Stress Detection with a Lightweight and Explainable Model for IoMT Devices

José L. López Ruiz[1(✉)], Carlos Montoya Peña[1], David Díaz Jiménez[1], Juan F. Gaitán Guerrero[1], Pedro A. Palomino Moral[2], and Macarena Espinilla Estévez[1]

[1] Department of Computer Science, University of Jaén, Jaén, Spain
llopez@ujaen.es
[2] Department of Nursing, University of Jaén, Jaén, Spain

Abstract. Stress is an increasingly serious public health problem, as prolonged exposure to stress is associated with cognitive, cardiovascular, and emotional disorders. However, despite its impact, stress frequently goes undetected in everyday life, emphasising the necessity for objective and continuous monitoring systems. The present paper proposes a lightweight and interpretable stress detection model based solely on physiological signals. The proposed solution is based on a deep Multilayer Perceptron architecture optimized for execution on resource-constrained devices, typically used in medical Internet of Things environments. The model was trained using the SWELL-KW dataset, and it has been demonstrated to achieve high performance in the classification of stress. In order to enhance explainability, SHAP values were calculated to identify the most relevant input features. This process enabled the development of a simplified version of the model using only 9 features. It has been demonstrated that both the full and reduced models achieve excellent results and demonstrate the feasibility of implementing accurate, explainable deep learning solutions in real-time stress monitoring systems for clinical and wearable environments.

Keywords: Health monitoring · Internet of Medical Things · Stress detection · Explainable AI

1 Introduction

According to the World Health Organization (WHO), stress [24] can be defined as a state of mental tension or worry caused by a difficult situation. We experience stress every day, as it is an innate reaction of the body to threatening or demanding situations.

This result has been partially supported by grant PID2021-127275OB-I00 funded by MICIU/AEI/10.13039/501100011033 and by "ERDF A way of making Europe" and grant PDC2023-145863-I00 funded by MICIU/AEI/10.13039/501100011033 and by "European Union NextGenerationEU/PRTR".

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2026
E. Corchado et al. (Eds.): SOCO 2025, CCIS 2806, pp. 237–246, 2026.
https://doi.org/10.1007/978-3-032-19763-4_22

Low levels of stress can be useful in helping us cope with everyday activities. However, when stress levels are excessive and sustained over time, they can trigger adverse physical and psychological effects [25]. People who are continuously exposed to stressful environments are at greater risk of central nervous system disorders, which affect the brain, memory, and cognitive functions. Similarly, negative effects on the immune, cardiovascular, and endocrine systems can also be observed.

Workplace environments significantly influence stress levels, which can vary depending on the nature of the job, the context in which professional activities take place, and the socioeconomic status of workers [21]. Several studies indicate that workers in the social and healthcare sector [2] are among the most exposed, especially those working in emergency services, intensive care units, or emergency vehicles. Similarly, caregivers of patients with cognitive impairment, including family members, are often under constant emotional strain, placing them in a situation of chronic stress that affects their physical and mental health.

Although stress can be mitigated through individual and organizational strategies, both in daily life and in the workplace, we are often not fully conscious of the specific levels of stress to which we are exposed, which makes it difficult to manage effectively and can delay the adoption of preventive or corrective measures.

Thanks to advances in artificial intelligence (AI) and the progressive reduction in the cost of microchips, the integration of AI and the Internet of Things (IoT) in the social and healthcare sector [15] represents a technological revolution with profound implications for diagnosis, treatment, and personalized care [4]. In this context, IoT has evolved into the Internet of Medical Things (IoMT), encompassing devices and systems exclusively for healthcare environments that combine large-scale data processing with intelligent automation, supporting healthcare professionals and enhancing evidence-based clinical decision-making [3]. AI plays a central role in precision medicine by enabling the analysis of clinical, genetic, and lifestyle data to tailor individualized treatments [9], while IoMT-connected devices [11,21] facilitate continuous remote patient monitoring, allowing earlier interventions and reduced healthcare costs, particularly for chronic conditions and long-term care.

These technologies take on strategic value in the context of stress. Given that individuals are often not fully aware of their stress levels, an intelligent monitoring system based on IoMT can be essential for detecting physiological patterns associated with this problem, facilitating objective measurement in real time and enabling active alerts to be sent to the person being monitored or clinical information to be provided to healthcare personnel, allowing them to assess the person's state of health and provide individualized treatment.

This paper presents a prototype designed to classify physiological signal samples and determine whether a person is in a state of stress. The proposed solution is based on a deep Multilayer Perceptron (MLP) model, optimized for execution on devices with limited resources, such as those used in IoMT-based environ-

ments. In addition, the model incorporates explainability mechanisms that allow its predictions to be interpreted and facilitate its validation in clinical contexts.

The novelties of the work are as follows: the design of a lightweight and efficient model based on deep MLP, specifically adapted to operate in IoMT environments with limited computational resources; the exclusive use of physiological signals, deliberately discarding multimodal data sources to enhance real applicability in wearable devices; and the application of the SHAP method as a model explainability technique, enabling analysis of the individual impact of each variable on the prediction and reinforcing system transparency.

This article is organized as follows. Section 2 reviews related work on stress monitoring using IoMT-based intelligent systems. Section 3 describes the dataset used, the objectives of the proposal, and the preprocessing and architecture of the developed model. Next, Sect. 4 presents the details of the experimentation and the analysis of the results obtained. Finally, Sect. 5 presents the conclusions of the work and proposes possible lines of future work.

2 Related Works

Elevated levels of stress have become a widespread public health problem, with significant effects on both mental and physical health. Even if a person appears to be physically healthy, undetected internal stress can trigger chronic disorders and seriously affect emotional stability [25]. This problem has driven AI and IoMT-based solutions that enable continuous monitoring of physiological parameters to objectively assess stress levels.

Recent advances in IoMT have revolutionized health monitoring, especially through the use of wearable devices that enable remote, real-time recording of critical physiological variables [17]. These sensor-enabled devices, integrated with IoMT capabilities, enable the collection of critical data that can be analyzed to identify signs of stress in the body [1,22].

Currently, one of the main challenges in IoMT-based stress monitoring systems is to develop solutions that adapt to end devices with limited resources in terms of processing power, memory, and energy consumption [8]. Despite these technical limitations, the potential of this technology to provide affordable and accessible health monitoring, as well as to promote well-being and disease prevention, makes it a field with great potential for future innovation [1].

In order to assess a person's stress level, IoMT-based systems focus on detecting variations in physiological signals, which tend to change during episodes of stress. These signals provide objective information that, after being analyzed by algorithms, can be used to determine the person's stress level. Among the most commonly used variables are galvanic skin response (GSR), heart rate (HR), skin temperature, and breathing patterns.

GSR is one of the most widely used indicators in stress measurement, as it directly reflects the activity of the nervous system and the degree of emotional arousal of the person being monitored [7,16,19]. HR and HR variability, obtained using photoplethysmography (PPG) or electrocardiography (ECG) sensors, are

also commonly incorporated, as they provide a clear insight into how the cardiovascular system responds to stressful situations.

On the other hand, variations in skin temperature and breathing patterns, registered by sensors located in the chest or abdominal area, or even blood oxygen saturation (SpO2) together with the basic physiological signals mentioned above, create a more complete framework for stress monitoring [13,23].

However, some of the physiological variables mentioned above require sensors that can be intrusive or uncomfortable in everyday use. Therefore, despite the robustness of many existing models, it is essential to move towards solutions that are discreet, accessible, and applicable in a variety of contexts, allowing continuous stress monitoring without interfering with people's daily activities.

Regarding the methodologies used for stress detection in recent years, machine learning algorithms have been fundamental to the development of these systems [18]. With the emergence of deep learning-based techniques [6], more approaches have been designed. Highly used datasets such as WASED [20] and SWELL-KW [10] have been used for training and testing these types of methodologies.

However, the solution to this problem is still in constant development, and some researchers propose what an ideal IoMT-based system for stress or epileptic attack detection would look like [5].

3 Materials and Methods

In this section, the methodological approach followed in the development of the proposed system for stress detection based on physiological signals is detailed. First, the dataset used is described, including its main characteristics. Subsequently, the objective of the proposed model is specified, followed by the preprocessing procedure applied to the dataset. Finally, the architecture of the proposed MLP model is presented, as well as the mechanisms used for its optimization and explainability.

3.1 Purpose

Our main proposal is to design a lightweight prototype that can classify physiological samples and determine whether a person has experienced moderate or high levels of stress over a certain period of time. In addition, we want the model to provide explainable results, an aspect that is increasingly valued in the specialized literature. This feature allows healthcare professionals, in this case, to trust the model and validate its performance.

Although there are studies that distinguish multiple levels of stress, this study employs a binary classification: stressed or not stressed, considering that both medium and high stress, if prolonged, represent a significant risk to a person's health.

3.2 Material

The SWELL-KW dataset [10] was used to test the effectiveness of the proposal. This dataset were collected to facilitate research on stress and is based on an experiment in which 25 participants performed office tasks under normal conditions and also in stressful situations for three hours. Factors such as interruptions via email and time pressure were introduced to induce stress.

The data obtained is diverse and sourced from multimodal sources, including:

- User interactions with the computer: mouse clicks, scrolling, switching between applications, and typing patterns.
- Video recordings of the face and upper body, used to analyze facial activation, head orientation, global facial features, and emotional estimates.
- Body postures recorded using a Kinect depth camera.
- Physiological signals such as HR and its variability, obtained using ECG and skin conductance sensors.

Although all of these data can provide valuable information, physiological signals are particularly useful in real-world settings with IoMT-based systems. Therefore, we have chosen to use a reduced and preprocessed version of the dataset, developed by Nkurikiyeyezu et al. [14]. This version contains 29 features aggregated by one-minute intervals. More details on these variables can be found in [14].

The dataset has a total of 410,322 samples, of which 222,240 correspond to non-stress situations and 188,082 to induced stress conditions.

3.3 Proposal

The proposal is detailed below. The procedure is described in detail, starting with preprocessing, the model architecture, and the subsequent feature filtering and architecture remodeling process.

Preprocessing. First, the initial stage of preprocessing the input data is described. As a reminder, the dataset has already been processed, generating new features derived from two fundamental features: HR and the interval between two peaks of the ECG signal (RR). In addition, these features are aggregated into one-minute periods.

On the other hand, as all are numerical variables, the features are normalized using the Min-Max scaling defined by the following formula.

$$x_{scaled} = \frac{x - x_{\min}}{x_{\max} - x_{\min}}, \quad x_{scaled} \in [0, 1] \subset \mathbb{R}$$

This normalization process is applied once the dataset has been divided into three subsets: training, validation, and testing. Thus, scaling is performed using only the values in the training set, and then the same transformer is applied to the validation and testing sets, ensuring consistency.

The original dataset includes three different classes: non-stressed, stressed by interruption, and stressed by time pressure. In this proposal, both conditions

associated with stress are grouped under a single class, which helps simplify the classification problem and improves the balance between classes.

Architecture. The proposed architecture is based on a deep MLP with a sequential structure, designed to operate in a single processing flow. This choice responds to the need to implement the model on devices with limited computational resources, such as single-board computers or microcontrollers, which are common in IoMT environments. Unlike more complex architectures such as convolutional networks, this configuration significantly reduces computation time, facilitating real-time execution on these types of devices.

The model consists of three consecutive dense blocks, with a decreasing number of neurons per layer, which allows complex patterns to be captured first and the representation to be progressively refined. Each block includes a PReLU activation layer, followed by batch normalization and a regularization layer using *dropout*, whose value is higher in the first block in order to mitigate overfitting in the early stages of learning.

Figure 1 shows the complete architecture.

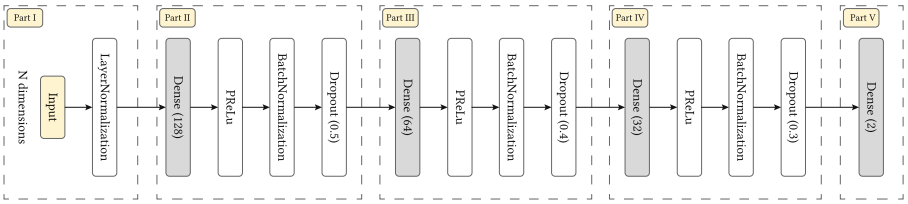


Fig. 1. Architecture of the proposed deep MLP model.

Explainability and Simplification. Finally, in order to provide the model with explicability, the SHapley Additive exPlanations (SHAP) method has been incorporated, an explainability technique based on cooperative game theory. This method allows for the accurate and consistent interpretation of the contribution of each individual feature to the model output by calculating Shapley values that quantify their impact on each prediction. In recent years, SHAP has proven to be particularly effective in complex models such as deep neural networks, as it allows the model's decision to be broken down into components that are understandable to experts, facilitating its validation by professionals and promoting trust in AI-based systems [12].

In this work, SHAP has been used not only for interpretative purposes, but also as an analysis tool to identify the physiological variables with the highest influence on predictions. This information has allowed us to generate a simplified version of the model, optimized to run in environments with limited computational resources, retaining only the most relevant features.

4 Experiments and Results

This section details the experimentation process and the results obtained from the generated models.

4.1 Experimentation

As indicated in the previous section, a reduced version of the SWELL-KW dataset [14] was used for the experimentation phase, focusing exclusively on the physiological signals recorded in the participants. After eliminating the descriptive variables that did not provide useful information for training the model, 27 of the 29 available features were used.

The dataset was randomly divided into three subsets: training (70%), validation (15%), and testing (15%), thus ensuring a balanced distribution of classes in each partition. Next, the model described in Sect. 3.3 was trained. The training process was limited to a maximum of 300 epochs, also applying an early stopping criterion with a patience of ten epochs on the accuracy value in the validation set, in order to avoid overfitting and favor the model's generalization performance.

Once the model was generated, the SHAP values were calculated to estimate the relative importance of each feature. Based on these results, the nine variables with the greatest impact on the predictions were selected, and a reduced subset of the dataset was constructed. The model was then re-trained using only these relevant features in order to validate a simplified version of the model, reducing both the complexity of the architecture and the dimensionality of the input data.

Table 1 below shows all the results obtained, and Fig. 2 shows the most relevant features when applying SHAP.

Table 1. Performance of the model using the full and optimized feature sets. Abbreviations: Ep. (Epochs), Train. (Training), Val. (Validation), Acc. (Accuracy), Prec. (Precision), Rec. (Recall) and F1 (F1-score).

Model configuration	Ep.	Train.	Val.	Test			
		Acc.	Acc.	Acc.	Prec.	Recall	F1
Full dataset (27 features)	58	0.9792	**0.9991**	**0.9995**	**0.9995**	**0.9995**	**0.9995**
Optimized dataset (9 features)	101	0.9382	0.9886	0.9887	0.9886	0.9886	0.9886

4.2 Results

The results obtained during the experimental phase demonstrate the effectiveness of the proposed model in both its complete and optimized versions. First, the model trained with the 27 features from the SWELL-KW dataset achieved



Fig. 2. Feature importance values based on SHAP analysis. The top ten variables with the greatest impact on model predictions are shown.

perfect accuracy on the test set and very high values on the training and validation subsets, indicating adequate generalization ability with no significant evidence of overfitting.

On the other hand, the simplified model, based solely on the nine features selected through SHAP analysis, also performed well, with an accuracy of 0.99 on the validation and test sets and 0.94 on the training set. The slight reduction in training set accuracy, compared to validation or test accuracy, may reflect the effect of regularization and early stopping, which contribute to improved generalization and help prevent overfitting.

In addition, all complementary evaluation metrics, such as accuracy, recall, and F1-score, achieved values of 1.0 or 0.99, reinforcing the robustness of the proposed approach. This behavior suggests that the reduced subset preserves most of the discriminative information necessary to detect stress states, which is especially valuable in contexts where computational resources are limited, such as in IoMT environments.

5 Conclusions and Future Work

The experimental results show that the proposed model, both in its full and optimized versions, performs excellently, with precision, completeness, and F1-score values above 99%. These findings demonstrate that it is possible to design deep learning models that maintain high performance even under resource constraints, which is key for their integration into IoMT devices aimed at continuous health status monitoring.

On the other hand, variable selection using SHAP has not only resulted in a lighter and more efficient architecture, but has also contributed to the inter-

pretability of the model, providing healthcare professionals with a reliable, transparent, and viable tool for informed decision-making.

Future lines of research include extending the evaluation of the model in real-world settings using data collected from wearable sensors in real time, which would allow its performance to be validated under uncontrolled conditions. The possibility of incorporating incremental or federated learning techniques is also being considered, which would allow the model to be adapted to the end user without compromising privacy or requiring intensive centralized training.

References

1. Al-Atawi, A.A., et al.: Stress monitoring using machine learning, IOT and wearable sensors. Sensors **23**(21), 8875 (2023)
2. Alawage, H.M.A., Zaidi, U.: Traumatic stress, psychological well-being, and sociodemographic correlates in high-stress environments among healthcare professionals. Open Psychology J. **17**(1) (2024)
3. Bajwa, J., Munir, U., Nori, A., Williams, B.: Artificial intelligence in healthcare: transforming the practice of medicine. Future Healthc. J. **8**(2), e188–e194 (2021)
4. Belbase, P., Bhusal, R., Ghimire, S.S., Sharma, S., Banskota, B.: Assuring assistance to healthcare and medicine: internet of things, artificial intelligence, and artificial intelligence of things. Front. Artif. Intell. **7** (2024)
5. Bhatt, M.W., Sharma, S.: An IOMT-based approach for real-time monitoring using wearable neuro-sensors. J. Healthc. Eng. **2023**(1) (Jan 2023)
6. Gupta, A., Kansal, D., Gupta, V., Shetty, M.K., Girish, M.P., Gupta, M.D.: X-ecgnet: an interpretable dl model for stress detection using ECG in covid-19 healthcare workers. In: 2021 4th International Conference on Bio-Engineering for Smart Technologies (BioSMART). IEEE (2021)
7. Hadhri, S., Hadiji, M., Labidi, W.: An intelligent stress detection and monitoring system using the IOT environment. In: 2022 International Conference on Technology Innovations for Healthcare (ICTIH), pp. 20–25. IEEE (2022)
8. Jiang, S., Firouzi, F., Chakrabarty, K., Elbogen, E.B.: A resilient and hierarchical IOT-based solution for stress monitoring in everyday settings. IEEE Internet Things J. **9**(12), 10224–10243 (2022)
9. Johnson, K.B., et al.: Precision medicine, ai, and the future of personalized health care. Clin. Transl. Sci. **14**(1), 86–93 (2020)
10. Koldijk, S., Sappelli, M., Verberne, S., Neerincx, M.A., Kraaij, W.: The swell knowledge work dataset for stress and user modeling research. In: Proceedings of the 16th International Conference on Multimodal Interaction. ICMI '14, ACM (2014)
11. Lu, Z.x., et al.: Application of ai and IOT in clinical medicine: summary and challenges. Current Med. Sci. **41**(6), 1134–1150 (2021)
12. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. In: Proceedings of the 31st International Conference on Neural Information Processing Systems, pp. 4768–4777. NIPS'17, Curran Associates Inc., Red Hook, NY, USA (2017)
13. Mozafari, M., Firouzi, F., Farahani, B.: Towards IOT-enabled multimodal mental stress monitoring. In: 2020 International Conference on Omni-layer Intelligent Systems (COINS), pp. 1–8. IEEE (2020)

14. Nkurikiyeyezu, K., Shoji, K., Yokokubo, A., Lopez, G.: Thermal comfort and stress recognition in office environment. In: Proceedings of the 12th International Joint Conference on Biomedical Engineering Systems and Technologies. SCITEPRESS - Science and Technology Publications (2019)
15. Pise, A.A., et al.: Enabling artificial intelligence of things (AIOT) healthcare architectures and listing security issues. Comput. Intell. Neurosci. **2022**, 1–14 (2022)
16. Raju, A.R., Ramadevi, R., Babu, P.R., D.V.: Galvanic skin response based stress detection system using machine learning and IOT. In: 2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), pp. 709–714. IEEE (2023)
17. Rashid, N., Mortlock, T., Faruque, M.A.A.: Stress detection using context-aware sensor fusion from wearable devices. IEEE Internet Things J. **10**(16), 14114–14127 (2023)
18. Rawat, S.K.: Stress detection using machine learning and deep learning. Int. J. Res. Appl. Sci. Eng. Technol. **12**(12), 1399–1405 (2024)
19. Scherz, W.D., Baun, J., Seepold, R., Madrid, N.M., Ortega, J.A.: A portable ECG for recording and flexible development of algorithms and stress detection. Proc. Comput. Sci. **176**, 2886–2893 (2020)
20. Schmidt, P., Reiss, A., Duerichen, R., Marberger, C., Van Laerhoven, K.: Introducing wesad, a multimodal dataset for wearable stress and affect detection. In: Proceedings of the 20th ACM International Conference on Multimodal Interaction, pp. 400–408. ICMI '18, ACM (2018)
21. Smith, M.D., Wesselbaum, D.: Global evidence on the prevalence of and risk factors associated with stress. J. Affect. Disord. **374**, 179–183 (2025)
22. Sun, W., et al.: A review of recent advances in vital signals monitoring of sports and health via flexible wearable sensors. Sensors **22**(20), 7784 (2022)
23. Tooolkar, S., Kampali, A., Bhagwat, S., Katre, K., Tahalyani, J.: An IOT-based stress level detection system for emergency medical attention personnel. In: 2024 International Conference on Intelligent Systems and Advanced Applications (ICISAA), pp. 1–4. IEEE (Oct 2024)
24. World health organization: stress (2023). https://www.who.int/en/news-room/questions-and-answers/item/stress. Accessed 30 Oct 2025
25. Yaribeygi, H., Panahi, Y., Sahraei, H., Johnston, T.P., Sahebkar, A.: The impact of stress on body function: a review. EXCLI Journal (2017)