Enhancement CNN based on LSTM for vital sign classification

Mina H. Madhi[1] *, Abbas M. Al-Bakry[2], Alaa Kadhim Farhan[3], El-Sayed M. El-kenawy[4]

[1]Information Institute Postgraduate Student, Iraqi Commission for Computers and Informatics,
Baghdad, Iraq;
[2]University of Information Technology and Communication, Baghdad, Iraq;
[3]University of Technology, Computer Sciences, Baghdad, Iraq;
[4]Department of Communications and Electronics, Delta Higher Institute of Engineering and
Technology, Mansoura 35111, Egypt

Emails: phd202010552@iips.icci.edu.iq; abbasm.albakry@uoitc.edu.iq; 110030@uotechnology.edu.iq ;
skenawy@ieee.org

Abstract

Monitoring vital signs is essential for tracking patient health and detecting changes in their condition. However, in aging cultures with overburdened healthcare staff, accurately and efficiently monitoring vital signs poses a challenge. To address this issue, an autonomous system for vital sign control is proposed, offering improved accuracy, real-time monitoring, alert systems, remote monitoring, and reduced staff labor costs. This paper presents a deep learning architecture using a publicly accessible dataset of 25,494 patients and five numerical characteristics to classify vital signs. A CNN-LSTM model is introduced, outperforming a traditional CNN model in terms of performance, parameter efficiency, and training time. The CNN-LSTM model effectively captures both spatial and temporal features from the input data, resulting in superior representation and improved accuracy compared to the CNN model, which only extracts spatial data. The suggested model achieved a remarkable accuracy of 98%, surpassing previous models. The findings demonstrate the potential of the CNN-LSTM model for early identification of medical issues, enabling prompt actions and enhanced patient outcomes. Overall, this research highlights the significance of implementing an autonomous system for vital sign control in healthcare organizations, offering substantial benefits in patient care and healthcare management.

Keywords: Vital Signs; Healthcare; deep learning; CNN; LSTM;

1. Introduction

In the fast-paced world of today, it is essential to maintain good health. Numerous factors, including a busy lifestyle, rising pollution [1,] and the emergence of epidemic and pandemic illnesses, have all contributed to the poor and unhealthy quality of life that many people are experiencing. The use of technology for disease prevention and control has grown in significance as healthcare professionals and companies struggle with these issues [2]. Monitoring vital signs is an important part of illness management. The significance of monitoring vital signs has grown more and more important with the growth of diseases like COVID-19, chronic diseases, cardiac diseases, and Alzheimer's [3]. The use of technology to monitor vital signs can significantly improve patient outcomes.

Vital signs, including temperature, blood pressure, and heart rate [4], play an important part in the early identification and diagnosis of numerous illnesses. They offer crucial details about a patient's health and may indicate the existence of particular disorders. For instance, a high fever could be a sign of an infection, but low blood pressure might be a sign of shock. Therefore, monitoring a patient's health and taking necessary action in case of changes require the measurement of vital signs. However, manually documenting vital signs can be a

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023

difficult task, particularly in aging cultures where overburdened medical staff find it difficult to keep up with demand. Here is where an automated system for controlling vital signs comes in. Healthcare companies may enhance patient outcomes by automating the vital sign monitoring process. This (a) increases accuracy, (b) offers real-time monitoring, (c) raises alerts in the event of crucial changes, and (d) reduces the burden on healthcare staff. As a response to these issues, smart health monitoring systems have arisen, giving healthcare companies a way to efficiently monitor patients' vital signs [5]. Innovative medical technologies such as the Internet of Medical Things (IoMT) [6] and Smart Health Monitoring (SHM) [3] enable medical professionals to provide patients with a complete and interconnected medical service. These technologies rely on a network of interconnected electronic devices, such as screening and conditioning systems, remote and telemedicine care systems, patient monitoring systems, disease and anomaly detection systems, and medical nursing and rehabilitation systems, which work together to provide a service that is more accurate and efficient than traditional methods. Healthcare professionals can remotely monitor patients and gather vital signs using IoT devices, which enables them to see possible health issues early and take the necessary measures. Healthcare professionals may proactively monitor patients' health thanks to SHM systems. The quality of care given by healthcare professionals might be significantly improved due to this technology. Furthermore, the use of deep learning (DL) [7] and machine learning (ML) technologies in healthcare systems can improve the quality of care for patients, doctors, and other health professionals. These techniques can distinguish between a patient's normal and aberrant data, which can help doctors determine the patient's condition [3]. DL [7] is a sort of ML that allows computers to perform tasks similar to those performed by humans. Its multi-level hierarchical architecture and subsequent information processing stages are what distinguish it. DL can be divided into two main categories: recurrent neural networks (RNNs) and convolutional neural networks (CNNs) [7].

The purpose of this study is to explore the benefits and capabilities of utilizing a hybrid deep learning system for monitoring vital signs in patients. The research is divided into four sections, beginning with a summary of deep learning architecture. The second section examines prior research on using DL and ML for patient monitoring and control. The third section provides a detailed explanation of the dataset used in the study as well as the steps involved in the proposed architecture. The fourth section presents the results of experiments conducted on the proposed models. The conclusion summarizes the main findings of the study, including the potential benefits of using DL techniques for vital sign classification, the effectiveness of the proposed CNN-LSTM model, and future areas of research in this field.

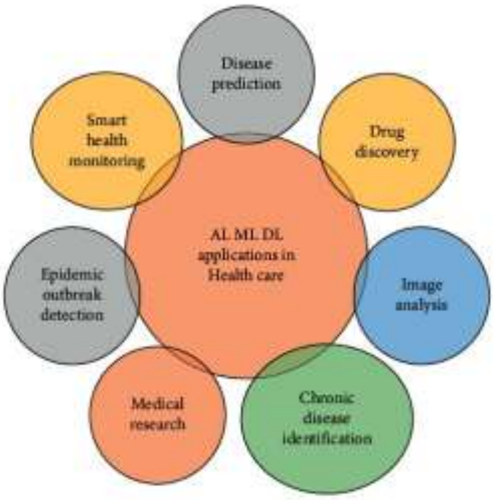


Figure 1: ML and DL applications in the healthcare system [2].

2. Deep Learning

In order to automatically extract characteristics from data, deep learning, a subset of machine learning (ML), focuses on building artificial neural networks with multiple layers [7]. Deep learning models like convolutional neural networks (CNNs), recurrent neural networks (RNNs), and long short-term memory (LSTM) networks [7] can learn complex patterns and representations from raw data, which enables them to achieve high accuracy in tasks like image and speech recognition, natural language processing, and time series forecasting [8]. In contrast to traditional machine learning techniques, which rely on manual feature extraction.

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023

2.1. Convolutional Neural Networks (CNNs)

CNNs are commonly used for handling 2D data, such as images, and consist of a fully connected layer and one or more hidden layers that usually include convolution, pooling, and fully connected layers [7]. Recently, a modified version of 2D CNNs called 1D CNNs has been developed as an alternative [9]. In various scenarios, 1D CNNs have an advantage over their 2D counterparts when dealing with 1D signals. This is because 1D CNNs have (a) lower computational complexity, (b) can handle complex tasks with shallower architectures, and (c) are suitable for real-time and low-cost applications due to their low computational demands [10]. The first layer of a convolutional network, the input layer, takes time-series data as input. The convolution and pooling layers that come next are primarily used for extracting features from the input. Each convolution layer comprises multiple kernels of the same size, followed by a pooling layer that performs either average or max pooling and then sends the output to the prediction or fully connected layer [7], [11].

2.2. Recurrent neural networks (RNNs)

Another common architecture for DL models is the RNN [7]. However, RNNs often encounter the issue of the vanishing gradient problem when working with long-time series data [12]. To address the gradient disappearance through memory cells problem, Hochreiter and Schmidhuber [8] developed LSTM, a variant of RNN. In order to identify time series information through the signals of the earlier time series, it uses a cyclic model that transfers the output of one hidden layer to the same hidden layer [13]. This paradigm is helpful when it's necessary to record long-term dependencies in sequential data. The LSTM consists of three gates (input, forget, and output) that control the flow of information into and out of the memory cells as well as memory cells that serve as information storage. [14] The input gate (1) decides which fresh information will be incorporated as inputs and which information will be used to update the current block state. The forget gate (2) determines which information should be saved or deleted. Which outputs to produce are decided by the output gate (3). The output of these gates, which are governed by learned weights, is utilized to update the state of the cell (6).

Because of this updating mechanism, LSTM is better suited for time-series data because it can keep track of long-term dependencies in the data.

Input gate Γi

$$\Gamma i = \delta(Wi[c\langle t-1\rangle, x\langle t\rangle] + bi) \tag{1}$$

Forget gate Γ_f:

$$\Gamma f = \delta(Wf\ [c\langle t-1\rangle, x\langle t\rangle] + bf \tag{2}$$

Output gate Γ_o:

$$\Gamma o = \delta(Wo[c\langle t-1\rangle, x\langle t\rangle] + bo) \tag{3}$$

Cell: $c_{\langle t\rangle}$

$$c\langle t\rangle = \Gamma i C\langle t\rangle c + \Gamma f\ c\langle t-1\rangle \tag{4}$$

$$Cc\langle t\rangle = tanh(Wc[a\langle t\rangle, x\langle t\rangle] + bc) \tag{5}$$

Output vector $a_{\langle t\rangle}$:

$$a\langle t\rangle = \Gamma o * tanh(c\langle t\rangle) \tag{6}$$

where $x_{\langle t\rangle}$ is the input at current step t, $C_{c\langle t\rangle}$ is the cell state vector, b and W denote the bias and weight matrices, respectively, and δ is an element-wise non-linear activation function.

3. Related Work

A summary of the existing research on the use of DL and ML techniques for human vital sign prediction is presented in this section. The approach described in this work [15] uses channel state information (CSI) to monitor a patient's respiration rate as they sleep without having them wear any wearables. The study explains how to extract useful features from the respiratory data using three different feature extraction approaches and three feature selection algorithms. Four ML classification algorithms (KNN, SVM, DT, and RF) then employ the chosen characteristics to forecast the patient's state of health. The availability of highly detailed and fused data for feature selection is ensured by the use of CSI. Using the relief feature selection technique, the KNN model obtained the best accuracy of 85.12%.

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023

This study [16] discusses the development of an automated system that examines radar readings and finds biological signals using a CNN architecture (ResNet50) that has been trained on ImageNet. The model was fine-tuned by the scientists' using data from real radar readings, and they discovered that it could approximately translate radargrams to heart-rate readings. According to the study, researchers conclude that CNNs can be a useful tool for analyzing and predicting a person's vital signs based on a radar reading. This paper [17] suggested a technique for identifying four respiratory patterns (eupnea, bradypnea, tachypnea, and apnea) from signals obtained by UWB radar using a 1D convolutional neural network (CNN). The procedure generates a dataset and trains a 1D CNN on it. The suggested technique has an accuracy of 95.8% to 15% greater than traditional classification algorithms like PCA and SVM, according to measurements of the recognition accuracy.

This study [18] examined the application of DL to enhance the interpretation of electrical activity of the diaphragm (Edi) signals, a new vital sign for monitoring breathing patterns and attempts in preterm newborns receiving ventilator care. Reliable breathing analysis is challenging because irregular noise interferes with the Edi signal. To classify candidates into final Edi peaks, the study used a local maximal detection approach and CNN. With a f1-score of 0.956 for respiratory Edi peak detection performance and an R2 value of 0.823 for respiratory rates based on the number of Edi peaks, the results demonstrated that this approach outperforms the recording technique currently used in the machine in terms of respiratory Edi peak detection and neural breathing analysis.

In this study [19], a semi-supervised deep clustering (DC) method for contact-free HR estimation utilizing Doppler data is suggested. The technique uses embeddings of the spectrogram of the Doppler signals using a deep LSTM. In the test phase, spectrogram segmentation, which distinguishes the mixed sources of heartbeat, breathing, and movements, uses K-means clustering to infer masks. For relatively static sitting and typewriting with movements, respectively, the technique can lower the average absolute error (AAE) of heartbeat measurement to 3.28 and 3.56 beats per minute (BPM).

This paper [20] outlines a vital sign monitoring system that continually monitors and gathers data on muscle activity, heart rate, and sleep apnea using a microprocessor, ECG sensor, and EMG sensor. The data is then shown on a mobile device, where it may be used to gauge the likelihood of developing heart failure. A notification mechanism is also part of the system to alert medical authorities when needed. The study suggests utilizing random forests to predict sleep apnea and CNN models to categorize heartbeats into abnormalities. The models' accuracy and loss were evaluated using ECG data and cardiac arrhythmia situations.

In order to forecast the beginning of sepsis in ICU patients, this study [21] introduces a method called the Smart Sepsis Predictor (SSP). SSP is made up of LSTM, convolutional, and fully connected layers. It operates in two modes, using demographic information and vital signs in Mode 1 and lab test findings in Mode 2. With an accuracy of 0.75 and 0.0.69 for 4 and 8 hours before sepsis start, respectively, and 0.81 and 0.66 for 12 hours before sepsis onset for modes 1 and 2, the strategy outperformed existing techniques when evaluated on a dataset of 40,366 ICU patients.

The authors of this study [22] suggest a method that focuses on the prediction of vital signs as an essential factor in the early and precise prediction of sepsis. They accomplish this by using a LSTM network, which is trained to perform two tasks: predicting sepsis and vital signs at time t+6. The authors argue that the use of this auxiliary task improves the training of the network, given the low prevalence of sepsis. The network is composed of three modules: an embedding module for compact input representation, a recurrent module with three LSTM layers connected with highway connections, and prediction modules with linear layers for the two tasks. In a repeated cross-fold validation on the network, the result was a utility score of 0.39.

Over 40,000 patients' vital sign readings, laboratory test results, procedures, and prescription information were gathered by the authors for this study [23]. They produced a clean dataset by preprocessing the data, balancing the bias, extracting features, and choosing useful and discriminating characteristics. Then, they predicted key vital signs of ICU patients one hour in advance using a variety of ML methods, including Random Forest, XGBoost, Artificial Neural Networks (ANN), and Long Short-Term Memory (LSTM). The accuracy of the generated predictive models was 91.3 percent for heart rate, 80.1 percent for blood oxygen level (SpO2), 75 percent for mean arterial pressure (MAP), 82.2 percent for respiration rate (RR), and 78.8 percent for systolic blood pressure (SBP).

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023

This paper [24] proposes a machine learning-based approach for the early detection of cardiac arrest using vital signs. The authors used the MIMIC-III database, which is a publicly available dataset of deidentified health data. The dataset includes electrocardiogram (ECG) signals, blood pressure, and respiration rate data for patients in intensive care units (ICUs) who experienced cardiac arrest. The dataset contains 617 patients and over 11,000 hours of data to train a support vector machine (SVM) classifier. The results show that the SVM classifier can accurately classify normal and abnormal vital signs, with an overall accuracy of 94%.

The authors in [25] suggest an ensemble of deep convolutional neural networks (CNNs) for the classification of vital signs. The authors used the PhysioNet/CinC Challenge 2017 dataset, which is a publicly available dataset of physiological signals recorded from patients with sepsis, a life-threatening condition caused by an overwhelming response to infection. The dataset includes electrocardiogram (ECG) signals, blood pressure, and oxygen saturation data from 40 patients to train the CNNs. The results show that the ensemble of CNNs outperforms individual CNNs and other machine learning algorithms, with an overall accuracy of 95%.

This paper [26] proposes a hybrid approach of improved artificial bee colonies (IABC) and extreme learning machines (ELM) for the classification of vital signs. The authors use the MIT-BIH Arrhythmia Database, which is a publicly available dataset of ECG signals recorded from patients with a variety of arrhythmias. The dataset includes over 100,000 ECG recordings from 48 patients to train the IABC-ELM classifier. The results show that the IABC-ELM classifier can accurately classify normal and abnormal vital signs, with an overall accuracy of 93%.

In [27], the authors offer a real-time vital sign classification system for automatic medical alert. The system uses a wearable device to collect data from 15 healthy volunteers. The dataset includes ECG signals, blood pressure, and respiration rate data collected from a wearable device to train a support vector machine (SVM) classifier to classify the vital signs. The system can detect abnormal vital signs and automatically send a medical alert to a healthcare provider, improving the early detection and diagnosis of medical conditions.

Table 1: Previews works resume

Reference	**Method**	**Dataset**	**Results**
[15]	KNN, SVM, DT, RF	CSI	accuracy= 85.12%
[16]	ResNet50	private data	-
[17]	1D-CNN	URB radar	accuracy= 95.8%
[18]	CNN	private data	F1-sore: 0.956 R2:0.823
[19]	deep LSTM	Dopler data	AAE: 3.28 BPM:3.56
[21]	SSP: CNN-LSTM	2019 PhysioNet/ computing in Cardiology Challenge	accuracy: 81%
[23]	RF, XGBoost, ANN, LSTM	ICU	accuracy = 91.3%: Heart rate 80.1%: spO2 75%: MAP 82.2%: RR 78.8%: SBP
[24]	SVM	MIMIC-III	accuracy = 94%
[25]	CNN	PhysioNet/CinC Challenge 2017	accuracy = 95%

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023

[26]	IABC-ELM	MIT-BIH Arrhythmia Database	accuracy = 93%
[27]	SVM	Collect from wearable device	-

4. Methodology

In this research, deep learning algorithms are used for classification tasks. As shown in Figure 2, the input data was prepared and divided into training and testing data to train and evaluate the implemented algorithms. Specifically, two algorithms are utilized: 1D-CNN and CNN-LSTM

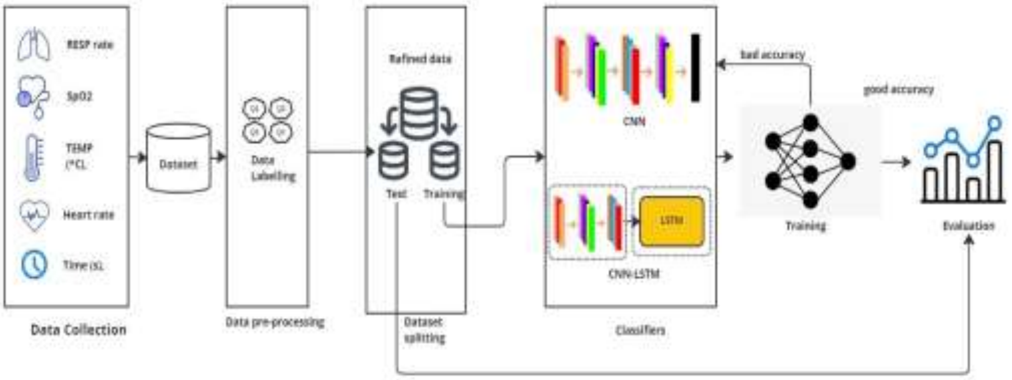
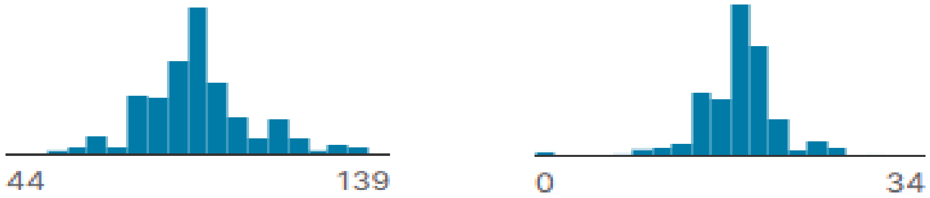


Figure 2: The proposed human vital signs classification system

The reason for using these algorithms is that 1D-CNN is commonly used to detect patterns in data consisting of 1D signals, as it uses the convolution layer to extract optimal features. However, for time series data, LSTM (long short-term memory) networks are also commonly used in conjunction with 1D-CNNs to capture the temporal dependencies in the data.

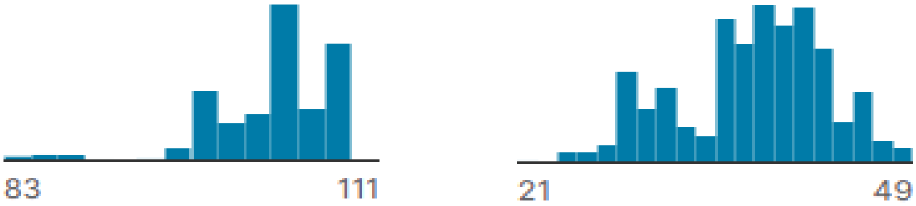
4.1. Dataset

The dataset contained input human signals made up of 1D vectors. These data are from a publicly available dataset with 25,494 patients and five numerical features: The time (s), expressed in seconds, at which the vital sign was taken Heart rate (HR) is expressed as a number of beats per minute (BPM). RESP (BPM), or respiratory rate, is a unit of measurement similar to BPM. SpO2, or blood oxygen saturation, is expressed as a percentage. The body temperature is TEMP (*C), which is expressed in degrees Celsius. For individuals in good health, the output labels are “normal," while for patients with any medical issue, they are “abnormal." On the basis of the five input characteristics, the 1D-CNN and CNN-LSTM models are trained on this dataset to identify patterns in the vital sign data and forecast the health condition of patients. 77% of the data has an abnormal label, whereas 23% has a normal label. Figure 3 shows the description and ranges of each vital sign in the dataset, where (a) shows the heart rate range, (b) shows the respiration range, (c) shows the SPO2 range, and (d) shows the temperature range.



(a) (b)

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023



(c) (d)

Figure 3: Ranges of dataset a: Heartrate b: Respiration c: SPO2 d: Temperature.

4.2 Preprocessing

Several preprocessing processes are executed to ensure accurate output label prediction. First, the output labels are defined and assigned the numeric values 0 and 1, representing distinct classes, which are then assigned. The dataset is then divided into training sets (70%) and testing sets (30%) to facilitate the training and evaluation of models. In addition, in order to prepare the labels for classification, they are transformed into categorical variables, allowing the model to effectively differentiate between the various classes. In addition, scaling or normalization techniques may be used to ensure that the input data is within a consistent range and to eliminate biases during model training. These preprocessing procedures are essential for establishing a trustworthy and robust dataset for training the deep learning model. By appropriately designating labels, slicing the data, and transforming labels into categorical variables, the model can be trained and evaluated accurately, resulting in accurate vital sign classification predictions. Several preprocessing steps are conducted in this subsection to ensure accurate output label prediction. First, the output labels are defined and assigned the numeric values 0 and 1, representing distinct classes, which are then assigned. The dataset is then divided into training sets (70%) and testing sets (30%) to facilitate the training and evaluation of models. In addition, in order to prepare the labels for classification, they are transformed into categorical variables, allowing the model to effectively differentiate between the various classes. In addition, scaling or normalization techniques may be used to ensure that the input data is within a consistent range and to eliminate biases during model training. These preprocessing procedures are essential for establishing a trustworthy and robust dataset for training the deep learning model. By appropriately designating labels, slicing the data, and transforming labels into categorical variables, the model can be trained and evaluated accurately, resulting in accurate vital sign classification predictions.

4.3. Convolution Neural Network (CNN)

In this paper, a deep CNN model was implemented in Keras, consisting of several 1D convolutional layers, max pooling layers, and dropout to prevent overfitting. The input shape is specified as the shape of the input training data, with an additional dimension of 1 to indicate the number of channels. The model starts with a 1D convolutional layer with 128 filters and a kernel size of 1, followed by a second convolutional layer with 64 filters. It then uses 1D max pooling layers, dropout, additional convolutional layers with 64 and 32 filters, and a max pooling layer. The model ends with a flattening layer, a dense layer with 2 units, and a softmax activation function, which is used for multi-class classification. In our case, we have two outputs: normal and abnormal. The implemented CNN architecture is illustrated in Figure 4.

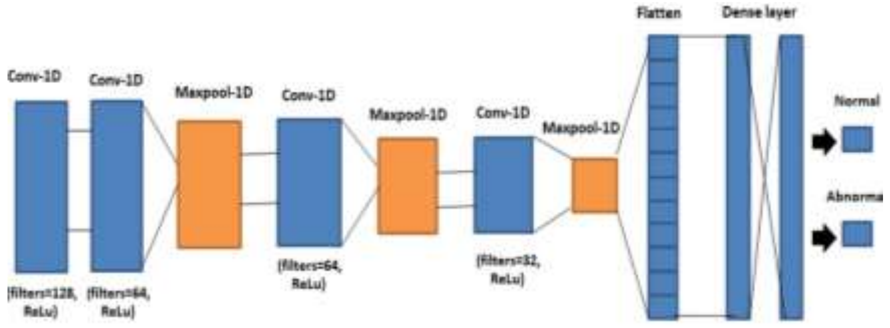


Figure 4: CNN architecture.

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023

4.4. Hybrid CNN-LSTM model

In this study, the human vital signs were predicted using a combination of the CNN and 190 LSTM networks, as illustrated in Figure 4. This CNN-LSTM model combines the strengths of CNNs and LSTMs to process time-series data. The model starts with a Conv1D layer for feature extraction, followed by a MaxPooling1D layer for dimensionality reduction. The model then has two LSTM layers for capturing temporal dependencies in the data, which are refined by additional LSTM layers and MaxPooling1D layers. Dropout layers are used throughout the model to prevent overfitting. Finally, the model has an LSTM layer and an output layer with a softmax activation function for classification.

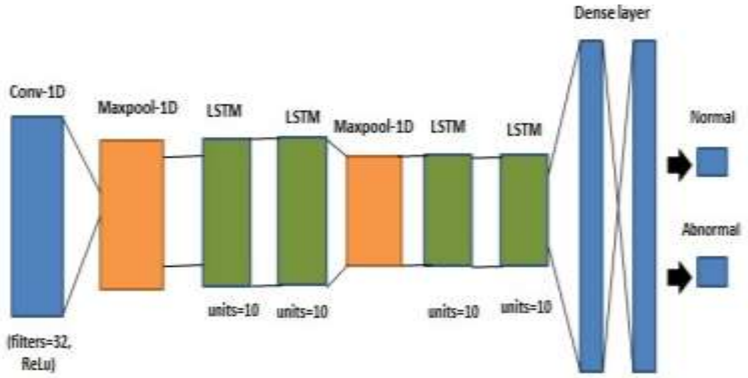
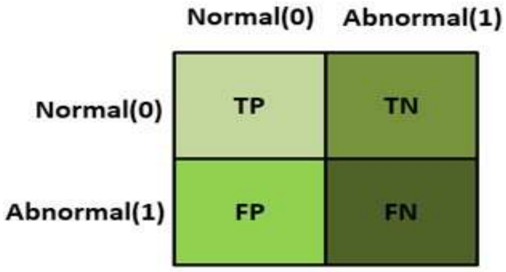


Figure 5: CNN-LSTM architecture.

4.5. Evaluation metrics

To evaluate the proposed models, we utilized the confusion matrix and the classification report, which include the accuracy (acc), precision (pr), recall (rec), and F1-score (F-s). The following confusion matrix (Figure 5) represents the performance of a classification model for human vital signs. The matrix compares the predicted class labels with the true class labels for two classes: "normal" and "abnormal.". The entries in the first row and first column of the matrix correspond to the “normal" class, and the entries in the second row and second column correspond to the “abnormal" class.



	Normal(0)	Abnormal(1)
Normal(0)	TP	TN
Abnormal(1)	FP	FN

Figure 6: Confusion matrix

The first entry in the first row represents the true positive (TP) for the" Normal" class. It presents the instances that were actually “normal" and correctly classified as “normal" by the model. The second entry in the first row represents the false positive (FP) for the “normal" class. It presents the instances that were actually “abnormal" and incorrectly classified as “normal" by the model. The first entry in the second row represents the false negative (FN) for the “abnormal" class. It presents the instances that were actually “abnormal" and incorrectly classified as “normal" by the model. The second entry in the second row represents the true negative (TN) for the “abnormal" class. It presents the instances that were actually abnormal" and were correctly classified as “abnormal" by the model. The Acc (7), PR (8), REC (9), and F-S (10) metrics equations are as following:

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023

$$\text{ACC} = \frac{TP+TN}{TP+TN+FP+FN} \quad (7)$$

$$\text{PR} = \frac{TP}{TP+FP} \quad (8)$$

$$\text{REC} = \frac{TP}{TP+FN} \quad (9)$$

$$\text{F} - \text{S} = \frac{2+PR*REC}{PR+REC} \quad (10)$$

5. Results

The proposed models will be trained on a dataset that is divided into a 7:3 ratio, with 70% of the data used for training and 30% for testing. The Adam optimizer will be employed to optimize the model during the training process. The loss function employed is categorical cross-entropy. The models will be trained for 50 epochs, which means the algorithm will iterate over the entire dataset 50 times to learn the best parameters for the model.

5.1. CNN Results

Initially presented the performance results of the basic model before testing the suggested hybrid architecture. After training the model, we evaluated its performance on the test set. Figure 6 shows the CNN confusion matrix. Table 2 shows that the CNN obtained 96% ACC, 88% PR, 99% REC, and an F-S of 93%.

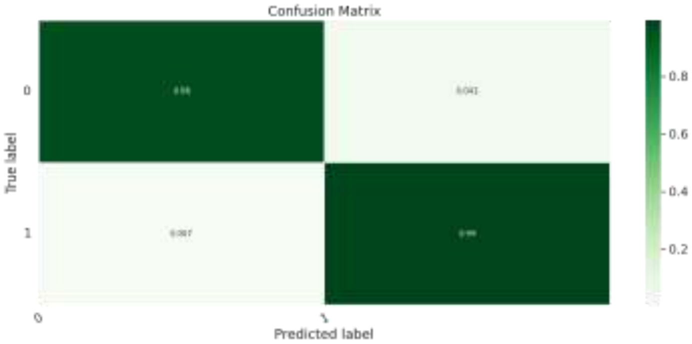


Figure 7: CNN confusion matrix

5.2. Hybrid CNN-LSTM Results

Figure 7 shows the confusion matrix obtained from the hybrid approach. Table 2 illustrates that the hybrid classifier demonstrates superior performance in classifying patients' health states, with an ACC, PREC, REC, and F1-S of 98%, 96%, 97%, and 97% respectively.

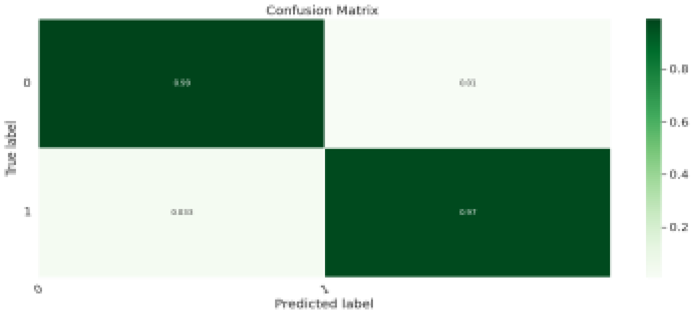
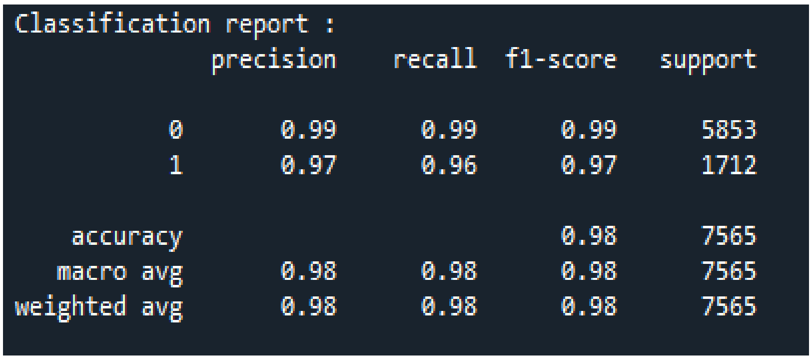


Figure 8: LSTM confusion matrix

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023



```
Classification report :
              precision    recall  f1-score   support

           0       0.99      0.99      0.99      5853
           1       0.97      0.96      0.97      1712

    accuracy                           0.98      7565
   macro avg       0.98      0.98      0.98      7565
weighted avg       0.98      0.98      0.98      7565
```

Figure 9: CNN-LSTM Classification report

5.3. Comparison Results

Table 2 compares the empirical outcomes of our proposed hybrid classification system's varied output targets to CNN's base classifier and other methodologies in the literature. Experiments indicate that the proposed model yields the highest ACC at 98%, outperforming the CNN model and previous models [15, 17, 21]. Furthermore, the CNN-LSTM model produces good results with a relatively low number of parameters and less training time when compared to a traditional CNN model. Specifically, the total number of parameters in the CNN model is 15,074, while in the CNN-LSTM model it is 4,550, which is significantly less. This indicates that the CNN-LSTM model is able to achieve high performance with fewer parameters, making it more computationally efficient and faster to train. By combining 1D-CNN and LSTM, we conclude that the CNN-LSTM model is able to extract both spatial and temporal features from the input data, leading to a better representation of the data and thus improved performance compared with the CNN model that could extract only the spatial data.

Table 2: The implemented model's comparison results

Method	Labels	Acc	Pr	REC	F-S
1D-CNN	Normal Abnormal	96% -	100% 88%	96% 99%	98% 93%
CNN-LSTM	Normal Abnormal	98%	99% 96%	99% 97%	99% 97%
1D-CNN [17]	-	95.5%	-	-	-
CNN [15]	-	85.12%	-	-	-
CNN-LSTM [21]	-	81%	-	-	-

6. Conclusion

In conclusion, this study has proposed a deep learning architecture for the classification of vital signs using a publicly available dataset with 25,494 patients and five numerical features. The proposed CNN-LSTM model was able to achieve high performance with a relatively low number of parameters and less training time when compared to a traditional CNN model. Specifically, the proposed model yielded the highest accuracy at 98%, outperforming previous models. These results explain that the CNN-LSTM model captures more nuanced and complex relationships between the input data and target predictions, which could help improve accuracy compared with CNN. Additionally, the number of parameters in the CNN-LSTM model is significantly less than the traditional CNN model, making it more computationally efficient and faster to train. These results demonstrate the potential of using deep learning techniques for vital sign classification and suggest that the proposed CNN-LSTM model can be a useful tool for healthcare professionals in monitoring patients' vital signs. One potential future direction for this research could be to incorporate additional features such as demographic information or patient history to improve the performance of the model. Another area of investigation could be to adapt the proposed model for real-time monitoring of vital signs in clinical settings by implementing the model on wearable devices or other portable devices.

Funding: “This research received no external funding”

Conflicts of Interest: “The authors declare no conflict of interest.”

References

[1] M. Dhimal et al., “Impact of Air Pollution on Global Burden of Disease in 2019,” Processes, vol. 9, no. 10, p. 1719, Sep. 2021, doi: 10.3390/pr9101719

[2] Dibyahash Bordoloi, et al., "Deep Learning in Healthcare System for Quality of Service", Journal of Healthcare Engineering, vol. 2022, Article ID 8169203, 11 pages, 2022. https://doi.org/10.1155/2022/8169203

[3] A.V.L.N. Sujith, Guna Sekhar Sajja, V. Mahalakshmi, Shibili Nuhmani, B. Prasanalakshmi,Systematic review of smart health monitoring using deep learning and Artificial intelligence, Neuroscience Informatics,Volume 2, Issue 3,2022,100028,ISSN 2772-5286,https://doi.org/10.1016/j.neuri.2021.100028.

[4] Cardona-Morrell, M., et al. "Vital signs monitoring and nurse–patient interaction: A qualitative observational study of hospital practice." International journal of nursing studies 56 (2016): 9-16 doi: 10.1016/j.ijnurstu.2015.12.007.

[5] H. Rohmetra, N. Raghunath, P. Narang, V. Chamola, M. Guizani, and N. R. Lakkaniga, "AI enabled remote monitoring of vital signs for COVID-19: methods, prospects and challenges," Computing, pp. 1-27, 2021.

[6] S. Razdan and S. Sharma, "Internet of medical things (IoMT): Overview, emerging technologies, and case studies," *IETE technical review,* vol. 39, no. 4, pp. 775-788, 2022.

[7] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *nature,* vol. 521, no. 7553, pp. 436-444, 2015.

[8] J. Schmidhuber and S. Hochreiter, "Long short-term memory," *Neural Comput,* vol. 9, no. 8, pp. 1735-1780, 1997.

[9] S. Huang, J. Tang, J. Dai, and Y. Wang, "Signal status recognition based on 1DCNN and its feature extraction mechanism analysis," *Sensors,* vol. 19, no. 9, p. 2018, 2019.

[10] M. G. Ragab *et al.*, "A novel one-dimensional cnn with exponential adaptive gradients for air pollution index prediction," *Sustainability,* vol. 12, no. 23, p. 10090, 2020.

[11] G. Swapna, S. Kp, and R. Vinayakumar, "Automated detection of diabetes using CNN and CNN-LSTM network and heart rate signals," *Procedia computer science,* vol. 132, pp. 1253-1262, 2018.

[12] R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *International conference on machine learning,* 2013: Pmlr, pp. 1310-1318.

[13] H. Salehinejad, S. Sankar, J. Barfett, E. Colak, and S. Valaee, "Recent advances in recurrent neural networks," *arXiv preprint arXiv:1801.01078,* 2017.

[14] K. Kawakami, "Supervised sequence labelling with recurrent neural networks," Technical University of Munich, 2008.

[15] M. I. Khan *et al.*, "Tracking vital signs of a patient using channel state information and machine learning for a smart healthcare system," *Neural Computing and Applications,* pp. 1-15, 2021.

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023

[16] R. Gibadullin and N. Marushkai, "Development of predictive CNN based model for vital signs alerts," in *2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, 2021: IEEE, pp. 404-409.

[17] S.-H. Kim and G.-T. Han, "1D CNN based human respiration pattern recognition using ultra wideband radar," in *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2019: IEEE, pp. 411-414.

[18] H.-G. Lee, G. Lee, J. Lee, Convolutional neural network-based respiration analysis of electrical activities of the diaphragm, Scientific Reports 12 (1) (2022) 1–8.

[19] C. Ye, G. Gui, and T. Ohtsuki, "Deep clustering with lstm for vital signs separation in contact-free heart rate estimation," in *ICC 2020-2020 IEEE International Conference on Communications (ICC)*, 2020: IEEE, pp. 1-6.

[20] G. T. M., V. M., Vital sign monitoring system and analysis with cnn, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH TECHNOLOGY (IJERT) 11 (06) (June 2022).

[21] A. Rafiei, A. Rezaee, F. Hajati, S. Gheisari, and M. Golzan, "SSP: Early prediction of sepsis using fully connected LSTM-CNN model," *Computers in biology and medicine,* vol. 128, p. 104110, 2021.

[22] B. Roussel, J. Behar, and J. Oster, "A recurrent neural network for the prediction of vital sign evolution and sepsis in icu," in *2019 Computing in Cardiology (CinC)*, 2019: IEEE, pp. Page 1-Page 4.

[23] D. Chang, D. Chang, and M. Pourhomayoun, "Risk prediction of critical vital signs for ICU patients using recurrent neural network," in *2019 International Conference on* Computational Science and Computational Intelligence (CSCI), 2019: IEEE, pp. 1003-1006.

[24] M. R. M. Al Jadda, M. T. Ali, and S. S. S. Hassan, "Vital Sign Classification for Early Detection of Cardiac Arrest," in 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC), 2021, pp. 0549-0554.

[25] S. K. Saha, A. Roy, and S. Kundu, "Classification of Vital Signs Using Ensemble of Deep Convolutional Neural Network," in 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE), 2021, pp. 408-413.

[26] S. R. E. Setiawan and M. A. Widyanto, "Vital Sign Classification Using Hybrid of Improved Artificial Bee Colony and Extreme Learning Machine," in 2021 International Conference on Advanced Computer Science and Information Systems (ICACSIS), 2021, pp. 384-389.

[27] J. W. Lee, J. Y. Kim, and J. Kim, "Real-Time Vital Sign Classification System for Automatic Medical Alert," in 2021 IEEE International Conference on Consumer Electronics (ICCE), 2021, pp. 1-2.

Doi: https://doi.org/10.54216/FPA.130202
Received: April 05, 2023 Revised: July 07, 2023 Accepted: September 09, 2023