

# CNN-LSTM Based Stress Recognition Using Wearables

Ritu Tanwar<sup>1,\*,\dagger</sup>, Orchid Chetia Phukan<sup>2,\dagger</sup>, Ghanapriya Singh<sup>1</sup> and Sanju Tiwari<sup>3,\dagger</sup>

<sup>1</sup>Department of Electronics Engineering, National Institute of Technology, Uttarakhand, India

<sup>2</sup>P.E.S University, Bengaluru, Karnataka, India

<sup>3</sup>Universidade Autonoma de Tamaulipas, Mexico

## Abstract

Mental health diseases such as depression, anxiety, and heart-related disorders are a big concern in today's life. As a result, it is now crucial to monitor stress levels. Here, we propose and suggest a deep learning hybrid approach convolution neural network-long short-term memory (CNN-LSTM) model that can recognize stress. Electrocardiogram (ECG), electromyography (EMG), body temperature, electrodermal activity (EDA), and respiration biomedical signals vary, when a human being is exposed to stress. Physiological signals can be measured using wearable devices, which can help measure stress. Therefore, it is possible to recognize stress using wearables-based physiological signals. For stress detection, the deep learning approach CNN-LSTM was implemented for feature learning and three-class stress classification (baseline, stress, amusement). To evaluate the model performance on stress detection, confusion matrix, and accuracy were used. The accuracy obtained by the proposed model was 90.20 % for stress classification, which is higher as compared to other previous studies. Consequently, the proposed model can help people in stress management in office working environments, in driving conditions, or everyday life. Furthermore, integrating the proposed model with other mechanisms such as attention and explainability may become more accurate and capable of stress detection and healthcare systems development.

## Keywords

Hybrid Deep Learning, Stress Recognition, Wearables, CNN-LSTM.

## 1. Introduction

The mental states of a human being, such as stress, are reflected in physiological signals [1]. Wearable sensors such as wrist-worn or chest-worn devices can be used to measure physiological signal variations [2]. Many researchers have contributed to studying the correlation existence between the physiological signals and the stress levels of a human being[3],[4],[5]. Stress levels of a human being can be recognized by observing the physiological signals pattern. The stress

---

*DLQ-2022: International Workshop on Deep Learning for Question Answering, Co-located with the KGSWC-2022, November 21-23, 2022, Madrid, Spain.*

\*Corresponding author.

<sup>\dagger</sup>These authors contributed equally.

✉ ritu.tanwar2012@gmail.com (R. Tanwar); orchidchetiaphukan1@gmail.com (O. C. Phukan);

ghanapriya@nitkkr.ac.in (G. Singh); tiwarisanju18@ieee.org (S. Tiwari)

ORCID 0000-0003-0110-4197 (R. Tanwar)



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

CEUR Workshop Proceedings (CEUR-WS.org)

levels are observed to be correlated with biomedical signals such as movement, respiration, and body temperature [6],[7]. Therefore, it is required to detect the stress levels for health monitoring of human beings with affected states such as depression, anxiety, and stress.

Previously, measuring the physiological signals involved cumbersome procedures with wearing very uncomfortable body sensors and was also expensive. But advancements in wearable technologies have made it easy and comfortable to measure physiological signal variations [8]. Wearables-based devices help monitor the human mental state and are also helpful for other purposes and applications [9], [10], [11]. In this study, various human beings' biomedical signals are used to recognize the stress levels in a very effective and accurate manner. A deep learning approach based on CNN-LSTM models was implemented to detect stress levels. The main contributions of this work are:

1. Multi-physiological signals (ECG, EMG, body temperature, EDA, and respiration) are considered for stress monitoring,
2. Wearable devices are used for collecting physiological signals in the dataset, and
3. A hybrid CNN-LSTM deep learning approach is proposed to classify three-class stress levels (baseline, stress, amusement).

The proposed methodology helps pass the information through different layers for pre-processing, feature learning and classification purposes. The paper is structured in four sections: the first section presents the related work in the stress recognition domain using machine learning, deep neural networks, and wearables approaches. Second section provides the dataset and methodology used in this study. The third section details the results of the proposed approach, and fourth section provides the conclusion.

## 2. Related work

Wearables-based stress recognition uses physiological signals such as a EMG, EDA, and ECG to determine the stress, neutral and amusement levels of human beings. Wearables based stress recognition approach exploits the possibilities of machine learning and deep neural networks' potential in healthcare [12, 13]. According to a survey [14], wearables or sensors do not interfere with daily life activities and help monitor stress levels quickly, leading researchers to innovate computational stress recognition by optimizing sensor modalities and using machine learning or deep learning approaches. Using uncomfortable sensors in laboratories to check stress levels might be frustrating.

Enormous implementation and successful models based on machine learning and deep neural networks method such as CNN in stress recognition have been introduced in recent years. For stress detection, a study presented a publically available dataset: wearable stress and affect detection (WESAD). Chest worn and wrist-worn based sensors named RespiBAN professional and Empatica E4 were selected to collect the physiological signals such as three-axis accelerometer, EDA, ECG, EMG, temperature, and respiration. Fifteen participants were chosen to collect the data and were exposed to stress conditions. Three-class classification (stress, amusement, and baseline) was evaluated using machine learning methods (e.g., k-Nearest Neighbor (kNN), Decision Tree (DT), and Random Forest (RF)) [12].

Montesinos et al. [15] detected acute stress using physiological signals from wrists-based Empatica E4 and chest-based Shimmer3 ECG wearable sensors, exploiting the machine learning algorithms such as kNN, DT, and RF to detect the stress levels and found accuracy of 84.13 %. Can et al. [16] also introduced an automatic stress recognition system using Samsung Gear S family wearables and Empatica E4 wearables. They used machine learning methods, e.g., kNN, linear regression (LR), RF, support vector machine (SVM), and multilayer perceptron to recognize stress in 21 participants. Cosoli et al. [17] also implemented machine learning techniques on the WESAD dataset for assessing human reactions with LR and SVM approaches and found 75 and 72.62 % accuracy.

The researchers presented a stress-recognition model using a deep learning approach and found 78.8 % accuracy. Here, stress based datasets were used to detect the stress and implemented the CNN model with ECG and EDA physiological signals [18]. A study proposed a stress recognition model using the WESAD dataset and implemented machine learning and deep neural networks to recognize the three class stress levels. They implemented DT, RF, kNN, SVM, linear discriminant analysis (LDA), and artificial neural network (ANN) [19, 20] for stress level classification with an accuracy rate of 68.16, 75.95, 74.71, 81.65, 74.82, and 84.32 % respectively [21].

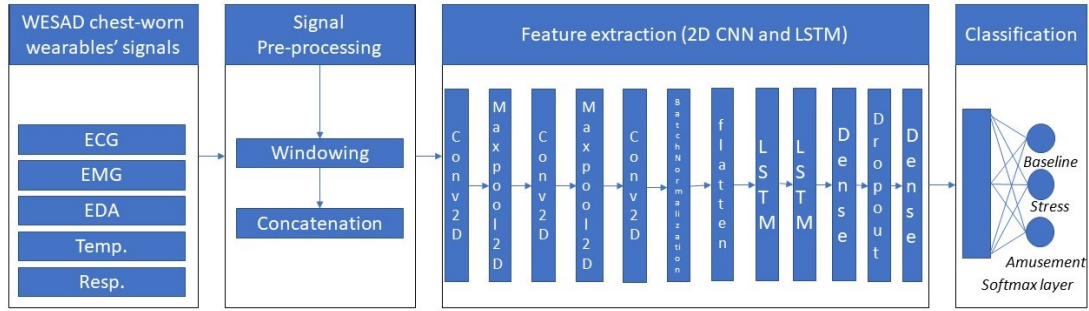
A recent study proposed a hybrid CNN deep neural network method for stress recognition with WESAD dataset and found 75.21 accuracy with three-class stress classification. They compared the results with conventional machine learning approaches and found low accuracy [22]. The researchers attempted to detect stress on the StudentLife dataset and implemented LSTM, CNN, and CNN-LSTM deep learning algorithms [23]. They found an accuracy rate of 62.83% with LSTM, 60.43 % with CNN, and 60% with CNN-LSTM approaches [24]. These findings suggest that wearables data and stress issues are related. The outcomes of stress recognition models vary depending on the wearables used, application areas, available stress datasets, and methodological approach. Researchers use different methods to recognize stress with wearable sensor data.

### **3. Materials and methods**

This part outlines and details of the dataset used in this study, preprocessing of the data and the methodology implemented to recognize the stress based on wearables.

#### **3.1. Dataset used**

WESAD, a publically available stress dataset introduced by Schmidt, was used to detect stress levels. This dataset involved 15 participants and is multivariate time-series data. The chest-worn and wrist-worn RespiBAN professional and Empatica E4-based wearables were selected to collect the biomedical signals data. Chest worn device collects the ECG, EDA, EMG, Respiration, body temperature, and three-axis accelerometer physiological signals from the participants with a sampling frequency of 700 Hz. The wrist-worn wearable collects BVP, EDA, body temperature, and accelerometer signals with sampling frequencies of 64 Hz, 4 Hz, 4 Hz, and 32 Hz, respectively. This dataset aims at inducing three mental states: baseline, stress, and amusement [12]. From Schimdt et al.[12], it was observed that chest-worn devices provided the best results. Thus, only



**Figure 1:** The proposed CNN-LSTM model for wearables-based stress detection. Chest worn wearables signals (ECG, EMG, Body Temperature, EDA, and Respiration) are input and stress, baseline, and amusement are the output.

chest-worn device signals were considered in this study. The proposed architecture includes wearable data selection, pre-processing, modalities fusion, feature extraction, and classification (Figure 1).

### 3.2. Pre-processing

For data pre-processing, chest-worn signals sampled at 700 Hz were selected. The different modalities, ECG, EMG, body temperature, EDA, and respiration, were pre-processed to make them ready to feed into the model. The windowing technique where data signals were segmented into 5-second window lengths with 2-second window shift was implemented for pre-processing. After pre-processing, each participant data was concatenated for constructing test data and training data for stress recognition.

### 3.3. Hybrid deep learning approach: Convolution neural network- Long short-term memory (CNN-LSTM)

Figure 1 shows the hybrid CNN-LSTM model used in this work. This model classifies stress level by fusing the chest-worn device modalities to classify baseline, stress, and amusement mental states. The configuration details of layers used in the proposed model are given in Table 1. CNN-LSTM model has three 2D convolution layers, two max-pooling layers, one batch normalization layer, two dense layers, one dropout, and one flatten layer. The model was selected with the best accuracy rate through trial and error. The convolution layer extracted the important features with its sliding filters. Rectified linear unit (ReLU) function was selected as an activation function, enhancing the model's convergence speed and robustness. The max pooling layer minimizes the data by 50 % to minimize the computational complexity. The batch normalization layer is implemented for data normalization and speeding up the training and feature learning. After normalization, flatten layer is used to create a one-dimensional feature vector. A one-dimensional vector is given as input to the two LSTM layers.

LSTM contains a forget gate, input, and output gate to process the input time series. These gates can manage the addition or deletion of information to facilitate forgetting and recollection.

**Table 1**

The proposed CNN-LSTM model configuration details

| Layer               | Output shape  | No. of parameters | Activation function | Others                            |
|---------------------|---------------|-------------------|---------------------|-----------------------------------|
| Input               | (-,5,3500,10) | -                 | -                   | -                                 |
| 2D conv             | (-,5,3500,10) | 260               | ReLU                | No. of kernels=5, padding= 'same' |
| Max pooling         | (-,2,1750,10) | -                 | -                   | Pool size = (2,2)                 |
| 2D conv             | (-,2,1750,20) | 820               | ReLU                | No. of kernels=5, padding= 'same' |
| Max pooling         | (-,1,875,20)  | -                 | -                   | Pool size = (2,2)                 |
| 2D conv             | (-,1,875,30)  | 2430              | ReLU                | No. of kernels=5, padding= 'same' |
| Batch normalization | (-,1,875,30)  | 120               | -                   | -                                 |
| Flatten             | (-,1,26250)   | -                 | -                   | -                                 |
| LSTM                | (-,1,128)     | 13506048          | Tanh                | -                                 |
| LSTM                | (-,60)        | 45360             | Tanh                | -                                 |
| Dense               | (-,512)       | 31232             | ReLU                | No of neurons, initializer = ''   |
| Dropout             | (-,512)       | 0                 | -                   | Portion=0.3                       |
| Dense               | (-,3)         | 1539              | Softmax             | -                                 |

In the LSTM network, the feature maps are transformed into matching hidden states. The output from LSTM layers is given to the dense layer, which is a completely connected layer with the neurons from all the previous layers. The dropout layer is also used in between two dense layers to prevent the overfitting of the fully connected neural networks. The output from the dense layer is given as an input vector to the last phase of classification and passed to the softmax layer to classify the stress. The suggested framework is applicable for processing various multimodal time series data. Without the requirement for handcrafted features, the suggested framework employs an end-to-end training methodology.

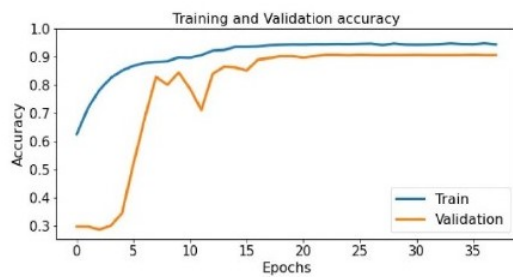
## 4. Results

The experiments were performed on the google platform "Google colab" and used 'Tensorflow' library for python coding. Out of 15 participants' data, 14 participants data was used for training the model and one for testing. The proposed model was trained with a cross-entropy loss function. Here, performance metrics, confusion matrix, accuracy, and cross-entropy loss function curves for training and validation of the model were selected to determine the proposed model performance for stress recognition. The results for baseline, stress, and amusement classification obtained with the proposed model using performance metrics are given in Table 2. Accuracy, precision, and F1-score are used as performance metrics. The definitions of performance metrics used are given in the Table 2. It is found in Table 2 that the precision for baseline, stress, and amusement conditions is 92.28, 90.57, and 84.72, respectively. Further, the F1 score for baseline, stress, and amusement conditions is 94.41, 90.39, and 77.91, respectively. Figure 2 shows the accuracy rate and cross-entropy loss with an increasing number of epochs for training and validation of the model.

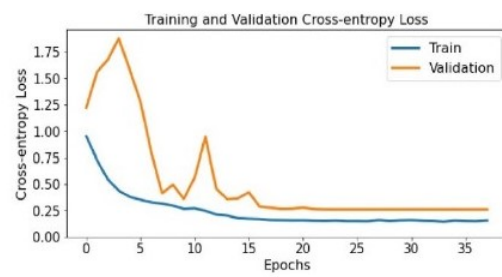
**Table 2**

Details of performance metrics and values achieved by the CNN-LSTM model

| Performance metrics | Definition   | Baseline (%) | Stress (%) | Amusement (%) |
|---------------------|--|--------------|------------|---------------|
| Accuracy            | Ratio of correctly classified samples to sum of samples  | 96.63        | 90.20      | 72.12         |
| Precision           | Ratio of true positive prediction samples to sum of true samples   | 92.28        | 90.57      | 84.72         |
| F1-score            | Measure of model accuracy. Mean of precision and recall, where recall is ratio of no. of true positive predicted samples to sum of samples | 94.41        | 90.39      | 77.91         |



(a)



(b)

**Figure 2:** Training and validation accuracy rate (a) and cross-entropy loss (b) for the model**Table 3**

Comparison with other previous studies on wearables-based stress recognition

| S.No. | Previous Studies       | Wearables used           | Method used   | Accuracy (%) |
|-------|------------------------|--------------------------|---|--------------|
| 1.    | Schimdt et al.[12]     | RespiBAN and Empatica E4 | Machine learning models<br>DT, RF, LDA, kNN<br>AdaBoost | 76.50        |
| 2.    | Chakraborty et al.[25] | RespiBAN and Empatica E4 | Deep learning model<br>(CNN)                            | 77.06        |
| 3.    | Montesinos [15]        | Shimmer3 and Empatica E4 | Machine learning models<br>(kNN, DT, RF)                | 84.13        |
| 4.    | Gil-Martin et al.[26]  | RespiBAN and Empatica E4 | Deep learning model<br>(CNN)                            | 77.21        |
| 5.    | Cosoli et al.[17]      | RespiBAN and Empatica E4 | Machine learning models<br>(LR, SVM)                    | 75           |
| 6.    | Proposed model         | RespiBAN and Empatica E4 | Deep learning model<br>(CNN-LSTM)                       | 90.20        |

Table 3 compares the proposed model with other stress recognition studies. Some previous studies are there that are based on stress recognition using wearables [12], [25], [15], [17]. Schmidt et al. [12] used machine learning models (e.g., DT, RF, LDA, and kNN) and accuracy obtained with these models was 76.50%, when chest worn signals were considered. Another study by Montesinos et al. [15] also used machine learning methods and obtained the accuracy of 84.13% with Shimmer3 and Empatica E4 wearables. Further, Cosoli et al. [17] used these models and obtained the accuracy of 75%. Thus, the studies that have used machine learning models for stress detection using wearables have obtained accuracy upto 85% approximately [12], [15], [17]. In contrast to machine learning approaches, deep learning methods implemented on physiological signals acquired through wearables had accuracy of approximately 77% [25], [26]. In this proposed hybrid model of CNN-LSTM deep learning approach, 90% accuracy is achieved.

## 5. Conclusion

In this study, a hybrid CNN-LSTM based model for stress recognition is proposed. We achieved an accuracy of 90.20% by using chest-worn device signals such as ECG, EMG, body temperature, EDA, and respiration. The results show accuracy improvements compared to previous studies that classified stress levels. This is mostly attributable to their capacity to benefit from the extensive hierarchical and temporal data dependence between different physiological signals. This may help to explain the reason behind higher accuracy achieved with CNN-LSTM model than model that solely employ CNN and LSTM structures when ECG, EMG, EDA, RESP, and Temp are simultaneously collected multivariate signals for stress identification. In future work, we plan to integrate an attention mechanism to extract the stress-related information effectively and improve the stress recognition accuracy rate. The proposed stress recognition model would be helpful in stress management in modern society's everyday life. Moreover, it is expected to assist people with depression, anxiety, and other stress-related problems.

## References

- [1] I. B. Mauss, R. W. Levenson, L. McCarter, F. H. Wilhelm, J. J. Gross, The tie that binds? coherence among emotion experience, behavior, and physiology., *Emotion* 5 (2005) 175.
- [2] D. P. Tobón Vallejo, A. El Saddik, Emotional states detection approaches based on physiological signals for healthcare applications: a review, *Connected Health in Smart Cities* (2020) 47–74.
- [3] A. Soni, K. Rawal, A review on physiological signals: Heart rate variability and skin conductance, in: *Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019)*, Springer, 2020, pp. 387–399.
- [4] M. Zanetti, L. Faes, M. De Cecco, A. Fornaser, M. Valente, G. Guandalini, G. Nollo, Assessment of mental stress through the analysis of physiological signals acquired from wearable devices, in: *Italian Forum of Ambient Assisted Living*, Springer, 2018, pp. 243–256.
- [5] A. Arza, J. M. Garzón-Rey, J. Lázaro, E. Gil, R. Lopez-Anton, C. de la Camara, P. Laguna, R. Bailon, J. Aguiló, Measuring acute stress response through physiological signals: towards

- a quantitative assessment of stress, *Medical & biological engineering & computing* 57 (2019) 271–287.
- [6] C. Tsigos, G. P. Chrousos, Hypothalamic–pituitary–adrenal axis, neuroendocrine factors and stress, *Journal of psychosomatic research* 53 (2002) 865–871.
  - [7] C. Schubert, M. Lambertz, R. Nelesen, W. Bardwell, J.-B. Choi, J. Dimsdale, Effects of stress on heart rate complexity—a comparison between short-term and chronic stress, *Biological psychology* 80 (2009) 325–332.
  - [8] M. Ragot, N. Martin, S. Em, N. Pallamin, J.-M. Diverrez, Emotion recognition using physiological signals: laboratory vs. wearable sensors, in: *International Conference on Applied Human Factors and Ergonomics*, Springer, 2017, pp. 15–22.
  - [9] H. Jebelli, B. Choi, S. Lee, Application of wearable biosensors to construction sites. i: Assessing workers’ stress, *Journal of Construction Engineering and Management* 145 (2019) 04019079.
  - [10] L. Han, Q. Zhang, X. Chen, Q. Zhan, T. Yang, Z. Zhao, Detecting work-related stress with a wearable device, *Computers in Industry* 90 (2017) 42–49.
  - [11] F. de Arriba-Pérez, J. M. Santos-Gago, M. Caeiro-Rodríguez, M. Ramos-Merino, Study of stress detection and proposal of stress-related features using commercial-off-the-shelf wrist wearables, *Journal of Ambient Intelligence and Humanized Computing* 10 (2019) 4925–4945.
  - [12] P. Schmidt, A. Reiss, R. Duerichen, C. Marberger, K. Van Laerhoven, Introducing wesad, a multimodal dataset for wearable stress and affect detection, in: *Proceedings of the 20th ACM international conference on multimodal interaction*, 2018, pp. 400–408.
  - [13] M. Reza, G. Hossain, A. Goyal, S. Tiwari, A. Tripathi, A. Bhan, P. Dash, et al., Automatic diabetes and liver disease diagnosis and prediction through svm and knn algorithms, in: *Emerging Technologies in Data Mining and Information Security*, Springer, 2021, pp. 589–599.
  - [14] N. Sharma, T. Gedeon, Objective measures, sensors and computational techniques for stress recognition and classification: A survey, *Computer methods and programs in biomedicine* 108 (2012) 1287–1301.
  - [15] V. Montesinos, F. Dell’Agnola, A. Arza, A. Aminifar, D. Atienza, Multi-modal acute stress recognition using off-the-shelf wearable devices, in: *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2019, pp. 2196–2201.
  - [16] Y. S. Can, N. Chalabianloo, D. Ekiz, C. Ersoy, Continuous stress detection using wearable sensors in real life: Algorithmic programming contest case study, *Sensors* 19 (2019) 1849.
  - [17] G. Cosoli, A. Poli, L. Scalise, S. Spinsante, Measurement of multimodal physiological signals for stimulation detection by wearable devices, *Measurement* 184 (2021) 109966.
  - [18] K. Radhika, V. R. M. Oruganti, Transfer learning for subject-independent stress detection using physiological signals, in: *2020 IEEE 17th India Council International Conference (INDICON)*, IEEE, 2020, pp. 1–6.
  - [19] S. Tiwari, O. Dogan, M. Jabbar, S. K. Shandilya, F. Ortiz-Rodriguez, S. Bajpai, S. Banerjee, Applications of machine learning approaches to combat covid-19: A survey, *Lessons from COVID-19* (2022) 263–287.
  - [20] D. Gaurav, F. O. Rodriguez, S. Tiwari, M. Jabbar, Review of machine learning approach for



drug development process, in: *Deep Learning in Biomedical and Health Informatics*, CRC Press, 2021, pp. 53–77.

- [21] P. Bobade, M. Vani, Stress detection with machine learning and deep learning using multimodal physiological data, in: *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, IEEE, 2020, pp. 51–57.
- [22] N. Rashid, L. Chen, M. Dautta, A. Jimenez, P. Tseng, M. A. Al Faruque, Feature augmented hybrid cnn for stress recognition using wrist-based photoplethysmography sensor, in: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 2021, pp. 2374–2377.
- [23] N. Kanagaraj, D. Hicks, A. Goyal, S. Tiwari, G. Singh, Deep learning using computer vision in self driving cars for lane and traffic sign detection, *International Journal of System Assurance Engineering and Management* 12 (2021) 1011–1025.
- [24] Y. Acikmese, S. E. Alptekin, Prediction of stress levels with lstm and passive mobile sensors, *Procedia Computer Science* 159 (2019) 658–667.
- [25] S. Chakraborty, S. Aich, M.-i. Joo, M. Sain, H.-C. Kim, A multichannel convolutional neural network architecture for the detection of the state of mind using physiological signals from wearable devices, *Journal of healthcare engineering* 2019 (2019).
- [26] M. Gil-Martin, R. San-Segundo, A. Mateos, J. Ferreiros-Lopez, Human stress detection with wearable sensors using convolutional neural networks, *IEEE Aerospace and Electronic Systems Magazine* 37 (2022) 60–70.