LSTM Model for Sepsis Detection and Classification Using PPG Signals

Michael A. Alvarez-Navarro[1(✉)], Luis Huallparimachi[1], Sebastián A. Cruz-Romero[2], and Heidy Sierra[2]

[1] Electrical and Computer Engineering, University of Puerto Rico at Mayagüez, Mayagüez, Puerto Rico, USA
{michael.alvarez2,luisenrique.huallparimachi}@upr.edu
[2] Computer Science and Engineering, University of Puerto Rico at Mayagüez, Mayagüez, Puerto Rico, USA
{sebastian.cruz6,heidy.sierra1}@upr.edu

Abstract. Sepsis is a severe medical illness with over 1.7 million cases reported each year in the United States. This condition is the result of an inflated immunological response to infection. Early diagnosis of sepsis remains a major challenge in healthcare due to initial symptoms being nonspecific and the lack of currently available biomarkers that demonstrate sufficient specificity or sensitivity suitable for clinical practice. Wearable technologies, such as photoplethysmography (PPG), have led to the development of improved diagnostic instruments. PPG uses optical technology to measure changes in blood volume in peripheral tissues, enabling continuous monitoring. Identifying modest physiological changes that indicate sepsis can be challenging since they occur without a body reaction. Deep Learning (DL) models can help overcome the diagnostic gap in sepsis diagnosis and intervention. This study analyzes sepsis-related characteristics in PPG signals utilizing a collection of waveform records from both sepsis and control cases. The proposed model consists of five layers: input sequence, long short-term memory (LSTM), fully-connected, softmax, and classification. The LSTM layer is chosen to extract and filter features from cycles of PPG signals; then, the features pass through a fully-connected layer to be classified. We tested our LSTM-based model on 915 one-second intervals to identify and classify sepsis severity. Our LSTM-based model accurately detected sepsis (91.30% for training and 89.74% for testing). The sepsis severity categorization achieved an accuracy of 85.9% in training and 81.4% in testing. Multiple training attempts were conducted to validate the model's detecting capabilities. Preliminary results show that a deep learning model utilizing an LSTM layer can detect and categorize sepsis using PPG data, potentially allowing for real-time diagnosis and monitoring within a single cycle.

Keywords: Sepsis · Classification · Deep Learning · LSTM · PPG signal

M. Kadoch et al. (Eds.): ISICN 2024, LNNS 1094, pp. 3–16, 2024.
https://doi.org/10.1007/978-3-031-67447-1_1

1 Introduction

Sepsis is a life-threatening medical condition affecting approximately 1.7 million adults in the United States annually, with nearly 270,000 resulting in fatalities [5]. Sepsis stems from an exaggerated immune response triggered by infection, releasing immune chemicals into the bloodstream. While initially intended to combat the infection, these chemicals instead incite widespread inflammation and blood clot formation, obstructing blood flow and depriving vital organs of oxygen and nutrients. This often results in significant organ damage. In healthcare, sepsis poses a considerable challenge, contributing significantly to mortality and hospital readmissions. The early diagnosis of sepsis remains a formidable task, primarily relying on standard clinical measures with limited sensitivity and specificity [19]. This involves evaluating vital signs, taking a patient's medical history, and conducting a physical examination. Subsequently, blood tests are performed to examine biomarkers like white blood cell count, C-reactive protein (CRP) levels, and procalcitonin. Imaging tests, such as chest X-rays or CT scans, may be used to identify infection sources or potential organ dysfunction.

However, these tests alone may not confirm sepsis [4]. Clinicians often refer to established clinical criteria like the SIRS (Systemic Inflammatory Response Syndrome) criteria to aid sepsis identification [14], but these criteria have limitations. Challenges in early diagnosis arise from the non-specific nature of initial sepsis symptoms, variability in biomarkers, reliance on clinical judgment, time constraints, and patient diversity [17].

Micro-circulatory dysfunctions observed in sepsis patients are reflected in parameters that can be conveniently evaluated at the skin level, such as the photoplethysmogram (PPG). The PPG signal is commonly monitored using devices like the pulse oximeter, which is widely used, user-friendly, and affordable [12]. PPG signals are a non-invasive optical method that monitors changes in blood volume within peripheral tissues, providing a wealth of information, including heart rate, oxygen saturation, and blood flow dynamics [20]. It allows continuous monitoring and early detection of circulatory issues, even before severe symptoms emerge. Wearable devices equipped with PPG sensors enable remote monitoring, which is especially beneficial for patients with chronic conditions.

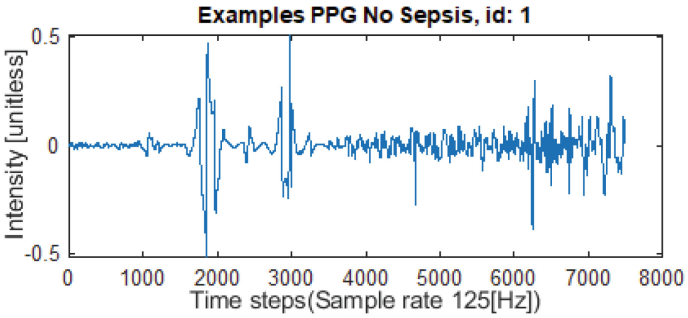
Additionally, PPG plays a pivotal role in diagnosing various medical conditions, making it an invaluable tool in modern healthcare [3]. Medical applications are related to the use of PPG signals and combine with algorithms and techniques in different ways; for example, Cardiovascular Disorders (CVD) use a fuzzy model, sleep disorders require peripheral oxygen saturation signals, and Diabetes is tracked at the same time that PPG signals to train machine learning models [1]. For the CVD, Prabhakar et al. in [18] optimize the CVD diagnosis from PPG signals by utilizing a fuzzy-based approach with classification. The first step consists of extracting parameters from the PPG signal: Energy, Variance, Approximate Entropy (ApEn), Mean, Standard Deviation (STD), Skewness, Kurtosis, and Peak Maximum. Later, optimization algorithms such as Differential Search (DS), Shuffled Frog Leaping Algorithm (SFLA), Wolf Search (WS), and Animal Migration Optimization (AMO) are applied to a fuzzy model

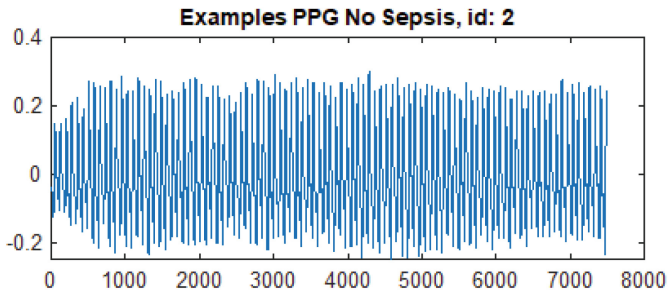
for classification. Another application in which the PPG signals are widely used is detecting and classifying sleep apnea. Lazazzera et al. in [10] develop a detector consisting of two parts: one that detects reductions in amplitude fluctuation of PPG and one that catches oxygen desaturations. The classification was performed to discriminate between central and obstructive events, apneas, and hypopneas.

Enhancing sepsis diagnostic capabilities through PPG and wearable technologies is imperative. These tools offer continuous monitoring, enabling the identification of subtle physiological changes indicative of sepsis before severe symptoms manifest. Integrating computational models, particularly Deep Learning (DL) models as Long Short-Term Memory (LSTM) models, enhances the potential to bridge the diagnostic gap in sepsis detection and intervention [9]. Kam and Kim in [9] present a comparative study of deep learning models for feature extraction from PPG signals that are used for the early detection of sepsis. This study's higher accuracy is 93%; for 100 features as input for an LSTM-based model, the features correspond to multiple physiological parameters like heart rate, blood pressure, blood oxygen saturation, and pH, among others. LSTMs excel in analyzing biomedical signals, capturing intricate temporal dependencies, and identifying subtle variations. Their ability to extract vital features from continuous physiological signals monitored by PPG and wearables significantly contributes to early sepsis detection and improved patient care.

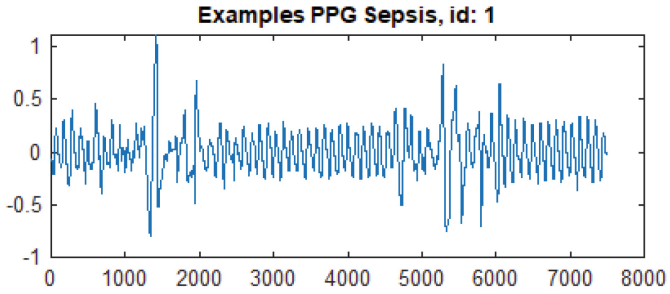
As sepsis diagnosis and intervention become increasingly data-driven, there is a growing opportunity to fuse medical science and computational approaches. This confluence promises to improve sepsis preventive diagnosis and advance our understanding of other complex medical conditions. The transformative potential of PPG, wearables, and DL models in healthcare is a testament to the evolving medical research and technology landscape. By harnessing these innovative tools and computational strategies, we are better equipped to confront the challenges of sepsis and other critical medical conditions, offering hope for more timely interventions and improved patient outcomes [16].

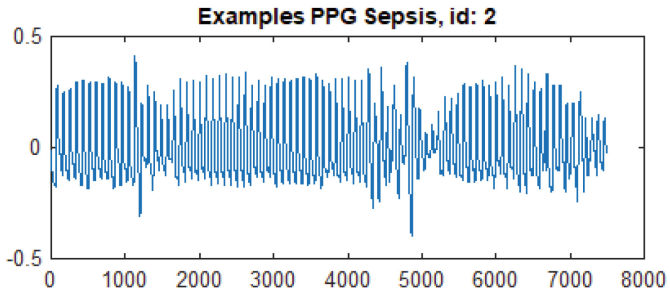
This document continues with the Methodology, Results, and Conclusion sections. In the Methodology section, the steps to generate the dataset, the preprocessing of the signal, the sepsis detection, and classification models are presented. The LSTM layer is the base of the model for the detection and classification tasks; see Fig. 4. The Results section shows the detection and classification experiments, with a comparison of the performance in terms of the random trials used in the training and testing stages. Finally, the Conclusion section presents the contribution and future steps of our work.





(a) PPG signals with Non-sepsis label.





(b) PPG signals with sepsis label.

Fig. 1: Example of PPG signals in MIMIC III database. Reduce to 1 min.

2 Methodology

2.1 Dataset

The *Medical Information Mart for Intensive Care III (MIMIC-III)* database [7], is a large, publicly available database containing de-identified healthcare data of

patients admitted to critical care units. From *MIMIC-III*, the subset *MIMIC-III Clinical Database* is available as a group of comma-separated value files that are imported into a PostgreSQL relational database system [8]. We obtain PPG waveforms records from *MIMIC-III Waveform Database Matched Subset* [15], a part of *MIMIC-III*. The common information between the databases is guarantee by the id of the patient.

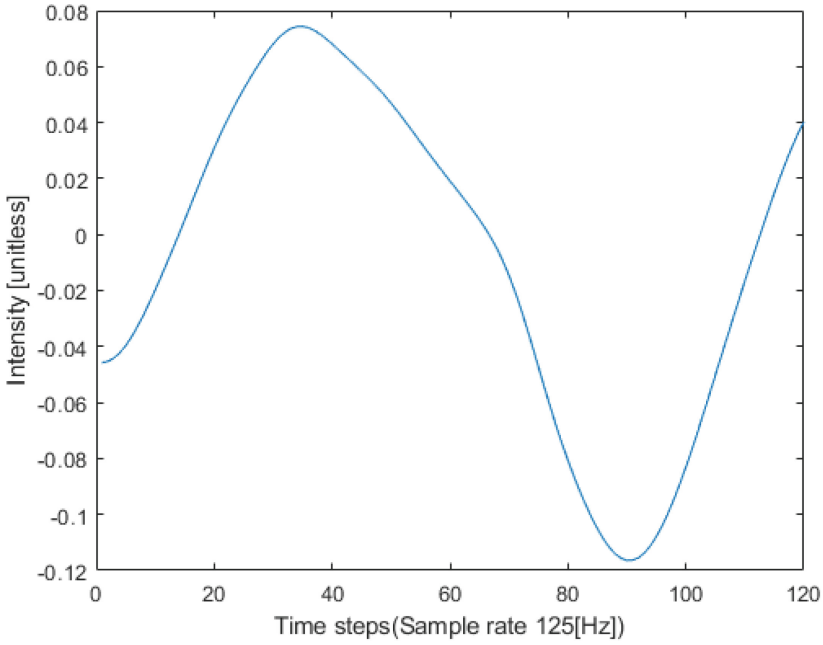
Diagnoses within these datasets are primarily documented as ICD-9 codes, typically generated upon the conclusion of a hospitalization, lacking any temporal information regarding the date of diagnosis. Consequently, we opted to select a subset of subjects, categorized as either "sepsis" (cases) or "non-sepsis" (controls) patients defined by their ICD-9 Codes, respectively. The selection criteria for distinguishing sepsis and non-sepsis subjects are summarized in Table 1, the initial benchmark is presented by Lombardi et al. in [12]. We employed a customized Structured Query Language (SQL) query during this phase to perform the necessary data extraction and selection. As a result, the group of patients containing a PPG signal with sepsis had 812 records, while the control group had 1248 records. Figure 1 shows the sepsis and control PPG signals.

Table 1: Criteria for the definition of control and case patients

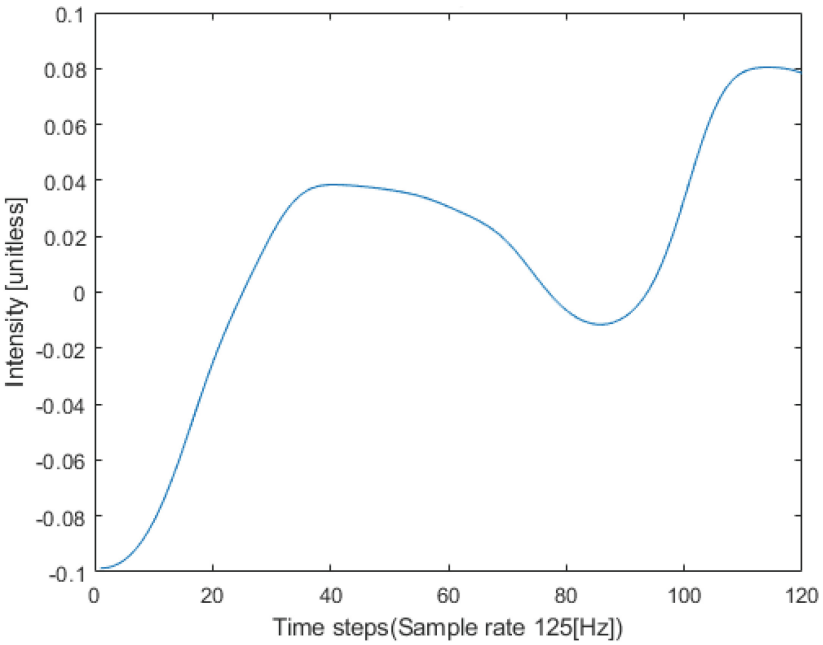
Control patients	Case patients
One or more following ICD-9 codes: 311 (Depressive disorder NEC), 3051 (Tobacco Use Disorder), 30,000 (Anxiety State NOS), 2498 (Other persistent mental disorders due to conditions classified elsewhere), 3004 (Dysthymic Disorder)	One or more following ICD-9 codes: 99,591 (sepsis), 99,592 (Severe sepsis), 78,552 (Septic Shock)
Subset was present in the Matched Subset of MIMIC-III Waveform Database	Subset was present in the Matched Subset of MIMIC-III Waveform Database

2.2 Preprocessing and Quality Estimation

We retrieved the recordings from the MIMIC-III Waveform Database Matched Subset and acquired the PPG data of the specified patients by utilizing the Waveform Data Base (WFDB) Python package [21]. Subsequently, the chosen signals were segmented into intervals of 1 to 3 h, and the rejection criteria consisted of signals with only NAN, or only 0 or amplitudes lower than 1e−5 are



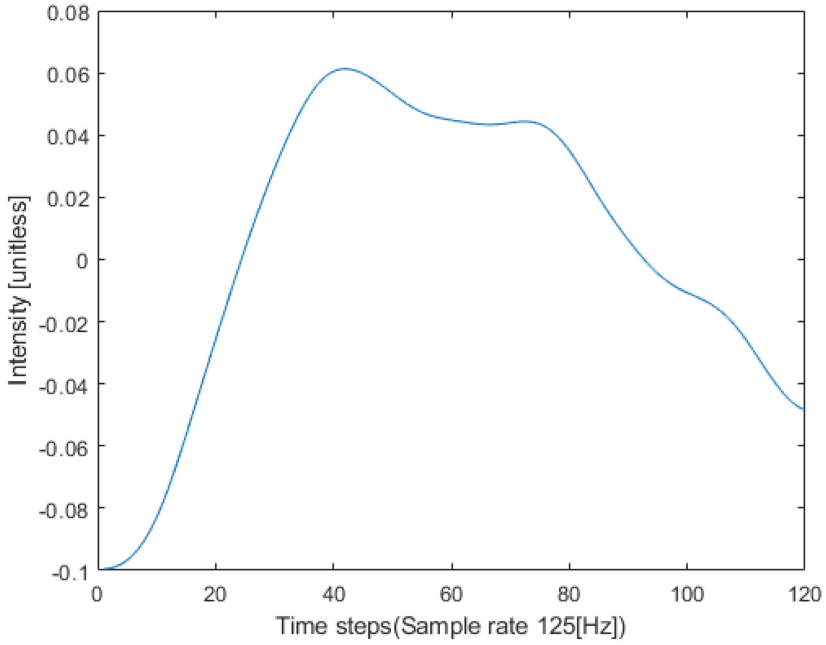
(a) PPG signals with Non-sepsis label.



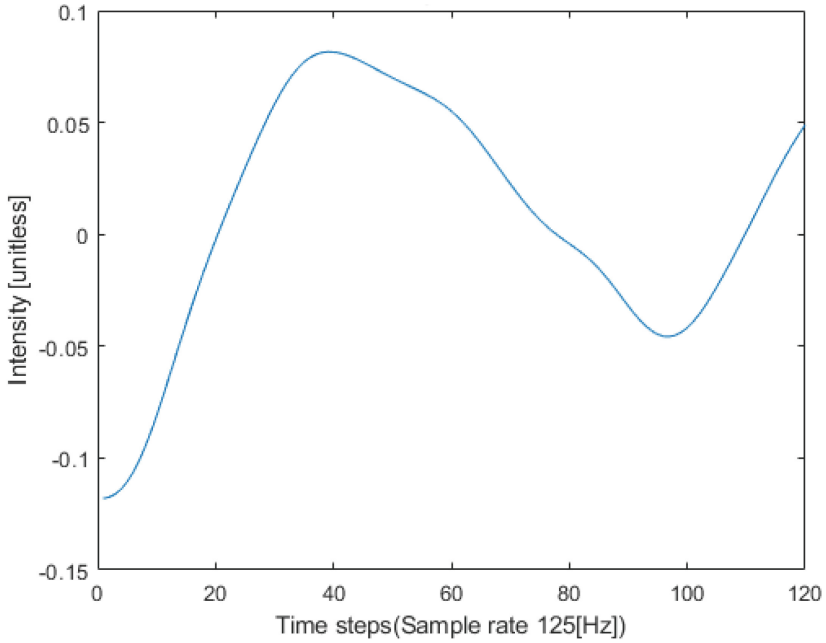
(b) PPG signals with sepsis label.

Fig. 2: Comparison of a single PPG cycle for (a) Non-sepsis and (b) sepsis labels.

removed. In our analysis, we specifically selected the initial minute of the signals that remained after the filter performed. Next, we take the first minute of the signal; the consistency and quality of each 1-min sample were evaluated using a *template* matching method. This method is based on the Elgendi paper [2] for analyzing PPG signals. This quality assessment involved the implementation of



(a) PPG signal with Severe sepsis label.



(b) PPG signal with Septic Shock label.

Fig. 3: Comparison of a single PPG cycle for (a) Severe sepsis and (b) Septic Shock labels.

a 10-s running window over the 1-min segment. We categorized each window by comparing the patient-acquired signal with an ideal template PPG signal. The degree of similarity between these time series was quantified using Pearson's correlation coefficient. We limited PPG records to 40 subjects per ICD-9 code

to avoid over representing a category in our dataset. These stringent condition reduced our dataset to 915 waveforms for our control and case groups.

The *template* was generated using the NeuroKit2 Python toolbox [13], a package designed for processing neurophysiological signals. The reference PPG signal was simulated without noise and motion artifacts. In this simulation process, input parameters such as the signal's sampling frequency and the mean heart rate within each window were essential. The sampling frequency was established at 125 Hz, consistent with the frequency utilized for all signals in the waveform database. Additionally, we subdivided our signals based on cardiac cycles, determining the quantity of systolic cardiac peaks for each second per hertz of the waveform. To pinpoint the locations of these peaks, we initially filtered the signal and subsequently employed NeuroKit's peak detection method.

Moreover, systolic peak detection was conducted by identifying systolic peaks within each 10-s segment. The diastolic counterpart was sampled for every identified systolic point as our cardiac cycle's start and end points. This approach allowed us to pinpoint the maximum amplitude within each segment, thereby successfully delineating each distinct signal cycle. We extracted the count of segments and the amplitudes associated with each cycle for further evaluation. Specifically, cycles within the size range of 110 to 120 were selected. We perform a resizing approach for cycles smaller than 120 samples, employing bicubic interpolation. This method facilitates the prediction of values at intermediate positions within the cycle. As a result, our final cohort comprises 915 signals, labeled as 235 for Non-sepsis, 210 for sepsis, 190 for Severe sepsis, and 280 for Septic Shock. Figures 2 and 3 show examples of the Non-sepsis, Sepsis, Severe Sepsis, and Septic Shock PPG signals.

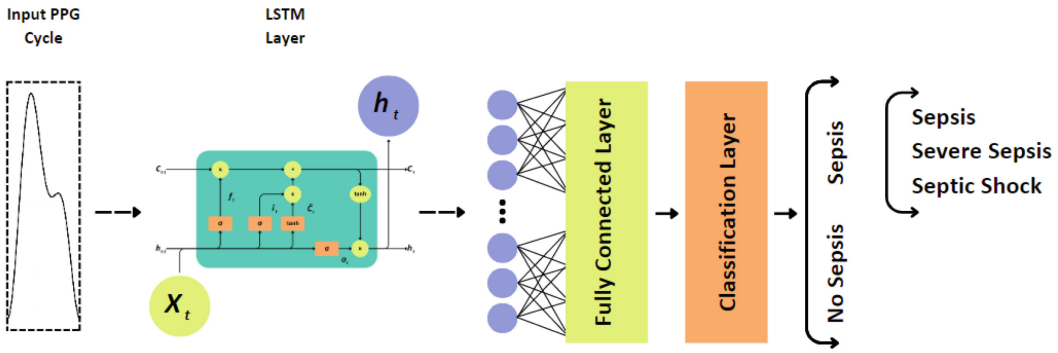


Fig. 4: Detection and classification model based on LSTM layer. The Input is a single cycle of a PPG signal. The Output a category label of Not-sepsis, sepsis, Severe sepsis and Septic Shock.

2.3 Classification

We explore different network architectures for sepsis classification, including Long Short-Term Memory (LSTM) [6], Convolutional Neural Networks (CNN)

[11], and fully-connected networks. The LSTM-based models show promising performance in sepsis detection and classification from PPG signals.

The DL model presented is formed by a graph of five layers: Sequence input, LSTM, fully-connected, Softmax, and Classification. The input/output diagram of the model is presented in Fig. 4. In particular, our binary classification study, which confined 915 PPG signals of sepsis and control patients, has an unbalanced splitting for the detection approach because the Sepsis case contains Severe sepsis and Septic Shock signals. Within this framework, the LSTM model was trained on 70%, validated on 10%, and tested on 20% of the data.

We compute the common metrics for a binary classifier and the Area Under the Curve (AUC), i.e.:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{2}$$

$$\text{Specificity} = \frac{TN}{FP + TN} \tag{3}$$

$$\text{AUC} = \text{Sensitivity}(TPR) - (1 - \text{Specificity})(FPR) \tag{4}$$

For the multi-class case, we present the confusion matrix and focus on the Accuracy metric because the correct predictions are the interest in sepsis detection and classification.

3 Results

In this section, we present the classification results in two scenarios: Sect. 3.1 for the detection of sepsis (case and control), and Sect. 3.2 the classification of the signals based patient condition: sepsis, Severe sepsis, and Septic Shock.

3.1 Detection Sepsis

In this case, the LSTM model classifies the PPG cycle in two categories, i.e. sepsis, and non-sepsis. The training process takes 200 epochs using a learning rate of 0.001; due to limited data, we perform several training trials as an alternative to validate the classifier. Table 2 shows the model's performance for the trials. The repeatability of the results is guaranteed using a *random number generator* (rng) in each case. The seed is randomly chosen as $k = [22, 5, 1, 6, 7, 9, 10, 15, 57, 87]$. The mean value of the overall accuracy for training data is 91.30%, and for testing data, it is 89.74%. Figure 5 presents an example of the confusion matrix for sepsis detection. We compare the metrics of the model in terms of the PPG length, showing in Table 3 that the PPG cycle alternative obtains the best results.

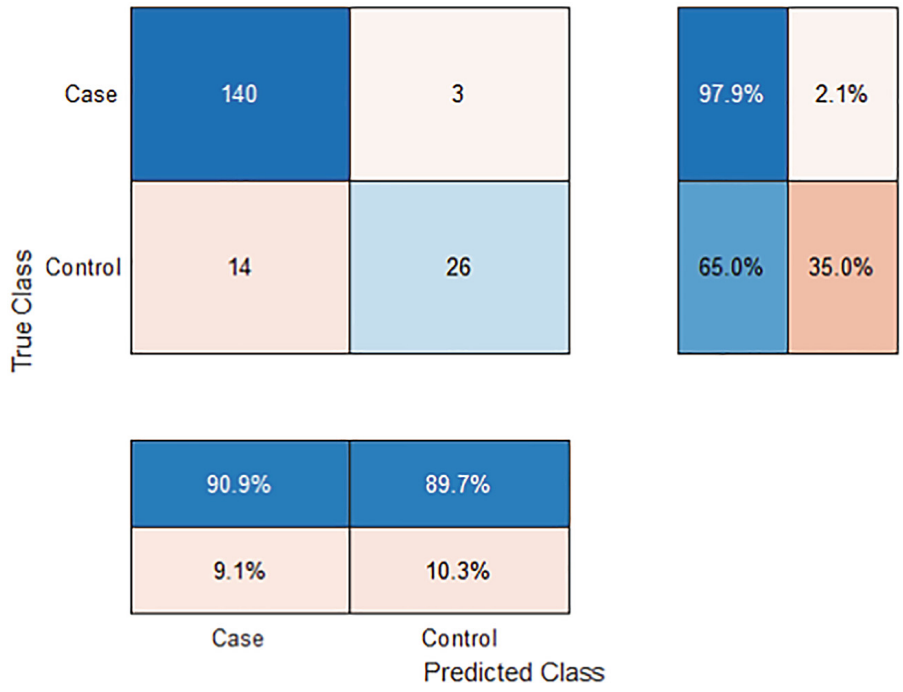


Fig. 5: Detection of sepsis using LSTM model. Testing data: Overall Accuracy = 90.71%, rng $k = 6$.

3.2 Classification of Sepsis, Severe Sepsis, and Septic Shock

For this case, we consider the ICD codes 99591, 99592, and 98552 as the labels for sepsis, Severe sepsis, and Septic shock, respectively. The Control patients are

Table 2: Detection performance of the DL model in terms of the seed k

k	Training	Test
22	91.82	90.16
5	90.65	89.62
1	91.59	90.16
6	92.06	90.71
7	89.72	91.26
9	92.29	89.07
10	93.22	87.98
15	89.95	88.52
57	90.42	90.16
87	90.19	87.98
mean	91.30	89.74
std	1.18	1.05

Table 3: Table of Binary classification metrics

	Accuracy	Sensitivity	Specificity	AUC
1 min	77.97	02.17	91.97	47.07
3 s	88.35	25.88	91.17	58.52
Cycle approach (<1 s)	90.71	90.91	89.66	90.28

the fourth class. Similarly to the detection approach, we perform several trials for training to validate the classifier. For training and testing data, we show the model's performance for the trials in Tables 4 and 5. The mean value of the overall accuracy for training data is 85.28, and for testing data, it is 85.79%. Figure 6 presents an example of the confusion matrix for the testing data.

Table 4: Performance of Accuracy Metric on Training data in terms of the seed k

k	Training data				
	Sepsis	Severe sepsis	Septic shock	Non sepsis	Overall
22	94.9	75.9	85.0	86.7	85.3
5	94.9	77.8	83.8	90.8	86.2
1	95.8	75.0	88.3	89.2	86.2
6	93.1	74.0	89.3	88.0	85.5
7	94.8	71.4	87.9	87.1	83.9
9	96.0	77.3	87.5	94.8	88.1
10	93.5	82.1	87.7	91.8	88.3
15	94.3	71.0	83.0	89.9	83.4
57	94.6	78.3	89.0	91.8	87.4
87	94.7	74.0	83.5	92.1	85.0
mean	94.7	75.7	86.5	90.2	85.9
std	0.9	3.4	2.4	2.5	2.0

We did the classification experiments in a Windows 10 Education Desktop computer, using an Intel(R) Core(TM) i7-8700 CPU 3.20 (GHz) with 16 GB RAM, in the Laboratory for Applied Remote Sensing and Image Processing (LARSIP) at the University of Puerto Rico, Mayaguez Campus. We employ the Deep Learning Toolbox of MATLAB.

Table 5: Performance of Accuracy Metric on Training data in terms of the seed k

k	Testing data				
	Sepsis	Severe sepsis	Septic shock	Not sepsis	Overall
22	97.0	79.4	80.5	90.2	85.2
5	88.1	71.2	91.5	81.0	82.5
1	100.0	64.2	77.6	72.7	77.0
6	85.1	84.6	77.8	87.2	83.6
7	95.0	64.3	88.2	86.1	82.0
9	79.2	74.5	76.6	87.9	78.7
10	93.2	72.4	81.4	84.2	82.0
15	90.0	70.8	75.9	81.5	79.2
57	97.7	64.3	83.3	86.1	81.4
87	90.5	72.9	85.7	85.1	82.5
mean	91.6	71.9	81.6	84.2	81.4
std	6.3	6.7	5.3	4.9	2.4

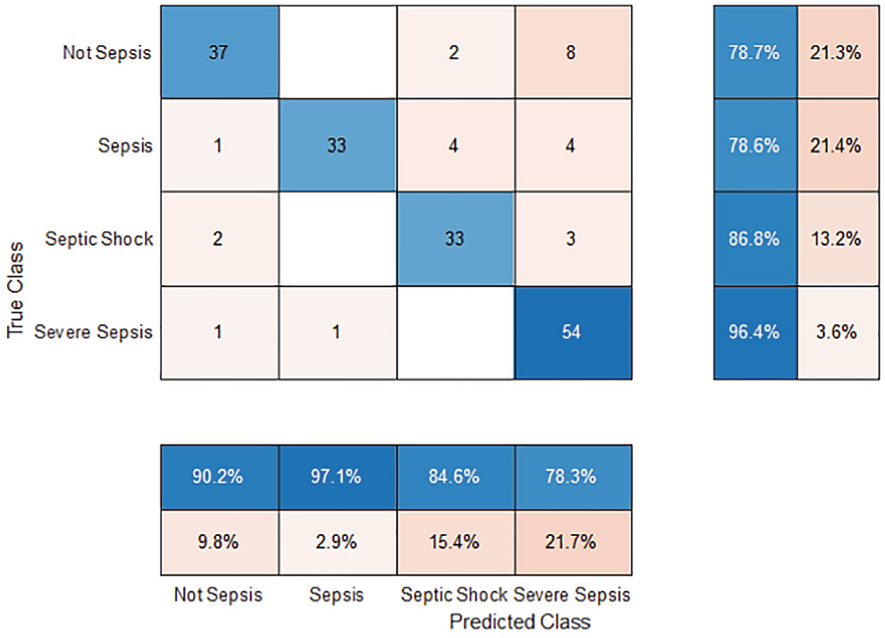


Fig. 6: Classification results of each sepsis condition and No sepsis type using LSTM model. Testing: Overall Accuracy = 85.79%

4 Conclusion and Future Work

This article presents a DL model based on the LSTM layer to detect and classify PPG signals concerning the sepsis condition. Our model has the advantage of

requiring a single cycle of PPG to classify the sepsis condition, making it possible for almost real-time diagnosis and monitoring.

We believe that the LSTM layer is appropriate for the classification task because the PPG signal has a structure of time sequences in which the LSTM can operate. Therefore, extracting features from the single cycle of PPG (i.e., the input of the model) is well performed, allowing the fully-connected layer to find an optimized solution for the classification.

The dataset utilized in this study can be generated from the publicly available data of the *MIMIC-III Waveform Database Matched Subset* repository. For replication of our findings, please refer to the methodology outlined in Sect. 2.1. Utilizing the *MIMIC-III Clinical Database* enabled us to ascertain patient records with increased certainty.

Taking into account that an analysis of the imported tables in Postgresql must be carried out to have the data set for more precise information on patient admissions at the ICU entrance, which guarantees the accuracy of the data used for our analysis. The codes necessary for reproducing our results are available upon request.

PPG technology is advancing in precision, sensitivity, and detection of subtle signals. We have portable and wearable devices that allow continuous health monitoring. With these advances, it can be seen that PPG signals can be used to detect cardiovascular, and respiratory diseases, diabetes, sepsis, and stress.

The advantages of PPG signals include their non-invasive nature, low cost, ease of use, and portability; these challenges include sensitivity to motion, environmental interference, and inter-individual variability. With this ongoing research, the use of PPG signals in disease detection may be validated. In future work, we aim to test our model on a larger patient group, consisting of both control and case groups with varying levels of sepsis severity. Our goal is to assess the effectiveness of our method in predicting the onset time of sepsis. To achieve this, we plan to train and test the model on a dataset where diagnosis times are known.

We recognize the importance of accurate diagnosis in saving lives and providing timely treatment. We believe that this work contributes to the early detection of sepsis with deep learning, presenting an approach that requires a single cycle of PPG signal for the input of an LSTM model. As a part of our future work, we aim to compare the accuracy of our model with a physician's precision in predicting sepsis.

Acknowledgment. This work is partially supported by: NSF Grant No. OAC-1750970, NSF Award No. OIA-1849243, and MARC-UPRM Program T34GM008419.

References

1. Almarshad, M.A., Islam, M.S., Al-Ahmadi, S., BaHammam, A.S.: Diagnostic features and potential applications of PPG signal in healthcare: a systematic review. Healthcare **10**, 547 (2022)
2. Elgendi, M.: On the analysis of fingertip photoplethysmogram signals. Curr. Cardiol. Rev. **8**(1), 14–25 (2012)
3. Esgalhado, F., Fernandes, B., Vassilenko, V., Batista, A., Russo, S.: The application of deep learning algorithms for PPG signal processing and classification. Computers **10**(12), 158 (2021)
4. Fan, S.L., Miller, N.S., Lee, J., Remick, D.G.: Diagnosing sepsis - the role of laboratory medicine. Clin. Chim. Acta **460**, 203–210 (2016)
5. Hajj, J., Blaine, N., Salavaci, J., Jacoby, D.: The "centrality of sepsis": a review on incidence, mortality, and cost of care. Healthcare (Basel) **6**(3), 90 (2018)
6. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **8**, 1735–1780 (1997)
7. Johnson, A., Pollard, T., Mark, R., III.: MIMIC-III clinical database (version 1.4). PhysioNet (2016)
8. Johnson, A.E., et al.: MIMIC-III, a freely accessible critical care database. Sci. Data **3**(1), 1–9 (2016)
9. Kam, H.J., Kim, H.Y.: Learning representations for the early detection of sepsis with deep neural networks. Comput. Biol. Med. **89**, 248–255 (2017)
10. Lazazzera, R., et al.: Detection and classification of sleep apnea and hypopnea using PPG and SpO_2 signals. IEEE Trans. Biomed. Eng. **68**(5), 1496–1506 (2020)
11. Lecun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proc. IEEE **86**(11), 2278–2324 (1998). https://doi.org/10.1109/5.726791
12. Lombardi, S., et al.: Classifying sepsis from photoplethysmography. Health Inf. Sci. Syst. **10**(1), 30 (2022)
13. Makowski, D., et al.: NeuroKit2: a Python toolbox for neurophysiological signal processing. Behav. Res. Meth. **53**(4), 1689–1696 (2021). https://doi.org/10.3758/s13428-020-01516-y
14. Marik, P.E., Taeb, A.M.: SIRS, qSOFA and new sepsis definition. J. Thorac. Dis. **9**(4), 943–945 (2017)
15. Moody, B., Moody, G., Villarroel, M., Clifford, G., Silva, I.: MIMIC-III waveform database matched subset (version1. 0). PhysioNet (2020)
16. Moor, M., Rieck, B., Horn, M., Jutzeler, C.R., Borgwardt, K.: Early prediction of sepsis in the ICU using machine learning: a systematic review. Front. Med. **8**, 607952 (2021)
17. Newman-Toker, D.E., et al.: Diagnostic errors in the emergency department: a systematic review (2022)
18. Prabhakar, S.K., Rajaguru, H., Kim, S.H.: Fuzzy-inspired photoplethysmography signal classification with bio-inspired optimization for analyzing cardiovascular disorders. Diagnostics **10**(10), 763 (2020)
19. Sapra, A., Malik, A., Bhandari, P.: Vital sign assessment. https://www.ncbi.nlm.nih.gov/books/NBK553213/
20. de Trafford, J., Lafferty, K.: What does photoplethysmography measure? Med. Biol. Eng. Comput. **22**(5), 479–480 (1984)
21. Xie, C., McCullum, L., Johnson, A., Pollard, T., Gow, B., Moody, B.: Waveform database software package (WFDB) for Python (2023)