2023 IEEE Symposium Series on Computational Intelligence (SSCI)
Mexico City, Mexico. December 5-8, 2023

Convolving Emotions: A compact CNN for EEG-based Emotion Recognition

Marco A. Cardoso-Moreno
Computational Cognitive Sciences Laboratory - CIC
Instituto Politécnico Nacional
Mexico City, Mexico
mcardosom2021@cic.ipn.mx

Cesar Macias
Computational Cognitive Sciences Laboratory - CIC
Instituto Politécnico Nacional
Mexico City, Mexico
cmaciass2021@cic.ipn.mx

Tania Alcantara
Computational Cognitive Sciences Laboratory - CIC
Instituto Politécnico Nacional
Mexico City, Mexico
talcantaram2020@cic.ipn.mx

Miguel Soto
Computational Cognitive Sciences Laboratory - CIC
Instituto Politécnico Nacional
Mexico City, Mexico
msotoh2021@cic.ipn.mx

Hiram Calvo
Computational Cognitive Sciences Laboratory - CIC
Instituto Politécnico Nacional
Mexico City, Mexico
hcalvo@cic.ipn.mx

Cornelio Yáñez-Márquez
Intelligent Computing Laboratory - CIC
Instituto Politécnico Nacional
Mexico City, Mexico
cyanez@cic.ipn.mx

***Abstract*—Emotion Recognition is a research area that has had a surge in interest, since areas such as mental health, psychological diagnosis, e-learning and assistance for people who are not capable of communicating their feelings, depend on certain level, on the capacities of computer systems to reliably identify emotions. There are several approaches to this task, for instance, analyzing facial expressions, speech, and physiological signals (electrocardiogram, galvanic skin response, electroencephalogram, among others). Electroencephalogram is one of the preferred methods due, in part, to is great temporal resolution. Therefore, in this paper we used the EEG Brainwave Dataset as benchmark to test our model, which is a four layer, one dimensional convolutional neural network. After the preprocessing pipeline, consisting on considering certain features of the dataset as signals and processing them accordingly, by creating several channels by two decomposition methods, our model achieved accuracy values of 98.36% and 95.31%, which is competitive with what is found on the state of the art, while being a significantly less complex model.**

***Index Terms*—Emotion Recognition, EEG, Convolutional Neural Network, CNN, Classification, Deep Learning**

I. Introduction

Biosignals are becoming—in recent years—increasingly studied for the task of emotion recognition [1], [2], among the most commonly used types of signals there are: Electrocardiogram (ECG), Galvanic Skin Response (GSR), Magneto Encephalography (MEG), Near-infrared Spectroscopy (NIRS), functional Magnetic Resonance Imaging (fMRI) and Electroencephalography (EEG); nevertheless, EEG is the one system providing better temporal resolution [2], [3], making it more suitable for real time applications; additionally it has been observed that EEG signals present certain characteristics related to emotional states [2].

Determining under what emotion an individual is at a given moment has an impact in several areas, for instance, it can help in managing mental health, promoting effective communication between pears [4] and even in Human-Computer Interaction (HCI), in areas such as online learning, psychological diagnosis and by communicating the emotional state of people who are not capable of expressing such information on their own, for example, newborns and elderly people [4], [5]. This research area, focused specifically in determining and understanding a subject's emotional state in real time, is relatively new subfield of HCI and Brain-Computer Interface (BCI) systems, is called Affective Computing (AC) [6].

In order to develop an emotion recognition pipeline, it is important to define what constitutes an emotion; according to Oatley and Jenkins [7], emotions have both psychological and physiological effects, that are reflected in muscular, as well as neurological, expressions. Based on this definition, one can infer that emotional states can be detected in readings from different kinds of sensors, such as the previously stated. Specifically, it has been observed that emotional states can be detected when reading electrical activity from the brain, procedure performed with EEG systems [2], [4].

Nevertheless, EEG presents certain drawbacks regarding its readings, since this signals tend to be complex, nonlinear, nonstationary and sensitive to interference [2], [6], [8]–[10].

In this work, we present a small one dimensional Convolutional Neural Network (CNN), consisting of three 1D convolutional layers, followed by a dense layer as classifier, for the emotion recognition task based on EEG signals. To this purpose, we selected the "EEG Brainwave Dataset: Feeling Emotions" presented in [8] by Bird and his colleagues, and publicly available at Kaggle[1]. Additionally, two different signal preprocessing pipelines were proposed and tested on the aforementioned model.

The rest of the paper is structured as follows: Section II provides an overview of related works in the field and, specifically, works that have used the EEG Brainwave Dataset; Section III presents the dataset and our methodology, from data preprocessing to validation procedures; Section IV is the cornerstone of this paper, since it is where our proposed model is introduced; in Section V we present our empirical results, as well as additional information on the training and testing setups; and, lastly, in Section VI we offer a conclusion from our results and observations.

II. Related Works

Emotion recognition through AC is a research area studied from several perspectives, such as facial expressions [11], [12], through speech [13], and physiological signals, for instance, ECG [14], [15], GSR [16], [17] and EEG [10]. From these methods EEG is the preferred since it is capable of reflecting human brain activity [18].

In this regard, several EEG based datasets have been published, for instance: the one presented in [19], the DEAP dataset [20], the DER-VREEG [21] and the EEG Brainwave Dataset [8].

With respect to the EEG Brainwave dataset, the original paper by Bird and his colleagues [8] used an EEG feature extraction approach based on the proposal of [22], based on this procedure they created four new datasets and tested several methods on them; the best result being a 97.89% of accuracy with a Random Forest (RF) model. In [23] several models from traditional Machine Learning (ML), like Support Vector Machines (SVMs) and Logistic Regression (LR), as well as Deep Learning (DL) architectures, such as Multi-Layer Perceptron (MLP), CNN, Recurrent Neural Networks (RNNs) and Long-Short Term Memory (LSTM) architectures were tested; the highest accuracy value obtained was 97.65% by the RNN. Moreover, in [24] statistical feature were extracted from the EEG signals and rearranged in the form of visual characteristics so that these data can be used as input for 2D and 3D CNNs; here, the 2D CNN obtained the best accuracy, 98.22%.

Similarly, in [25], an RNN, an LSTM and a Gated Recurrent Unit (GRU) with a number of parameters of around one million were tested, where each model obtained 95%, 96% and 96% of accuracy, respectively.

Finally, the work presented in [2] consists of an ensemble model of two levels. The first level consists of three models: RF, Light Gradient Boosting Machine (LightGBM) and Gradient Boosting Classifier (GBC); the second level consists of a RF model used as a meta classifier, where the predictions of the three models from the previous level are passed as inputs to this meta classifier. The classification accuracy obtained by the model was 99.55%.

[1] https://www.kaggle.com/datasets/birdy654/eeg-brainwave-dataset-feeling-emotions

III. Methodology

This section summarizes the main aspects of the dataset, as well as the pipeline proposed from signal preprocessing, to emotion classification based on such signals.

A. EEG Brainwave Dataset: Feeling Emotions

In order to collect the EEG signals, Bird and colleagues [8] used a MUSE EEG headband, consisting on four electrodes, corresponding to the TP9, AF7, AF8 and TP10 electrodes. In the experiments, two subjects were involved (one male and one female), to whom visual stimuli was applied in the form of movie clips, in order to elicit emotions on the participants while the MUSE band was collecting their brain activity. Six movies—three eliciting negative emotions and three eliciting positive ones— were selected for this task, each one of them having a corresponding clip of one minute of duration; since two subjects were involved, the dataset is formed by 12 minutes of EEG readings (six per subject). Additionally, six minutes of neutral brain activity were collected, i.e., brain activity without stimuli, that serves as a resting emotional state of the participants. The collected data was furtherly downsampled to a frequency of 150 Hz.

Since the movies elicited two different classes of emotions: negative and positive, and there is additionally neutral EEG recording, the dataset constitutes a three class classification problem.

B. Signal Preprocessing

Here, we decided to treat the features marked as *FFT*—750 for each signal data— in the dataset as EEG signals. For signal preprocessing two variants were proposed, consisting on the following steps:

1) Digital Filter or Empirical Mode Decomposition (EMD)
2) Normalization via z-score

From the previous list, it follows that our approach consists on a minimalist and streamlined preprocessing stage. Moreover, the only difference in our two variants is in step 1, which consists on the signal decomposition method used, since data is presented as *single channel*.

For the digital filter procedure, since data was resampled at 150 Hz, we decided to use a 4th degree Butterworth bandpass filter with several bands, so to decompose the signal into several channels, the bands utilized are:

- 0.1 - 10.1 Hz
- 9.9 - 20.1 Hz
- 19.9 - 30.1 Hz
- 29.9 - 40.1 Hz
- 39.9 - 50.1 Hz
- 49.9 - 60.1 Hz
- 59.9 - 74.9 Hz

having now, for each signal, a 6 frequency band representation.

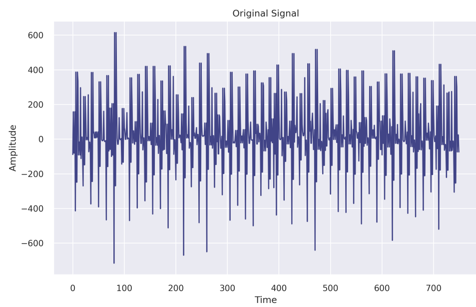

Fig. 1: Original Signal.

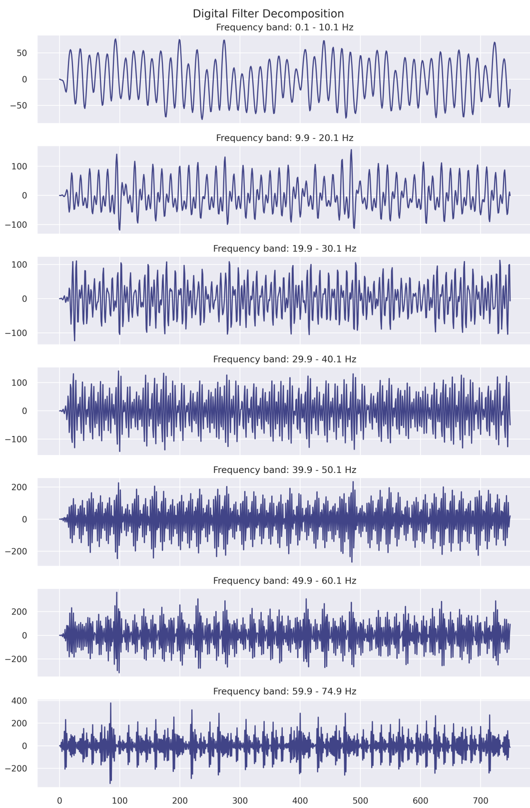

Fig. 2: Digital filter decomposition.

Similarly, for the EMD procedure, we decomposed the signal into five Intrinsic Mode Functions (IMFs). Among the advantages of using EMD there is the fact that it is a data-driven and unsupervised decompostiion method that satisfies the perfect reconstruction property, i.e., summing all IMFs and the residual reconstructs the original signal with no loss or distortion [9]. In this variant, we decomposed the signals into five IMFs, going from a single channel representation to a five IMFs one. Fig. 1 shows an example of a signal as provided in the dataset, whereas Figs. 2 and 3 show the digital filter and EMD decomposition results, respectively.

Lastly, a normalization procedure was carried on on each frequency band or IMF channel, according to which decomposition method was applied. The normalization method selected was Z-Score normalization, a procedure that centers the mean of a distribution at zero, and scales the value of the standard deviation to one, and is computed by means of the following equation:

Fig. 3: EMD decomposition.

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

where x is the original value, μ is the original mean of the distribution and σ is the corresponding standard deviation.

C. Validation Splits

In order to test our models, we used Hold-out as a validation method. At first instance, we split the data in a 80-20 fashion, 80% of the data was used as the training set, while the resting 20% was used as the testing set. Subsequentally, the training set was furtherly split in 90-10, where this 90% (72% of the original data) was used for training the network while the other 10% (8% of the original data) was used for validation during training. Table I shows the data distribution after the partition.

TABLE I: Dataset partition into training, validation and testing sets

Dataset partition	Percentage [%]	Number of signals
Training	72	1536
Validation	8	170
Testing	20	426

IV. Proposed Model

As previously stated, our proposal consists on a one-dimensional CNN, consisting of only 4 layers, three of them convolutional layers, while the last one is a linear layer with 3 neurons, that serves as the classifier.

The first convolutional layer consist of eight filters with a kernel size of one and no activation function; the purpose of this layer is to learn linear combinations of the channels created during preprocessing (frequency bands or IMFs), therefore, a kernel of size one is adequate since it will focus on only one point at a time and the network will learn the corresponding contribution of each channel.

Following the first layer, the model consists of two more convolutional layers, both with a kernel size of three, a max pooling procedure with window size of 2 and Hyperbolic Tangent (TanH) as activation function; these layer only differ on the number of filters: 16 for the second layer and eight for the third one.

Lastly, before the fourth layer a flatten procedure is applied and the resulting vector is passed to a linear layer consisting of three neurons; the classifier. Here a SoftMax activation function is applied in order to get probability values for each class.

A graphical representation of our model is shown in Fig. 4

V. Results

For training, we use the following setup: a learning rate of 3.8×10^{-4}, a batch size of 16 and 45 epochs, with a Checkpoint to save the model with the lowest validation loss value. Additionally, Adam was used as optimizer, with a weight decay value of 0.05.

Although the network architecture is essentially the same, since we proposed two preprocessing pipelines, the model varies in its number of parameters depending on the number of channels that the input has. Table II shows the number of parameters of each configuration. As one can see, both configurations end up with a network of—virtually—the same number of parameters; furthermore, this value is significantly small compared to what is commonly found in DL applications.

TABLE II: Number of parameters of the CNN depending on the preprocessing pipeline

Preprocessing Pipeline	Number of parameters
Frequency Bands	5347
IMFs	5331

After training, we used the testing partition to make predictions for each signal to be classified as positive, negative or neutral. Table III shows the accuracy values obtained on the testing dataset for both, Frequency Bands and IMFs configurations. Additionally, Figs. 5 and 6 show the confusion matrix for the Frequency Band and IMFs configurations, respectively; the colorscale values have been normalized with respect to the ground truth.

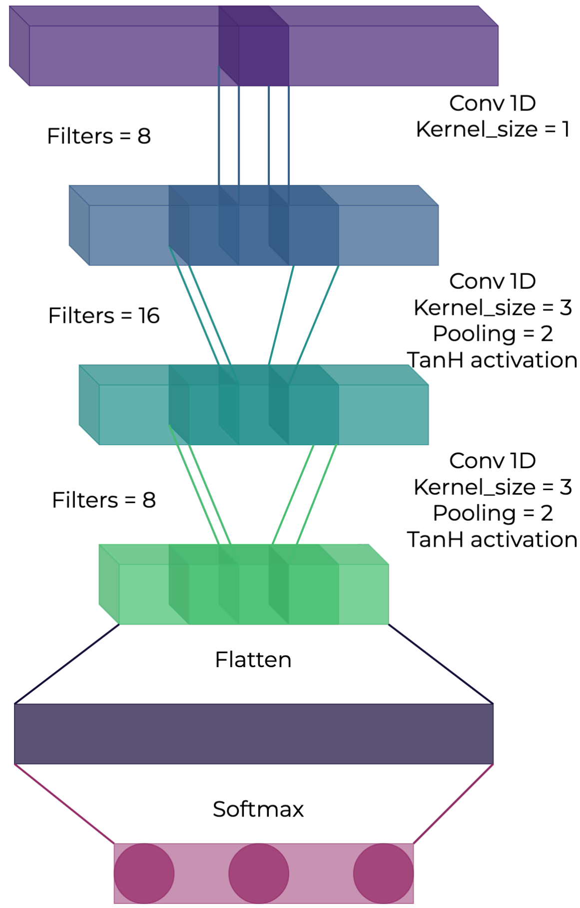

Fig. 4: One dimensional CNN proposal for EEG based emotion recognition.

TABLE III: Accuracy of the model for both preprocessing configurations.

Preprocessing Pipeline	Accuracy [%]
Frequency Bands	98.36
IMFs	95.31

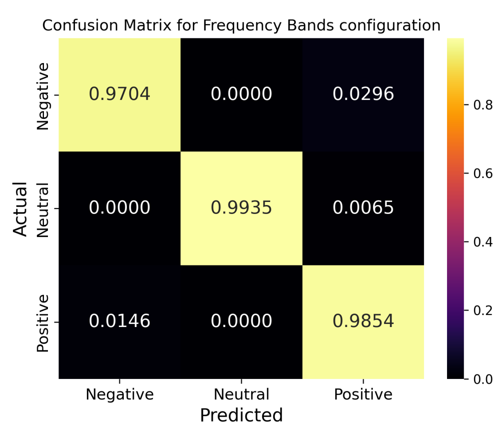

Fig. 5: Confusion matrix for Frequency Bands configuration.

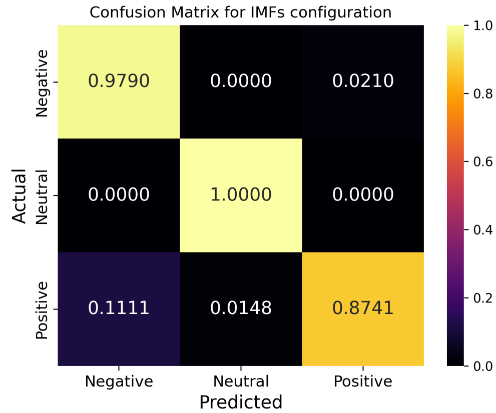

Fig. 6: Confusion matrix for IMFs configuration.

VI. Conclusions

We have proposed a small and simple 1D CNN model useful to classify emotions from EEG signals. The performance of this model is competitive with what is found on the literature, finding itself just behind an ensemble model of RF, LightGBM and GBC, followed by another RF model. Even though this model outperforms our proposal, the CNN of this paper distinguishes itself from the other proposal thanks to its simplicity, size and the lack of ensemble techniques. Moreover, given that our model consists of only four layers, the number of parameters is significantly small, around 5 thousand, in contrast with what was found in the literature review (approximately one million) and in other research areas where DL architectures are implemented. Moreover, the size, simplicity and number of parameters together, make for our CNN a model that is easy to train and retrain, an aspect of great importance in BCI systems, making it feasible to use in real-time applications.

Acknowledgements

The authors wish to thank the support of the Instituto Politécnico Nacional (COFAA, SIP-IPN, Grant SIP 20230140) and the Mexican Government (CONAHCyT, SNI).

References

[1] Z. Yan, J. Zhou, and W.-F. Wong, "Eeg classification with spiking neural network: Smaller, better, more energy efficient," *Smart Health*, vol. 24, p. 100261, 2022.

[2] S. Chatterjee and Y.-C. Byun, "Eeg-based emotion classification using stacking ensemble approach," *Sensors*, vol. 22, no. 21, 2022.

[3] A. Apicella, P. Arpaia, F. Isgrò, G. Mastrati, and N. Moccaldi, "A survey on eeg-based solutions for emotion recognition with a low number of channels," *IEEE Access*, vol. 10, pp. 117411–117428, 2022.

[4] E. P. Torres, E. A. Torres, M. Hernández-Álvarez, and S. G. Yoo, "Eeg-based bci emotion recognition: A survey," *Sensors*, vol. 20, no. 18, 2020.

[5] S. Jerritta, M. Murugappan, R. Nagarajan, and K. Wan, "Physiological signals based human emotion recognition: a review," in *2011 IEEE 7th International Colloquium on Signal Processing and its Applications*, pp. 410–415, 2011.

[6] E. H. Houssein, A. Hammad, and A. A. Ali, "Human emotion recognition from eeg-based brain–computer interface using machine learning: a comprehensive review," *Neural Computing and Applications*, vol. 34, no. 15, pp. 12527–12557, 2022.

[7] K. Oatley, D. Keltner, and J. M. Jenkins, *Understanding emotions*. Blackwell publishing, 2006.

[8] J. J. Bird, A. Ekart, C. D. Buckingham, and D. R. Faria, "Mental emotional sentiment classification with an eeg-based brain-machine interface," in *Proceedings of the International Conference on Digital Image and Signal Processing (DISP'19)*, 2019.

[9] A. Zeiler, R. Faltermeier, I. R. Keck, A. M. Tomé, C. G. Puntonet, and E. W. Lang, "Empirical mode decomposition-an introduction," in *The 2010 international joint conference on neural networks (IJCNN)*, pp. 1–8, IEEE, 2010.

[10] D. W. Prabowo, H. A. Nugroho, N. A. Setiawan, and J. Debayle, "A systematic literature review of emotion recognition using eeg signals," *Cognitive Systems Research*, vol. 82, p. 101152, 2023.

[11] A.-L. Cîrneanu, D. Popescu, and D. Iordache, "New trends in emotion recognition using image analysis by neural networks, a systematic review," *Sensors*, vol. 23, no. 16, 2023.

[12] F. Z. Canal, T. R. Müller, J. C. Matias, G. G. Scotton, A. R. de Sa Junior, E. Pozzebon, and A. C. Sobieranski, "A survey on facial emotion recognition techniques: A state-of-the-art literature review," *Information Sciences*, vol. 582, pp. 593–617, 2022.

[13] R. A. Khalil, E. Jones, M. I. Babar, T. Jan, M. H. Zafar, and T. Alhussain, "Speech emotion recognition using deep learning techniques: A review," *IEEE Access*, vol. 7, pp. 117327–117345, 2019.

[14] A. Goshvarpour and A. Abbasi, "An emotion recognition approach based on wavelet transform and second-order difference plot of ecg," *Journal of AI and Data Mining*, vol. 5, no. 2, pp. 211–221, 2017.

[15] T. Fan, S. Qiu, Z. Wang, H. Zhao, J. Jiang, Y. Wang, J. Xu, T. Sun, and N. Jiang, "A new deep convolutional neural network incorporating attentional mechanisms for ecg emotion recognition," *Computers in Biology and Medicine*, vol. 159, p. 106938, 2023.

[16] J. A. Domínguez-Jiménez, K. C. Campo-Landines, J. C. Martínez-Santos, E. J. Delahoz, and S. H. Contreras-Ortiz, "A machine learning model for emotion recognition from physiological signals," *Biomedical signal processing and control*, vol. 55, p. 101646, 2020.

[17] V. K. Patil, V. R. Pawar, S. Randive, R. R. Bankar, D. Yende, and A. K. Patil, "From face detection to emotion recognition on the framework of raspberry pi and galvanic skin response sensor for visual and physiological biosignals," *Journal of Electrical Systems and Information Technology*, vol. 10, no. 1, pp. 1–27, 2023.

[18] T. Paus, P. Sipila, and A. Strafella, "Synchronization of neuronal activity in the human primary motor cortex by transcranial magnetic stimulation: an eeg study," *Journal of Neurophysiology*, vol. 86, no. 4, pp. 1983–1990, 2001.

[19] A. Babayan, M. Erbey, D. Kumral, J. D. Reinelt, A. M. Reiter, J. Röbbig, H. L. Schaare, M. Uhlig, A. Anwander, P.-L. Bazin, *et al.*, "A mind-brain-body dataset of mri, eeg, cognition, emotion, and peripheral physiology in young and old adults," *Scientific data*, vol. 6, no. 1, pp. 1–21, 2019.

[20] S. Koelstra, C. Muhl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, and I. Patras, "Deap: A database for emotion analysis; using physiological signals," *IEEE transactions on affective computing*, vol. 3, no. 1, pp. 18–31, 2011.

[21] N. S. Suhaimi, J. Mountstephens, and J. Teo, "A dataset for emotion recognition using virtual reality and eeg (der-vreeg): Emotional state classification using low-cost wearable vr-eeg headsets," *Big Data and Cognitive Computing*, vol. 6, no. 1, 2022.

[22] J. J. Bird, L. J. Manso, E. P. Ribeiro, A. Ekart, and D. R. Faria, "A study on mental state classification using eeg-based brain-machine interface," in *2018 international conference on intelligent systems (IS)*, pp. 795–800, IEEE, 2018.

[23] A. Sarkar, A. Singh, and R. Chakraborty, "A deep learning-based comparative study to track mental depression from eeg data," *Neuroscience Informatics*, vol. 2, no. 4, p. 100039, 2022.

[24] J. J. Bird, D. R. Faria, L. J. Manso, P. P. Ayrosa, and A. Ekart, "A study on cnn image classification of eeg signals represented in 2d and 3d," *Journal of Neural Engineering*, vol. 18, no. 2, p. 026005, 2021.

[25] M. K. Chowdary, J. Anitha, and D. J. Hemanth, "Emotion recognition from eeg signals using recurrent neural networks," *Electronics*, vol. 11, no. 15, 2022.