

EMOTION CLASSIFICATION USING 1D-CNN AND RNN BASED ON DEAP DATASET

Farhad Zamani and Retno Wulansari

Telkom Corporate University Center, Telkom Indonesia, Bandung, Indonesia

ABSTRACT

Recently, emotion recognition began to be implemented in the industry and human resource field. In the time we can perceive the emotional state of the employee, the employer could gain benefits from it as they could improve the quality of decision makings regarding their employee. Hence, this subject would become an embryo for emotion recognition tasks in the human resource field. In a fact, emotion recognition has become an important topic of research, especially one based on physiological signals, such as EEG. One of the reasons is due to the availability of EEG datasets that can be widely used by researchers. Moreover, the development of many machine learning methods has been significantly contributed to this research topic over time. Here, we investigated the classification method for emotion and propose two models to address this task, which are a hybrid of two deep learning architectures: One-Dimensional Convolutional Neural Network (CNN-1D) and Recurrent Neural Network (RNN). We implement Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM) in the RNN architecture, that specifically designed to address the vanishing gradient problem which usually becomes an issue in the time-series dataset. We use this model to classify four emotional regions from the valence-arousal plane: High Valence High Arousal (HVHA), High Valence Low Arousal (HVLA), Low Valence High Arousal (LVHA), and Low Valence Low Arousal (LVLA). This experiment was implemented on the well-known DEAP dataset. Experimental results show that proposed methods achieve a training accuracy of 96.3% and 97.8% in the 1DCNN-GRU model and 1DCNN-LSTM model, respectively. Therefore, both models are quite robust to perform this emotion classification task.

KEYWORDS

Emotion Recognition, 1D Convolutional Neural Network, LSTM, GRU, DEAP.

1. INTRODUCTION

Many industries and human resource fields began to implement emotional recognition of the employee in their organization. When they can assess the emotional state of the employee, the human resource could gain advantages from it as they could improve the quality of decision makings regarding their employee. This subject could become a base for emotion recognition tasks in human resource cases. For this reason, this task will become important widely used shortly.

Emotions play a crucial role in human beings as these are associated with neuro physiological aspects that also correspond to a coordinated set of responses, which may include verbal, behavioral, physiological, and neural mechanisms. In another perspective, the emotion can be mapped into the Valence, Arousal, and Dominance (VAD) dimensions. Valence represents a dimension that corresponds to the level of pleasantness that goes from very positive (pleasure) to very negative (displeasure). On the other hand, arousal is the intensity level of emotion that an

event creates, ranging from excited (positive) to calm (negative). Lastly, dominance is the degree of control exerted by a stimulus[1]. The most common model used is the Circumplex Model of Affect, which only describes emotion in the valence and arousal dimension [2].

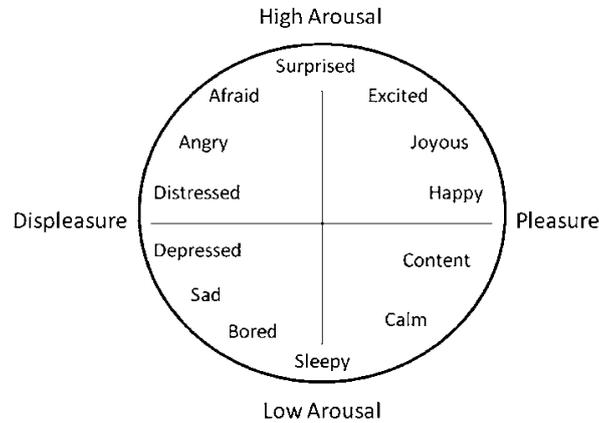


Figure 1. Circumplex Model of Affect Graphical Representation.

In recent studies, human emotion can be recognized based on their peripheral physiological signal, such as electroencephalogram. The electroencephalogram (EEG) is defined as the electrical activity of an alternating type recorded from the scalp surface after being picked up by metal electrodes and conductive media [3]. A typical amplitude of EEG ranged from 0.5 to 100uV. In general, this EEG signal is divided into a specific range of frequencies, namely delta, theta, alpha, beta, and gamma. Each range is associated with a particular condition or state of the brain.

Recently, numerous researchers have been using EEG signals to aid them to classify emotional states. Moreover, various EEG datasets that can be used in emotion recognition studies have been published by many researchers. DEAP dataset is commonly used for this study. Several signal processing methods, as well as machine learning algorithms, are applied in many studies.

In this study, we suggest simpler signal processing methods to prepare the DEAP dataset before we feed it into our models. We also propose a combination of two neural network architectures as our model, which is CNN and RNN. One-Dimensional CNN is chosen as our CNN architecture, and for the RNN, we will evaluate the performance of GRU and LSTM which we will put together with 1D-CNN. Subsequently, the accuracy and loss value in the training, validation, and testing process of both models, 1D-CNN + GRU and 1D-CNN + LSTM, will be compared.

The layout of the paper is as follows. In section 2, we overview several approaches related to our research topic. The materials and methods of the experiment are described in detail in section 3. The results of the experiment and discussions are covered in Section 4. The conclusion, limitation and future attempt of this work follows in Section 5.

2. RELATED WORKS

Xiang Li et al. (2016) proposed a hybrid deep learning model, C-RNN, which integrates CNN and RNN, for emotion recognition [4]. The data is pre-processed using continuous wavelet transform and frame construction before being trained in the model. This experiment performance is 74.12% and 72.06% for arousal and valence dimensions, respectively. Alhagry et al. (2017) suggested LSTM as its deep learning method for emotion recognition tasks based on

the DEAP dataset [5]. The average accuracy given from this method is 85.65%, 85.45%, and 87.99% with arousal, valence, and liking classes, respectively.

Lin et al. (2017) presented an approach to perform emotional states classification by end-to-end learning of deep CNN based on the DEAP dataset [6]. The datasets were transformed into six gray images which contain time and frequency domain information, then the feature extracted before trained using the AlexNet model. The model accuracy of this study is 87.30% and 85.50% for arousal and valence, respectively. Li et al. (2017) applied a combined CNN and LSTM RNN (CLRNN) in their emotion recognition task and carried it out with the DEAP dataset [7]. The dataset is transformed into a Multidimensional Feature Image sequence beforehand. The average emotion classification accuracy of each subject with the hybrid neural networks proposed in the study is 75.21%.

Acharya et al. (2020) compared the performance of LSTM and CNN models to carry out emotion classification tasks [8]. The dataset used in this study is the DEAP dataset and the feature extraction method implemented is Fast Fourier Transform. The best result obtained from the experiment was from the LSTM model with 88.6% accuracy on the liking emotion and 87.2% accuracy when using the CNN model. Zhang et al. (2020) investigated the application of several deep learning models to the research field of EEG-based emotion recognition: DNN, CNN, LSTM, and a hybrid model of CNN and LSTM [9]. This research also used the DEAP dataset, and several features were extracted from it, such as mean, standard deviation, max value, min value, skewness, and kurtosis. The CNN and CNN-LSTM models displayed the best performance to perform this task with 90.12% and 94.17% accuracy, respectively.

Anubhav et al. (2020) studied the EEG signals in scope for developing a headgear model for real-time monitoring of emotions [10]. Band power and frequency domain features were extracted from the DEAP dataset and compared the classification accuracies for valence and arousal dimensions. In this study, the LSTM model achieves the best classification accuracy of 94.69% and 93.13% for valence and arousal, respectively. Dar et al. (2020) proposed 2D CNN architecture to process EEG signals for emotion recognition tasks [11]. The datasets used in this experiment are DREAMER and AMIGOS, in which both data are converted into a 2D feature matrix (PNG format) before the data is fed into CNN. Aside from EEG, other peripheral physiological signals such as ECG and GSR are also utilized for the multi-modal emotion recognition process. Accuracy acquired using EEG modality only is 76.65%, and the overall highest accuracy of 99.0% for the AMIGOS dataset and 90.8% for the DREAMER dataset is achieved with multi-modal fusion.

3. MATERIALS AND METHODS

Firstly, this section will elaborate on the details of datasets used in this study as well as how the data will be classified. Then, the proposed methodology for the experiment will be discussed in detail.

3.1. Diagram Block

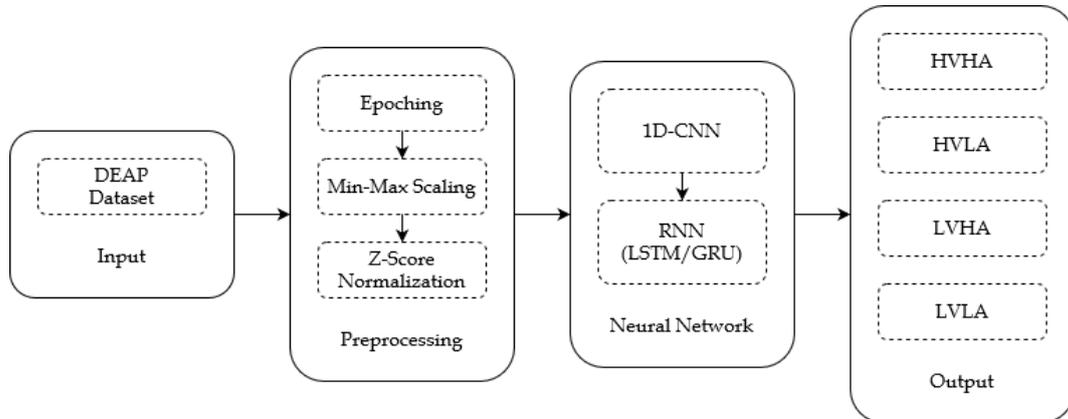


Figure 2. Diagram Block of Our Proposed Method

As shown in Figure 1, the framework for this study is made up of two main processes, the data pre-processing part, and the neural network part. The details of each process will be explained in the following section.

3.2. Datasets and Label Classification

The dataset used in this study is DEAP which was made publicly available for researchers to test their affective state estimation method [12]. This dataset contains EEG and other peripheral physiological signals, such as EOG and EMG. The data was acquired from 32 participants as each watched 40 one-minute-long music videos. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance, and familiarity.

The DEAP dataset contains 32 groups of EEG data in total, corresponding to the experimental data of 32 subjects (s01–s32). For each subject, EEG signals are acquired from 32 channels and 8 additional channels are from other peripheral signals, which makes 40 channels in total. The data of each subject contains two arrays: data and labels. The structure for each file in this dataset is shown in Table 1.

Table 1. Dataset Structure.

Array Name	Array Shape	Array Contents
data	40 x 40 x 8064	video/trial x channel x data
labels	40 x 4	video/trial x label (valence, arousal, dominance, liking)

In this study, we used a pre-processed version of the DEAP dataset. This version of the signal has been down sampled to 128 Hz and filtered by a band-pass filter with a frequency bandwidth of 4-45 Hz. EOG artifacts were also removed, and the data were averaged to the common reference.

Each video/trial was rated on a scale from 1 to 9. The lower the rating, the weaker the emotion, and the higher the rating, the stronger the emotion. The dataset is divided into 4 classes based on the threshold of valence and arousal value, then they will be classified as the following: High-Valence High-Arousal (HVHA), High-Valence Low-Arousal (HVLA), Low-Valence High-

Arousal (LVHA), and Low-Valence Low-Arousal (LVLA). The threshold used in this classification is 5. If the value is greater than 5, it will be classified as high, otherwise is low.

Table 2. Label Classification.

Label	HVHA	HVLA	LVHA	LVLA
Valence	>5	>5	<=5	<=5
Arousal	>5	<=5	>5	<=5

To reduce the computational cost for this study, not all channels are selected. Only 14 channels which we further process for emotion classification task [13]. This channel selection is based on the significance of the brain region which represent emotional states. Those channels are listed below in Table 3. This task was done after we classify the label into four categories: HVHA, HVLA, LVHA, LVLA. As a result, it will increase the training accuracy as well as decrease the validation loss.

Table 3. Channel Selection.

Channel	Index	Channel	Index
AF3	1	AF4	17
F3	2	F4	19
F7	3	F8	20
FC5	4	FC6	21
T7	7	T8	25
P7	11	P8	29
O1	13	O2	31

3.3. Data Pre-Processing

To prepare the DEAP dataset before we feed it into our proposed model, it must be processed first to get a better training result. In the following sections, our data pre-processing methods will be described in detail.

3.3.1. Epoching

Epoching is a procedure in which specific time windows are extracted from the continuous EEG signal. To begin with, this task is performed by extracting 2 seconds wavelengths of each signal with 2 second time step. Therefore, it will give us 31 signals for each channel, with 256 data points for each epoch that represents 2 seconds of the signal.

Figure 3 is a sample of the signal after the epoching process.

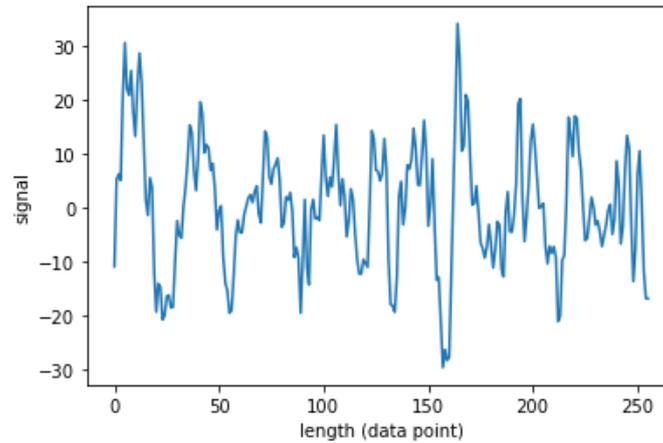


Figure 3. Sample Signal after Epoching

3.3.2. Standardization

Every signal in the dataset that is used in this study will be standardized first through two methods: z-score normalization and followed by min-max scaling. These two processes are carried out to prevent overfitting as well as increase the accuracy of the model.

3.3.2.1. Min-Max Scaling

Min-Max Scaling (often also simply called normalization) is a standardization method to scale data to a fixed range: 0 to 1. The min-max scaling is calculated using the following equation.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (1)$$

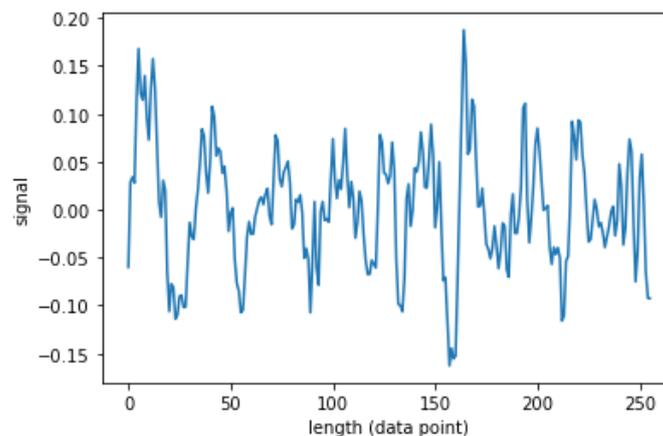


Figure 4. Signal Sample after Min-Max Scaling Normalization

3.3.2.2. Z-Score Normalization

The result of standardization (also called z-score normalization) is that the feature will be rescaled so that they will have the properties of a standard normal distribution with:

$$\mu = 0 \text{ and } \sigma = 1 \quad (2)$$

where μ is the mean (average) and σ is the standard deviation from the mean; standard scores (also called z scores) of the samples are calculated as follows:

$$z = \frac{x - \mu}{\sigma} \quad (3)$$

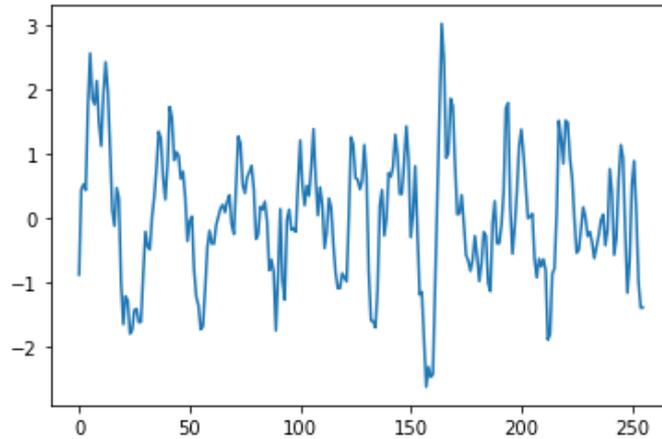


Figure 5. Sample Signal after Z-Score Normalization

3.4. Neural Network Architecture

The neural network architecture that will be used in the proposed method is One-Dimensional Convolutional Neural Network and Recurrent Neural Network. The overview of each architecture will be portrayed in brief in the following section.

3.4.1. 1D-CNN

In this study, all signals used are one-dimensional. One-Dimensional Convolutional Neural Network (1D-CNN) has become a very powerful tool to be used upon time-series data like EEG signals. This model can extract and distinguish several features of EEG data which represent significant features for emotion classification tasks.

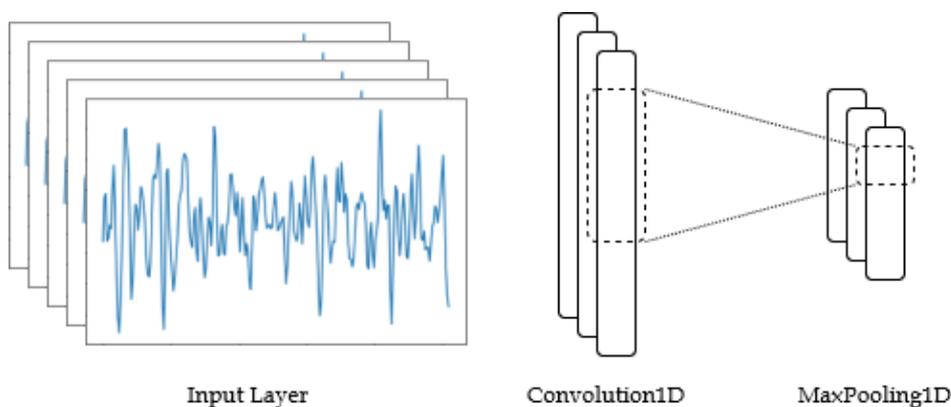


Figure 6. 1D-CNN Structure.

3.4.2. LSTM

Simple Recurrent Neural Network suffers from short-term memory and is often called vanishing gradient problem. LSTM is a specialized version of RNN that was designed to overcome this shortcoming. LSTM was first introduced by Hochreiter (1997)[14], and this solution works well on a large variety of problems. The core concept of LSTM's are the cell state and its various gates: forget gate, input gate, and output gate.

Cell state act as a memory of the network which carries relevant information throughout the processing of the sequence from earlier steps to later steps. Information gets added via gates composed out of the sigmoid layer and point wise multiplication operation. LSTM model can learn which information is relevant to keep or forget during the training.

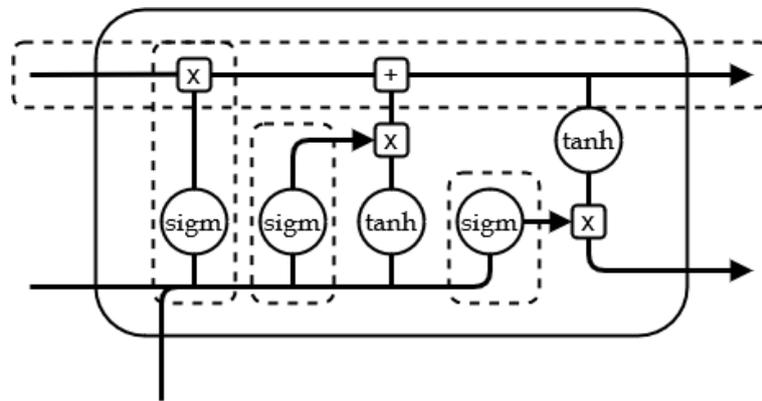


Figure 7. LSTM Cell Structure.

3.4.3. GRU

GRU is the newer generation of Recurrent Neural Network and is similar to an LSTM[15]. GRU's got rid of the cell state and used the hidden state to transfer information within the network. It only has two gates. The first is the reset gate, which is used to decide how much past information to forget. The second is, instead of use forget gate and input gate, GRU uses a single update gate. It also merges the cell state and hidden state. As a result, this model is much simpler than LSTM model.

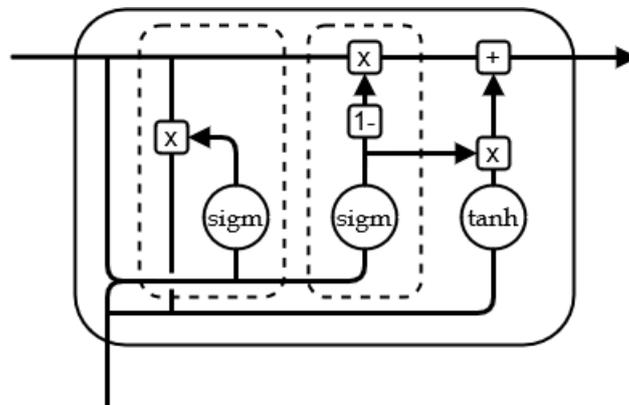


Figure 8. GRU Cell Structure.

3.5. Proposed Model

3.5.1. Data Shape

For this study, we use all of the preprocessed DEAP dataset from 32 subjects. The preprocessed dataset went through further preprocessing procedures and label classification. Then, we randomly split the dataset into train, validation, and test data with 60:20:20 ratio. The data shape for our experiment is shown in Table 4.

Table 4. Data Shape.

Set	Data	Target	Percentage
Train Set	(333312, 256, 1)	(333312, 4)	60%
Validation Set	(111104, 256, 1)	(111104, 4)	20%
Test Set	(111104, 256, 1)	(111104, 4)	20%

3.5.2. Hyperparameter

To get a good result, hyperparameter must be tuned adequately. The batch size finalized for our architectures is 256 and the optimizer used is Adam. The loss function implemented for updating the weights during backpropagation is categorical cross-entropy since we use a multi-labeled dataset. Furthermore, we use the callbacks method with a patience value of 10 which monitors validation loss value. As a result, the epoch for the two models will be varied depends on the validation loss value.

Table 5. Hyperparameter.

Parameters	Value
Batch Size	256
Optimizer	Adam
Loss Function	Categorical cross entropy

3.5.3. 1D-CNN – GRU

This model is a composite of two kinds of deep learning structures: 1D-CNN and GRU. The structure of the model is presented in Table 6. The input size of the networks is 256 x 1, that correspond with two seconds of the epoched signal which contains 256 data point. We set the number of convolutional filters as 128 with ReLU as its activation layer. Following the first convolutional layer is a max-pooling layer with a pool size of 2. We set a dropout of 0.2 after the max-pooling layer to reduce the probability of overfitting. The second convolutional layer is the same as the first.

The convolutional layers are followed by the GRU layer with 256 units and 32 units, subsequently. After every GRU layer, we also set a dropout layer of 0.2. Before connecting to the dense unit, a flattening operation is implemented to transform features into a one-dimensional feature vector. The dense layer units are set to 32 with the ReLU activation function. The last layer is 4 units dense layer that represents four labels of classification. The activation function of this dense layer is softmax. The number of trainable parameters of this model is 378.820. The overview of this model configuration is illustrated in Table 6.

Table 6. Configuration of 1D-CNN – GRU.

Layer	Parameter
InputLayer	input_shape = (256,1)
Conv1D + ReLU	units = 128, kernel = 3
MaxPooling1D	pool_size = 2
Dropout	rate = 0.2
Conv1D + ReLU	units = 128, kernel = 3
MaxPooling1D	pool_size = 2
Dropout	rate = 0.2
GRU	units = 256, return_sequences = True
Dropout	rate = 0.2
GRU	units = 32
Dropout	rate = 0.2
Flatten	none
Dense + ReLU	units = 128
Dense + Softmax	units = 4

3.5.4. 1D-CNN – LSTM

The configuration of this model is identical to 1D-CNN – GRU. The only difference is, we substitute the GRU layer with the LSTM layer. The number of trainable parameters of this model is 485.764, which is larger than that of 1DCNN-GRU model. The overview of this model configuration is summarized as shown in Table 7.

Table 7. Configuration of 1D-CNN – LSTM.

Layer	Parameter
InputLayer	input_shape = (256,1)
Conv1D + ReLU	units = 128, kernel = 3
MaxPooling1D	pool_size = 2
Dropout	rate = 0.2
Conv1D + ReLU	units = 128, kernel = 3
MaxPooling1D	pool_size = 2
Dropout	rate = 0.2
LSTM	units = 256, return_sequences = True
Dropout	rate = 0.2
LSTM	units = 32
Dropout	rate = 0.2
Flatten	none
Dense + ReLU	units = 128
Dense + Softmax	units = 4

4. EXPERIMENTAL RESULTS

In this section, we present the results of our experiments. Firstly, we will elaborate on the performance of both models used in this study based on the accuracy and loss value of the model. Following this, we present the comparison of the results between GRU and LSTM that is used in this experiment. Later, the overall results and analysis of the experiment will be explained in detail. Lastly, we discussed limitations of the proposed method.

All the experiments are implemented by using the Tensorflow framework version 2.5.0 and Python version 3.7.10. The workstation used is MSI GF65 Thin 10SDR, which included Intel i7-10750H (12 CPUs @ 2.6 GHz), NVIDIA GeForce GTX 1660 Ti for GPU acceleration, 500 GB SSD, and 8 GB RAM.

The first experiment was designed to evaluate the performance of the 1D-CNN – GRU model for the emotion classification task. The result is shown in Table 8. We obtained training accuracy of 96.3%, along with validation and test accuracy above 99%. Additionally, we also acquire training, validation, and test loss of 10.7%, 0.7%, and 0.6%, respectively. Compared with previous relevant works, this model shows a very good performance.

As illustrated in Figure 9, train and validation graphs show the desired value of accuracy and loss change per epoch. There is no overfitting occurred and the training process stopped at 165 epochs. The average time consumed for training per epoch is 50 seconds. Furthermore, we also made a confusion matrix to analyze the difference between the predicted and the actual value (Figure 10). Moreover, this technique is good to be applied to measure the multilabel classification performance. We got F1-Score and recall value of 1 which infer that the quality of this model is good.

Table 8. 1D-CNN – GRU experimental results.

GRU	Training	Validation	Test
Accuracy	96.3%	99.9%	99.9%
Loss	10.7%	0.7%	0.6%

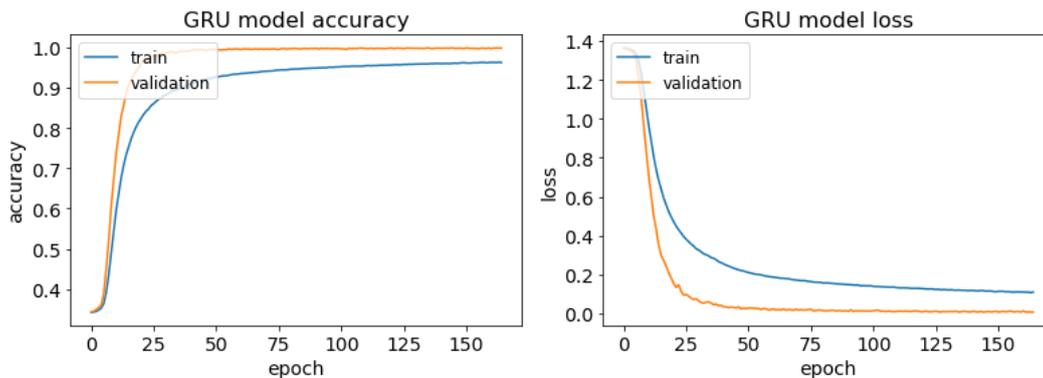


Figure 9. Train and Validation Graph of 1D-CNN – GRU model.



Figure 10. Confusion Matrix for 1D-CNN – GRU model

The second experiment was designed to evaluate the performance of the 1D-CNN – LSTM model to carry out the emotion classification task. Table 9 illustrates the result of this experiment. We obtained a training accuracy of 97.8%. The validation and test accuracy values of this model are 99.9% and 99.9% respectively which is above the first model. Aside from that, we acquire a training loss of 6.7%, with validation and test loss of 0.1% and 0.1%.

Table 9. 1D-CNN – LSTM experimental results.

LSTM	Training	Validation	Test
Accuracy	97.8%	99.9%	99.9%
Loss	6.7%	0.1%	0.1%

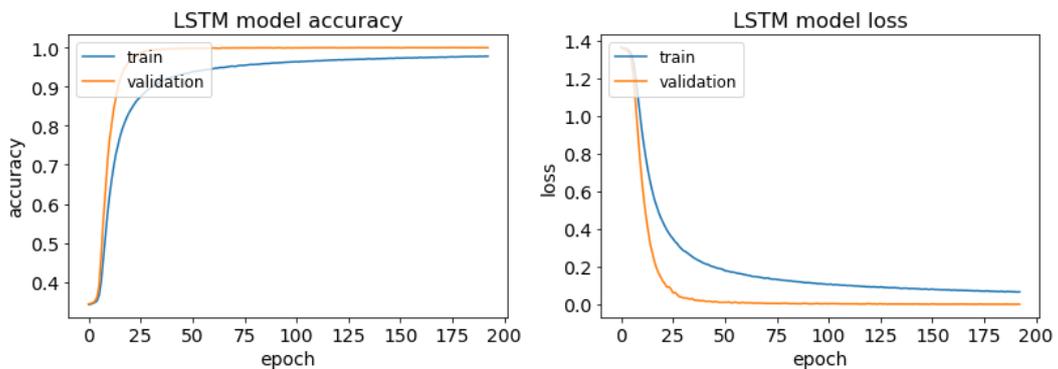


Figure 11. Train and Validation Graph of 1D-CNN – LSTM model.

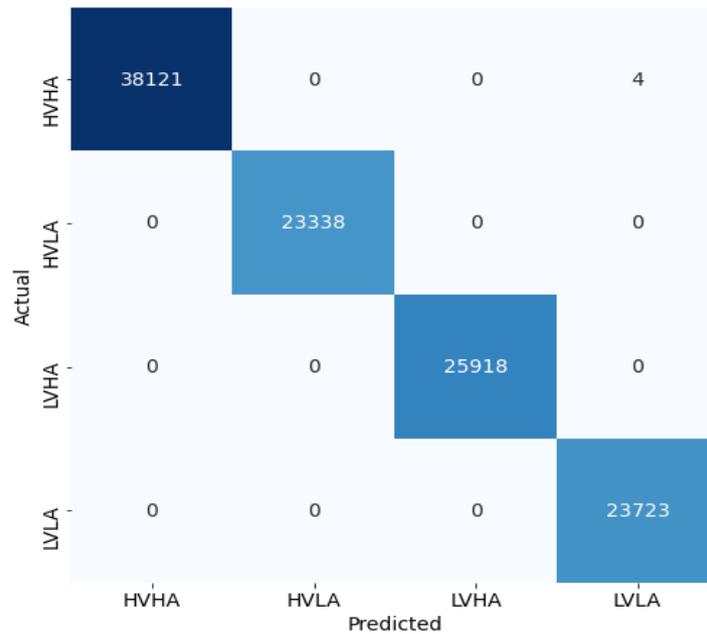


Figure 12. Confusion Matrix for 1D-CNN – LSTM model.

In Figure 11, we can see that train and validation graphs also show the desired value change per epoch. In this model, we can avoid overfitting as well. The training process stopped at 193 epochs with an average training time per epoch is 69 seconds, more than the 1D-CNN – GRU model. We also analyze the difference between the predicted value and actual value using a confusion matrix (Figure 12). We got an F1-Score and a recall value of 1. It concludes that the overall quality of this model is better than the previous model, albeit the training time is longer.

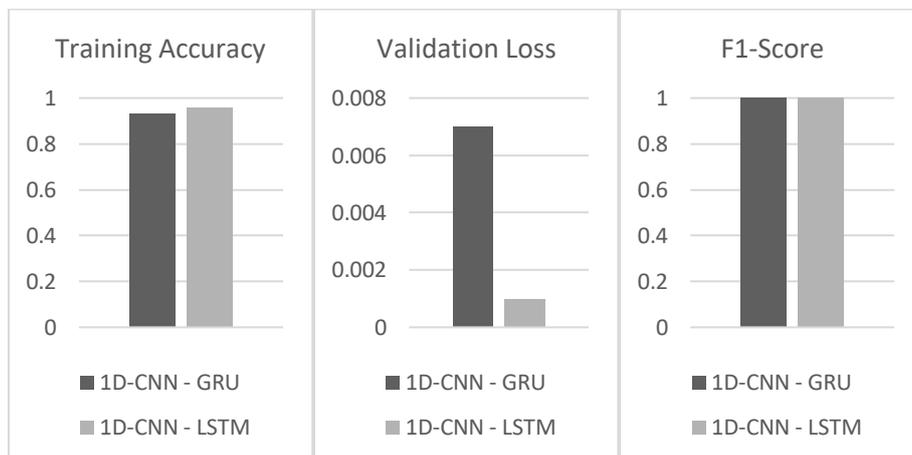


Figure 13. Performance chart comparison based on training accuracy, validation loss, and f1-score value.

If we compare the performance of two models, 1D-CNN – GRU and 1D-CNN – LSTM, it is clear the latter is better than the former. More specifically, if we take account of the training accuracy and validation loss, the 1D-CNN – LSTM model performance surpasses the other one. The F1-score of the 1D-CNN – LSTM model is the same with 1D-CNN – GRU, with a score of 1. The comparison of both models is shown in Figure 13.

Through this experiments, it was proven that One-Dimensional Convolutional Neural Network can be regarded as a feature extractor that can recognize a hidden patterns or features of the EEG signal. Even though we do not perform Fast Fourier Transform (FFT), Continuous Wavelet Transform (CWT), or other feature extraction methods to acquire a particular features of the signal, we can conclude that this methods still could execute this task very well. Furthermore, the proposed methods, from the preprocessing to the classification task, do not require large memory to be executed since the methods is relatively simple. Besides that, since we only select 14 channels from DEAP dataset, the overall training speed was quite fast, and the model can achieve convergence before 100 epochs.

The overall performance of our proposed method outperformed the previous approaches that have been explained before in related works section. Each of these methods implemented various preprocessing and deep learning technique to perform the EEG feature extraction and emotion classification tasks. All the methods we compared use the DEAP dataset, and the classification accuracy comparison for each method is illustrated in Table 10.

Table 10. Methods and classification accuracy comparison with previous related works which used DEAP dataset.

Study	Methodology	Classification Accuracy
[5]	LSTM	85.65% (A), 85.45% (V)
[6]	CWT, CNN	87.30% (A), 85.30% (V)
[7]	CNN-LSTM	75.21%
[8]	FFT, CNN	84.7% (A), 85.01% (V)
[8]	FFT, LSTM	81.91% (A), 84.39% (V)
[9]	CNN-LSTM	94.17%
[10]	Welch, LSTM	93.13% (A), 94.69% (V)
Proposed study	1DCNN-GRU	96.3%
Proposed study	1DCNN-LSTM	97.8%

5. CONCLUSION, LIMITATION, AND FUTURE WORKS

In this paper, we present simpler signal processing methods as well as 1D-CNN – RNN hybrid model to perform emotion classification tasks on the DEAP dataset. The results of both models show that proposed models can achieve very high accuracy in classifying EEG signals based on four emotion states: HVHA, HVLA, LVHA, and LVLA. The 1DCNN-LSTM architecture with 97.8% classification accuracy is slightly better than the 1DCNN-GRU architecture that has 96.3% classification accuracy. The 1D-CNN architecture has shown its quality to extract EEG signal features and the RNN architectures, GRU and LSTM, are able to perform well in learning and processing the sequence data of time-series signals.

Although the proposed model has achieved significant results in emotion classification task, we still need to evaluate it further in other datasets and improve the robustness of our model. One of the advantages of our proposed model is its simplicity. Because of that, this model do not require data with numerous signal preprocessing before fed into the model. However, there are still limitation that need to be addressed in the next work. More specifically, this model only trained using DEAP dataset and still not tested in other subjects. Therefore, in our future work, we intend to re-train and test this model using of another EEG datasets that used in emotion recognition study, such as DREAMER and AMIGOS. Moreover, we plan to implement a k-fold cross validation in the next work in order to give more accurate estimation of performance of our proposed method. Lastly, we aim to use this model to assess emotional states of a subject and

integrate it in a real-time application. The results of the experiment suggest that it is a good start to begin implementing this work in the human resources field, especially to gain insights from employee states of emotion.

ACKNOWLEDGMENT

This research work is performed under the Bio signal and AI Research Team of Telkom Corporate University Center, Telkom Indonesia, Bandung, Indonesia.

REFERENCES

- [1] S. M. Alarcão and M. J. Fonseca, "Emotions Recognition Using EEG Signals: A Survey," in *IEEE Transactions on Affective Computing*, vol. 10, no. 3, pp. 374-393, 1 July-Sept. 2019, doi: 10.1109/TAFFC.2017.2714671.
- [2] Tseng, Angela & Bansal, Ravi & Liu, Jun & Gerber, Andrew & Goh, Suzanne & Posner, Jonathan & Colibazzi, Tiziano & Algermissen, Molly & Chiang, I-Chin & Russell, James & Peterson, Bradley. (2013). Using the Circumplex Model of Affect to Study Valence and Arousal Ratings of Emotional Faces by Children and Adults with Autism Spectrum Disorders. *Journal of autism and developmental disorders*. 44. 10.1007/s10803-013-1993-6.
- [3] M. Teplan, "Fundamentals of EEG Measurement," *IEEE Measurement Science Review*, Vol. 2, 2002, pp. 1-11.
- [4] X. Li, D. Song, P. Zhang, G. Yu, Y. Hou and B. Hu, "Emotion recognition from multi-channel EEG data through Convolutional Recurrent Neural Network," 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2016, pp. 352-359, doi: 10.1109/BIBM.2016.7822545.
- [5] Salma Alhagry, Aly Aly Fahmy and Reda A. El-Khoribi, "Emotion Recognition based on EEG using LSTM Recurrent Neural Network" *International Journal of Advanced Computer Science and Applications*(ijacsa), 8(10), 2017.
- [6] Lin W., Li C., Sun S. (2017) Deep Convolutional Neural Network for Emotion Recognition Using EEG and Peripheral Physiological Signal. In: Zhao Y., Kong X., Taubman D. (eds) *Image and Graphics. ICIG 2017. Lecture Notes in Computer Science*, vol 10667. Springer, Cham.
- [7] Li Y, Huang J, Zhou H, Zhong N. Human Emotion Recognition with Electroencephalographic Multidimensional Features by Hybrid Deep Neural Networks. *Applied Sciences*. 2017; 7(10):1060.
- [8] Acharya D. et al. (2021) Multi-class Emotion Classification Using EEG Signals. In: Garg D., Wong K., Sarangapani J., Gupta S.K. (eds) *Advanced Computing. IACC 2020. Communications in Computer and Information Science*, vol 1367. Springer, Singapore.
- [9] Zhang Y, Chen J, Tan JH, Chen Y, Chen Y, Li D, Yang L, Su J, Huang X and Che W (2020) An Investigation of Deep Learning Models for EEG-Based Emotion Recognition. *Front. Neurosci*. 14:622759. doi: 10.3389/fnins.2020.622759
- [10] Anubhav, D. Nath, M. Singh, D. Sethia, D. Kalra and S. Indu, "An Efficient Approach to EEG-Based Emotion Recognition using LSTM Network," 2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), 2020, pp. 88-92, doi: 10.1109/CSPA48992.2020.9068691.
- [11] Dar MN, Akram MU, Khawaja SG, Pujari AN. CNN and LSTM-Based Emotion Charting Using Physiological Signals. *Sensors*. 2020; 20(16):4551.
- [12] S. Koelstra et al., "DEAP: A Database for Emotion Analysis ;Using Physiological Signals," in *IEEE Transactions on Affective Computing*, vol. 3, no. 1, pp. 18-31, Jan.-March 2012, doi: 10.1109/TAFFC.2011.15.
- [13] N. K. Al-Qazzaz, M. K. Sabir, S. Ali, S. A. Ahmad and K. Grammer, "Effective EEG Channels for Emotion Identification over the Brain Regions using Differential Evolution Algorithm," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 2019, pp. 4703-4706, doi: 10.1109/EMBC.2019.8856854.
- [14] Hochreiter, S. & Schmidhuber, J"urgen, 1997. Long short-term memory. *Neural computation*, 9(8), pp.1735–1780.
- [15] Chung, J., Gülçehre, Ç., Cho, K., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. ArXiv, abs/1412.3

AUTHORS

Farhad Zamani is a biomedical engineering graduate from Bandung Institute of Technology. Currently he works as a researcher in Biosignal and Artificial Intelligent Research Team at Telkom Indonesia, Bandung. His research work focuses specifically on the biosignal and machine learning. He also has interests on medical informatics and bioinformatics. In his free time, Farhad ride his bicycle or read a novel.



Retno Wulansari is a researcher for Indonesia Telecommunication and Digital Research Institute at Telkom Corporate University Center, Bandung. Before that, she worked as researcher for advanced technology focused on Artificial Intelligent, Natural Language Processing and Bio Signal Research Team at Telkom Digital Service Division. Her favorite things to do in her free time is gardening and cooking.

