Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 207 (2022) 3087–3096

Procedia Computer Science

www.elsevier.com/locate/procedia

26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Unsupervised emotional state recognition based on clustering of EEG features

Aleksandra Dura[a], Agnieszka Wosiak[a,*]

[a]Lodz University of Technology, Institute of Information Technology,Wólczańska 215, Lodz, Poland

Abstract

Efficient information retrieval from the EEG sensors is a complex and challenging task, particularly in the context of psychology, including emotional states. Therefore, different machine learning strategies are considered to improve the processes based on EEG signal analysis. Most of them use supervised approaches since EEG datasets usually include metadata and descriptions that can be used for learning. However, these descriptions are mainly based on self-reports of emotional states, which means that they may not be reliable or objective. The paper proposes an approach that incorporates unsupervised learning techniques as a solution supporting classification where classification labels may be uncertain. The research proved that our approach improves the recognition of emotions and gives results with an average accuracy greater by five percentage points.

Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)

Keywords: emotion recognition; EEG analysis; feature extraction; clustering; unsupervised learning;

1. Introduction

Emotion recognition is gaining popularity in a number of domains, including human-computer interaction and affective computing [1]. By examining physiological and non-physiological emotion-related information, the technique tries to automate the identification of human emotion. An electroencephalogram (EEG) can be used to collect physiological signals from the central nervous system. EEG data, which are non-invasive, directly indicate the strength and location of brain activity with excellent temporal resolution. As a result, EEG-based emotion recognition is an important topic [2, 3].

Although it has received considerable research interest, adequate identification of the level of the emotional state remains a significant challenge. It involves a number of aspects that still need to be considered in research:

- the number of electrodes from which the EEG signal originates,

* Corresponding author.
E-mail address: agnieszka.wosiak@p.lodz.pl

1877-0509
Peer-review under responsibility of the scientific committee of the 26th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems (KES 2022)
10.1016/j.procs.2022.09.367

- feature extraction from the original EEG signal, and
- the availability of training data that includes correct labels indicating the level of emotional state.

EEG signal analysis is directly related to the optimal choice of measuring electrodes. Electrical signals are recorded using electrodes typically placed on the scalp according to the international 10-20 electrode placement system. However, the selection of specific electrodes can be made depending on the purpose of the application and can have the benefits of reducing computational complexity and ignoring non-relative noise. The optimal electrode placement is usually determined by statistical factors such as correlation coefficient or feature selection methods [4, 5].

Feature engineering is another key step in the processing and analysis of EEG signals, affecting the results of further analysis. It involves EEG signal preparation, including downsampling, band-pass filtering, windowing, artifact handling, and feature extraction. Many researchers emphasize automated feature learning as a motivation for using deep learning-based approaches [6, 7, 8] - however, a significant number of research articles use at least one preprocessing method. Studies focusing on emotion recognition using the DEAP dataset [2] have typically used the same preprocessing methodology proposed by the dataset authors as a first step [8]. However, the next step, feature extraction by transforming them to the time domain or frequency domain and then selecting the key signal representation, is equally important [9].

Defining reliable labels for psychological states is a challenging task [10]. It is usually based on a human expert assessment or self-assessment. Given a labeled dataset, which can be taken to further analysis, a typical task in emotion recognition is therefore concentrated on classification or regression issues. However, self-assessment is burdened with a dose of uncertainty, which may provide a reason to validate the approach by using unsupervised learning methods [11].

We propose to use unsupervised machine learning i.e. k-means algorithm to recognize the high or low level of emotional state based on extracted features from EEG signal considering full and limited number of channels. For automatic electrode selection, we used the RCA algorithm based on inverse correlation analysis (Reversed Correlation Algorithm [12]), applied in between-subjects analysis to identify band-channel combinations most correlated with emotional states, and in within-subjects analysis to make the results universal [13]. RCA was validated by classifying emotions using the Support Vector Machine (SVM) technique. In addition - based on previous studies described in [14] - we have shown that a limited number of input channels also enhances the deep learning approach. Nevertheless, previous studies have not considered the additional use of feature extraction. In this study, we used several features dedicated to real-time data to represent the original signal. Furthermore, given that supervised learning techniques require training data, which may be inherently more difficult to acquire in the context of emotion recognition, we demonstrate that the use of unsupervised learning may be beneficial for the quality of emotion recognition. Thus, our approach may contribute to advancing research work with EEG signals in the context of affective computing.

The remainder of the paper is organized as follows. The methodology, including emotion recognition problem in terms of the EEG test, dataset description, electrode selection, feature engineering and supervised and unsupervised approach to machine learning, is described in the following section. Next, the experiments carried out are depicted, and the obtained results are discussed. The final section presents the study's conclusions and indicates future research.

2. Materials and Methods

2.1. Emotional State Recognition using EEG Signal

As the emotional state is reflected by brain activity, which in turn can be monitored by electroencephalography, EEG analysis has been incorporated into emotion recognition. Different emotion categorization scales have been proposed to interpret emotions automatically, but Russell's valence-arousal scale is one of the most extensively utilized [15]. Russell's scale uses a two-dimensional plane representation of each psychological state, with arousal and valence serving as the vertical and horizontal axes. However, automated emotion classification is challenging because of the subjectivity of the method, which usually entails collecting EEG data and comparing them to expert classification or self-assessment.

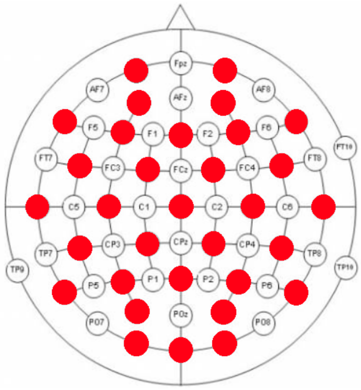

Fig. 1. EEG initial channels.

2.2. *Dataset Description*

A Database for Emotion Analysis using Physiological Signals (DEAP) is one of the most popular publicly available emotional EEG datasets [2]. It includes the EEG and peripheral physiological signals of 32 participants when watching 40 one-minute music videos. The dataset consists of 32 channels originally preprocessed by the authors of the dataset by using bandpass filter set between 4Hz and 45Hz, and the EOG artifacts removal. Except for channels, data contains participants' rate of each video regarding the levels of arousal, valence, like/dislike, dominance, and familiarity on a scale between 1 and 9.

In our research, we consider only 32 electrodes placed on the scalp. The original distribution of EEG channels is shown in Fig. 1.

Furthermore, our study concentrated on the two most objective emotions, valence and arousal. The valence, according to Russell's scale, represents how the viewer feels while watching the film (happy or sad). Arousal, on the other hand, indicates how strong a film's impression is (calm or enthusiastic).

The original ratings were converted to a binary scale, with 0 indicating low valence/arousal (less than 4.5 points on the original scale) and 1 indicating strong valence/arousal (greater than 4.5 points). Because the scale includes a 9-point rating, a score of 4.5 was used as a cutoff.

2.3. *EEG Channel Selection*

EEG signals are collected from numerous locations across the scalp and often involve more than 100 electrodes. When analyzing EEG signals, too large number of electrodes may have a negative impact on computational complexity. It also raises the potential of signal overlapping, which can also lead to recognition problems. As a result, effective channel selection is critical for further analysis. According to the literature, a smaller group of channels could achieve similar or even identical performance [4]. What is more, electrode selection could be critical in overcoming issues with complicated and high-intrusive devices. However, experts disagree on the exact number and placement of EEG electrodes.

To reduce the number of electrodes needed for emotion categorization, we propose adopting the RCA feature selection approach. The RCA method for reducing the number of EEG electrodes for emotion recognition was described in [13]. The authors excluded channels least correlated with emotional state applying three-stage procedure:

1. Selecting bands of frequencies based on calculating average frequencies for each second of a trial.
2. Selecting band-electrode combinations based on a statistical analysis of correlation coefficients using an intra-subject approach.
3. Building an inter-subject subset of electrodes by eliminating all the electrodes that did not appear in any user's subset and summing up the occurrences.

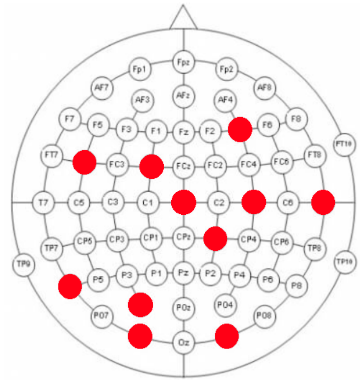

Fig. 2. Selected EEG channels.

2.4. EEG Feature Extraction

Due to EEG signals' high complexity and nonlinearity, wavelet decomposition is a typical method to extract the most significant features for further analysis.

Based on the literature studies and our investigation, we propose using six parameters for each channel as described below. All parameters assume data segmentation into 4 seconds with 0.5 seconds overlapping. One EEG signal's length lasts 60 seconds with 128Hz downsampling, so the total number of rows is 7680.

Power Spectral Density. Power spectral density (PSD) is the most recommended statistical feature in EEG-based emotional analysis [16]. It estimates the power of a signal at different frequencies for any temporal signal [17].

PSD features are extracted from five frequency bands:

- theta (4-8Hz),
- alpha (8-12Hz),
- small beta (12-16Hz),
- large beta (16-25Hz),
- gamma (25-45Hz).

Hjorth's Parameters. Hjorth Parameters - also known as normalized slope descriptors (NSDs) - are based on statistical calculations and describe the characteristics of EEG signals in the time domain. Hjorth Parameters include activity, mobility, and complexity [18]. In our experiments, we consider only the complexity measure, which provides the information on how the obtained signal is similar to a pure sine wave and estimates the signal's bandwidth.

Differential Entropy. Differential entropies (DEs) are widely used in EEG-based analysis, including emotion recognition. They are based on Shannon's entropy [19]. We consider spectral entropy (SE), which measures complex signals in the frequency domain after transforming the signal into Power Spectral Density. It has already been proved that SE gives good results [9, 20].

Detrended Fluctuation Analysis. Detrended Fluctuation Analysis compares fluctuations of the selected window of time series with longer fluctuations [21]. The main application of this parameter is disease detection, such as Alzheimer's, epilepsy, or depressive disorder [22]. However, it may also be helpful for emotion recognition taking into account the high variability of states produced in the brain.

Fractal Dimension. Fractal dimension (FD) is another measure of a signal's complexity, as FD of any signal has an inverse relation to the complexity of that signal. Time domain fractal dimension can be calculated using Higuchi's method. It was proved to give good results for mental disorders recognition, e.g., schizophrenia, depression, or anxiousness [23, 24]. All the diseases mentioned earlier and emotional state analysis share the same background, so using HFD in emotion recognition seems reasonable.

Mean of Time Series. The last feature - mean of time series - is taken as a statistical measure.

2.5. Classification using Support Vector Machine

Due to the popularity and importance of data classification, research on the quality of the classification process is often undertaken in research projects. The selection of an appropriate classification algorithm is one of the critical elements affecting the success of the whole task. When deciding on a specific classification algorithm, besides measurable indicators, i.e., validation measures that confirm the correctness of the classifier's performance, one should also consider factors that facilitate human experts' understanding of the algorithm's performance. These factors primarily include:

- clear graphical representation [25],
- operation consistent with human performance, rather than so called "black box" approach [26].

The Support Vector Machine (SVM) method divides the data vector space into two disjoint regions corresponding to classes. Thus, the task can be summarized as finding the decision boundary between classes and is related to the concept of linear separability, according to which two classes are linearly separable when there exists a hyperplane H expressed by $g(x)$ according to (1)

$$g(\mathbf{x}) = \mathbf{w}^T\mathbf{x} + b, \tag{1}$$

taking values as in (2)

$$\begin{cases} g(\mathbf{x}_i) > 0 & \text{for } \mathbf{x}_i \in 1, \\ g(\mathbf{x}_i) < 0 & \text{for } \mathbf{x}_i \in -1. \end{cases} \tag{2}$$

where: $\mathbf{x}$ - denotes the data vector, while $\mathbf{w}$ and b are parameters of the model.

As a result, there is a set of multiple possible solutions (i.e., hyperplanes), from which the hyperplane that maximizes the margin of the linear classifier is chosen. The model parameters should be adjusted so that the maximum margins, defined as the distance between the boundary hyperplanes m_{i1} and m_{i2}, are the geometric locations of the points $\mathbf{x}$ satisfying the conditions expressed by (3).

$$\begin{array}{ll} \mathbf{w} * \mathbf{x} + b = 1 & \text{for } m_{i1}, \\ \mathbf{w} * \mathbf{x} + b = -1 & \text{for } m_{i2}. \end{array} \tag{3}$$

Before the SVM classification, data was normalized to get better performance. The classifier was learned using 70% - 30% data split. The experiment was repeated ten times for the evaluation purpose, and mean accuracy was taken as a metric.

2.6. Classification Using Deep Learning Approach

Nowadays, Convolutional Neural Networks (CNN) are one of the most commonly used deep learning methods. The main application of CNNs is two-dimensional image data analysis. However, they are also increasingly applied for real-time data such as EEG signals [27]. What is more, the usage of CNNs for electroencephalography data has been growing exponentially in the last few years. The very high complexity of the data is the main reason that convolutional neural networks are of great interest in EEG data analysis.

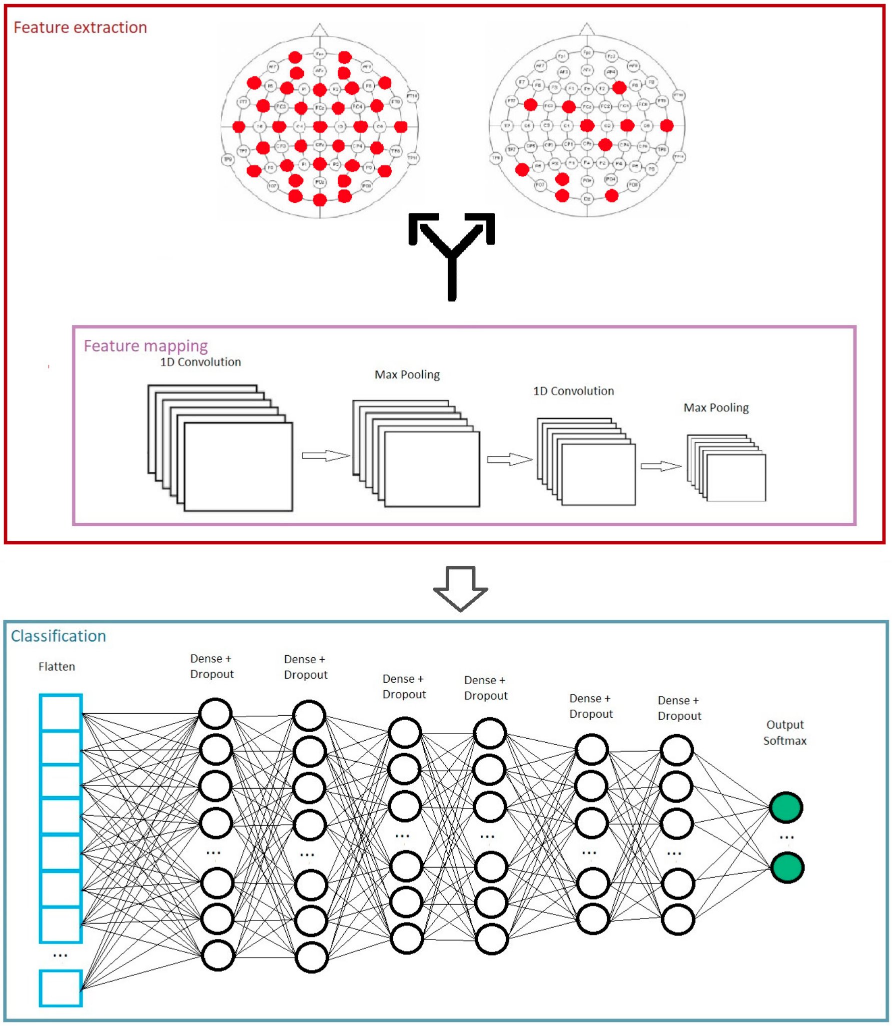

Fig. 3. Architecture of CNN.

Our proposed algorithm using the sequential architecture of CNN consists of an input layer followed by hidden layers, and a final output layer. It is illustrated in Fig. 3 and consists of the following actions:

1. The algorithm starts with one-dimensional convolutional input to extract features from the dataset with 128 neurons and a kernel size equal to 10.
2. Then, max-pooling is added to reduce dimensionality and pick the most significant features.
3. Due to the complexity of features, one more convolutional layer with pooling operation is implemented.
4. Next, flatten layer converts the data into a single feature vector to prepare input for final classification.
5. The next stage of CNN is a fully connected layer that consists of six dense layers. These structures aggregate the most meaningful information from all extracted features and are configured as follows:

    - A dropout of 0.2 is placed, i.e., 20% of random nodes are excluded from training data to prevent an overfitting problem.

- ReLU function is implemented as it is recommended for signals processing by many researchers [28, 29, 30]. In comparison to other activation functions, ReLU is more efficient when avoiding gradient disappearance [29].
- The number of neurons in dense layers is reduced twice, starting from 64.

6. The last layer uses ten neurons and a Softmax classifier for carrying out probability estimation as an output, as proposed in [31].

The dataset is split according to 80% - 20% training-testing schema. Moreover, we use 200 training epochs and a batch size set of 400 in our configuration.

2.7. Clustering with k-Means Algorithm

Cluster analysis is an unsupervised classification technique used to group complex multidimensional data. Opposite to supervised methods, the profiles of obtained groups cannot be stated precisely, and using additional techniques for discovering the meaning of clustering is required in many cases [32].

In further investigations, which aim to evaluate the presented technique regarding its efficiency on EEG data, a deterministic k-means clustering algorithm is considered. Compared to other techniques, it demonstrated good performance for medical data regarding the accuracy and lower root mean square error [33]. The k-means algorithm is one of the most popular partitioning methods, where clusters are built around k centers by minimizing a distance function. The goal of the algorithm is to find the set of clusters for which the sum of the squared distance values between their points and respective centers is minimal. The Euclidean metric, applied in most cases [34, 35], is used as the distance function. First, k centers are usually chosen randomly, which does not guarantee to find optimal clusters. The algorithm is usually launched several times with different initial choices to increase the chance of finding the optimum, and the result of the smallest total squared distance is indicated [35].

The optimal number of clusters is one of the most crucial parts of the clustering process. We used the elbow technique based on the statement that the number of clusters should increase together with the quantity of information. The last number of clusters, for which a gain value is augmented, should be indicated as optimal. That cut-off point is presented as an angle in the graph illustrated validation measure plotted against the number of clusters. However, it is worth noticing that there are cases when angles cannot be unambiguously identified, and other methods should confirm the number of clusters indicated by the elbow technique. In emotion recognition, most research split labels into the ”low” level (”0” - values 1-5) and the ”high” level (”1” - values 5-9), and then, further analysis is carried out [9, 20, 28]. Therefore, there is a natural intent to split the whole dataset into two groups.

In our experiments the k-Means method was carried out with 300 iterations.

3. Results and Discussion

The experiments aimed to analyze the impact of unsupervised labeling of emotional states’ levels on emotion recognition using supervised classification.

A lack of literature in the context of unsupervised emotion recognition methods can signify that it is not effective due to the difficulty in analyzing the EEG signal and extracting features appropriately. On the other hand, the use of supervised learning methods may be impaired by users’ ambiguous self-assessment [14]. Therefore, in our current research, we decided to combine these two approaches by considering cluster analysis only in cases where the classification was not unambiguous. In this way, uncertain labels could be replaced by automatically generated ones based on cluster characteristics.

The experiments were conducted with the application of methods and a dataset introduced in Section 2. Three main procedures were performed as illustrated in Fig. 4:

A. The experimental procedure that includes classification with original labels.
B. The experimental procedure that assigns labels obtained from grouping for the ”uncertain” cases, i.e., cases misclassified by both SVM and CNN.

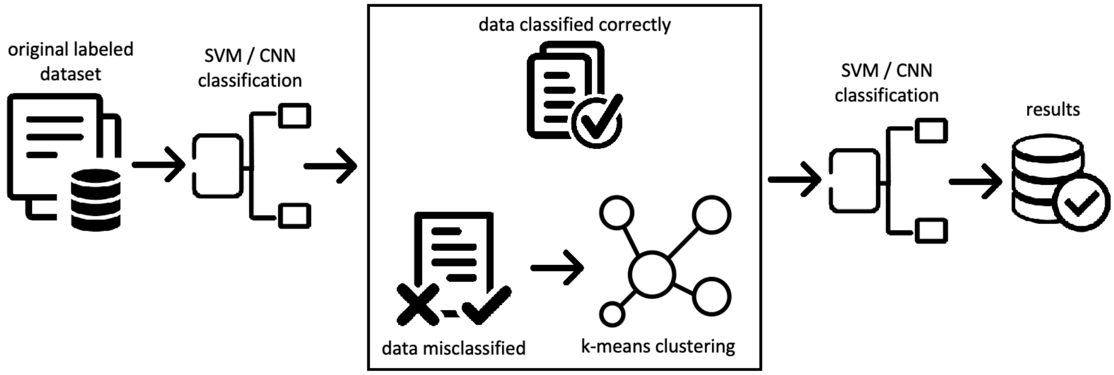

Fig. 4. Schema of the proposed experimental procedure

Table 1. Accuracy of the valence / arousal classification with and without grouping replacement for 32 and 11 channels

		32 channels				11 channels			
Procedure		Valence ACC [%]		Arousal ACC [%]		Valence ACC [%]		Arousal ACC [%]	
		SVM	**CNN**	**SVM**	**CNN**	**SVM**	**CNN**	**SVM**	**CNN**
original labels	min	76.62	49.23	75.88	47.68	71.83	49.23	70.50	47.68
	max	96.76	82.19	96.31	89.16	94.32	82.19	94.32	89.16
	mean	85.80	63.80	86.65	67.50	83.36	63.83	84.38	67.42
	median	86.50	63.22	86.95	68.31	83.41	63.22	84.99	68.25
after grouping replacement	min	72.62	62.38	71.35	71.46	72.35	62.80	69.89	62.42
	max	91.85	92.67	94.28	96.84	96.32	89.90	92.78	96.64
	mean	80.88	75.27	82.30	86.85	85.07	70.91	82.27	83.98
	median	80.66	73.75	81.63	88.71	85.48	68.86	82.47	88.24

C. The experimental procedure that includes classification with labels after grouping imputation.

We compared the classification process with and without grouping replacement with the initial 32 channel input and the selected subset of 11 electrodes. Only valence and arousal were analyzed. The experimental environment was based on Python programming language and its libraries.

The comparison results are presented in Table 1 and illustrated in Fig. 5. The accuracy was calculated as a similarity measure, and the mean value and the median were calculated for all users.

The first column of Table 1 indicates the procedure type. The second informs of the statistics type. The following eight columns contain the average statistics for valence and arousal, including classification by SVM and CNN. One can notice that replacing labels derived from clustering improves the classification of emotional states in most of the cases. On the other hand, the replacement process is complex, and some improvements in classification accuracy are not statistically important. Nevertheless, even the tiny gain obtained by label substitution shows that self-assessment of psychological states should be considered with great caution and that automating classifications or regressions on this basis may be affected to some extent.

To sum up, the comparison of accuracies justifies clustering analysis in terms of emotional states' level recognition.

4. Conclusions

The research presented in this paper addresses the problem of applying clustering in assessing the level of emotional state analyzed from EEG signals.

Our experiments showed that the difficulty in classifying emotional states might be related to the incorrect self-assessment of emotional states, and thus the correct assignment of labels is a crucial aspect to be verified before

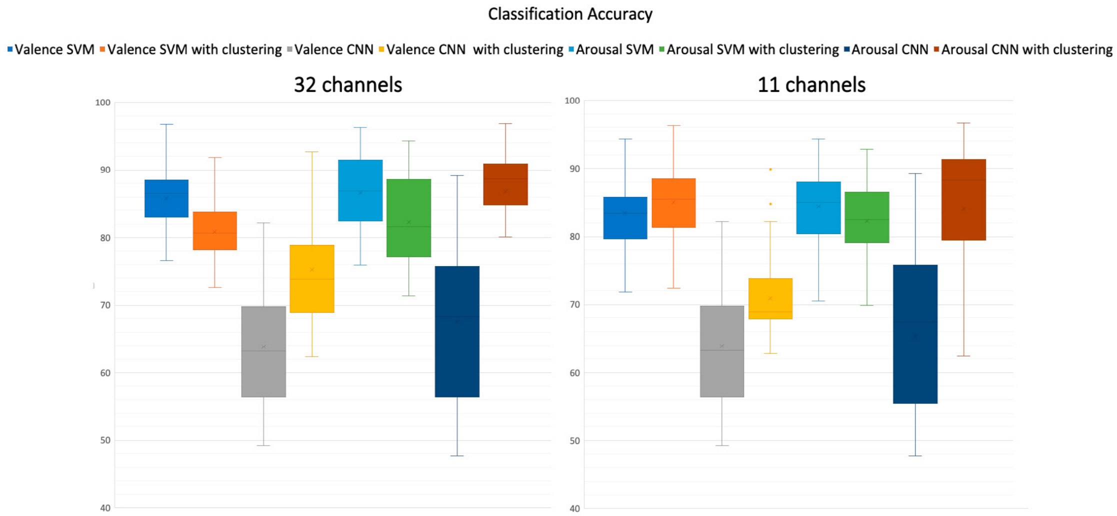

Fig. 5. Accuracy of the valence / arousal classification with and without grouping replacement

proceeding further. Unfortunately, this is a challenging issue to address, requiring multiple self-assessment tests and impossible in the context of existing reference sets. Thus, the use of clustering, which, as an unsupervised learning technique, does not require labels but instead discovers the intrinsic characteristics of the data itself and performs partitioning based on this, is a perfect solution that can also be applied to supervised learning, improving its accuracy, as demonstrated in the paper.

Nevertheless, additional research is still needed. Therefore, further research is planned to apply Jimmy methods for cluster analysis and extraction of additional features [36, 37, 38], which would extract better quality clusters and thus simplify the process of label substitution. Furthermore, more profound insights into the characteristics of better quality clusters could also benefit understanding emotional states.

References

[1] Marin-Morales, J., Higuera-Trujillo, J. L., Greco, A., Guixeres, J., Llinares, C., Scilingo, E. P., Alcaniz, M., Valenza, G. (2018) "Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors." *Scientific Reports*, 8(1), pp. 1–15.

[2] Koelstra, S., Muhl, C., Soleymani, M., Lee, J.S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., Patras, I. (2012) "DEAP: A database for emotion analysis; using physiological signals." *IEEE Transactions on Affective Computing*, 3(1), pp. 18–31.

[3] Shu, L., Xie, J., Yang, M., Li, Z., Liao, D., et al. (2018) "A Review of Emotion Recognition Using Physiological Signals." *Sensors*.

[4] Wosiak, A., and Dura, A. (2020) "Hybrid Method of Automated EEG Signals' Selection Using Reversed Correlation Algorithm for Improved Classification of Emotions." *Sensors*, 20(24), 7083.

[5] Al-Saegh, A., Dawwd, S. A., and Abdul-Jabbar, J. M. (2021) "Deep learning for motor imagery EEG-based classification: A review." *Biomedical Signal Processing and Control*, 63, 102172.

[6] Hasib, M. M., Nayak, T., and Huang, Y. (2018) "A hierarchical LSTM model with attention for modeling EEG non-stationarity for human decision prediction." *IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, pp. 104–107.

[7] Hussein, R., Palangi, H., Ward, R., and Wang, Z. J. (2018) "Epileptic seizure detection: A deep learning approach." *arXiv preprint*, arXiv:1803.09848.

[8] Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., and Faubert, J. (2019) "Deep learning-based electroencephalography analysis: a systematic review." *Journal of neural engineering*, 16(5), 051001.

[9] Li, X., Song, D., Zhang, P., Zhang, Y., Hou, Y., and Hu, B. (2018) "Exploring EEG features in cross-subject emotion recognition." *Frontiers in Neuroscience*, 12, DOI: 10.3389/fnins.2018.00162.

[10] Terzis, V., Moridis, C.N., Economides, A.A.(2010) "Measuring Instant Emotions during a Self-assessment Test: The Use of FaceReader.", *ACM International Conference Proceeding Series*, DOI:10.1145/1931344.1931362

[11] Betella, A., Verschure, P. F. (2016) "The Affective Slider: A Digital Self-Assessment Scale for the Measurement of Human Emotions." *PloS one*, 11(2), e0148037. https://doi.org/10.1371/journal.pone.0148037

[12] Wosiak, A.; Zakrzewska, D. (2018) "Integrating Correlation-Based Feature Selection and Clustering for Improved Cardiovascular Disease Diagnosis." *Complexity*, 250706, DOI:10.1155/2018/2520706.

[13] Dura, A., Wosiak, A., Stasiak, B., Wojciechowski, A., and Rogowski, J. (2021). "Reversed Correlation-Based Pairwised EEG Channel Selection in Emotional State Recognition." *IEEE International Conference on Computational Science*, Springer, Cham, 528–541, DOI: 10.1007/978-3-030-77967-2_44.

[14] Dura, A., Wosiak, A. (2021) "EEG channel selection strategy for deep learning in emotion recognition." *Procedia Computer Science*, 192, 2789–2796, DOI: 10.1016/j.procs.2021.09.049.

[15] Russell, J.A. (1980) "A circumplex model of affect." *Journal of Personality and Social Psychology*, 39(6), pp. 1161–1178.

[16] Pan, C., Shi, C., Mu, H., Li, J., Gao, X. (2020). "EEG-based emotion recognition using logistic regression with Gaussian kernel and Laplacian prior and investigation of critical frequency bands." *Applied Sciences*, 10(5), 1619, DOI: 10.3390/app10051619.

[17] Perez-Zapata, A. F., Cardona-Escobar, A. F., Jaramillo-Garzon, J. A., Diaz, G. M. (2018). "Deep convolutional neural networks and power spectral density features for motor imagery classification of EEG signals." In: *International Conference on Augmented Cognition*, Lecture Notes in Computer Science, 10915, pp. 158-169, DOI: 10.1007/978-3-319-91470-1_14.

[18] Hjorth, B. (1970) "EEG Analysis Based on Time Domain Properties. Electroencephalograph." *Clinical Neurophysiology* 29, 306—310, DOI: 10.1016/0013-4694(70)90143-4.

[19] Shannon, C.E. (1949) "Communication Theory of Secrecy Systems." *The Bell System Technical Journal*, 28, 656—715.

[20] Galvao, F., Alarcao, S. M., Fonserca, M. J. (2021) "Predicting Exact Valence and Arousal Values from EEG." *Sensors*, 21, 3414, DOI: 10.3390/s21103414.

[21] Marton, L. F., Brassai, S., T., Bako, L., Lasonczi, L. (2013) "Detrended fluctuation analysis of EEG signals." *Elsevier*, DOI: 10.1016/j.protcy.2013.12.465

[22] Mumtaz, W., Malik, A. S., Ali, S. S. A., Yasin, M. A. M., Amin, H. (2015) "Detrended fluctuation analysis for major depressive disorder." *IEEE*, pp. 4162-4165, DOI: 10.1109/EMBC.2015.7319311.

[23] Bachmann, M., Lass, J., Suhhova, A., Hinrikus, H. (2013) "Spectral Asymmetry and Higuchi's Fractal Dimension Measures of Depression Electroencephalogram." *Computational and Mathematical Methods in Medicine*, DOI: 10.1155/2013/251638.

[24] Kawe, T.N.J., Shadli, S.M., McNaughton, N., (2019) "Higuchi's fractal dimension, but not frontal or posterior alpha asymmetry, predicts PID-5 anxiousness more than depressivity." *Sci Rep 9*, 19666, DOI: 10.1038/s41598-019-56229-w

[25] Cerquides, J., López-Sánchez, M., Ontañón, S., Puertas, E., Puig, A., Pujol, O., Tost, D. (2005) "Classification algorithms for biomedical volume datasets." *Conference of the Spanish Association for Artificial Intelligence*, pp. 143–152, Springer.

[26] Czajkowski, M., Grześ, M., Kretowski, M. (2014) "Multi-test decision tree and its application to microarray data classification." *Artificial Intelligence in Medicine*, 61(1), pp. 35–44, Elsevier, DOI:10.1016/j.artmed.2014.01.005.

[27] Mahmud, M., Kaiser, M.S., McGinnity, T.M., Hussain, A. (2021) "Deep Learning in Mining Biological Data", *Cognitive Computation* 13, 1–33, https://doi.org/10.1007/s12559-020-09773-x

[28] Zhong, Q., Zhu, Y., Cai, D., Xiao, L., Zhang, H. (2020) "Electroencephalogram access for emotion recognition based on a deep hybrid network." *Frontiers in Human Neuroscience*, DOI: 10.3389/fnhum.2020.589001

[29] Liu, J., Wu, G., Luo, Y., Qiu, S., Yang, S., Li, W., Bi, Y. (2020) "EEG-Based Emotion Classification Using a Deep Neural Network and Sparse Autoencoder", *Frontiers in Systems Neuroscience*, 14(43), DOI: 10.3389/fnsys.2020.00043.

[30] Chao, H., Dong, L., Liu, Y., Lu, B. (2019) "Emotion Recognition from Multiband EEG Signals Using CapsNet", *Sensors*, 19(9), 2212, DOI: 10.3390/s19092212.

[31] Jirayucharoensak, S., Pan-Ngum, S., Israsena, P. (2014) "EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation", *The Scientific World Journal*, 627892, DOI: 10.1155/2014/627892.

[32] Han, J., Kamber, M., Pei, J. (2011) "Data Mining: Concepts and Techniques", Elsevier, USA, 2011.

[33] Liu, H., Yu, L. (2005) "Toward Integrating Feature Selection Algorithms for Classification and Clustering", *IEEE Transactions on Knowledge and Data Engineering*, 17, 491–502.

[34] Looney, S.W., Hagan, J.L. (2011) "Statistical Methods for Assessing Biomarkers and Analyzing Biomarkers Data." In: C.R. Rao, J.P. Miller, D.C. Rao (eds): *Essential Statistical Methods for Medical Statistics*, Elsevier, 27–65.

[35] Witten, I.H., Frank, E., Hall, M.A. (2011) "Data Mining. Practical machine learning tools and techniques", Morgan Kaufmann, San Francisco, USA, 2011.

[36] Paja, W., Pancerz, K., Grochowalski, P. (2018) "Generational feature elimination and some other ranking feature selection methods." *Advances in Feature Selection for Data and Pattern Recognition*, pp. 97–112.

[37] Paja, W., Knap M. (2020) "A Constructive Induction of Feature using Random Forest Approach." *Procedia Computer Science*, 176, 3318–3323, DOI: 10.1016/j.procs.2020.09.066.

[38] Paja, W., Pancerz, K., Pekala, B., Sarzynski, J. (2021) "Application of the Fuzzy Logic to Evaluation and Selection of Attribute Ranges in Machine Learning." In: *IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, 1–6.