (1D) as input.

The confusion matrix for subject 1 (session 1) is depicted in Figure 5. It is the summation of the five confusion matrices obtained over 5-folds. From Figure 5, it can be observed that the performance of the CNN model is well balanced over the three classes of emotion. For neutral, positive, and negative classes, it has correctly identified 1048, 1139, and 1012 samples, respectively. However, MLP model is performing better for positive class than for negative and neutral classes.

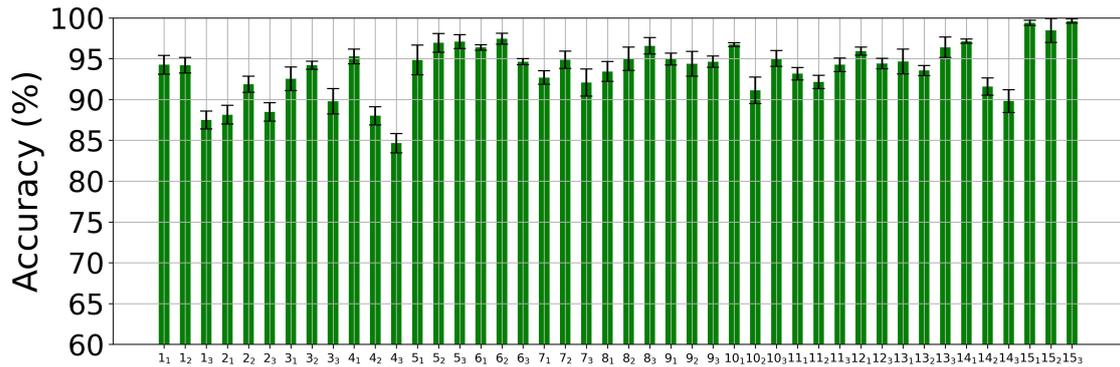


Figure 3: Accuracy (average over 5-folds) for all 15 subjects for SEED dataset using CNN classifier. The X-axis represents subjects and the subscript represents the session.

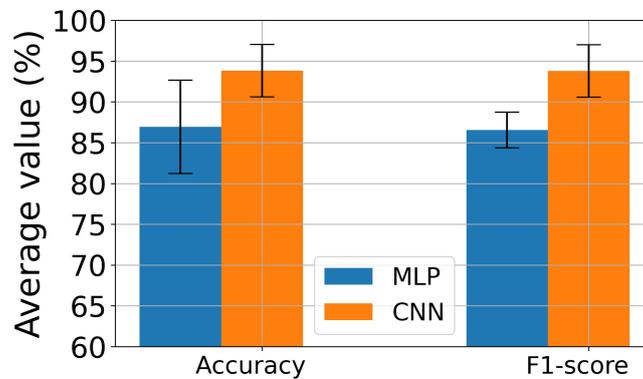


Figure 4: Average results over 15 subjects for SEED dataset using MLP and CNN classifier

4.2. Results on DEAP dataset

For the DEAP dataset, the participants' rating are available on a scale of 1 to 9. Hence, ratings greater than 5 are considered as high valence (HV) and high arousal (HA), and others are

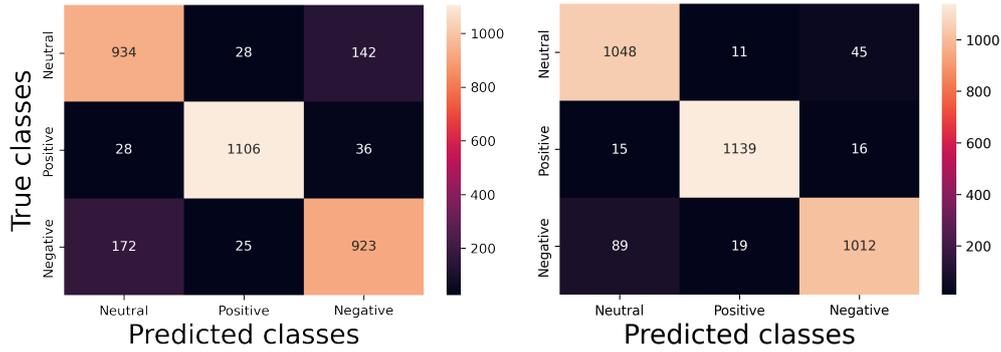


Figure 5: Confusion matrix for subject 1 (session 1) for SEED dataset using (a) MLP classifier (b) CNN classifier.

considered as low valence (LV) and low arousal (LA). The classification performance of the proposed models is evaluated for two different sets of classes which are summarised as follows:

4.2.1. High and low valence classes

The accuracy of MLP classifier for DEAP dataset is shown in Figure 6 for low and high valence classes. Subjects 15 and 22 have shown the highest and lowest accuracy of 97.2% and 86%, respectively. The accuracy is higher than or equal to 95% for the following 10 subjects, 1,6,7,10,13,15,16,18,23 and 27. The results of the CNN classifier in terms of accuracy are shown in Figure 7. Once again, the highest accuracy of 97.29% is observed for 15th subject. The CNN classifier shows the lowest accuracy of 86.4% for 5th subject. However, for subject 22, the CNN classifier shows 88.5% accuracy for which MLP classifier shows the lowest accuracy. It should be noted that accuracy above 95% is observed for the same 10 subjects as mentioned for the MLP classifier. Hence, it can be perceived that there is not any significant difference between MLP and CNN classifiers' performance for the DEAP dataset for high and low valence classes. The average accuracy and F1-score over 32 subjects are depicted in Figure 11. The classification results are not significantly different for MLP and CNN classifiers (p -value > 0.05). For subject 1, the summation of the five confusion matrices obtained over 5-folds is shown in Figure 8. From Figure 8, it can be perceived that the two models are performing well for both high and low valence cases. Also, from the confusion matrix, it can be inferred that performance of the CNN and MLP models are nearly closed to each other. The reason for this is that the SEED used 62 channels while the DEAP has only 32 channels and therefore the CNN does not benefit much from only 32 channels compared to 62 channels that provides more information about possible correlations among nearby channels.

4.2.2. High and low arousal classes

The accuracy of MLP classifier for DEAP dataset is shown in Figure 9 for low and high arousal classes. Subjects 13 and 22 have shown the highest and lowest accuracy of 97.33% and 86.7%, respectively. The accuracy is higher than 95% for the following subjects, 1,3,7,10,12,13,15,16,18,19,20,21,23,24,25 and 29. The results of CNN classifier in terms of accuracy

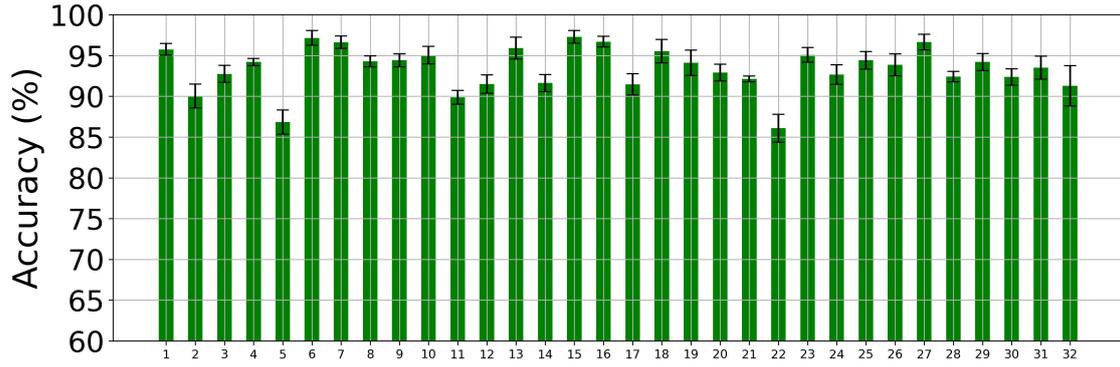


Figure 6: Accuracy (average over 5-folds) for all 32 subjects for DEAP dataset using MLP classifier. The X-axis represents subjects. High and low valence classes.

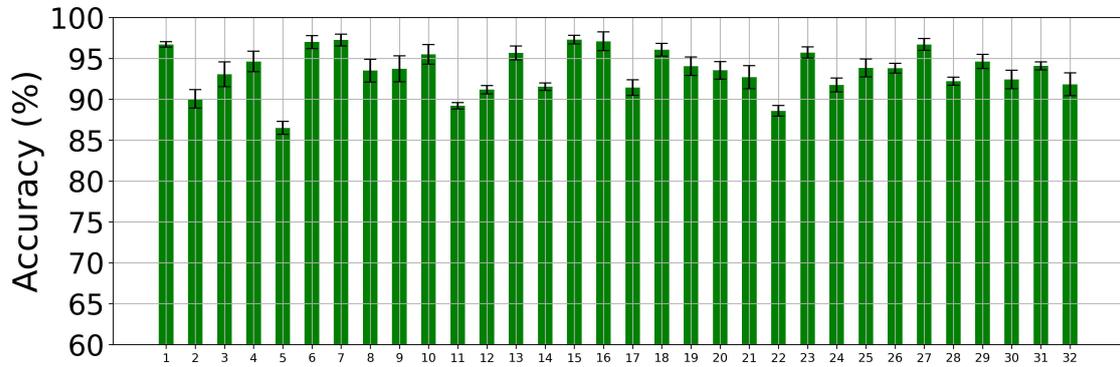


Figure 7: Accuracy (average over 5-folds) for all 32 subjects for DEAP dataset using CNN classifier. The X-axis represents subjects. High and low valence classes.

are shown in Figure 10. Once again, the highest accuracy of 97.75% is observed for the 13th subject. The CNN classifier shows the lowest accuracy of 86.4% for 5th subject. However, for subject 22 the CNN classifier shows 89.4% accuracy for which MLP classifier shows the lowest accuracy. Subjects 1,3,7,12,13,15,16,18,19,20,21,23,25 and 27 have shown more than 95% accuracy for the CNN classifier. There is not any significant difference (p -value>0.05) between MLP and CNN classifier performance for high and low arousal classes. The average result over 32 subjects are shown in Figure 11, which are also not significantly different for MLP and CNN classifiers. The same reason as mentioned for high and low valence classes indicated for similar results of MLP and CNN classifiers for high and low arousal classes. The extra strength of CNN is not effectively used when the number of channels is 32. In Figure 12, the summation of the five confusion matrices obtained over 5-folds for subject 1 is demonstrated. It can be seen in Figure 12, that the two models are equally good to detect the high and low arousal classes.

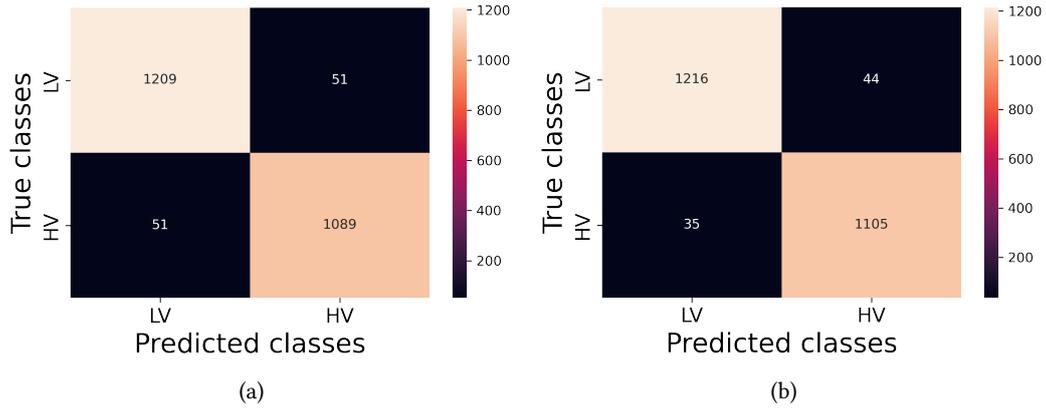


Figure 8: Confusion matrix for subject 1 (High and low valence classes) for DEAP dataset using (a) MLP classifier and (b) CNN classifier.

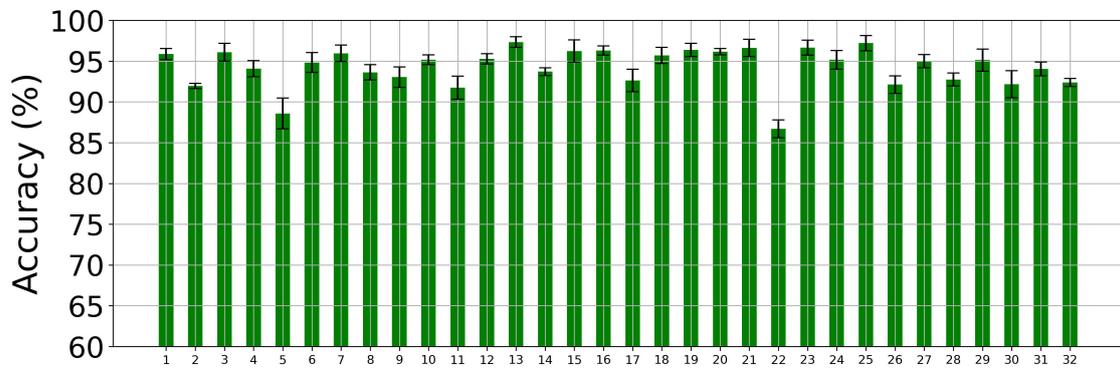


Figure 9: Accuracy (average over 5-folds) for all 32 subjects for DEAP dataset using MLP classifier. The X-axis represents subjects. High and low arousal classes.

4.3. Comparison with other works

The performance of the proposed MLP and CNN models are also compared with other state of the art methods. The comparison is summarised in Table 1. In [14], a parallel convolution recurrent neural network (PCRNN) is proposed. This method has achieved an average accuracy of 90.80% and 91.03% for high/low valence classes and high/low arousal classes respectively, with a 10-fold cross-validation method on DEAP dataset. The continuous convolutional neural network (CCNN) method proposed in [21] yielded 89.45% and 90.24% accuracy for high/low valence classes and high/low arousal classes respectively, with a 10-fold cross-validation method on DEAP dataset. In [15], the 4D-convolution recurrent neural network (4D-CRNN) method is proposed. It has achieved 94.22% and 94.58% accuracy for high/low valence classes and high/low arousal classes respectively, with a 5-fold cross-validation method on DEAP dataset. The same method is also applied to the SEED dataset and achieved 94.74% accuracy. The performance of our method is also comparable to the above-mentioned state-of-the-art methods. The architecture of the classification model in the above-mentioned methods [14, 15] exhibits

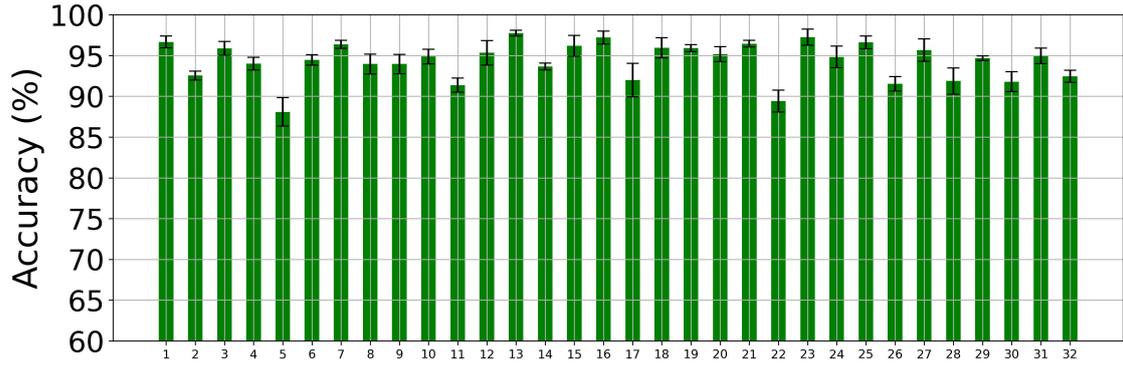


Figure 10: Accuracy (average over 5-folds) for all 32 subjects for DEAP dataset using CNN classifier. High and low arousal classes. The X-axis represents subjects.

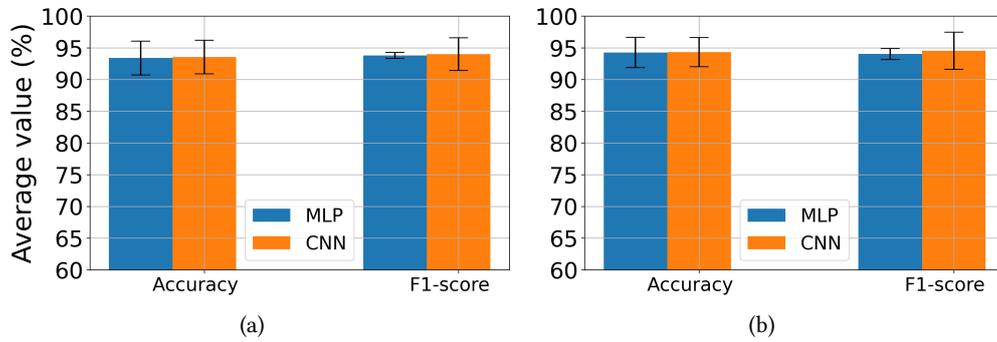


Figure 11: Average results over 32 subjects for DEAP dataset using MLP and CNN classifier. (a) High/low valence classes, and (b) High/low arousal classes.

higher complexity than the ones presented in this paper. However, our results are not directly comparable to the results of the method proposed in [14, 21], as they have used a 10-fold cross-validation method. The obtained results indicate that comparable accuracy can be achieved with less complex methods and also highlight the importance of the spatial features when using CNN. The spatial features acquired relevance only when the number of channels was 62 compared to when the number of channels was 32.

5. Conclusion

In this work, two different classification approaches which are based on MLP and CNN models are analysed and their performance is investigated on two publicly available EEG emotion datasets, SEED and DEAP. The DE feature is used to feed the input of the classifiers. DE feature is computed from the different EEG rhythms namely, delta, theta, alpha, beta, and gamma. The experimental results show that CNN-based method outperforms the MLP-based method for the SEED dataset. The average accuracy for CNN-based method for SEED dataset is 93.8%.

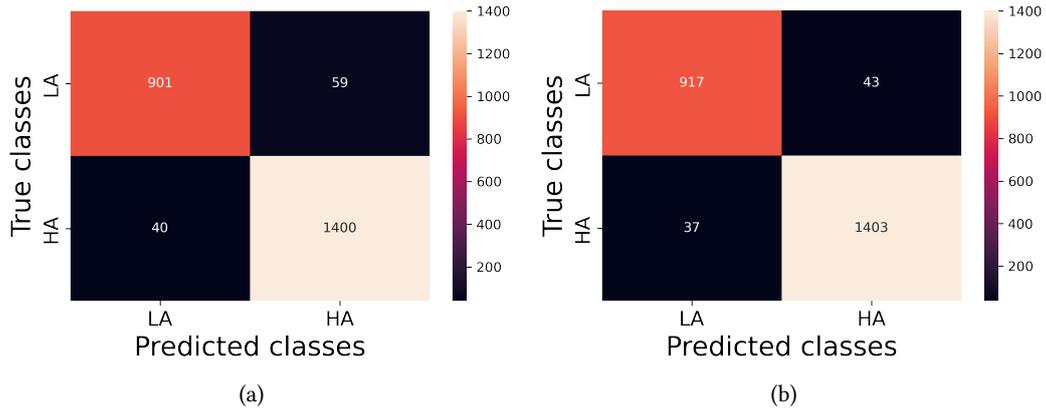


Figure 12: Confusion matrix for subject 1 (High and low arousal classes) for DEAP dataset using (a) MLP classifier, and (b) CNN classifier.

Table 1

Comparison of the results (average accuracy (ACC) and standard deviation (STD)) of the present work with other works.

Authors	Method	Cross-validation	DEAP (32 channels)		SEED
			High and low valence	High and low arousal	(62 channels)
			ACC/STD	ACC/STD	ACC/STD
Yang et al. [14]	PCRNN	10-fold	90.8/3.08	91.03/2.99	
Yang et al. [21]	CCNN	10-fold	89.45	90.24	
Shen et al. [15]	4D-CRNN	5-fold	94.22/2.61	94.58/3.69	94.74/2.32
Present work	MLP	5-fold	93.39/2.66	94.25/2.37	86.8/6.4
	CNN	5-fold	93.53/2.65	94.33/2.29	93.81/3.21

However, for DEAP dataset, no significant difference is observed in the performance of MLP and CNN-based approaches. The average accuracy for DEAP dataset is found to be 94.33% and 93.53% for high/low arousal and high/low valence classes, respectively. The obtained results show that it is possible to achieve state-of-the-art performance with less complex network models. In future, our aim is to improve the obtained results and to achieve this, we plan to implement a channel selection approach to select the channels which are more relevant for emotion detection. We will also work to develop new features that can achieve better results.

6. Acknowledgments

This work is supported by the European Research Consortium for Informatics and Mathematics (ERCIM) fellowship.

References

- [1] W.-L. Zheng, B.-L. Lu, Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks, *IEEE Transactions on Autonomous Mental Development* 7 (2015) 162–175.
- [2] S. M. Alarcão, M. J. Fonseca, Emotions recognition using EEG signals: a survey, *IEEE Transactions on Affective Computing* 10 (2019) 374–393.
- [3] D. Sammler, M. Grigutsch, T. Fritz, S. Koelsch, Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music, *Psychophysiology* 44 (2007) 293–304.
- [4] X. Li, D. Song, P. Zhang, Y. Zhang, Y. Hou, B. Hu, Exploring EEG features in cross-subject emotion recognition, *Frontiers in Neuroscience* 12 (2018). doi:10.3389/fnins.2018.00162.
- [5] Y. Li, W. Zheng, Y. Zong, Z. Cui, T. Zhang, X. Zhou, A bi-hemisphere domain adversarial neural network model for EEG emotion recognition, *IEEE Transactions on Affective Computing* 12 (2021) 494–504.
- [6] Y.-P. Lin, C.-H. Wang, T.-L. Wu, S.-K. Jeng, J.-H. Chen, EEG-based emotion recognition in music listening: a comparison of schemes for multiclass support vector machine, in: *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 489–492.
- [7] M. Li, B.-L. Lu, Emotion classification based on gamma-band EEG, in: *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2009, pp. 1323–1326.
- [8] R.-N. Duan, J.-Y. Zhu, B.-L. Lu, Differential entropy feature for EEG-based emotion classification, in: *6th International IEEE/EMBS Conference on Neural Engineering (NER)*, 2013, pp. 81–84.
- [9] Z. Yin, Y. Wang, L. Liu, W. Zhang, J. Zhang, Cross-subject EEG feature selection for emotion recognition using transfer recursive feature elimination., *Front Neurobot* (2017).
- [10] A. Craik, Y. He, J. L. Contreras-Vidal, Deep learning for electroencephalogram (EEG) classification tasks: a review, *Journal of Neural Engineering* 16 (2019) 031001.
- [11] W.-L. Zheng, J.-Y. Zhu, Y. Peng, B.-L. Lu, EEG-based emotion classification using deep belief networks, in: *2014 IEEE International Conference on Multimedia and Expo (ICME)*, 2014, pp. 1–6.
- [12] P. Zhong, D. Wang, C. Miao, EEG-based emotion recognition using regularized graph neural networks, *IEEE Transactions on Affective Computing* 13 (2022) 1290–1301.
- [13] L.-Y. Tao, B.-L. Lu, Emotion recognition under sleep deprivation using a multimodal residual LSTM network, in: *International Joint Conference on Neural Networks (IJCNN)*, 2020, pp. 1–8.
- [14] Y. Yang, Q. Wu, M. Qiu, Y. Wang, X. Chen, Emotion recognition from multi-Channel EEG through parallel convolutional recurrent neural network, in: *2018 International Joint Conference on Neural Networks (IJCNN)*, 2018, pp. 1–7.
- [15] F. Shen, et al., EEG-based emotion recognition using 4D convolutional recurrent neural network, *Cognitive Neurodynamics* 14 (2020) 815–828.
- [16] S. Koelstra, et al., DEAP: a database for emotion analysis using physiological signals, *IEEE Transactions on Affective Computing* 3 (2012) 18–31.
- [17] DEAP Dataset, <https://www.eecs.qmul.ac.uk/mmv/datasets/deap/readme.html>, . Accessed:

2022-08-30.

- [18] SEED Dataset, <https://bcmi.sjtu.edu.cn/home/seed/seed.html>, . Accessed: 2022-08-30.
- [19] X.-W. Wang, D. Nie, B.-L. Lu, Emotional state classification from EEG data using machine learning approach, *Neurocomputing* 129 (2014) 94–106.
- [20] L.-C. Shi, Y.-Y. Jiao, B.-L. Lu, Differential entropy feature for EEG-based vigilance estimation, in: 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2013, pp. 6627–6630.
- [21] Y. Yang, Q. Wu, Y. Fu, X. Chen, Continuous convolutional neural network with 3D input for EEG-based emotion recognition, in: L. Cheng, A. C. S. Leung, S. Ozawa (Eds.), *Neural Information Processing*, 2018, pp. 433–443.
- [22] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006.
- [23] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*, MIT Press, 2016.
- [24] F. Chollet, et al., Keras, <https://keras.io>, 2015.
- [25] D. Kingma, J. Ba, Adam: a method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).