


Human Affects Classification During Human-Robot Interaction: Insights from Physiological Data



Wael Hasan Ali Almohammed^{1*}, Hussein Ibrahim Hussein^{2,3}, Sinan Adnan Muhisn⁴, Zahraa Abed Aljasim Muhisn⁵

¹ Department of Computer Science, College of Computer Science and Information Technology, University of Kerbala, Kerbala 56001, Iraq

² Department of Computer Engineering Techniques, Alsafwa University College, Kerbala 56001, Iraq

³ Department of Information Security, College of Information Technology, University of Babylon, Babylon 51001, Iraq

⁴ College of Biotechnology, Al-Qasim Green University, Babylon 51001, Iraq

⁵ Computer Center, Al-Qasim Green University, Babylon 51001, Iraq

Corresponding Author Email: wael.h@uokerbala.edu.iq

Copyright: ©2026 The authors. This article is published by IETA and is licensed under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>).

<https://doi.org/10.18280/isi.310211>

ABSTRACT

Received: 19 September 2025

Revised: 1 December 2025

Accepted: 20 February 2026

Available online: 28 February 2026

Keywords:

emotion classification, human-robot interaction, physiological signals, Neural Network, AFFECT-HRI dataset, time-domain features

Affects are something, mostly, everybody feels and expresses daily. Nevertheless, human affects are complex psychological and physiological phenomenon. that are expressed by speech, facial expression, and body gestures. In human-robot interaction (HRI), understanding these emotional cues is crucial for formulating appropriate robotic behavior, enhancing the quality of interaction. In some circumstances, humans intentionally mask their emotions, making it difficult to recognize their true feelings through conventional means. In contrast, physiological signals are trustworthy data that cannot be simulated or manipulated by humans. It responds immediately to changes in emotions; therefore, it generates more precise and reliable emotional results. Also, physiological activations (e.g., heartbeat and skin temperature) were recognized as closely associated with emotions. Hence, different works have studied emotion understanding from physiological datasets. Yet, there is ongoing work to improve accuracy. Also, there are limited datasets in this regard, especially for HRI. Therefore, this study proposes a neural network-based model for emotion classification during HRI. AFFECT-HRI dataset has been utilized for training and testing the proposed model. Accordingly, time domain features were extracted and the dataset went through pre-processing steps to simplify data, and enhance emotion detectability and reliability of recognition. Several evaluation metrics have been used for understanding model performance. The empirical result demonstrated that the proposed network reached a high accuracy of 89.04%. This work contributes to advancing emotion classification in HRI, offering a more reliable and efficient approach by leveraging physiological data. It establishes a competitive benchmark for affective computing in real-time robotic interactions and provides a foundation for further improvements in emotion recognition systems.

1. INTRODUCTION

Human emotion plays an essential role in human-human interpersonal interaction (HHI). Accordingly, affective computing and personality are key factors in human-centered computing and human-machine interaction fields. In particular, improvement of machines' ability to recognize and understand human affect and other individualities contributes to an impactful and successful human-machine interaction [1, 2]. To emphasize, close and extensive interaction with robots may result in a stressful environment, which may impair task achievement. However, social presence in a robot is the key to making individuals feel that they are communicating with another social entity. Therefore, humans are anticipated to engage in emotionally intelligent human-robot interaction (HRI), which means robots can understand humans' affective cues

and behave correspondingly. Typically, this is related to the ability of robots to perceive, recognize, interpret, and mimic human emotional states [3, 4].

In addition, such features allow robots to formulate their behavior in accordance with the interaction context, which increases the transparency of the interaction. Affects recognition promotes a convenient and emotional HRI. Particularly, social robots shall simulate humans' capabilities by understanding the emotions of interaction's parties to enforce the communication. An automated emotion recognition feature is a crucial advantage toward naturalistic and smooth interaction. Hence, emotion recognition has gained remarkable significance as an advanced human interface technology [5]. Human emotions are a complicated psychological and physiological phenomenon that has a part in everyday activities. Emotions are expressed by different

cues such as body gestures, facial expressions, and audio tones. Nevertheless, humans might conceal their emotions deliberately due to many different uncomfortable situations (fear, judgment, etc.), which is so-called social masking. On the other hand, physiological data are involuntary, so they are not subjected to human control, which makes them difficult to manipulate consciously [6, 7].

Therefore, physiological signals provide an accurate and reliable features for inner emotional states. As a result, physiological emotion recognition would offer an effective and precise understanding in real-time interaction. Thus, it has garnered much attention in HRI and affective computing lately by the advancement in physiological data sensors. Additionally, this trend has thrived with the accelerated evolution in touch-based technologies such as wearable sensors, which become convenient and attainable [8-10]. Consequently, such devices have been developed to read the physiological changes in a human's body. The collected data involves machine learning and deep learning to analyze the patterns and understand the inner emotional states. Recently, there are several works that employed machine learning and deep learning algorithms for emotion recognition research [11-14]. However, improvement of recognition accuracy is an ongoing interest that would never be wane.

Besides accuracy, which is a substantial concern, there is a notable lack of physiological datasets, whereas most studies rely on two public datasets, SEED and DEAP. Likewise, different datasets are available for human-human and human-computer interaction, but often overlook emotions in HRI [9, 15-18]. On other hand, a limited work focused on multi-modal emotion recognition that combined both physiological and identity data in real-time. Current works utilized datasets that depended on physiological data for the emotion recognition only. While, the research has begun to draw attention around metadata and personal information. Lately, there is an open argument toward using such information as additional input to enforce the performance of recognition [19].

Furthermore, this study seeks the answer of following research questions. Firstly, how to develop a model to improve emotion recognition accuracy during HRI using physiological data. Secondly, what is an alternative dataset to be utilized to train and evaluate the propose model. Finally, how to involve personal information and metadata to enhance recognition accuracy. Accordingly, this study addresses these issues by leveraging a neural network to develop classification model optimized for understanding human emotions in real-time HRI. However, features of physiological data cannot be generalized to the big dataset [20]. Hence, we use a fully connected supervised deep learning approach to discover the latent features. Additionally, we employed a recently released AFFECT-HRI physiological dataset which combined emotional data and personal and metadata for training and testing our model.

2. BACKGROUND AND RELATED STUDIES

In this section, we discuss the key concepts and background of related topics. We also present related works on emotion recognition using machine learning ML and deep learning DL.

2.1 Affective computing in human-robot interaction

Affective computing is aiming to identify human emotional

states and adapt computational systems to these states [21]. In applied contexts, the desired is to build intelligent system that can distinguish and understand people's affects, meanwhile, making sensitive and friendly responses promptly [22]. For example, emotion recognition can be adopted in virous areas like public safety, human-computer interaction, and retailing.

Emotions and feelings are essential constituents of which human intelligence consists. However, they deeply influence human behavior and decisions, unlike the long-held thoughts of considering logical reasoning and rationality only. Typically, emotion is profoundly intertwined with cognitive processes, whereas generating crucial evaluative data could be used to reorient behaviors towards the ultimate goal [23]. Consequently, emotion has been demonstrated its key role in manipulating human responses. Emotional states took a significant position in human-human interaction. Moreover, there is a remarkable rise in HRI applications. Hence, it is mandatory to recognize and understand human emotions and these technologies should understand human emotions in accordance to the interaction scenario. Thus, the affective computing research field has emerged and become an area of interest [3, 24].

As HRI grows lately, robots are required to identify human preferences conveyed by emotion, so they can assist in human decisions and improve user experience [25]. As a subfield of affective computing, affects recognition research intends to grant a better understanding of human emotional signals in HRI. Recent works have demonstrated, indeed, the potential of using affective robots in diverse sectors [26]. So, numerous researchers trust that emotion recognition is the way to promote and advance the development of human emotion understanding and human intelligence during HRI.

Emotion classification models are still in debate, yet in terms of affective computing, emotions are categorized into two main categories, discrete and continuous. Discrete emotion states include the main six emotions: anger, fear, disgust, sadness, happiness, and surprise and some other intermediate states in between. While, continuous states, as depicted in Figure 1, could be categorized as valence (pleasure), arousal, and dominance. At first, valence describes emotion, whether positive or negative. Second, arousal represents the activity level: active (excited) or passive (calm). At last, dominance indicates the level of control, including high dominance (pride) or low dominance (fear) [27, 28]. Emotions could be mainly recognized and understood by their physical signals, cognitive signals, or physiological signals. These models have been devoted to emotion recognition purposes individually, where it is called unimodal emotion recognition, or collectively as multimodal [29, 30].

Formerly, visual data was the dominant model for recognition since it is the primary form of human expression, accessible, and perceivable in terms of tools and users. As well as, they have demonstrated a remarkable effectiveness in terms of recognition accuracy [7, 32]. However, humans might suppress or mask their physical emotions deliberately due to several reasons during interaction. So, it is very complicated to perceive and understand humans' internal emotions. Therefore, cognitive and physiological signals have begun to be employed pervasively, as they expose more reliable results. Humans have no self-control over changes in these signals [33]. Accordingly, physiological signals offer subtle emotion recognition results regardless of whether they are generated and expressed explicitly or implicitly. Even so, such signals are limited in practical applications due to the difficulty of

collecting them from humans [34].

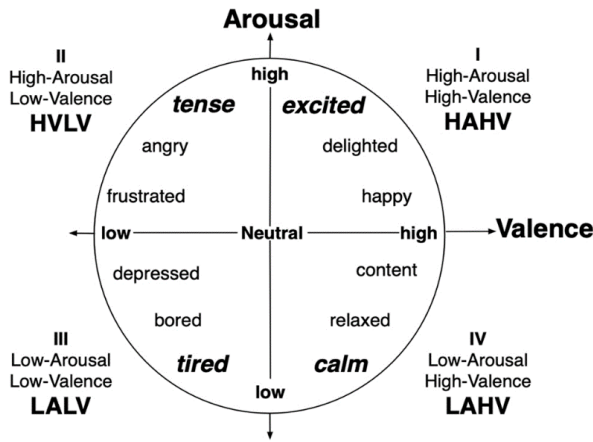


Figure 1. Continuous emotion (Valence-Arousal) models adopted from [31]

The process of physiological data acquisition requires attaching a specialized body-worn device to bodies, which is surely intrusive and discomfort, especially for children. Moreover, it is inapplicable in public and non-restricted HRI environments. Due to prompt advancement of physical-touching technologies such as wearable devices, there is a notable rise in attention to the use physiological data for emotion recognition. Wearable devices have built-in biosensors to acquire fine-grained physiological signals in real-time, which provide advanced understanding for complex emotions in different scenarios [35, 36].

Following that, ML and DL models have a pivotal role in recognition due to their ability to process multimodal, high-dimensional, and complex behavioral and physiological data. They are utilized to pre-process, extract features, and eventually capture the complicated patterns of human emotion states. These features might not be generalizable implicitly for big datasets; therefore, supervised models are one of the methods to overcome such difficulty [20]. Several ML and DL models have been used for emotion recognition. Yet, each has its own strengths and drawbacks; hence, each study commonly designs recognition model in accordance to the characteristics of the data and task [37]. In this study we employed deep neural network for emotion classification. In next section, we are going to discuss some previous related studies. Then compare the archived results of this study with the existing results and highlight the contributions.

2.2 Related studies

Previous studies have concentrated on verbal and visual emotion recognition [32]. With the latest developments of physiological signal technologies, researchers tend to utilize this type of data for recognition approaches involving ML and DL models [38, 39]. Regardless of their promising results, there is room for improvement of accuracy.

Firstly, Zhang and his associates have developed an emotional recognition model based on Deep Siamese Networks (EmoDSN). The model has been trained and tested on three datasets that were collected by the researchers. It has the ability to recognize fine-grained valence and arousal of five different emotional models. The model has the advantage of learning on a small amount of data; despite that, the experiments achieved a relatively low average recognition

accuracy of 76.62%. Secondly, the work conducted by Li and other researchers utilized a couple of public datasets, which are DEAP and SEED, to train and test Electroencephalography (EEG)-based emotion recognition using an ensemble learning model. The experiments demonstrated the effective classification rate for arousal and valence emotional models reached 84.44%. The proposed method has the ability to convert a recognition single model from discrete to continuous value for a better recognition result. Thirdly, the researchers employed the same dataset of previous work with spatio-temporal and self-adaptive Graph Convolutional Networks (GCNs) to classify emotional states into valence and arousal classes. The experiments showed an outperforming result, reaching 85.65% in terms of accuracy.

Fourthly, a complex-valued neural network has been proposed for a real-time recognition system by Ean-Gyu Han and his colleagues. The system used Photoplethysmography and Electromyography signals from real-time data. Results showed accuracy of 81.78%. However, the proposed system has some drawbacks, like the experiments conducted in closed laboratory settings that might not grab a wide variety of emotions. As well as, the findings cannot be generalized to a wider scale of population. Another study introduced a hybrid deep learning framework that utilized Gated Recurrent Unit (GRU) and Convolutional Neural Networks (CNN) algorithms for emotion recognition tasks. Two public datasets were used to train and test the proposed framework, and the results showed an improvement by achieving an accuracy rate of 87.04%. Table 1 summarized some of recent related works.

Table 1. Summary of related studies

Year	Ref.	Models	Emotions	Acc (%)
2022	[40]	Deep Siamese Networks	Two-dimensional 5-class classification	76.62
2022	[41]	Ensemble Learning	Arousal and Valence	84.44
2022	[17]	GCN network	Arousal and Valence	85.65
2023	[42]	Complex-valued Neural Network	Arousal and Valence.	81.78
2023	[43]	Hybrid GRU and CNN	Arousal and Valence	87.04

3. EXPERIMENTAL METHOD

3.1 Research framework

In this section, we specifically focus on the conceptual framework of developing deep neural network for affects classification based on physiological signals during HRI. It consists of integrated phases that have been designed to handle the recognition issue. The framework starts with collecting the raw signals data from its resources. Then, the raw data passes through feature extraction phase to extract the meaningful description, which may depict associated patterns. Later on, the data of extracted features undergoes to several pre-processing steps to clean data, encoding, scaling, and so on in order to make the network's performance robust. Next, the network has been carefully crafted to be trained on part of dataset. Eventually, the network is tested on an undiscovered part of the dataset to measure network performance. Several

evaluation metrics have been used including recall, precision, and accuracy, and later highlighting the network’s efficiency in overcoming the recognition of complex patterns. The whole research framework is depicted in Figure 2.

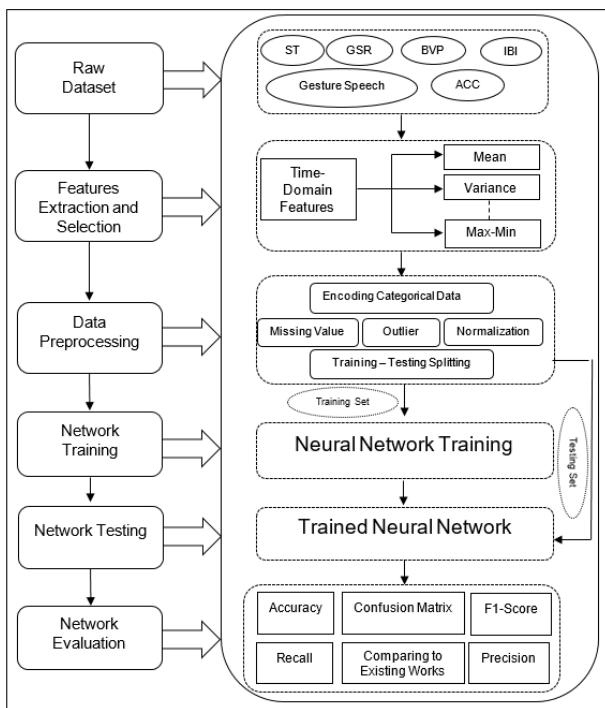


Figure 2. Research framework

3.2 Dataset description

The dataset for this work serves as a base for the network to learn and analyze the emotional states and patterns. AFFECT-HRI, which is a comprehensive physiological dataset that was lately published in 2024, has been selected for understanding the temporal emotional dynamics during HRI. The experiments for data collection have been conducted in a laboratory equipped with an anthropomorphic service robot and an Android robot. It covered five different conditions: moral, immoral, neutral, liability, and transparency. The dataset has been collected from 146 participants. It consists of demographic information, physiological data collected by sensors, which are self-human affect assessments, self-assessment questionnaire data, robot speech, and robot gestures.

Table 2. Dataset description

Feature	Details
ACC	Acceleration
BVP	Blood Volume Pulse
EDA	Electrodermal Activity
IBI	Inter-Beat Interval
ST	Skin Temperature
Label	Human Affects

However, physiological data is the key for studying the emotional classification and understanding. It has been collected by using wristband smartwatches in terms of weight and comfort that offer us continuous and imperceptible measurement for such data. It contains data of electrical variation in the skin, blood volume pulse, inter-beat interval, peripheral skin temperature, and physical movement as

described in Table 2. The data has been labeled using Valence-Arousal models by five different affective states (HNVHA/Stressed, HPVHA/Happy, HNVLA/Depressed, HPVLA/Relaxed, and Neutral) based on the questionnaire result before and after the interaction [44].

3.3 Features extraction and selection

The features extraction phase concerns choosing relevant features for the emotion recognition model. As Figure 2 depicted, a set of active and informative features have been extracted integrated with the raw dataset to contribute to the analysis and classification process. At first, raw data of multimodal physiological have been preprocessed to conserve the temporal nature. Then, additional features were extracted based on time-domain to introduce compact representations of short-term dynamics. These features represent the fluctuation in the sensors’ reading over time to understand the emotional state changes during interaction.

There are several approaches to analyzing physiological data in the time domain. Numerous features have been extracted such as mean, standard deviation, power, maximum, minimum, median, skewness, relative band energy, variance, and so on. The proposed network has been trained utilizing a hybrid strategy, where raw physiological data and their corresponding extracted features were jointly used for training purpose. Such design offers the proposed network a way of learning from raw data the low-level temporal patterns directly, concurrently it benefits from higher-level statistical descriptors.

Furthermore, selection of features is essential step to improve the interpretability and network performance. However, such features have been extracted and calculated to a one-dimensional vector and later used as input data. Besides, Principal Component Analysis (PCA) was utilized to eliminate the unrelated features, reduce feature dimensionality, and select the pertinent features for emotion recognition purpose. It transformed the original features set into a smaller set that support noise reduction and enhance learning efficiency. Particularly, the following formulas are example that have been performed for feature extraction [39]:

Power:

$$Px = \frac{1}{T} \sum_{t=1}^T |x(t)|^2 \quad (1)$$

Mean:

$$Mx = \frac{1}{T} \sum_{t=1}^T x(t) \quad (2)$$

Standard deviation:

$$\sqrt{\frac{1}{T-1} \sum_{t=1}^T (x(t) - m_x)^2} \quad (3)$$

3.4 Data pre-processing

Pre-processing phase is crucial step for multimodal affect classification in order to prepare the data for efficient training results in a deep neural network. We have applied different pre-processing steps to secure suitability and consistency of

the dataset to reach the ultimate network’s performance. To discuss some, the null values were addressed and manipulated using imputation by mean method to avoid the network’s bias. In the same context, the dataset underwent the process of outlier identification and elimination, so, there was neutralization of outliers that may distort network performance. In addition, categorical data was converted into numerical values that fit the computational model.

Thereafter, min–max normalization process is applied for scaling the numerical values. It ensures that the dataset is compatible and effect equally during training, which is important to the neural network’s best result. Eventually, dataset is divided into training and testing sets to train the neural network and evaluate it with unseen data. This phase ensures effective learning and improves network classification accuracy. Furthermore, due to overcome the issue of classes imbalance during training phase, a class-weighted loss function has been utilized. Each class weight is calculated as the inverse frequency of each class in training dataset. This strategy promotes the proposed network to penalize misclassification of minor classes more strongly while maintaining overall convergence stability.

3.5 Network architecture and evaluation

Generally, this study proposed a fully connected deep neural network to capture the temporal nature of the physiological signals. It used statistical summarization and short-time window segmentation instead of utilizing other explicitly temporal network architectures. The design was dictated by the target real-time HRI applications, where computational efficiency, architectural simplicity, and low-latency inference. So, this architecture offers stable behavior and alleviate computational overhead, making it appropriate for interactive system under limited conditions. Also, the design overcomes the issue related to small and complex affective datasets in HRI and achieve accurate multi-class prediction.

Indeed, the input layer starts with node for each feature of the dataset. Afterward, multiple hidden layers are set for figuring out complicated patterns and connections embedded in the affective dataset. The number of nodes and activation functions for hidden layers were chosen based on empirical trials to achieve balance in network recognition accuracy and computational effectiveness. Therefore, each hidden layer employed LeakyReLU activation function in order to deal with mitigate vanishing gradient and non-linearity problems.

On the other hand, the output layer employed softmax activation function to fulfill multi-class classification. The network concludes in five output nodes representing emotional states in the dataset. Dropout regularization has been used resolve overfitting issue which may raise by fully connected architecture, as well as, L2 weight regularization was included into loss function. AdamW optimizer selected to train the network with learning rate 0.0015. Overall, the architecture has been carefully crafted to guarantee the performance and accuracy as detailed parameters and hyperparameters listed in Table 3. The overview of the network’s architecture is depicted in Figure 3 below.

Later on, the evaluation phase has been conducted by utilizing different evaluation metrics to verify the network’s efficiency. However, the evaluation process is another primary phase in deep learning application to assess the effectiveness of the network with an unseen dataset. As earlier stated, the

dataset is divided into training and testing portions in 80:20 percent as empirically validated. Thus, the network was trained iteratively for 100 epochs to reach the most optimal parameters and reduce the loss function. However, results mainly reflect within session–level generalization. Consequently, it does not promote subject independence and it may allow these patterns to be learned partially.

Table 3. Training parameters and hyperparameters details

Parameter / Hyperparameter	Details
Architecture	Fully Connected
Activation functions	LeakyReLU and Softmax
Regularization	Dropout & L2
Optimizer	AdamW
Learning rate	0.0015
Early stopping	Yes (Patience = 5)
Epochs	100
Batch size	20

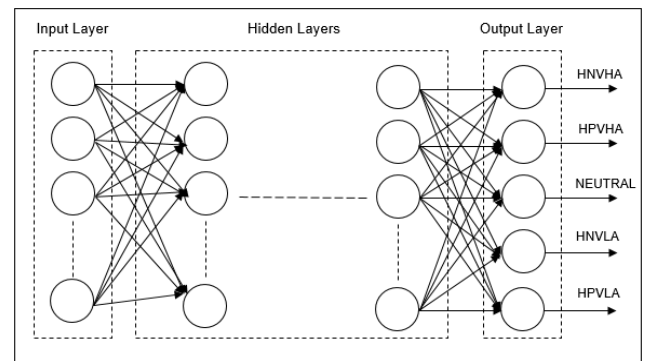


Figure 3. Architecture overview of proposed neural network

However, accuracy was calculated as the key metric for assessing the correct recognition percent. Furthermore, the confusion matrix (true positives, false positives, true negatives, and false negatives of each state) is used to obtain detailed insight about the network’s performance and highlight network underperformance. Additionally, F1-score, precision, and recall have been used to evaluate the network’s performance to avoid any misleads in accuracy evaluation. Lastly, the result of this study is compared to the result of some existing studies. Collectively, they provide a comprehensive view on the network’s performance.

4. RESULT AND DISCUSSION

An affect classification pipeline using multimodal physiological data was designed, including two main representations, which are raw data and extracted features, followed by preprocessing. Such design is employed to train DNN classifier for multi-class classification and enables the network to learn complicated relationships among multiple physiological modalities. The results of the proposed neural network model demonstrated an effective performance in classifying human emotional states. Overall, it has achieved a promising accuracy that reached 89.04%. Figure 4 illustrates the model’s ability to recognize among the emotional states across diverse threshold settings, and highlights its overall performance.

Moreover, this study employed a substitutional dataset favorable for affects classification in HRI settings, consisting

of multimodal physiological signals (e.g., BVP, GSR, ACC, and skin temperature). In contrast to other common datasets which are non-HRI scenario, this data has been gathered during interactive and real conditions in HRI sessions. The results proved that the used dataset has adequate affects-related information to support affects classification during interaction. Notably, HRI-related metadata and personal information contain a crucial contextual impact that enhance affects classification performance and may support personalization in HRI. Nonetheless, in this experiment, the metadata was used for contextual encouragement. Hence, while the results demonstrated the feasibility of affects classification based on physiology data, the current findings do not quantify the performance obtain attributable to metadata.

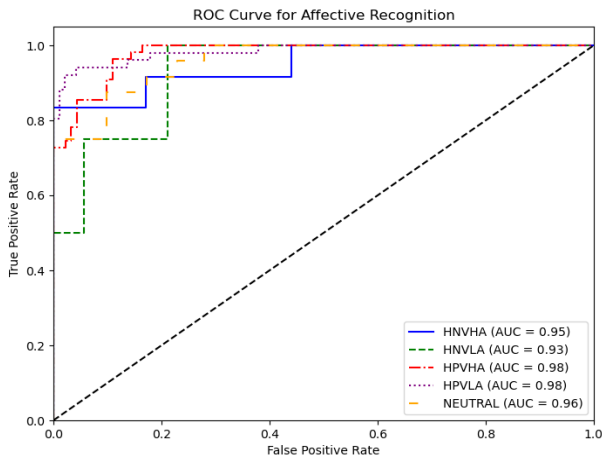


Figure 4. Receiver Operating Characteristic (ROC) curve for proposed model

The precision, recall, and F1-score have proven the consistency and robustness of the proposed model performance over all emotions as shown in Table 4. Moreover, the model reduced false negatives and false positives, which makes it appropriate for real-time HRI. However, a closer examination by using confusion matrices reveals that a certain emotion, NEUTRAL was classified lower compared to other states by overlapping with the HAPPY state. As a comparative analysis with the existing works, the proposed network enhances the classification accuracy. Yet, a minor misclassification has been figured out, indicating the necessity of the model for additional improvement. However, results of existing studies are mentioned to compare with proposed network in contextual reference in regard to the comprehensive performance that witnessed for physiological data affect recognition, instead of head-to-head to the state of art. Figure 5 visualizes the result of the proposed neural network model compared to existing studies, these mentioned in Table 1 above, in terms of accuracy.

Table 4. Results of evaluation metrics

Class	Precision	Recall	F1-Score
HNVHL	95%	91%	93%
HNVLA	90%	93%	91%
HPVHA	91%	94%	92%
HPVLA	86%	89%	87%
NEUTRAL	82%	76%	78%
Overall	89%	89%	88%

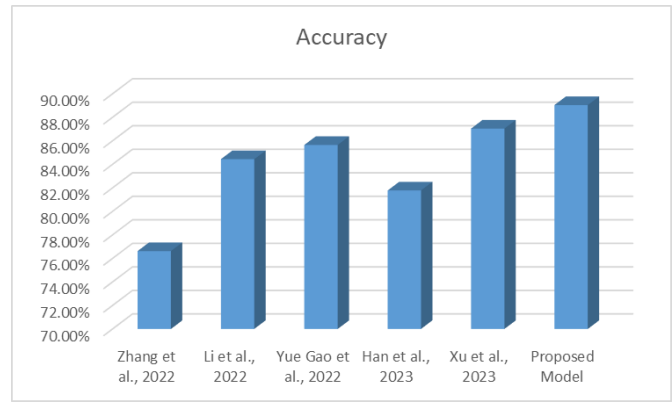


Figure 5. Contextual comparative analysis with some existing works

5. CONCLUSION AND FUTURE DIRECTIONS

In this work, a neural network model has been developed for human affect classification based on physiological data during HRI. It addressed the gap of lack in HRI affects datasets and involved the personal information to personalize the recognition model. However, the AFFECT-HRI dataset has been used to train and test the proposed model. The result demonstrates that the proposed model outperforms the existing works in its performance effectiveness and classification accuracy. Yet, some limitations have been observed, which highlight future improvement directions. The future direction of work will pay attention to enhance model’s generalization by involving a larger dataset with higher diversity. Also, employment for an explicit physiological index for interpretability and will be integrated in future work to additional generalization and explainability. Additionally, enhancing the feature extraction phase and adding multimodal fusion techniques. Finally, real-time deployment and investigation of users’ adaptation techniques for additional improvement in HRI and affective computing fields.

ACKNOWLEDGMENT

The authors would like to thank the Faculty of Computer Science and Information Technology, University of Kerbala, Alsafwa University, Al-Qasim Green University for their support. The authors have no financial support for our research.

REFERENCES

[1] Liu, J.B., Ang, M.C., Chaw, J.K., Kor, A.L., Ng, K.W. (2023). Emotion assessment and application in human–computer interaction interface based on backpropagation neural network and artificial bee colony algorithm. *Expert Systems with Applications*, 232: 120857. <https://doi.org/10.1016/j.eswa.2023.120857>

[2] Zhou, Z.J., Asghar, M.A., Nazir, D., Siddique, K., Shorfuzzaman, M., Mehmood, R.M. (2023). An AI-empowered affect recognition model for healthcare and emotional well-being using physiological signals. *Cluster Computing*, 26: 1253-1266. <https://doi.org/10.1007/s10586-022-03705-0>

- [3] Gervasi, R., Barravecchia, F., Mastrogiacomo, L., Franceschini, F. (2023). Applications of affective computing in human-robot interaction: State-of-art and challenges for manufacturing. *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, 237(6-7): 815-832. <https://doi.org/10.1177/09544054221121888>
- [4] Wirtz, J., Patterson, P.G., Kunz, W.H., Gruber, T., Lu, V.N., Paluch, S., Martins, A. (2018). Brave new world: Service robots in the frontline. *Journal of Service Management*, 29(5): 907-931. <https://doi.org/10.1108/JOSM-04-2018-0119>
- [5] Cárdenas, P., García, J., Begazo, R., Aguilera, A., Dongo, I., Cardinale, Y. (2024). Evaluation of robot emotion expressions for human-robot interaction. *International Journal of Social Robotics*, 16: 2019-2041. <https://doi.org/10.1007/s12369-024-01167-5>
- [6] Ju, X.Y., Li, M., Tian, W.L., Hu, D.W. (2024). EEG-based emotion recognition using a temporal-difference minimizing neural network. *Cognitive Neurodynamics*, 18: 405-416. <https://doi.org/10.1007/s11571-023-10004-w>
- [7] Zhang, J.H., Yin, Z., Chen, P., Nichele, S. (2020). Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Information Fusion*, 59: 103-126. <https://doi.org/10.1016/j.inffus.2020.01.011>
- [8] Sarkar, P., Etemad, A. (2020). Self-supervised ECG representation learning for emotion recognition. *arXiv preprint arXiv:2002.03898*. <https://doi.org/10.48550/arXiv.2002.03898>
- [9] Chen, J.L., Lin, X.F., Ma, W.F., Wang, Y.C., Tang, W. (2024). EEG-based emotion recognition for road accidents in a simulated driving environment. *Biomedical Signal Processing and Control*, 87: 105411. <https://doi.org/10.1016/j.bspc.2023.105411>
- [10] Ba, S., Hu, X. (2023). Measuring emotions in education using wearable devices: A systematic review. *Computers & Education*, 200: 104797. <https://doi.org/10.1016/j.compedu.2023.104797>
- [11] Pan, C., Shi, C., Mu, H.L., Li, J., Gao, X.B. (2020). EEG-based emotion recognition using logistic regression with Gaussian kernel and Laplacian prior and investigation of critical frequency bands. *Applied Sciences*, 10(5): 1619. <https://doi.org/10.3390/app10051619>
- [12] Wagh, K.P., Vasanth, K. (2022). Performance evaluation of multi-channel electroencephalogram signal (EEG) based time frequency analysis for human emotion recognition. *Biomedical Signal Processing and Control*, 78: 103966. <https://doi.org/10.1016/j.bspc.2022.103966>
- [13] Joshi, V.M., Ghongade, R.B., Joshi, A.M., Kulkarni, R.V. (2022). Deep BiLSTM neural network model for emotion detection using cross-dataset approach. *Biomedical Signal Processing and Control*, 73: 103407. <https://doi.org/10.1016/j.bspc.2021.103407>
- [14] Chen, L.F., Li, M., Wu, M., Pedrycz, W., Hirota, K. (2024). Coupled multimodal emotional feature analysis based on broad-deep fusion networks in human-robot interaction. *IEEE Transactions on Neural Networks and Learning Systems*, 35(7): 9663-9673. <https://doi.org/10.1109/TNNLS.2023.3236320>
- [15] Iyer, A., Das, S.S., Teotia, R., Maheshwari, S., Sharma, R.R. (2023). CNN and LSTM based ensemble learning for human emotion recognition using EEG recordings. *Multimedia Tools and Applications*, 82: 4883-4896. <https://doi.org/10.1007/s11042-022-12310-7>
- [16] Joshi, V.M., Ghongade, R.B. (2021). EEG based emotion detection using fourth order spectral moment and deep learning. *Biomedical Signal Processing and Control*, 68: 102755. <https://doi.org/10.1016/j.bspc.2021.102755>
- [17] Gao, Y., Fu, X.L., Ouyang, T.X., Wang, Y. (2022). EEG-GCN: Spatio-temporal and self-adaptive graph convolutional networks for single and multi-view EEG-based emotion recognition. *IEEE Signal Processing Letters*, 29: 1574-1578. <https://doi.org/10.1109/LSP.2022.3179946>
- [18] Xu, F.F., Pan, D., Zheng, H.H., Ouyang, Y., Jia, Z., Zeng, H. (2024). EESCN: A novel spiking neural network method for EEG-based emotion recognition. *Computer Methods and Programs in Biomedicine*, 243: 107927. <https://doi.org/10.1016/j.cmpb.2023.107927>
- [19] Kim, H., Hong, T. (2024). Enhancing emotion recognition using multimodal fusion of physiological, environmental, personal data. *Expert Systems with Applications*, 249: 123723. <https://doi.org/10.1016/j.eswa.2024.123723>
- [20] Ross, K., Hungler, P., Etemad, A. (2020). Unsupervised multi-modal representation learning for affective computing with multi-corpus wearable data. *arXiv preprint arXiv:2008.10726*. <https://doi.org/10.48550/arXiv.2008.10726>
- [21] Massaccesi, C., Korb, S., Willeit, M., Quednow, B.B., Silani, G. (2022). Effects of the mu-opioid receptor agonist morphine on facial mimicry and emotion recognition. *Psychoneuroendocrinology*, 142: 105801. <https://doi.org/10.1016/j.psyneuen.2022.105801>
- [22] McColl, D., Hong, A., Hatakeyama, N., Nejat, G., Benhabib, B. (2016). A survey of autonomous human affect detection methods for social robots engaged in natural HRI. *Journal of Intelligent & Robotic Systems*, 82: 101-133. <https://doi.org/10.1007/s10846-015-0259-2>
- [23] Stock-Homburg, R. (2022). Survey of emotions in human-robot interactions: Perspectives from robotic psychology on 20 years of research. *International Journal of Social Robotics*, 14: 389-411. <https://doi.org/10.1007/s12369-021-00778-6>
- [24] Bera, A., Randhavane, T., Prinja, R., Kapsaskis, K., Wang, A., Gray, K., Manocha, D. (2019). The emotionally intelligent robot: Improving social navigation in crowded environments. *arXiv preprint arXiv:1903.03217*. <https://doi.org/10.48550/arXiv.1903.03217>
- [25] Leong, S.C., Tang, Y.M., Lai, C.H., Lee, C.K.M. (2023). Facial expression and body gesture emotion recognition: A systematic review on the use of visual data in affective computing. *Computer Science Review*, 48: 100545. <https://doi.org/10.1016/j.cosrev.2023.100545>
- [26] Spitale, M., Gunes, H. (2023). Affective robotics for wellbeing: A scoping review. *arXiv preprint arXiv:2304.01902*. <https://doi.org/10.48550/arXiv.2304.01902>
- [27] Li, Q., Liu, Y.Q., Yan, F., Zhang, Q., Liu, C. (2023). Emotion recognition based on multiple physiological signals. *Biomedical Signal Processing and Control*, 85: 104989. <https://doi.org/10.1016/j.bspc.2023.104989>
- [28] Matsumoto, D., Wilson, M. (2023). Effects of multiple

- discrete emotions on risk-taking propensity. *Current Psychology*, 42: 15763-15772. <https://doi.org/10.1007/s12144-022-02868-8>
- [29] Shoumy, N.J., Ang, L.M., Seng, K.P., Rahaman, D.M.M., Zia, T. (2020). Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals. *Journal of Network and Computer Applications*, 149: 102447. <https://doi.org/10.1016/j.jnca.2019.102447>
- [30] Gandhi, A., Adhvaryu, K., Poria, S., Cambria, E., Hussain, A. (2023). Multimodal sentiment analysis: A systematic review of history, datasets, multimodal fusion methods, applications, challenges and future directions. *Information Fusion*, 91: 424-444. <https://doi.org/10.1016/j.inffus.2022.09.025>
- [31] Aly, L., Godinho, L., Bota, P., Bernardes, G., da Silva, H.P. (2024). Acting emotions: A comprehensive dataset of elicited emotions. *Scientific Data*, 11: 147. <https://doi.org/10.1038/s41597-024-02957-2>
- [32] A.V., G., T., M., D., P., E., U. (2024). Multimodal emotion recognition with deep learning: Advancements, challenges, and future directions. *Information Fusion*, 105: 102218. <https://doi.org/10.1016/j.inffus.2023.102218>
- [33] Mateos-García, N., Gil-González, A.B., Luis-Reboredo, A., Pérez-Lancho, B. (2023). Driver stress detection from physiological signals by virtual reality simulator. *Electronics*, 12(10): 2179. <https://doi.org/10.3390/electronics12102179>
- [34] Sajjad, M., Nasir, M., Ullah, F.U.M., Muhammad, K., Sangaiah, A.K., Baik, S.W. (2019). Raspberry Pi assisted facial expression recognition framework for smart security in law-enforcement services. *Information Sciences*, 479: 416-431. <https://doi.org/10.1016/j.ins.2018.07.027>
- [35] Liu, Z., Ren, Y.P., Kong, X., Liu, S. (2022). Learning analytics based on wearable devices: A systematic literature review from 2011 to 2021. *Journal of Educational Computing Research*, 60(6): 1514-1557. <https://doi.org/10.1177/07356331211064780>
- [36] Wang, Z.M., Zhou, X.X., Wang, W.L., Liang, C. (2020). Emotion recognition using multimodal deep learning in multiple psychophysiological signals and video. *International Journal of Machine Learning and Cybernetics*, 11: 923-934. <https://doi.org/10.1007/s13042-019-01056-8>
- [37] Pan, B., Hirota, K., Jia, Z.Y., Dai, Y.P. (2023). A review of multimodal emotion recognition from datasets, preprocessing, features, and fusion methods. *Neurocomputing*, 561: 126866. <https://doi.org/10.1016/j.neucom.2023.126866>
- [38] Ghosh, R., Sinha, D. (2024). Human emotion recognition by analyzing facial expressions, heart rate and blogs using deep learning method. *Innovations in Systems and Software Engineering*, 20: 499-507. <https://doi.org/10.1007/s11334-022-00471-5>
- [39] Samal, P., Hashmi, M.F. (2024). Role of machine learning and deep learning techniques in EEG-based BCI emotion recognition system: A review. *Artificial Intelligence Review*, 57: 50. <https://doi.org/10.1007/s10462-023-10690-2>
- [40] Zhang, T., El Ali, A., Hanjalic, A., Cesar, P. (2023). Few-shot learning for fine-grained emotion recognition using physiological signals. *IEEE Transactions on Multimedia*, 25: 3773-3787. <https://doi.org/10.1109/TMM.2022.3165715>
- [41] Li, R., Ren, C., Zhang, X.W., Hu, B. (2022). A novel ensemble learning method using multiple objective particle swarm optimization for subject-independent EEG-based emotion recognition. *Computers in Biology and Medicine*, 140: 105080. <https://doi.org/10.1016/j.compbiomed.2021.105080>
- [42] Han, E.G., Kang, T.K., Lim, M.T. (2023). Physiological signal-based real-time emotion recognition based on exploiting mutual information with physiologically common features. *Electronics*, 12(13): 2933. <https://doi.org/10.3390/electronics12132933>
- [43] Xu, G.X., Guo, W.H., Wang, Y.J. (2023). Subject-independent EEG emotion recognition with hybrid spatio-temporal GRU-Conv architecture. *Medical & Biological Engineering & Computing*, 61: 61-73. <https://doi.org/10.1007/s11517-022-02686-x>
- [44] Heinisch, J.S., Kirchhoff, J., Busch, P., Wendt, J., von Stryk, O., David, K. (2024). Physiological data for affective computing in HRI with anthropomorphic service robots: The AFFECT-HRI data set. *Scientific Data*, 11: 333. <https://doi.org/10.1038/s41597-024-03128-z>