Emotion recognition with deep learning using GAMEEMO data set

T. B. Alakus✉ and I. Turkoglu

Emotion recognition is actively used in brain–computer interface, health care, security, e-commerce, education and entertainment applications to increase and control human–machine interaction. Therefore, emotions affect people's lives and decision-making mechanisms throughout their lives. However, the fact that emotions vary from person to person, being an abstract concept and being dependent on internal and external factors makes the studies in this field difficult. In recent years, studies based on electroencephalography (EEG) signals, which perform emotion analysis in a more robust and reliable way, have gained momentum. In this article, emotion analysis based on EEG signals was performed to predict positive and negative emotions. The study consists of four parts. In the first part, EEG signals were obtained from the GAMEEMO data set. In the second stage, the spectral entropy values of the EEG signals of all channels were calculated and these values were classified by the bidirectional long-short term memory architecture in the third stage. In the last stage, the performance of the deep-learning architecture was evaluated with accuracy, sensitivity, specificity and receiver operating characteristic (ROC) curve. With the proposed method, an accuracy of 76.91% and a ROC value of 90% were obtained.

Introduction: Emotion can be defined as the voluntary or involuntary reaction of people against an external stimulus while performing actions such as talking, thinking, communicating, learning, making decisions etc. Since all these and similar actions are carried out through emotions, emotions have a great impact on daily life. While negative emotions affect people both physically and psychologically, positive emotions make people more successful in society and bring better living conditions [1, 2]. There are many different emotion analysis studies in order to comprehend the nature and behaviour of emotions. However, the fact that the concept of emotion is abstract and does not have an objective result makes it difficult to analyse emotions [3, 4]. In addition, a large number of methods to collect and process emotion data makes the analysis process more difficult and time-consuming. For these reasons, a computer-based system is needed [5].

Emotions can be obtained through physical and non-physical methods. Examples of these include voice signals, body language, facial expressions and physical activities. Since these methods are easy to apply, data can be obtained quickly and easily. However, during the data collection phase, emotions can be manipulated intentionally or unintentionally by the subjects. While the voice signals are collected, subjects can imitate their voices and similarly hide their facial expressions [6]. Therefore, the fact that the data obtained by these methods are both incomplete and untrustworthy caused the need for a more reliable system and increased the importance of physiological signals such as electroencephalography (EEG) [7]. EEG signals are the most widely used method in this area because of their ease of use, low cost and portable models [8].

There are two types of emotional patterns in the literature, discrete and dimensional. There are eight basic emotions (anger, joy, trust, fear, surprise, sadness, disgust and anticipation) in the discrete emotion model [9]. In the dimensional model, emotions are expressed not by their names but according to their positions in the arousal–valence plane [10]. In this plane, emotions are divided into four main areas. Valence axis refers to the x-axis, and this axis indicates whether the emotion is negative or positive. Y-axis expresses arousal and emotions are ordered from low to high according to the degree of activity. The arousal–valence plane is given in Fig. 1. The plane is divided into four different zones, as can be seen in Fig. 1. While there are high arousal positive valence emotions in the first zone, there are high valence negative emotions in the second zone. In the third and fourth zones, there are emotions of negative valence-low arousal and positive valence-low arousal, respectively. In this model, emotions are named according to their location in the coordinate plane rather than their names. For example, the emotion of happiness is expressed as high arousal positive valence. Similar inferences can be made for other types of emotions. In this study, dimensional emotion model was used and emotions were evaluated as positive-valence and negative-valence.

The study consists of four stages. In the first step, EEG signals were collected from the GAMEEMO data set. In the second step, the spectral entropy values of each EEG signal were calculated. Then these values were classified with the bidirectional long-short term memory (BiLSTM) deep-learning model and the prediction process was carried out. In the last stage, the performance of the BiLSTM model was measured with different evaluation metrics.

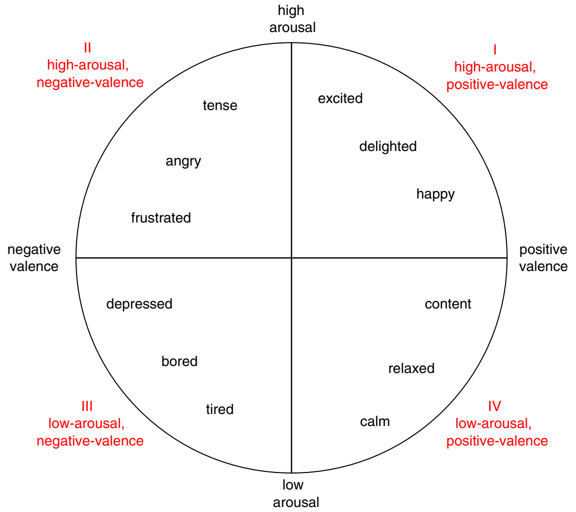

Fig. 1 *Example of arousal–valence dimensional emotion model*

The main contributions of the study can be summarised as follows:

- To the best of our knowledge, the GAMEEMO data set was analysed for the first time in this study with the BiLSTM deep-learning model.
- With this study, it was observed that EEG signals obtained from a portable device can also be used for emotion analysis.

The rest of the work is organized as follows: studies conducted with EEG signals are mentioned in the related works section. In the data and methods section, general information about the data set, the spectral entropy and BiLSTM model used in this study are given. In the application results section, the performance of the BiLSTM model was examined and the results were discussed. In the conclusion section, the study was examined and explanations were made based on possible future applications.

Related works: In this section, emotion analysis studies performed with EEG signals are examined. The authors in [11] used the LSTM model for emotion prediction and classified the EEG signals for it. The DEAP data set was used in the study and a classification process was performed for valence, arousal and liking classes. The signals have not been preprocessed and classified directly in the LSTM architecture. The study was validated with four-fold cross-validation and the performance of the LSTM model was evaluated with the accuracy metric only. At the end of the study, the accuracy of 0.8565 for arousal, 0.8545 for valence and 0.8799 for liking was obtained. Authors in [12] carried out emotion analysis using deep learning. The DEAP data set was used in the study and the feature extraction was carried out before classification. In the feature extraction phase, the signals were transformed with empirical model decomposition and variational model decomposition and the power spectral density, and the first difference values of intrinsic mode functions were collected from these converted signals. Then the signals were classified with both support vector machines (SVMs) and deep neural network. After the classification process, accuracies of 0.6125 for arousal and 0.6250 for valence were reached. The authors in [13] performed emotion analysis with principal component analysis and deep learning. As in other studies, the DEAP data set was used in this study. The signals were first transformed into five different bands with fast Fourier transform and power spectral values were obtained from each band. Then the values were normalised and classified with deep-learning network. Valence emotions were predicted with 0.5342 and arousal emotions with 0.5205 accuracies with the proposed method.

Data and methods: In this study, EEG signals belonging to the GAMEEMO [14] data set are used. The data set contains EEG signals of 28 people. Unlike conventional EEG collecting devices, the data were obtained with a portable EEG device

(Emotiv EPOC + 14-Channel Wireless EEG Headset). The EEG device used has 14 channels in total as AF3, AF4, F3, F4, F7, F8, FC5, FC6, O1, O2, P7, P8, T7 and T8. The sampling rate of the obtained signals is 128 Hz. The data set contains raw and preprocessed signals. Since noise-free data were used in this study, pre-processed data were considered.

In order to obtain emotions, the subjects played four computer games and each subject played games for 5 min. There are a total of 1568 ($4 \times 14 \times 28$) EEG data in the data set. The number 4 refers to the stimuli used. This value is 4 because 4 games were played in the data set. The number 14 indicates the number of EEG channels, while number 28 refers to the subjects. Sample length of each EEG data is 38,252. More technical and detailed information about the data set can be obtained from [14]. Researchers who want to use GAMEEMO data can access the data from the link provided (https://data.mendeley.com/datasets/b3pn4kwpmn/3).

In the study, feature extraction was carried out and spectral entropy values of EEG signals were calculated. Spectral entropy measures how sharp the spectrum of a signal is [15]. A signal with a sharp spectrum, such as the sum of sinusoids, has low spectral entropy. In contrast, a flat spectrum signal such as white noise has high spectral entropy. Spectral entropy treats the normalised power distribution in the frequency domain of the signal as a probability distribution and calculates the Shannon entropy. Shannon entropy in this context is the spectral entropy of the signal. Spectral entropy is effectively used in fault detection and diagnosis [16, 17], speech recognition [18] and biomedical signal processing [19]. In this study, the spectral entropy value of each EEG signal was calculated and these values were used to classify with BiLSTM.

In this study, recurrent neural network was used instead of traditional CNN architectures because of their success in time series applications [20–22]. For this reason, a recurrent neural network model – bidirectional LSTM, was used in the proposed study. Bidirectional LSTMs are an extension of the LSTM model and have been proposed to improve model performance in classification problems. In the BiLSTM architecture, input values train two LSTMs instead of one. Therefore, information flows both from the past to the future and from the future to the past. In traditional LSTM architectures, information from the future is evaluated and preserved, while in BLSTM architecture, information from both the past and the future is preserved and valued. Owing to this advantage, BiLSTM is more successful than LSTM [23]. Thus, BiLSTM was considered in the study. The graphical abstract of the study is given in Fig. 2.

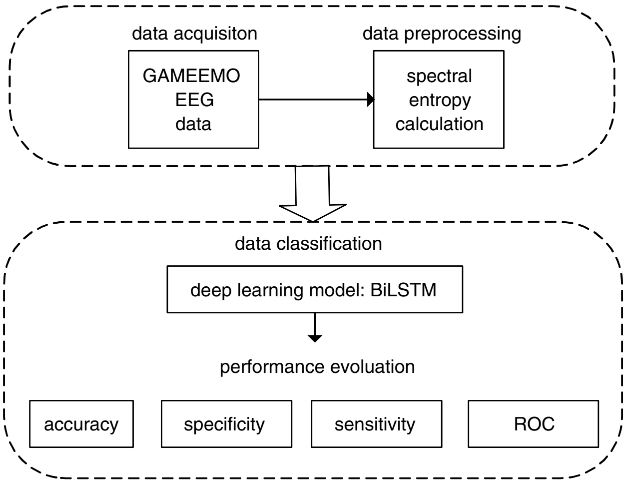

Fig. 2 *Flow chart of the study*

Application results: In this study, the EEG signals of the GAMEEMO data set were classified and positive–negative emotions were predicted. BiLSTM was used for the classification process and the performance of the deep-learning model was measured with accuracy, sensitivity, specificity and receiver operating characteristic (ROC) values. The parameters of the developed BiLSTM model can be summarised as follows:

- EEG data, whose spectral entropy values were calculated, were used in the input layer.
- Then the 128-unit BiLSTM layer was designed. ReLU function was used as an activation function.
- Then, the data were transformed into a one-dimensional vector by the flattening process.
- Later, the batch normalisation was performed and the data were normalised.
- Dropout was used to prevent overfitting problem and its degree was set to 0.25.
- Finally, a fully connected layer was designed and the number of neurons was determined as 512.
- In the classification layer, the sigmoid function has been used and the binary classification process has been made.
- Stochastic gradient descent was applied as an optimiser with default values.
- The loss of the model was calculated by binary cross-entropy.
- The epoch value was chosen to be 250.
- To validate the model, the train-test split approach was used and 80% of the data was used for training and 20% for testing.
- All of these parameters were determined by trial and error approach and the parameters giving the best result were used in the study.
- Table 1 shows the classification results of the BiLSTM model.

Table 1: Classification results of positive and negative emotions

Accuracy, %	Sensitivity, %	Specificity, %	ROC
76.91	76.93	76.89	0.90

As seen in Table 1, positive and negative emotions were classified with an accuracy rate of 0.7691 with the proposed BiLSTM model. In addition, the sensitivity value was 0.7693 and the specificity value was 0.7689. ROC value was measured as 0.90 for both classes. The graph of the ROC is given in Fig. 3. The area under the curve (AUC) score is used effectively in biomedical studies and is expressed as a better analysis [24]. In order for the classification process to be considered good, the AUC score must be >0.8 [24]. Furthermore, the AUC score between 0.9 and 1.0 indicates that the classification is excellent [24]. In this study, the AUC score was calculated as 0.9, indicating that the proposed method is effective and successful.

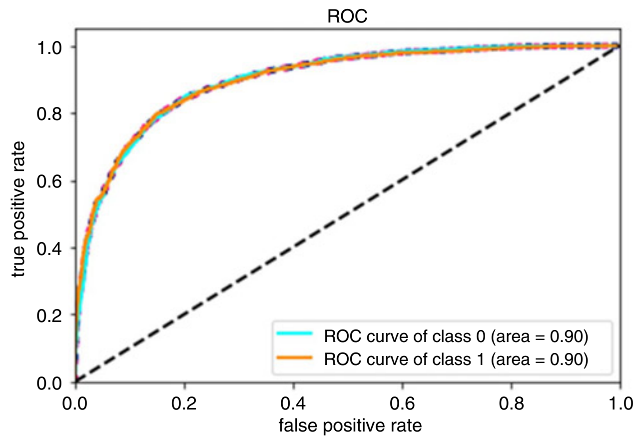

Fig. 3 *ROC curve of positive and negative emotions (class 0 refers to negative emotions, class 1 refers to positive emotions)*

These results were also compared with the machine-learning algorithm results used in the original article. The comparison results are given in Table 2. Since there is only one study in the literature with this data set, only the results in the original article could be examined. According to the results given in Table 2, it is seen that the proposed deep-learning method was better than the machine-learning algorithms used in the original study. While 73 and 66% accuracy values were achieved for SVM and KNN, respectively, this rate increased to ~77% with the BiLSTM model. According to these results, it can be inferred that the deep-learning method is at least as successful and even better as existing machine-learning methods.

Table 2: Comparison of classification results

Reference	SVM (accuracy)	KNN (accuracy)	BiLSTM (accuracy)
[14]	73%	66%	—
the proposed method	—	—	76.93%

Conclusion: In this study, positive and negative emotions were analysed using the EEG data of the GAMEEMO data set. In the first part of the study, pre-processed data were obtained from the data set. Then, spectral entropy values were collected from the data of each EEG channel and these values were used in the BiLSTM model. In the final phase, the classification process was made with BiLSTM and the performance of the deep-learning model was measured with accuracy, sensitivity, specificity and ROC values. With the proposed method, 76.91% accuracy, 76.93% sensitivity, 76.89% specificity and 90% ROC values were achieved. In addition, the proposed method was compared with the machine-learning algorithms used in the original article and it was observed that the proposed method was at least as successful as them. In the future, this data set will be examined in more detail and comparisons will be made using different deep-learning algorithms and signal processing methods. Emotions will be examined with both binary-class classification and multi-class classification. Emotions are of great importance in human life. In daily life, we use our emotions intentionally or unintentionally. Therefore, emotion analysis studies are important for understanding emotions and determining their behaviour.

Submitted: *23 August 2020* E-first: *22 October 2020*
doi: 10.1049/el.2020.2460
One or more of the Figures in this Letter are available in colour online.

T. B. Alakus (*Kirklareli University, Faculty of Engineering, Department of Software Engineering, Kirklareli, Turkey*)

✉ E-mail: talhaburakalakus@klu.edu.tr

I. Turkoglu (*Firat University, Faculty of Technology, Department of Software Engineering, Elazig, Turkey*)

References

1 Naji, M., Firoozabadi, M., and Azadfallah, P.: 'Emotion classification during music listening from forehead biosignals', *Signal. Image. Video. Process.*, 2015, **9**, pp. 1365–1375, doi: 10.1007/s11760-013-0591-6
2 Alakus, T.B., and Turkoglu, I.: 'EEG based emotion analysis systems', *TBV J. Comput. Sci. Eng.*, 2018, **11**, (1), pp. 26–39
3 Michalopoulos, K., and Bourbakis, N.: 'Application of multiscale entropy on EEG signals for emotion detection'. IEEE EMBS Int. Conf. Information Technology Applications in Biomedicine, Orlando, FL, USA, February 2017, pp. 341–344, doi: 10.1109/BHI.2017.7897275
4 Alakus, T.B., and Turkoglu, I.: 'Feature selection with sequential forward selection algorithm from emotion estimation based on EEG signals', *Sakarya Univ. J. Sci.*, 2019, **23**, (6), pp. 1096–1105, doi: 10.16984/saufenbilder.501799
5 Turnip, A., Simbolon, A.I., Amri, M.F., *et al.*: 'Backpropagation neural networks training for EEG-SSVEP classification of emotion recognition', *Internetwork. Indonesia J.*, 2017, **9**, (1), pp. 53–57
6 Alakus, T.B., and Turkoglu, I.: 'Determination of effective EEG channels for discrimination of positive and negative emotions with wavelet decomposition and support vector machines', *Int. J. Inform. Technol.*, 2019, **12**, (3), pp. 229–237
7 Yan, J., Chen, S., and Deng, S.: 'A EEG-based emotion recognition model with rhythm and time characteristics', *Brain. Inform.*, 2019, **6**, (1), pp. 1–8, doi: 10.1186/s40708-019-0100-y
8 Pan, J., Li, Y., and Wang, J.: 'An EEG-based brain-computer interface for emotion recognition'. Int. Joint Conf. Neural Networks, Vancouver, BC, Canada, July 2016, pp. 2063–2067, doi: 10.1109/IJCNN.2016.7727453
9 Mason, W.A., and Capitanio, J.P.: 'Basic emotions: a reconstruction', *Emot. Rev.*, 2012, **4**, (3), pp. 238–244, doi: 10.1177/1754073912439763
10 Russel, A.: 'Core affect and psychological construction of emotion', *Psychol. Rev.*, 2003, **110**, (1), pp. 145–150
11 Alhagry, S., Fahmy, A.A., and El-Khoribi, R.A.: 'Emotion recognition based on EEG using LSTM recurrent neural network', *Int. J. Adv. Comput. Sci. Appl.*, 2017, **8**, (10), pp. 355–358, doi: 10.14569/IJACSA.2017.081046
12 Pandey, P., and Seeja, K.R.: 'Subject independent emotion recognition from EEG using VMD and deep learning', *J. King Saud Univ. – Comput. Inform. Sci.*, 2019, **2019**, pp. 53–58, doi: 10.1016/j.jksuci.2019.11.003
13 Jirayucharoensak, S., Pan-Ngum, S., and Israsena, P.: 'EEG-based emotion recognition using deep learning network with principal component based covariate shift adaptation', *Sci. World J.*, 2014, **2014**, pp. 1–10, 627892, doi: 10.1155/2014/627892
14 Alakus, T.B., Gonen, M., and Turkoglu, I.: 'Database for an emotion recognition system based on EEG signals and various computer games – GAMEEMO', *Biomed. Signal Proc. Control*, 2020, **60**, pp. 1–12, doi: 10.1016/j.bspc.2020.101951
15 Vanluchene, A.L.G., Vereecke, H., Thas, O., *et al.*: 'Spectral entropy as an electroencephalographic measure of anesthetic drug effect: a comparison with bispectral index and processed midlatency auditory evoked response', *Anesthesiology*, 2004, **101**, (1), pp. 34–42
16 Pan, Y.N., Chen, J., and Li, X.L.: 'Spectral entropy: a complementary index for rolling element bearing performance degradation assessment', *Proc. Inst. Mech. Eng. C, J. Mech. Eng. Sci.*, 2009, **223**, (5), pp. 1223–1231, doi: 10.1243/09544062JMES1224
17 Sharma, V., and Parey, A.: 'A review of gear fault diagnosis using various condition indicators', *Procedia Eng.*, 2016, **144**, pp. 256–263, doi: 10.1016/j.proeng.2016.05.131
18 Majstorovic, N., Andric, M., and Mikluc, D.: 'Entropy-based algorithm for speech recognition in noisy environment'. Telecommunications Forum, Belgrade, Serbia, November 2011, pp. 667–670, doi: 10.1109/TELFOR.2011.6143635
19 Vakkuri, A., Yli-Hankala, A., Talja, P., *et al.*: 'Time-frequency balanced spectral entropy as a measure of anesthetic drug effect in central nervous system during sevoflurane, propofol, and thiopental anesthesia', *Acta Anaesthesiol. Scand.*, 2004, **48**, (2), pp. 145–153, doi: 10.1111/j.0001-5172.2004.00323.x
20 Hewamalage, H., Bergmeir, C., and Bandara, K.: 'Recurrent neural networks for time series forecasting: current status and future directions', *Int. J. Forecast.*, 2020, **2020**, pp. 1–40, doi: 10.1016/j.ijforecast.2020.06.008
21 Guo, T., Xu, Z., Yao, X., *et al.*: 'Robust online time series prediction with recurrent neural networks'. Int. Conf. Data Science and Advanced Analytics, Montreal, QC, Canada, October 2016, pp. 816–825, doi: 10.1109/DSAA.2016.92
22 Connor, J.T., Martin, R.D., and Atlas, L.E.: 'Recurrent neural networks and robust time series prediction', *IEEE Trans. Neural Netw.*, 1994, **5**, (2), pp. 240–254, doi: 10.1109/72.279188
23 Graves, A., Fernandez, S., and Schmidhuber, J.: 'Bidirectional LSTM networks for improved phoneme classification and recognition'. Int. Conf. Artificial Neural Networks: Formal Models and Their Applications, Warsaw, Poland, September 2005, doi: 10.1007/11550907_126
24 Mandrekar, J.N.: 'Receiver operating characteristic curve in diagnostic test assessment', *J. Thorac. Oncol.*, 2010, **5**, (9), pp. 1315–1316, doi: 10.1097/JTO.0b013e3181ec173d