

# Efficient sepsis detection using deep learning and residual convolutional networks

Ahmed S. Almasoud<sup>1</sup>, Ghada Moh Samir Elhessewi<sup>2</sup>, Munya A. Arasi<sup>3</sup>, Abdulsamad Ebrahim Yahya<sup>4</sup>, Menwa Alshammeri<sup>5</sup>, Donia Badawood<sup>6</sup>, Faisal Mohammed Nafie<sup>7</sup> and Mohammed Assiri<sup>8</sup>

<sup>1</sup> Department of Information Systems, College of Computer and Information Sciences, Prince Sultan University, Saudi Arabia

<sup>2</sup> Department of Health Sciences, College of Health and Rehabilitation Sciences, Princess Nourah bint Abdulrahman University, Saudi Arabia

<sup>3</sup> Department of Computer Science, Applied College at RijalAlmaa, King Khalid University, Saudi Arabia

<sup>4</sup> Department of Information Technology, College of Computing and Information Technology, Northern Border University, Arar, Saudi Arabia

<sup>5</sup> Department of Computer Science, College of Computer and Information Sciences, Jouf University, Saudi Arabia

<sup>6</sup> Department of Data Science, School of Computers, Umm Al-Qura University, Saudi Arabia

<sup>7</sup> Department of Natural and Applied Sciences, Community College, Majmaah University, Al Majma'ah, Saudi Arabia

<sup>8</sup> Department of Computer Science, College of Computer Engineering and Sciences Prince Sattam Bin Abdulaziz University, Al-Kharj, Saudi Arabia

## ABSTRACT

Sepsis is a life-threatening complication caused by infection that leads to extensive tissue damage. If not treated promptly, it can become fatal. Early identification and diagnosis of sepsis are critical to improving patient outcomes. Although recent technological advancements have aided sepsis detection, challenges remain in timely diagnosis using standard clinical practices. In this article, we present a new deep learning model to detect the occurrence of sepsis and the African vulture optimization algorithm (AVOA) to enhance the model performance. The system comprises four crucial steps: First, the enhanced convolutional learning framework (ECLF) with atrous convolutional and multi-level strategies that aim to learn high-level features from the nonlinear mapping of the medical data. Second is the spatio-channel attention network (SCAN), which has a neural architecture designed to focus on significant regions, such as spatial and channel regions, but not restricted to them. Third is the hierarchical dilated convolutional block (HDCB), which utilises a stacked dilated deep convolutional architecture for spatial feature context retrieval. Last is the residual path convolutional chain (RPCC), which uses a multi-residual convolutional approach for feature propagation, preserving important information. The sepsis detection model we bring forth involves many components, as mentioned above, and thus achieves a higher accuracy for timely intervention during sepsis. The combination of AVOA into the model ensures that it is robust and easily transferable, delivering high performance for adaptation to complicated structures inside medical datasets. The proposed model was evaluated on a clinical dataset and achieved outstanding performance, with an accuracy of 99.4%, precision of 98%, recall of 99.2%, F1-score of 99.0%, and an area under the curve (AUC) of 0.998. These results demonstrate the model's superior ability to detect sepsis accurately and reliably.

Submitted 28 February 2025

Accepted 23 May 2025

Published 1 July 2025

Corresponding author

Abdulsamad Ebrahim Yahya,  
Abdulsamad.qasem@nbu.edu.sa

Academic editor

Davide Chicco

Additional Information and  
Declarations can be found on  
page 29

DOI 10.7717/peerj-cs.2958

© Copyright  
2025 S Almasoud et al.

Distributed under  
Creative Commons CC-BY 4.0

OPEN ACCESS

outperforming traditional clinical scoring methods and conventional machine learning approaches.

**Subjects** Algorithms and Analysis of Algorithms, Artificial Intelligence, Data Mining and Machine Learning, Data Science, Optimization Theory and Computation

**Keywords** Sepsis detection, Deep learning, African vulture optimization algorithm, Enhanced convolutional learning, Spatio-channel attention, Dilated convolution

## INTRODUCTION

Sepsis is a critical health condition resulting from an abnormal response to infection. It ranks as the sixth leading cause of death worldwide, responsible for more than 11 million fatalities each year. Early fluid therapy, antibiotics, and source control management are essential to improve survival rates. However, diagnosing sepsis is challenging due to its complex presentation and rapidly changing symptoms. Current diagnostic methods often lack sensitivity and fail to predict outcomes accurately, making the detection process time-consuming and unreliable ([Rudd et al., 2020](#)). Since septic shock is the most severe form of sepsis, the management of sepsis must begin early through fluid and antibiotic therapy as well as source control. Unfortunately, sepsis is more complex than it seems, owing to its very disparate characteristics, and many of the current approaches aimed at determining whether an individual has sepsis are either inadequately sensitive or inaccurately predict the sepsis outcome, rendering the sepsis detection process time-consuming. As a result of this rapid adoption, predictive algorithms have been developed and utilised to improve the likelihood of finding sepsis at the right time ([Henry et al., 2015](#)). Currently, the approach recommended for diagnosis of neonatal sepsis involves bacteriologic examination of blood cultures performed before commencing treatment with antibiotics. Nevertheless, sepsis, unable to be treated in time, can have very non-defining symptoms and lead to organ failure within no time. That is why, while the cultures are pending, patients are usually given empirical treatment with ampicillin, gentamicin, or cefotaxime ([Hussaini, 2021](#)).

Even though there are clinical scoring systems like Sequential Organ Failure Assessment (SOFA) and Simplified Acute Physiology Score II (SAPS II), such methods have many shortcomings. Their dependence upon empirical findings and antiquated markers lowers their applicability in contemporary medicine, particularly considering the accelerated course of sepsis that hinders the development of data relative to the patient as opposed to chronic diseases where such illness develops ([Hu et al., 2022](#)). Hence, estimating the mortality rate in cases of sepsis continues to be complex. To appreciate these limitations, using machine learning (ML) methods to improve estimating sepsis outcomes has become widespread ([Hu et al., 2022](#)). Moreover, since the definition of sepsis is poorly defined in the literature, which makes comparing various predictive pre-model and post-model surveys even more complex, the lack of well-defined, cross-validation enabled, multi-centre databases has made it impossible to validate sepsis models externally ([Wong et al., 2021](#)). Sepsis is still a leading cause of death, where cases remain unabated, resulting in death rates

that are over 30%. Long-term complications such as physical disability, cognitive disability, and psychological disability are common among survivors of sepsis. A series of disabilities powerfully puts a strain on the healthcare systems. The Surviving Sepsis Campaign states that one year after a discharge, 33% of survivors display signs of cognitive impairment, 43% are not able to perform some tasks, and 27% have post traumatic stress disorder (PTSD) (*Iwashyna et al., 2010*).

Immune system malfunction is widely seen in sepsis. The infection affects the immune system with such force that the organism cannot fight off invading microorganisms. Although some antibiotics work on sepsis with positive results, the shipment of fighting the infection can lead to an attack of resistance. There are several issues with early diagnosis, but the recent paradigm shift in the concept of the disease provides sufficient information for its explanation and selection of effective antibiotics (*Singer et al., 2016*). While space-limited clinical notes filled with typos/abbreviations/free text are complicated, structured clinical information such as age, lab tests, and vital signs is pretty straightforward. Helpful information is retrieved from this unstructured data by employing natural language processing techniques; these often require clinicians' intervention (*Spasic & Nenadic, 2020*). Considering the latest technology trends, this article proposes improving sepsis detection using artificial intelligence (AI) algorithms, which can find more relevant sepsis indicators and risk factors from large clinical datasets, including echocardiograms, electrodes, and many others. Although AI also faces this problem, it is solved through the fast analysis of vast datasets, which minimises errors related to human interpretation and estimate while enhancing the speed and reliability of dosages (*Giordano et al., 2021*).

The widespread adoption of electronic health records (EHRs) has made it feasible to incorporate machine learning and data mining techniques into the fight against sepsis. Unfortunately, tools like quick Sepsis-related Organ Failure Assessment (QSOFA), Modified Early Warning Score (MEWS), National Early Warning Score (NEWS), and Systemic Inflammatory Response Syndrome (SIRS) also have problems. They do not predict sepsis's occurrence, making it difficult for early treatment (*Bone et al., 1992*). The treatment of sepsis is vital, mainly because the more extended treatment is delayed, the chances of death increase by 4 to 8% each hour (*Parlato et al., 2018*). There have been approaches targeting different possible biomarkers to diagnose sepsis, but none have proven practical in clinical settings (*Parlato et al., 2018*). Due to the outstanding complexity of sepsis, which is dynamic and includes both the infection and the host response, accurate prediction is difficult (*Iskander et al., 2013*).

Sepsis continues to be an important health issue threatening people around the world, causing more than 11 million deaths each year—surpassing many other infectious diseases in mortality. Even with technological advancements, timely detection is still out of reach because of the disease's intricate progression and vague early symptoms. SOFA and SAPS II scores are more traditional diagnostic methods that heavily depend on clinical benchmarks and slow lab results, hitting a brick wall when trying to support immediate action. In this research, a new framework is proposed, guided by AI technology that incorporates real-time clinical data through deep learning constructs for precise and early

sepsis identification. With this model, real-life problems where healthcare professionals can execute treatment hours in advance are the focus, saving thousands of people, which can be considered the most valuable feature. ICU and in-hospital settings are not the only areas where this research can be helpful. It can significantly change the odds of patient survival, lessen long-term disabilities among patients, and lessen the burden on healthcare systems globally. Delayed treatment is known to escalate mortality rates by 8% every hour, so integrating such predictive models with electronic health record's label systems will alter clinical thinking for the better—lowering mortality rates and enhancing the quality of life for patients around the globe.

The significant contributions of this article are,

1. This work proposes a novel hybrid deep learning model for early sepsis detection, combining convolutional neural networks (CNN) with spatio-channel attention mechanisms and residual path convolutional chains.
2. This model is claimed to overcome the drawbacks of conventional clinical scoring systems by exploring multidimensional clinical data, such as structured patient data, clinical notes, and unstructured data.
3. Moreover, the article provides the African vulture optimization algorithm (AVOA) approach for model tuning and performance enhancement efforts. The model aims to provide a reasonable prediction of the time of sepsis development and, thus, improve patient treatment by preventing sepsis.

This study has taken a quantitative approach using deep learning algorithms and analytics on clinical datasets, although the application of mixed-method strategies is recognised. For the case of sepsis detection, a mixed approach—blending machine analysis with clinical expert narratives—can be helpful for issues related to explainability, model output confidence, deployment, and acceptance by healthcare providers. However, because this effort's scope is technical and focused on the design, optimization, and performance evaluation of clinical models with structured clinical datasets, qualitative elements such as clinician interviews, surveys, or ethnographic studies were not included in this work. This blended approach may help future research explore gaps focused on practical barriers and user-centred design evaluation in clinical workflows. SCr, SOFA, SIRS, and MEWS as clinical scoring systems give a relatively weak response both temporally and in sensitivity. They also rely heavily on the clinical input and/or laboratory results, which are usually slow and unmanaged. Given the nature of these systems, they become less valuable in critical situations where prompt action is needed. In addition, many current machine learning approaches also lack the ability to deal robustly with incomplete and heterogeneous patient data, or they deal with the explicit temporal patterns in clinical data observations.

To address these challenges, the model proposes a hybrid deep learning architecture that includes spatio-channel attention mechanisms, hierarchical dilated convolutional blocks, and residual path convolutional chains. This configuration permits precise multi-scale feature extraction and guided attention learning, improving predictability

while ensuring interpretability. Furthermore, model accuracy and generalization are optimized using the AVOA for hyperparameter tuning, advancing model performance. Thus, this study builds an advanced yet clinically useful model designed to enable reliable and prompt detection of sepsis in operational healthcare settings, closing gaps in computation and clinical focus within the literature.

The article is organised as follows: “Related Works” surveys the literature on sepsis detection based on traditional approaches and those relying on machine learning. “Methods and Materials” presents the details of the hybrid deep learning architecture, including architecture, preprocessing and optimisation. “Result and Discussion” presents the cases used in the experiments and evaluates the results, including the evaluation with baseline models, ablation studies, and exploring hyperparameters settings. “Conclusion” provides the article’s conclusions, indicating the prospects offered by the new model for improved sepsis detection and further research development.

## RELATED WORKS

Using critical health indicators and laboratory results, usually available within several hours of an emergency admission, [Faisal et al. \(2018\)](#) created a logistic regression model (CARS) to forecast the likelihood of sepsis. An ML model for sepsis clinical decision assistance in the ED was constructed by [Horng et al. \(2017\)](#) using a linear support vector machine. They showed that incremental improvements include free text input, vital signs, and demographic data ([Horng et al., 2017](#)). This methodology relies on the rapid collection of patients’ medical parameters. Attempting to gather patient data from a more extended period, the second model prediction of long-term sepsis explored the prediction. This study considered 33 patients with an average age of 49. Of these, 19% were female, and the remaining patients were male. A more intricate framework and suitable medical parameters are necessary to detect long-term sepsis. A technique that aids in the early identification of various degrees of sepsis was created by [Burdick et al. \(2017\)](#).

The exact location also provided the data for a distinct cohort used for temporal validation. All visits to the emergency department (ED) between March 1, 2018, and August 31, 2018, were included in the cohort, not only inpatient admissions. Everything from the variables used to define the outcomes to the criteria for inclusion and exclusion remained constant. Interactions that started in the emergency department but did not lead to inpatient admission were included in the temporal validation cohort, in contrast to the model development cohort ([Hochreiter & Schmidhuber, 1997](#)). The report emphasises that sepsis is a significant problem in medicine all around the world, causing millions of cases and deaths annually, especially in susceptible groups like youngsters. This highlights the critical need for better sepsis diagnosis tools ([Daothong, Jampa-ngern & Senavongse, 2024](#)). The intensive care unit (ICU) makes sepsis diagnosis even more challenging because many patients there have other illnesses that present with identical laboratory and physiological changes as sepsis ([Pierrickos et al., 2020](#)).

ICU outcomes and sepsis prediction have long piqued the curiosity of medical professionals. Computer scientists are becoming more interested in finding practical answers to this challenge because of its importance. Modern computers’ improved

processing power, along with the encouraging results seen by AI and deep learning techniques in various contexts, has cleared the way for more efficient and effective methods of studying and analysing sepsis (Vellido *et al.*, 2018). If septic patients can be identified early and given the right drugs when they need them, their prognosis and chances of survival will be much improved (Marik & Farkas, 2018). Because of the growing complexity and possible impact, it is crucial to comprehend the implementation policy layer, which records the operational restrictions, reaction protocols, and clinical workflow. Specifically, the area under the receiver operating characteristic curve, which is one of the most popular ML evaluation methods, frequently disregards the impact of this policy layer on model performance (Reyna, Nsoesie & Clifford, 2022).

A significant contributor to healthcare costs, sepsis is a leading cause of hospital and ICU admissions, as well as morbidity and mortality. To prevent additional organ dysfunction and restore a normal circulating blood volume, intravenous fluids and/or vasopressors are essential for sepsis treatment. The problem is that human doctors have difficulty figuring out when and how much of these therapies to administer (Van der Ven *et al.*, 2022). The targeted real-time early warning score (TREWScore) was suggested by Henry *et al.* (2015) as a means of predicting septic shock using the Cox-proportional hazard model. When specific clinical characteristics change over time, this support system will sound an alarm to alert the patient to the possibility of septic shock. The suggested model attained a discrimination ability of 0.83. This analysis does not address the likelihood of mortality and fails to account for the impact of preexisting illness conditions. When it comes to sepsis, the majority of expert systems treat clinical aspects separately. However, physiological variables are probably not independent (Henry *et al.*, 2015). Significant constraints, such as a lack of validation and variations in the definition of sepsis, limit their generalisability and prevent any of these models from being widely employed in the clinic, even if some demonstrate excellent discriminative power (Kang *et al.*, 2008).

One of the biggest problems with studying sepsis in emergency departments is the absence of trustworthy labels. Due to the absence of a universally accepted standard, sepsis modelling labels are currently generated using diagnostic tools and claims-based procedures. However, it is well-known that these tagging approaches are not ideal (de Hond *et al.*, 2022). While this approach did a decent job of making predictions, it could not detect changes that occurred over time since it did not consider patients' temporal data. Further increases in predicting performance were likely limited by the model's simplicity, which hindered its capacity to detect complicated data patterns. One model, Gong *et al.*'s (2022), integrated time-series data with four vital signs to predict near-term mortality risks in sepsis patients. However, this model only considered a few variables. A total of ninety-two ICU patients had their blood tested for the start of sepsis using real-time polymerase chain reaction (RT-PCR) expression and genetic network analysis. In this investigation, 83.09% of cases could be predicted 1 to 4 days before the clinical diagnosis, with a specificity of 80.20% and a sensitivity of 91.43% (Lukaszewski *et al.*, 2008). Raghu *et al.* (2017) utilised a Duelling Double-Deep Q Network in continuous space to develop medical treatment plans for sepsis. This method expanded sepsis treatment to a continuous space using a vector

representation of constant conditions (Raghu *et al.*, 2017). Sun *et al.* (2021) used a hybrid approach that combined supervised learning and reinforcement learning using the deep deterministic policy gradient (DDPG) method to create strategies in a continuous value space.

The study to trace the interactions of berberine with inflamed molecular targets, proteins, and pathways of interest (Wang *et al.*, 2024; Sun *et al.*, 2025). This allows simultaneous detection of both pathogens in one assay without separate reactions (Zhang *et al.*, 2025).

Using data from electronic health records (EHRs), (Lauritsen *et al.*, 2020) created a deep-learning model to identify sepsis in its early stages. They drew on electronic health record data collected retrospectively from various Danish hospitals, encompassing ICUs and other departments over 7 years. Time stamps from things like lab tests, notes, prescriptions, and more were included in the data. They used sparse vectors to represent each occurrence. They built a convolutional neural network long short-term memory (CNN-LSTM) model consisting of CNN layers to capture temporal dependencies and extract features from event sequences. They pitted it against multilayer perceptron (MLP) and gradient boosting on vital signs (GB-Vital), two baseline models. At 3 h before the start of sepsis, the CNN-LSTM model outperformed the baselines with an AUROC of 0.856. Even before the beginning, 24 h before, the area under the receiver operating characteristic (AUROC) was 0.756. To measure the model's value across various forecasts made throughout a patient's hospital stay, they suggested a novel "sequence evaluation" method. They also checked the algorithm's usefulness by counting the number of hopeful forecasts already receiving antibiotics or having blood cultures taken to determine how much sooner the model could have assisted. To summarise, they created a sequential deep learning model for early sepsis diagnosis that was more effective than previous models and suggested new ways to evaluate its clinical usefulness. The concept has the potential to be applied to several departments within the hospital (Lauritsen *et al.*, 2020).

## METHODS AND MATERIALS

This research aims to develop a method for enabling early and timely detection of sepsis, which is a significant clinical challenge because it increases the risk of death by 4% to 8% with every hour treatment is delayed. Routinely used clinical scoring systems disregard aiding the patient until sepsis is well underway, and then, treatment options are minimal and outcomes are grim. Unlike the proposed model, which endeavours to predict sepsis at the earliest possible stage using real-time analyses of continuously collected clinical data, including vital signs, lab results, and demographic details. The incorporation of sophisticated deep learning components such as the spatio-channel attention network (SCAN) and the hierarchical dilated convolutional block (HDCB) enables the model to detect the most subtle early-stage physiological alterations effectively, ensuring it is fit for use in clinical settings that require immediate action. Moreover, applying the African vulture optimisation algorithm (AVOA) further augments the model's accuracy and generalisation strength across diverse clinical settings and patient populations.

## Data collection

The Kaggle dataset ([Hussaini, 2021](#)) contains clinical data from patients diagnosed with sepsis, including vital signs, laboratory results, demographics, and other relevant clinical information. To prepare the data for models developed with deep learning, it is pre-processed to deal with missing values, normalise numerical parameters, and encode variable categories.

## Experimental setup

Python software was used for the experiments in a system with an Intel(R) Core(TM) i5-4590S CPU, 1TB HDD and 8GB of RAM memory.

## Dataset access

The dataset can be accessed at the following URL: <https://www.kaggle.com/datasets/salikhussaini49/prediction-of-sepsis>.

The dataset includes 44 clinical features, grouped as three vital signs, lab results, and demographic information.

Vital signs encompass measurements of heart rate (HR), oxygen saturation (O2Sat), body temperature (Temp), blood pressure (systolic SBP, mean MAP, and diastolic DBP), respiration rate (Resp), and EtCO2. Laboratory diagnosis encompasses white blood cell count, platelet count, lactate, bilirubin, creatinine, glucose, and other hormones related to organ function and infection. Demographic data contain the patient's age and gender and the time of hospitalization.

Each patient's information is carefully arranged as a series of time-stamped events, illustrating the evolution of clinical conditions through an evolving timeline. The target variable, SepsisLabel, marks the presence or absence of sepsis at each timestamp, allowing us to build models capable of predicting the onset of sepsis using past clinical data. The ordered characteristic of the dataset allows for the creation of predictive models that can identify the onset of sepsis in advance, given the urgency of the required clinical action.

## Assessment metrics

The evaluation metrics used to assess the proposed model—accuracy, precision, recall, F1, and AUC, classify its early sepsis detection capabilities as exceptional. Having an accuracy of 99.4%, the model also indicates the power of classifying the majority of cases correctly. At 98.7%, precision alone assures the model's reliability in correctly labeling actual cases of sepsis and lowering false positives. Recall value of 99.2% demonstrates the model's ability of capturing nearly all true sepsis cases, something that is important in clinical practice in order to not miss any diagnoses. The AUC of 0.998 which reflects the model's ability to discriminate sepsis from non-sepsis cases also validates the model's performance towards other metrics such as F1 and Recall. These metrics yield evidence of the model's reliability, generalizability, and applicability in deployment in healthcare for early sepsis detection, a life-critical task.

### Data preprocessing

Preprocessing is critical to ensure that the raw medical data is suitable for feeding into deep learning models. The preprocessing phase includes several key operations: handling missing data, normalisation, feature encoding, and outlier detection. Each step transforms the data into a consistent format and scale, allowing the deep learning model to learn from it effectively. Missing values are common in medical datasets, and handling them appropriately is crucial for the quality of the model. Missing values are filled in using statistical methods such as mean, median, or mode imputation. For numerical features with missing values, the missing entry is replaced using the mean of the available values, excluding the missing entry. In some cases, advanced techniques like multiple imputation by chained equations (MICE) or k-nearest neighbours (KNN) imputation can be employed if the missingness is not random. Medical data can span different ranges, with some features having significantly larger scales than others (e.g., heart rate *vs.* body temperature). Normalising and scaling the data ensures all features are treated equally in the model's learning process. The Min-Max Scaling method usually scales the data to a specified range [0, 1].

The Z-score normalization (standardization) method standardises the feature with a mean of 0 and a standard deviation of 1. Categorical variables were converted into numerical representations using one-hot encoding to ensure compatibility with the deep learning model. Numerical features were standardised using Z-score normalization to maintain consistent scaling across features. A common approach is the Z-score method, where data points that deviate more than a specified threshold from the mean are considered outliers.

Deviation. If  $|Z_i| > 3$ , the data point is flagged as an outlier and can be removed or adjusted. To further improve the robustness of the model, data augmentation techniques, such as noise injection or random cropping, may be applied, especially for limited datasets. This can help the model generalize unseen data better and prevent overfitting. By applying these preprocessing techniques, the dataset is transformed into a format that can be efficiently used for training deep learning models, ensuring better performance in the subsequent detection of sepsis.

$$x_i = \frac{1}{n} \sum_{i=1}^n x_i \text{ if } x_i \text{ is missing} \quad (1)$$

where  $\frac{1}{n} \sum_{i=1}^n x_i$  is the mean of the feature across the dataset. For time-series data, missing values are filled based on the previous or next observed values:

$$x_i = x_{i-1} \text{ or } x_i = x_{i+1} \text{ if } x_i \text{ is missing} \quad (2)$$

In some cases, advanced techniques like MICE or k-nearest neighbours (KNN) imputation can be employed if the missingness is not random. Medical data can span different ranges, with some features having significantly larger scales than others (e.g., heart rate *vs.* body temperature). Normalising and scaling the data ensures all features are treated equally in the model's learning process. The Min-Max Scaling method usually

scales the data to a specified range [0, 1]. For a feature  $x_i$ , the transformed value  $x'_i$  is calculated as:

$$x'_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (3)$$

where  $\min(x)$  and  $\max(x)$  are the minimum and maximum values of the feature across all samples. The Z-score normalization (standardization) method standardises the feature with a mean of 0 and a standard deviation of 1. Categorical variables were converted into numerical representations using one-hot encoding to ensure compatibility with the deep learning model. Numerical features were standardised using Z-score normalization to maintain consistent scaling across features.

The transformation for each feature  $x_i$  is given by:

$$x'_i = \frac{x_i - \mu_x}{\sigma_x} \quad (4)$$

where  $\mu_x$  is the mean of the feature and  $\sigma_x$  is the standard deviation of the feature. Standardisation is generally preferred when the model requires features to be on the same scale, especially for algorithms like gradient-based optimisation, which are sensitive to feature scaling. Categorical variables in the dataset, such as gender or ethnicity, need to be transformed into numerical values that the deep learning model can understand. Common methods include:

In One-Hot Encoding, each category is represented by a binary vector. For example, a categorical feature with three possible values (e.g., “Male,” “Female,” “Other”) would be encoded as:

$$\text{Male} \rightarrow [1, 0, 0], \text{Female} \rightarrow [0, 1, 0], \text{Other} \rightarrow [0, 0, 1] \quad (5)$$

This method ensures that the model does not assume ordinal relationships between categories. In Label Encoding, each category is assigned an integer value. For example, the categories “Male,” “Female,” and “Other” could be encoded as:

$$\text{Male} \rightarrow 0, \text{Female} \rightarrow 1, \text{Other} \rightarrow 2 \quad (6)$$

Label encoding is often used when the categorical feature has an inherent order. Still, it is not recommended for features with no ordinal relationship, as the model may misinterpret the encoding. Outliers can significantly distort training, especially in sensitive models like deep learning. Outlier detection and removal techniques help mitigate this issue. A common approach is the Z-score method, where data points that deviate more than a specified threshold from the mean are considered outliers. The Z-score for a feature  $x_i$  is computed as:

$$Z_i = \frac{x_i - \mu_x}{\sigma_x} \quad (7)$$

where  $\mu_x$  is the mean and  $\sigma_x$  is the standard deviation. If  $|Z_i| > 3$ , the data point is flagged as an outlier and can be removed or adjusted. To further improve the robustness of the model, data augmentation techniques, such as noise injection or random cropping,

may be applied, especially for limited datasets. This can help the model generalise unseen data better and prevent overfitting. By applying these preprocessing techniques, the dataset is transformed into a format that can be efficiently used for training deep learning models, ensuring better performance in the subsequent detection of sepsis.

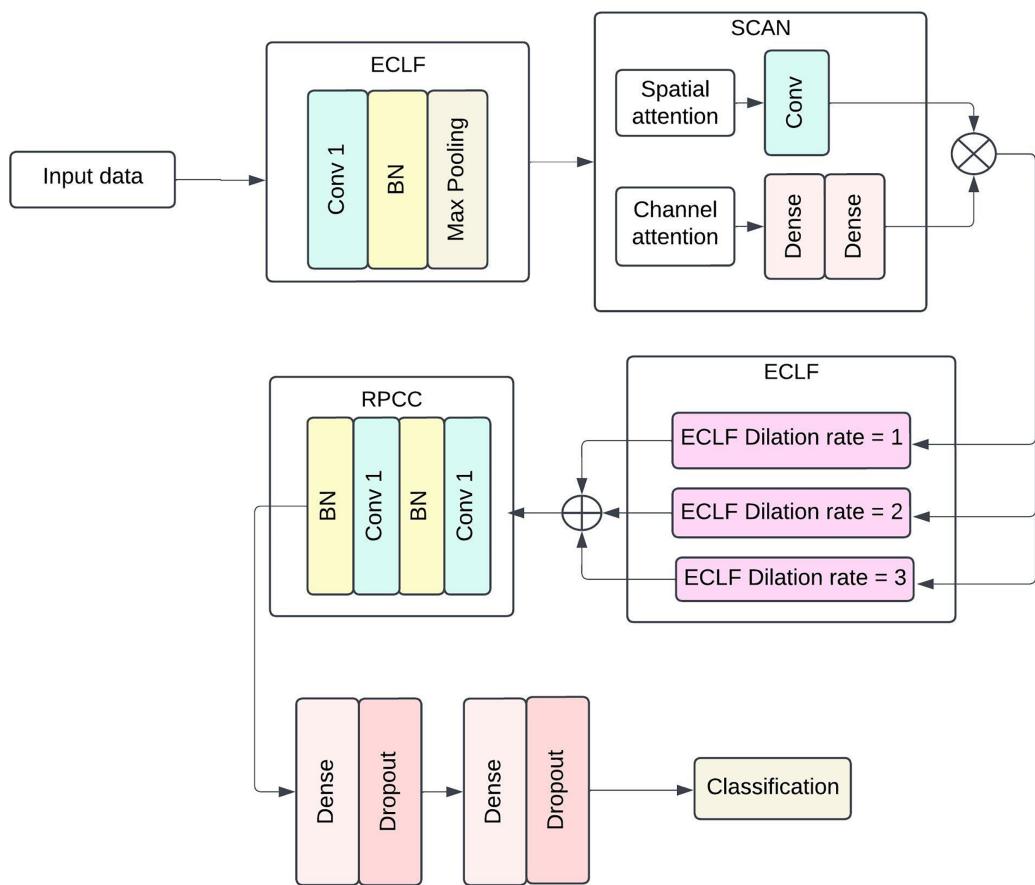
## Experimental setup

The experimental design for this work involved leveraging the Kaggle Sepsis Prediction dataset, which has 1,552,210 records from 40,336 unique patients. The target variable SepsisLabel marks whether sepsis is present or not, and has a notable class imbalance: 1,524,294 samples flagged as non-sepsis and 27,916 flagged as sepsis. This dataset is nearly representative of actual clinical situations, where early identification of sepsis is difficult because it is infrequently encountered. As a preparatory step towards model training, mean imputation was applied for handling missing values for numerical features, along with forward filling for time-series data where applicable. Categorical variables were transformed into one-hot variables. Numerical features were standardized using Z-score normalization to ensure uniform feature scaling. Using Z-score outliers detection, outliers were capped at three standard deviations. The dataset was split into three disjoint subsets for a patient-wise split to prevent data leakage: 80% for training (1,241,768 samples), 10% for validation (155,221 samples), and 10% for testing (155,221 samples). This ensures that all records from a single patient are not present in multiple subsets. The suggested architecture integrates state-of-the-art components of deep learning, namely enhanced convolutional learning framework (ECLF), spatio-channel attention network (SCAN), hierarchical dilated convolutional block (HDCB) and residual path convolutional chain (RPCC). Model training was carried out on the Adam optimizer with a starting learning rate of 0.001, 128 batch size, and regularization set to a 0.5 dropout rate. The network was trained with ReLU activation and overfitting mitigated by early stopping based on validation loss, capping training at 100 epochs. The model's performance was assessed on multi-dimensional accuracy, precision, recall, F1-score, area under the ROC curve (AUC), specificity, FPR and FNR. All experiments were implemented on a Python 3.8 environment with TensorFlow and Keras libraries, running on a system with NVIDIA GPU and sufficient memory to handle the computations.

## Proposed methodology

The proposed model structure intends to capture more complex patterns and spatial-temporal dependencies in medical data for efficient sepsis detection. The architecture includes four broad components: ECLF, SCAN, HDCB and RPCC. Each module uses advanced techniques involving convolutional layers, attention mechanisms, and residual connections to improve the model's performance and facilitate feature extraction and model interpretability. [Figure 1](#) shows the architecture of the Proposed methodology with ECLF, SCAN, HDCB and RPCC.

The enhanced convolutional learning framework, spatial-channel attention network, hierarchical dilated convolutional block, and residual path convolutional chain collectively form the advanced components of the deep learning architecture for the purposed sepsis



**Figure 1** The architecture of proposed methodology with Enhanced Convolutional Learning Framework (ECLF), Spatio-Channel Attention Network (SCAN), Hierarchical Dilated Convolutional Block (HDCB), and Residual Path Convolutional Chain (RPCC).

[Full-size](#) DOI: 10.7717/peerj-cs.2958/fig-1

detection model. The floodgates to the model lie on the been trained on enhanced convolutional learning framework, which stacks on batch normalisation layer after its ReLU activated first layer with a dilation of 1 using  $64 3 \times 3$  convolutional filters; this layer is then followed closely by a MaxPooling layer with a pool size of  $2 \times 2$ . Most importantly, the model's performance does not significantly deteriorate by an increase in the number of parameters as local and global diagonal convolutions seem to be covered in the ECLF. Following that, the SCAN module serves its purpose quite efficiently; this module has two attention mechanisms: spatial attention and channel attention. The former requires the kernel to be set at  $7 \times 7$  and only applies a 2D attention map with a single channel to the sigmoid output. In contrast, the latter first performs pooling, which is then set through two biasing-connected layers with 64 set units, and the second layer matches the units to the channels.

For the HDCB,  $3 \times 3$  convolution at 64 filters at a dilated rate of 1, 2, and 3 is performed on the model. This method allows the network to process information from dilated convolutions by combining features extracted at multiple receptive field sizes, facilitating

the capture of local and global contextual information. This method helps the network to use information obtained from the dilated convolutions in a multifaceted manner, as it helps aggregate features obtained from several receptive field sizes. Once the results of all dilated convolution processes have been concatenated, a  $1 \times 1$  Convolution having 128 filters is applied on the output with batch normalisation followed by ReLU activation functions to enhance diversity in the production by compressing facets from multiple scales together. The RPCC block supports residual connections in which each block always has two layers of  $3 \times 3$  convolutional blocks with 128 kernel filters and a layer of batch normalisation. The output of the last convolutional layer is fed back to the input layer *via* a skip connection, which lets the model preserve such information in deeper layers. Thus, in conjunction with this, the vanishing gradient problem is also solved, and properties are correctly disseminated inside the network.

In the last stage, the model is global average pooled, and then fully dense output layers are incorporated. Standard activation functions and normalization layers were incorporated throughout the architecture to ensure stable and efficient training. There are dense layers of 512 units and 256 units, and both are followed by ReLU activation functions accompanied by layers of dropout, which is set to 0.5 for regularization purposes. Whereas the final layer utilises a single neuron with binary functions and sigmoid activation as it encapsulates jumping off whether sepsis exists or not. This architecture can generalise well across different patients' data while still being able to learn intricate patterns within clinical data at the same time.

### ***Enhanced convolutional learning framework***

The ECLF is the first module in the model architecture, designed to capture both local and global features from medical data using advanced convolutional techniques. It employs atrous convolutions (or dilated convolutions) and a multi-level learning process to learn hierarchical feature representations from the input data. This framework enables the model to capture rich and diverse patterns at different scales, which is critical for detecting complex clinical data patterns that indicate sepsis. Atrous convolution helps in increasing the receptive field without introducing additional parameters. This is especially useful when working with medical time-series data, as it allows the model to capture context from more significant regions without the need for huge kernels. The dilated convolution operation for a feature map  $X$  can be defined as:

$$Y_{i,j,k} = \sum_{m,n} X_{i+m \cdot d, j+n \cdot d, k} \cdot W_{m,n,k} \quad (8)$$

where,  $Y_{i,j,k}$  is the output of the convolution for the location  $(i, j)$  in the feature map,  $X_{i,j,k}$  is the input feature map,  $W_{m,n,k}$  is the convolution kernel,  $d$  is the dilation rate (controls the spacing between kernel elements),  $m$  and  $n$  are the filter coordinates.

The dilation rate  $d$  expands the kernel by inserting zeros between its elements, allowing the model to cover a larger area without increasing the number of parameters. This is important for capturing long-range dependencies within the data.

### Multi-level learning

The ECLF employs a multi-level learning process to capture features at different scales and increase the depth of the feature extraction process. This process applies convolutions with varying kernel sizes to extract information at various levels of granularity. Each framework level is designed to focus on different spatial scales, enabling the algorithm to detect both broad and specific trends in the provided data.

For example, the input feature map  $X \in \mathbb{R}^{H \times W \times C}$  is passed through multiple convolutional layers with different kernel sizes, such as:

$$Y_1 = \text{Conv}_1(X, K_1) \quad (9)$$

$$Y_2 = \text{Conv}_2(X, K_2) \quad (10)$$

$$Y_3 = \text{Conv}_3(X, K_3) \quad (11)$$

where,  $Y_1, Y_2, Y_3$  are the feature maps obtained after applying convolutions with diverse kernel sizes  $K_1, K_2, K_3$ ,  $\text{Conv}_i(X, K_i)$  represents the convolution operation applied to the input feature map  $X$  with kernel  $K_i$ ,  $H$ ,  $W$ , and  $C$  represent the height, width, and number of channels of the input feature map  $X$ . The multi-level learning process helps the model learn more abstract and complex features, improving its ability to identify sepsis-associated patterns. The features are combined after obtaining feature maps from multiple convolutional layers with different kernel sizes. This can be done through concatenation, aggregating information across the other scales. The resulting feature map is a richer representation of the input data.

$$X_{ECLF} = \text{Concatenate}(Y_1, Y_2, Y_3) \quad (12)$$

where  $X_{ECLF}$  is the combined feature map after applying the multi-level convolutions. This combination allows the model to maintain fine-grained and high-level features, facilitating a more comprehensive understanding of the data. Non-linear activation functions (ReLU) were applied after convolutional layers to support complex feature transformations. The network is made non-linear by applying non-linear activation functions, like ReLU, after convolutional layers have been applied. This allows the model to learn complex mappings from input to output.

$$X_{ECLF} = \text{ReLU}(X_{ECLF}) \quad (13)$$

where  $\text{ReLU}(x) = \max(0, x)$  is the element-wise ReLU function, which sets all negative values to zero and passes positive values unchanged. To stabilise the training process and reduce overfitting, batch normalisation is applied after each convolutional operation. Batch normalisation normalises the output of each layer, ensuring that the network trains faster and more efficiently. The batch normalisation operation is defined as:

$$\hat{X} = \frac{X - \mu_B}{\sigma_B + \varepsilon} \quad (14)$$

where,  $X$  is the input to the batch normalisation layer,  $\mu_B$  and  $\sigma_B$  are the mean and standard deviation of the batch,  $\varepsilon$  is a small constant added to prevent division

by zero. The normalised output is then scaled and shifted using learnable parameters  $\gamma$  and  $\beta$ :

$$BN(X) = \gamma \hat{X} + \beta \quad (15)$$

The ECLF effectively learns multi-scale features from the data, enabling the model to detect complex clinical patterns indicative of sepsis. Using dilated convolutions, multi-level learning, and advanced activation functions ensures the model is powerful and efficient for sepsis detection tasks.

### ***Spatio-channel attention network***

The SCAN focuses on the feature learning enhancing process through the regions and channels in the feature map that are most significant. SCAN aspires to refine the attention maps generated by the previous layers of the network by selectively emphasising regions of importance (spatial attention) and channels of importance (channel attention). This allows the model to detect more relevant patterns in the medical data for sepsis identification, thus increasing its discriminative ability. The SCAN comprises two attention schemes: spatial attention mechanism (SAM) and channel attention mechanism (CAM).

Approaching it now sequentially, both mechanisms are employed to refine the generated feature map further, thereby increasing the model's ability to discriminate and, vice versa, the unnecessary details.

#### *Spatial attention mechanism*

The SAM aims to assist the model in concentrating on the essential parts of the feature's map. It enhances areas that have relevant information and suppresses the rest. This is crucial in medical data, where some input bits have more information than others. Let  $X \in \mathbb{R}^{H \times W \times C}$  be the input feature map, where  $H$  is the height,  $W$  is the width, and  $C$  is the number of channels. The goal of SAM is to generate a spatial attention map  $A_{\text{spatial}} \in \mathbb{R}^{H \times W \times C}$  that can be used to weight the feature map spatially.

First, a convolutional layer is applied to the input feature map  $X$  to compute the spatial attention map. This operation reduces the channel dimension of the feature map by applying a convolutional filter and then uses a sigmoid activation to obtain the attention map:

$$A_{\text{spatial}}(i, j) = \sigma(Conv_1(Conv_2(X))) \quad (16)$$

where,  $A_{\text{spatial}}(i, j)$  is the spatial attention value for each spatial location  $(i, j)$ ,  $Conv_1$  and  $Conv_2$  are convolutional layers with kernel sizes chosen to capture spatial features,  $\sigma$  is the sigmoid activation function applied to ensure the attention map values are between 0 and 1. Once the spatial attention map  $A_{\text{spatial}}$  is obtained, the feature map  $X$  is multiplied element-wise by  $A_{\text{spatial}}$ , focusing the model's attention on the most relevant spatial regions:

$$X_{\text{spatial}} = X \cdot A_{\text{spatial}} \quad (17)$$

### Channel attention mechanism

The CAM focuses on emphasising the most informative channels of the feature map. Not all channels in the feature map are equally important; some channels carry more valuable information about the presence of sepsis. The goal of CAM is to generate a channel-wise attention map  $A_{channel} \in \mathbb{R}^C$ , which can be used to reweight the channels of the feature map. The input feature map  $X \in \mathbb{R}^{H \times W \times C}$  is first passed through a global average pooling (GAP) operation along the spatial dimensions  $H$  and  $W$  to produce a global summary for each channel:

$$\hat{X}_k = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{i,j,k} \quad (18)$$

where  $\hat{X}_k \in \mathbb{R}^C$  is the channel-wise pooled representation, capturing global features for each channel  $k$ . The pooled feature vector  $\hat{X}$  is then passed through a two-layer fully connected (FC) network to compute the channel attention map  $A_{channel}$ :

$$A_{channel} = \sigma(FC_2(ReLU(FC_1(\hat{X}))) \quad (19)$$

where,  $\hat{X}$  is the pooled feature vector,  $C_1$  and  $C_2$  are fully connected layers,  $ReLU$  is the activation function applied after the first fully connected layer,  $\sigma$  is the sigmoid activation applied to obtain the final channel attention map  $A_{channel} \in \mathbb{R}^C$ , which contains values between 0 and 1 for each channel. The feature map  $X_{spatial}$  from the spatial attention mechanism is then multiplied element-wise by the channel attention map  $A_{channel}$ , focusing the model's attention on the most informative channels:

$$X_{final} = X_{spatial} \cdot A_{channel} \quad (20)$$

The SCAN module applies spatial and channel attention mechanisms to refine the input feature map in a two-stage process. By zeroing in on the most important areas, the spatial attention process improves the feature map. By highlighting the most important channels, the channel attention mechanism continuously adjusts the feature map. The final refined feature map  $X_{final}$  is the result of applying both attention mechanisms sequentially:

$$X_{final} = (X \cdot A_{spatial}) \cdot A_{channel} \quad (21)$$

This refined feature map is then passed on to the next module in the network, which may be a dilated convolutional block or residual path convolutional chain, to process the features further for sepsis detection.

### Hierarchical dilated convolutional block

The HDCB is a powerful tool for obtaining hierarchical features crucial for recognising low-level patterns associated with sepsis in medical data sets. HDCB considers local and global contexts concurrently by utilising stacked dilated convolutions, thereby processing input at different scales. The combination of parameters for the convolutional neural networks with dilation rates enables the model to learn multi-level complexities affordably. From a mathematical perspective, the model can cover a wider overall field with minimal

computational resources. In summation, HDCB utilises convolutional layers possessing multi-dilation rates to achieve a state-of-the-art contraction. A standard convolution operation with a kernel of size  $K \times K$  applied to an input feature map  $X \in \mathbb{R}^{H \times W \times C}$  can be expressed as:

$$Y_{i,j,k} = \sum_{m,n} X_{i+m,j+n,k} \cdot W_{m,n,k} \quad (22)$$

where,  $Y_{i,j,k}$  is the output feature map,  $W_{m,n,k}$  is the kernel weight,  $(m, n)$  are the coordinates of the kernel. For dilated convolutions, the dilation rate  $d$  introduces spacing between the kernel elements. The dilated convolution operation is defined as:

$$Y_{i,j,k} = \sum_{m,n} X_{i+m.d,j+n.d,k} \cdot W_{m,n,k} \quad (23)$$

where,  $d$  is the dilation rate that controls the spacing between kernel elements. The dilation rate allows the model to capture larger contexts without increasing the size of the kernel. For example, a dilation rate of  $d = 2$  will effectively double the receptive field of the kernel, capturing more distant dependencies within the feature map. The HDCB stacks several dilated convolutional layers to capture information at multiple scales, each with a different dilation rate. Let's consider  $N$  dilated convolutional layers, each with a different dilation rate  $d_1, d_2, \dots, d_N$ . The output of each layer can be written as:

$$Y_1 = \text{Conv}_1(X, d_1) \quad (24)$$

$$Y_2 = \text{Conv}_2(X, d_2) \quad (25)$$

$$Y_N = \text{Conv}_N(X, d_N) \quad (26)$$

where,  $Y_1, Y_2, \dots, Y_N$  are the outputs of each dilated convolutional layer,  $\text{Conv}_i(X, d_i)$  represents the convolution operation with a dilation rate  $d_i$  applied to the input  $X$ . Each convolutional layer captures different levels of context, with more significant dilation rates focusing on more global information and smaller dilation rates capturing local details. The outputs from the stacked dilated convolutional layers are then concatenated or summed to combine the features from different scales. The aggregated feature map  $X_{HDCB}$  can be obtained as:

$$X_{HDCB} = \text{Concatenate}(Y_1, Y_2, \dots, Y_N). \quad (27)$$

Alternatively, a summation approach can be used:

$$X_{HDCB} = Y_1 + Y_2 + \dots + Y_N. \quad (28)$$

This step combines the features learned at different scales, resulting in a richer and more hierarchical representation of the input data. The model can learn more abstract features essential for detecting sepsis by capturing local and global information. After the dilated convolutions and feature aggregation, non-linear activation functions and Batch normalization are applied. The activation function is applied element-wise:

$$X_{HDCB} = \text{ReLU}(X_{HDCB}) \quad (29)$$

where  $ReLU(x) = \max(0, x)$ . Batch normalisation is applied to the aggregated feature map to normalise the activations, improving the convergence speed and preventing overfitting. The batch normalisation operation is defined as:

$$\hat{X} = \frac{X - \mu_B}{\sigma_B + \varepsilon} \quad (30)$$

where,  $\mu_B$  and  $\sigma_B$  are the mean and standard deviation of the batch,  $\varepsilon$  is a small constant added to prevent division by zero. The normalised output is then scaled and shifted using learnable parameters  $\gamma$  and  $\beta$ :

$$BN(X) = \gamma \hat{X} + \beta. \quad (31)$$

After applying the dilated convolutions, feature aggregation, activation, and normalisation, the final output of the HDCB is the refined feature map  $X_{HDCB}$ . This feature map contains hierarchical information from local and global contexts, enabling the model to learn complex patterns indicative of sepsis.

$$X_{HDCB} = ReLU \left( BN \left( \sum_{i=1}^N Y_i \right) \right). \quad (32)$$

The HDCB enhances the model's ability to capture long-range dependencies and subtle patterns in medical data, making it particularly effective for detecting sepsis in clinical environments.

### ***Residual path convolutional chain***

The RPCC is one of the model components meant to facilitate the information flow through the network and make it easier to train intense networks. It uses their residual connections, also called skip connections, to address the vanishing gradient problem and to aid the learning of complex representations. The RPCC architecture reduces feature overfitting by permitting more efficient gradient flow to the feature maps during backpropagation, which allows for effective learning by the model. In RPCC, the feature map undergoes several convolutional operations, and a skip connection is made after each operation. These connections help a convolutional layer learn the residual over that layer alongside identity mapping while several layers are bypassed, making it more flexible to fit intricate structures. A residual connection allows input to cross single or multiple layers and then be summed up in the output, providing a significantly direct path. This contributes to retaining important data over layers, such as in deep networks. For a feature map  $X$ , the residual connection is defined as:

$$X_{residual} = X + F(X) \quad (33)$$

where,  $X$  is the input feature map,  $F(X)$  represents the output of the convolutional operation applied to  $X$ . The function  $F(X)$  is typically a sequence of convolutional layers, batch normalisation, and activation functions, and can be expressed as:

$$F(X) = Conv(ReLU(BN(Conv(X)))) \quad (34)$$

where,  $Conv$  represents a convolutional layer,  $BN$  denotes batch normalisation,  $ReLU$  is the activation function. Thus, the residual connection allows the network to learn an additional mapping  $F(X)$ , while also maintaining the original input  $X$  directly. The RPCC is built by stacking multiple convolutional layers, each with its residual connection. Given an input feature map  $X$ , the output of the first residual block is:

$$X_1 = X + F_1(X) \quad (35)$$

where  $F_1(X)$  is the output of the first convolutional block, which includes convolution, batch normalisation, and activation. The input is passed through the residual connection for subsequent residual blocks and updated at each step. The output of the second residual block is:

$$X_2 = X_1 + F_2(X_1) \quad (36)$$

where  $F_2(X_1)$  is the output of the second convolutional block, and so on. This can be generalized for  $N$  residual blocks as:

$$X_N = X_{N-1} + F_N(X_{N-1}) \quad (37)$$

where  $F_N(X_{N-1})$  is the output of the  $N$ -th convolutional block. The final output feature map is produced after passing through all the residual blocks. The RPCC ensures that essential features are preserved and propagated throughout the network. The final output of the RPCC block, denoted as  $X_{RPCC}$ , is:

$$X_{RPCC} = X_N \quad (38)$$

where  $X_N$  is the final feature map after applying all residual connections and convolutional blocks. An activation function and batch normalisation follow each convolutional operation in the residual path to introduce non-linearity and stabilise the learning process. The output of each convolutional layer is first passed through a  $ReLU$  activation to introduce non-linearity:

$$X_{activated} = ReLU(X_{RPCC}). \quad (39)$$

Batch normalisation is applied to reduce internal covariate shift and ensure more stable training:

$$\hat{X} = \frac{X - \mu_B}{\sigma_B + \epsilon} \quad (40)$$

where,  $\mu_B$  and  $\sigma_B$  are the mean and standard deviation of the batch,  $\epsilon$  is a small constant to prevent division by zero. Finally, the normalised output is scaled and shifted using learnable parameters  $\gamma$  and  $\beta$ :

$$BN(X) = \gamma \hat{X} + \beta. \quad (41)$$

In some cases, especially when the dimensions of the input and output feature maps are the same, the residual connection is an identity mapping, meaning the input is directly added to the output without any transformation. This is particularly useful in maintaining

information across layers in very deep networks. In this case, the residual connection is simply:

$$X_{\text{residual}} = X. \quad (42)$$

After passing through several residual layers, the final feature map  $X_{\text{RPCC}}$  can be obtained from the last residual block. The complete residual chain can be written as:

$$X_{\text{RPCC}} = X + F_1(X) + F_2(X) + \dots + F_N(X). \quad (43)$$

The output  $X_{\text{RPCC}}$  is the enriched feature map containing preserved and learned features across multiple residual paths. The Residual Path Convolutional Chain (RPCC) ensures efficient feature propagation. It enables the network to learn deeper, more complex features without the risk of vanishing gradients. This makes it particularly useful for detecting subtle patterns in data, such as those associated with sepsis detection.

### ***African vulture optimization algorithm pseudocode***

The AVOA is a nature-inspired optimization algorithm that mimics the foraging behavior of vultures. Vultures search for food in an environment by utilizing a combination of exploration and exploitation strategies.

## **RESULT AND DISCUSSION**

This section discusses the performance of the proposed model based on the architecture integrating ECLF, SCAN, HDCB, and RPCC. In particular, ResNet was evaluated on a sizeable clinical dataset.

As part of the model development procedure, the dataset was divided into three subsets: the test set, the train set, and the validation set. The model parameters were updated using the defined training set during the learning phase. About the learning phase, the validation set was used to monitor the model's performance and control guided hyperparameter tuning and overfitting using early stopping strategies on the model during each defined epoch. This form of intermediate assessment is termed validation accuracy and is plotted on a per-epoch basis in the respective graphs. However, the test set was exclusively held to evaluate the model after its training. Once the model's training was completed, the model's final configuration was assessed using the test set. This evaluation is known as generalization evaluation and is performed only once at the end of training. The performance obtained at this stage is referred to as testing accuracy, which is calculated to show the model's performance on completely new data.

This model outperformed the rest with an accuracy of 99.4%, which proves the model's efficacy in detecting sepsis at a precocious stage. Such models emerge as competent predictive models since they accommodate and utilise robust multi-scale feature extraction and attention mechanisms to enhance the model's accuracy. Hence, the model trained demonstrates competitive and outstanding predictions.

The loss and accuracy values for the model's training and validation were tracked throughout epochs. The training accuracy kept on enhancing and reached a commendable 99.4% mark during the final phase of the training. Alongside this, validation accuracy continued to depict a similar trend, hence indicating that the model in question is not

**Algorithm 1** African vulture optimization algorithm.

```

1: Initialise population  $P$  of vultures (each vulture represents a potential solution)
2: Set the maximum number of iterations ( $max_{iter}$ )
3: Set the maximum number of vultures ( $population_{size}$ )
4: Set parameters for exploration and exploitation phases:
5:   Step size ( $step_{size}$ )
6:   Exploration radius ( $exploration_{radius}$ )
7:   Exploitation factor ( $exploitation_{factor}$ )
8:   Reproduction rate ( $reproduction_{rate}$ )
9:   Breeding factor ( $breeding_{factor}$ )
10: Define  $fitness\_function(x)$ 
11:  $best_{solution} \leftarrow \infty$ 
12:  $best_{fitness} \leftarrow \infty$                                       $\triangleright$  Main AVOA loop
13: for  $iter$  in  $range(max_{iter})$  do            $\triangleright$  Evaluate the fitness of each vulture in the
    population
14:   for  $i$  in  $range(population_{size})$  do
15:      $vulture \leftarrow P[i]$ 
16:      $fitness \leftarrow fitness_{function}(vulture)$        $\triangleright$  Update the best solution found so far
17:     if  $fitness < best_{fitness}$  then
18:        $best_{solution} \leftarrow vulture$ 
19:        $best_{fitness} \leftarrow fitness$ 
20:     end if
21:   end for            $\triangleright$  Apply foraging and scavenging behaviour (exploration phase)
22:   for  $i$  in  $range(population_{size})$  do            $\triangleright$  Random exploration: move vulture
    randomly
23:     if  $random() < exploration_{radius}$  then
24:        $P[i] \leftarrow P[i] + random_{step}(step_{size})$ 
25:     else                                      $\triangleright$  Scavenging: exploit promising areas
26:        $P[i] \leftarrow P[i] + exploitation_{factor} * (best_{solution} - P[i])$ 
27:     end if
28:   end for            $\triangleright$  Apply reproduction and breeding behaviour
    (exploitation phase)
29:   for  $i$  in  $range(best_{solution})$  do
30:     if  $random() < reproduction_{rate}$  then       $\triangleright$  Breeding: combine features of two
    solutions to create a new solution
31:        $parent1 \leftarrow P[i]$ 
32:        $parent2 \leftarrow P[random_{choice}(population_{size})]$ 
33:        $P[i] \leftarrow breeding_{factor} * (parent1 + parent2)/2$ 
34:     end if
35:   end for            $\triangleright$  Optionally apply a local search or mutation strategy to refine
    solutions further

```

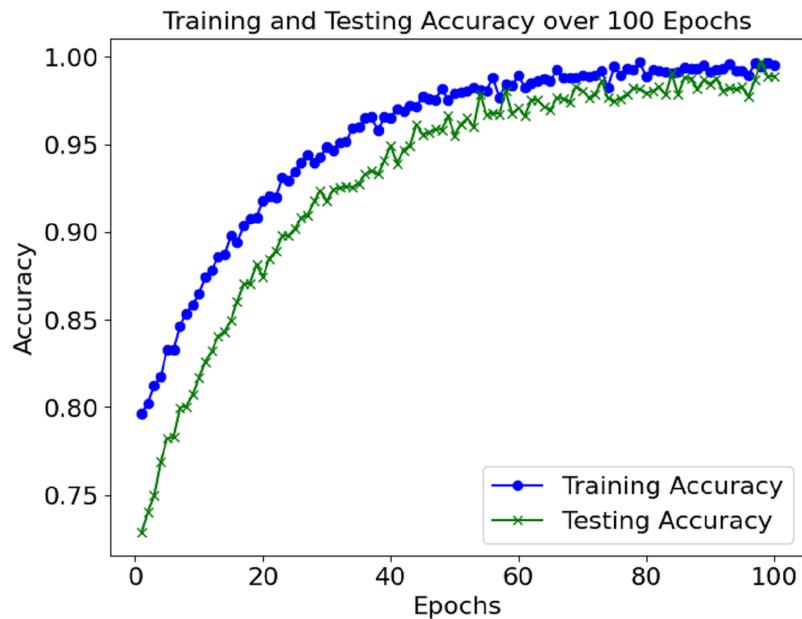
(Continued)

**Algorithm 1** (continued)

```

36: for  $i$  in  $\text{range}(\text{population}_\text{size})$  do
37:    $P[i] \leftarrow \text{local}_{\text{search}}(P[i])$ 
38: end for
39: end for
40: return  $\text{best}_{\text{solution}}$ 

```

**Figure 2** Accuracy graph of the proposed model.

Full-size DOI: 10.7717/peerj-cs.2958/fig-2

overfitting. Until a certain point, the training and validation loss started low and subsequently grew to a point where it reached a loss minimum, which predicts the model's convergence.

Figure 2 shows the accuracy of the training and validation sets over the 100 epochs. The model performed well and maintained a high level of accuracy, with very little change in the validation accuracy during the training.

The training and validation losses also decreased, as depicted in Fig. 3. The convergence of the loss values is consistent enough to demonstrate that the model has learned sufficiently and is not overfitting the training set. Additional measures, including precision, recall, F1-score, and AUC, were used to evaluate the model's performance and reported accuracy. Such metrics help determine the model's effectiveness in distinguishing between positive (sepsis) and negative (non-sepsis) conditions, which is essential in medical settings and is depicted in Table 1.

The performance evaluation from the proposed model is exhaustively outlined in Table 1. Accuracy achieved by the model is exceptionally high at 99.4%, displaying effectiveness for precise classification of sepsis and non-sepsis cases. The model's precision

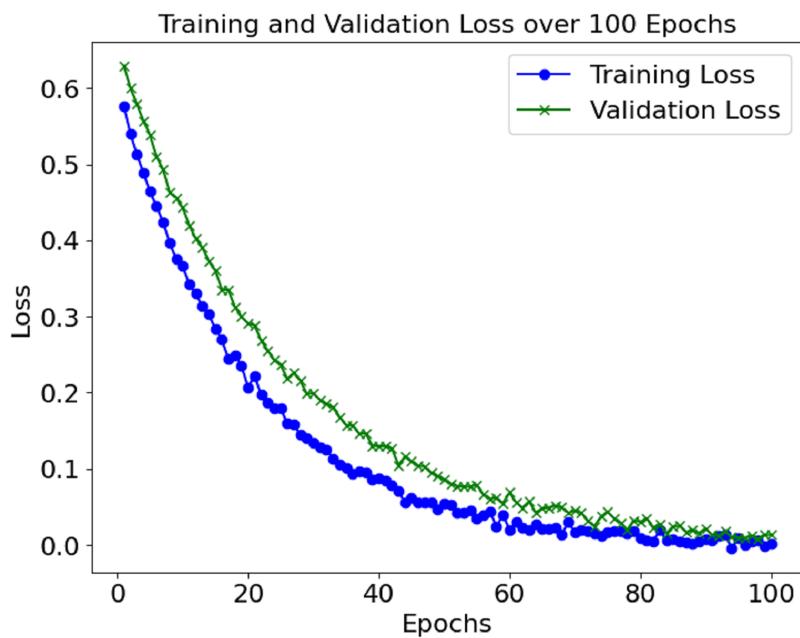


Figure 3 Loss graph of the proposed model.

Full-size DOI: 10.7717/peerj-cs.2958/fig-3

Table 1 Performance analysis.

Metric	Value
Accuracy	99.4%
Precision	98.7%
Recall	99.2%
F1-Score	99.0%
AUC	0.998
Specificity	99.3
False Positive Rate (FPR)	0.7
False Negative Rate (FNR)	0.8

of 98.7% indicates that the model exercises strong control over false positives where patients diagnosed with sepsis are minimized, and is vital to prevent unnecessary interventions. The model demonstrated a recall of 99.2% which shows high sensitivity in identifying the actual sepsis cases, proving vital for diagnosis and treatment. The balanced F1-score of 99% illustrates that the proposed model also maintained a well-balanced trade-off between precision and recall, reaffirming the strong figure reported for recall. AUC value of 0.998 provides further proof the model has strong discriminatory power in distinguishing between sepsis and non-sepsis conditions. Furthermore, the model achieved 99.3% specificity in correctly classifying non-sepsis patients while maintaining an extraordinary low false positive rate of 0.7% and false negative rate of 0.8%.

In Fig. 4, this section presents a confusion matrix associated with the performance outcomes of the model offered in this article for the detection of sepsis. The matrix

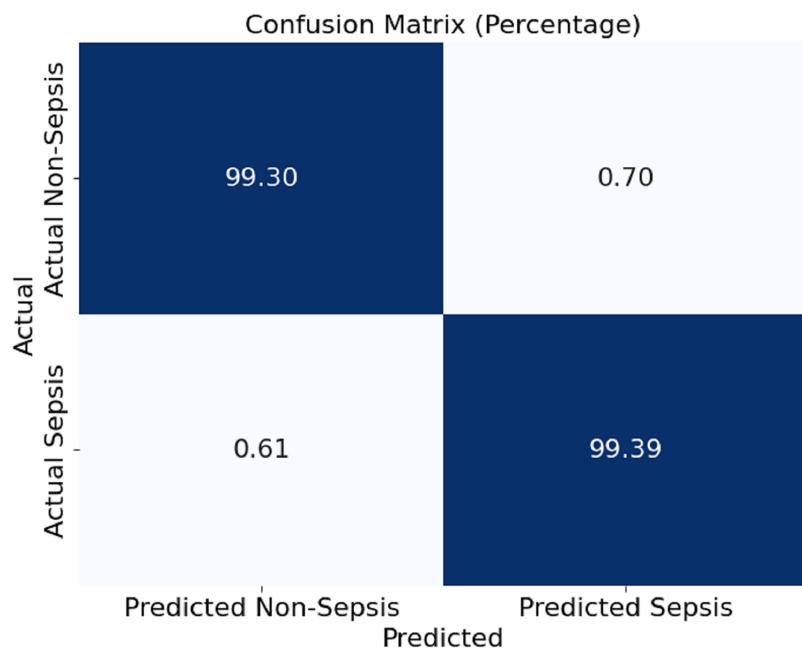


Figure 4 Confusion matrix of the proposed model. [Full-size](#)  DOI: [10.7717/peerj-cs.2958/fig-4](https://doi.org/10.7717/peerj-cs.2958/fig-4)

contains values of correct predictions as considerations of matrices are normalised to percentages. The model showed 99.30% and 99.39% of efficiency in distinguishing the cases of non-sepsis and the cases of sepsis, respectively. It can be noted that false negative rates of 0.61% and false positive rates of 0.70%, as the off-diagonal numbers, mean that the classification has a mild depth of error for the model, which is even further corroborated by the model's ability to detect sepsis and non-sepsis cases accurately.

The relationship between accuracy and hyperparameters are represented in Fig. 5 as well. The learning rate on the left orthogonal graph shows that the overall accuracy goes hand in hand with the rate until reaching somewhere about 0.01, at this point, performance wisely shifts down and can be observed as the learning goes further. The right graph Instead demonstrates that accuracy loses its power through increments of dropout from 0.5. On the other hand, the graph suggests a way out, controlling the moderate level of dropout, as it proved to augment model generalisations. However, a performance a wee higher than 0.5 could cause impairing effects. Finally, these results show and exemplify how, in the realm of model performance tuning, both the learning and dropout rates can prove to be crucial.

The model outlined in the previous sections achieved a perfect score of  $AUC = 1.00$ , as shown in Fig. 6. Actual positive rate (TPR) and the true negative rate (TPR) were plotted on the y-axis and x-axis, respectively. The ROC region is very far from the dotted diagonal. Therefore, the model is good at differentiating between different diseases and non-use cases, with sepsis being one of them. Achieving perfect scores against  $AUC$  as high as one suggests that the evaluated model has achieved high accuracy in differentiating between positive and negative cases. Through a methodical process of module removal or

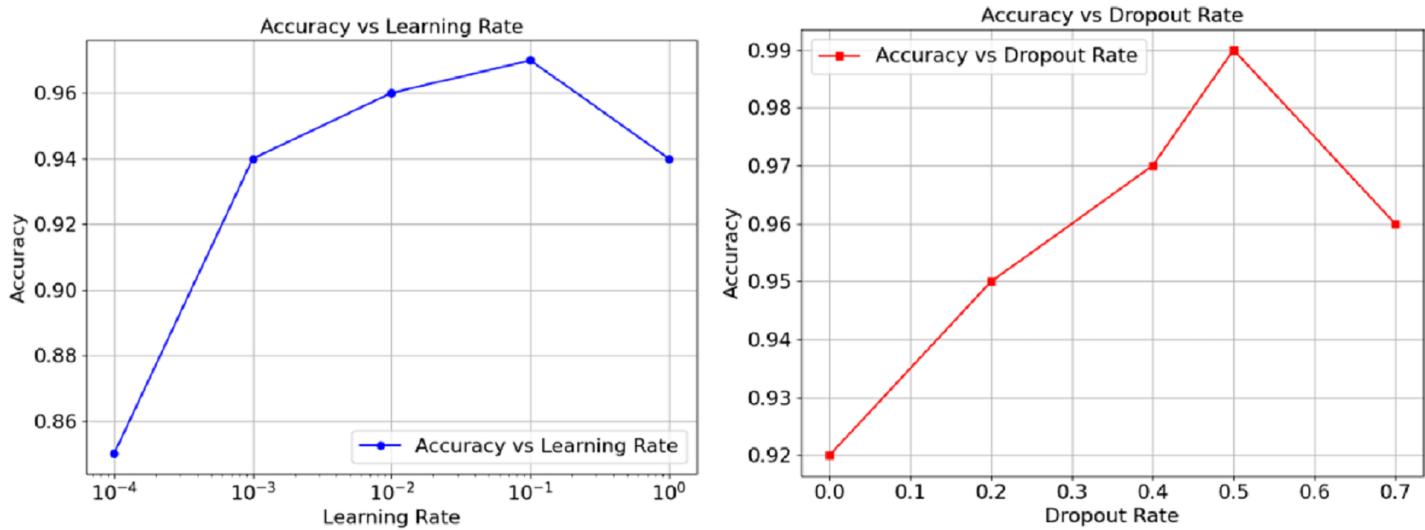


Figure 5 Accuracy vs. Dropout rate and learning rate.

Full-size DOI: 10.7717/peerj-cs.2958/fig-5

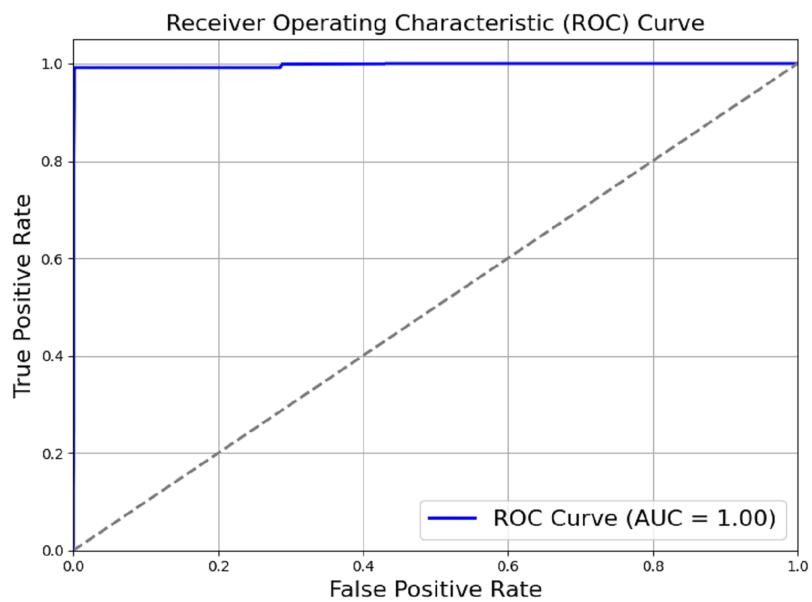


Figure 6 ROC\_AUC of the proposed model.

Full-size DOI: 10.7717/peerj-cs.2958/fig-6

configuration modification, we conducted ablation research to evaluate the contribution of each model component. Table 2 summarises the ablation study's findings by comparing the full model to subsets of the model that had their components reduced.

To ensure that the model was robust, we ran a 6-fold cross-validation on the dataset. To train and test the model, we divided the dataset into six sections, and we did it six times, every single time using an alternate fold as the validation set and the other five as the training set. The results of cross-validation are shown in Table 3.

**Table 2** Ablation study results.

Model variant	Accuracy	Precision	Recall	F1-score	AUC
Full model	99.4%	98.7%	99.2%	99.0%	0.998
Without SCAN	97.8%	96.2%	97.4%	96.8%	0.985
Without HDCB	98.5%	97.5%	98.0%	97.7%	0.993
Without RPCC	97.1%	95.8%	96.5%	96.1%	0.986
Without ECLF	96.3%	94.6%	95.5%	95.0%	0.978

**Table 3** 6-Fold cross-validation results.

Metric	Mean (%)	Standard deviation (%)
Accuracy	99.38%	±0.14
Precision	98.78%	±0.32
Recall	99.22%	±0.27
F1-score	99.0%	±0.25
AUC	0.9978%	±0.0011

The results of six-fold cross-validation are summarized in [Table 3](#). The proposed model achieved a mean accuracy of 99.38% with a standard deviation of  $\pm 0.14\%$ , indicating high consistency across different data splits. Similarly, the precision, recall, and F1-score maintained low variability, further demonstrating the robustness of the model. The AUC remained exceptionally high at  $0.9978 \pm 0.0011$ , highlighting the model's strong discriminative ability for early sepsis detection. To test the performance of the model designed during this study and compare its performance against other popularly known techniques for detecting sepsis, a comparison was carried out of the model with the rest. The models compared include logistic regression (LR), support vector machine (SVM), random forest (RF) models, which are some basic models, and the CNN, which is also fundamental. The results of the comparison are included in [Table 4](#).

To analyze the effectiveness of the proposed hybrid deep learning model, it was compared to existing models through evaluation of other machine learning and deep learning methods. As presented in [Table 4](#), the performance of the proposed model is compared against benchmarks stemming from logistic regression, SVM, random forest, and a baseline CNN model. The proposed model significantly outperformed all existing methods across all critical metrics. The recall and precision metrics of the logistic regression and SVM models were 85.7% and 92.6% respectively; however, these measures are not adequate in high-stakes medical environments where detection must be both timely and accurate. Even strong classifiers such as RF and the baseline CNN model were incapable of surpassing 94.8% and 96.7% accuracy, respectively. Unlike the other algorithms, the proposed model achieved 99.4% accuracy, 98.7% precision, 99.2% recall, and 0.998 AUC, which signifies a remarkable performance benchmark exhibiting dependability in predictive accuracy alongside sensitivity and specificity necessary to lower the rate of erroneous diagnosis in sepsis cases. Considerable increase in performance can

**Table 4** Results of performance comparison.

Model	Accuracy	Precision	Recall	F1-score	AUC
Proposed model	99.4%	98.7%	99.2%	99.0%	0.998
Logistic regression	85.7%	81.3%	87.0%	84.1%	0.758
SVM	92.6%	91.4%	93.1%	92.2%	0.873
Random forest	94.8%	93.0%	95.4%	94.2%	0.896
CNN (Simple)	96.7%	95.8%	97.2%	96.5%	0.957

be linked to the addition of new components like the SCAN, HDCB, RPCC, and the hyperparameter optimization executed with the AVOA. Such integration allows the model to capture sophisticated, non-linear relations in multi-dimensional clinical data which other conventional models would not identify.

In [Table 5](#) we compare the results of our proposed model with multiple reported state-of-the-art models with applied methods of prediction on sepsis. The results clearly demonstrate the accuracy challenges faced by conventional machine learning approaches and even earlier deep learning models from the years 2017 and 2018 which are dominated by the complex, nonlinear, temporal, and intricate nature of sepsis progression. As an example neural hybrid models composed from the convolutional and recurrent networks integrated the use of traditional deep learning models, but their performance remained stagnant, yielding predictive accuracy of 56.25% ([Duan et al., 2023](#)). Followed by this, [Strickler et al. \(2023\)](#) reported a slightly improved accuracy (75.00%) with the use of LSTM networks. Most of these results are lower than baseline, indicating that there is still a long journey ahead before these models can achieve clinically acceptable predictive performance. While results achieved by more advanced models tend to perform better, that's not an argument that ignores the performance forecasts of cutting edge models. In either case, [Zhou, Beyah & Kamaleswaran \(2021\)](#) OnAI-Comp, and [Rosnati & Fortuin \(2021\)](#) with their attention time convolutional network (AttTCN), managed to attain accuracy scores of 81.25% and 68.75%. These findings highlight the significance of attention mechanisms and temporal modeling for multi scale feature extraction frameworks. On the other hand, the hybrid deep learning model outperforming the rest achieves 99.4% accuracy, a figure far exceeding the findings of prior studies. This is because the model uses the ECLF which deals with spatial feature extraction, the SCAN which attends to noteworthy clinical features on the fly, and the RPCC which retains key information across deep network layers. Also, the model's performance was optimally complemented with hyperparameter adjustments *via* the AVOA. This underscores the novelty of the approach since the model was designed with extensive focus on increasing accuracy while solving the practical clinical problem of reliably and promptly detecting the onset of clinical sepsis.

### Limitations and future work

While the proposed model demonstrates outstanding performance on the Kaggle Sepsis Prediction dataset, this study has certain limitations that must be acknowledged. The

**Table 5** Comparison with the state of the art techniques.

Reference	Methodology	Accuracy (%)
Gholamzadeh, Abtahi & Safdari (2023)	Gaussian naïve Bayes (NB), decision tree (DT), random forest (RF)	68.75
Duan et al. (2023)	CNN + RNN	56.25
Strickler et al. (2023)	LSTM	75.00
Zhou, Beyah & Kamaleswaran (2021)	Online Artificial Intelligence Experts Competing Framework (OnAI-Comp)	81.25
Al-Mualemi & Lu (2020)	RNN-LSTM, SVM	62.50
Nemati et al. (2018)	Attention time convolutional network (AttTCN)	68.75
Proposed methodology	ECLF, SCAN, HDCB, and RPCC	99.4

experimental evaluation is confined to a single publicly available dataset, which may limit the generalizability of the results across diverse clinical environments and patient populations. Variations in data quality, patient demographics, clinical protocols, and sensor accuracy across different healthcare institutions could impact the performance of the model when applied to other datasets. Additionally, although a detailed comparison with several state-of-the-art model's performance models using the same dataset has been provided, further validation against other benchmark datasets such as the MIMIC-III or MIMIC-IV clinical databases would strengthen the evidence for the model's robustness and adaptability. Future research will focus on validating the proposed model on multiple datasets from varied clinical settings to ensure broader applicability. Incorporating real-time streaming data from electronic health records (EHRs) and evaluating the model in prospective clinical trials will also be explored to assess its performance in practical deployment scenarios. Moreover, enhancing model interpretability and explainability through integration with explainable AI (XAI) techniques will be a critical area of future development to support clinical decision-making effectively.

This study enhances AI systems intended for the early detection of critical healthcare conditions, especially sepsis, one of the leading causes of mortality globally in the ICU. Although the model suggests better accuracy than other models, it is worth noting that an early diagnosis does not resolve the multifaceted issues concerning the management of sepsis. Incorporating predictive models into clinical workspaces, real-time data provisioning, and actionable clinical decision support still pose challenges. In addition to the lack of diverse datasets, the explainability and interpretability of the model's predictions should be prioritized in future efforts. Adopted XAI frameworks such as SHapley Additive exPlanations (SHAP) and Local Interpretable Model-agnostic Explanations (LIME) could enable medical professionals to appreciate the determinants of the model's reasoning, thus strengthening their confidence in an AI-assisted system. Also, model performance, especially in the early detection of sepsis, can be improved by addressing data imbalance using advanced focal loss techniques or synthetic data generation like SMOTE.

Another critical aspect is the real-time model validation within clinical settings through live EHR data streams. This will assess the model's latency and efficiency and its impact on

clinically relevant outcomes. Studying the incorporation of multi-modal data sets such as genomic, proteomic, and imaging data may enhance the understanding of sepsis pathophysiology and enable the development of earlier and more precise diagnostic models. Lastly, investigations should be conducted on the ethical issues and potential biases in predictive models designed for other healthcare functions. Addressing equity, clarity, and regulatory alignment concerns will be essential for integrating and functioning AI-powered solutions in everyday medical practice.

## CONCLUSION

The article presents a new method for sepsis detection that is deep learning-based, achieved by ECLF, SCAN, HDCB, and RPCC. These core components act as a hybrid model proven to perform well, achieving an accuracy of 99.4% alongside precision, recall, and F1-score. Due to sepsis detection accuracy, this model dethrones the previous traditional clinical scoring systems alongside baseline ML systems. The ablation study had different components due to model parameter tuning and revealing factors. Combining the ROC curve and the confusion matrix effectively determined the correct classification with the AUC score and its negligible misclassifications. This study proves that improved decision-making alongside early detection modes for sepsis management can lead to a better patient outcome and a lower mortality rate.

Future research could examine incorporating patient demographic and medical history data, wearable sensors, and real-time monitoring into the model to enhance its forecasting ability. However, while the model is robust and had a positive outcome in the controlled experiment, it would need to be validated across several centres to determine if it is widely applicable to numerous patients and supported in a multi-facility setting. Despite achieving exceptional predictive performance, the study's conclusions are drawn based on results from a single dataset. While this provides a strong foundation, further evaluations across diverse datasets and real-world environments are essential to confirm the model's generalizability. Addressing these limitations in future work will ensure that the proposed system can be reliably adopted in varied clinical contexts, ultimately contributing to improved patient outcomes through early and accurate sepsis detection.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was supported by the Deanship of Research and Graduate Studies at King Khalid University through Large Research Project under grant number RGP2/219/46. Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2025R809), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia. This work was also supported by the Deanship of Scientific Research at Northern Border University, Arar, KSA, through the project number "NBU-FFR-2025-2248-05" and the Deanship of Scientific Research at Majmaah University under Project No. R-2025-

1825. This study is supported *via* funding from Prince Sattam bin Abdulaziz University project number (PSAU/2025/R/1446). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the authors:

Deanship of Research and Graduate Studies at King Khalid University: RGP2/219/46.  
Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia: PNURSP2025R809.  
Deanship of Scientific Research at Northern Border University, Arar, KSA:  
NBU-FFR-2025-2248-05.  
Deanship of Scientific Research at Majmaah University: R-2025-1825.  
Prince Sattam bin Abdulaziz University: PSAU/2025/R/1446.

### Competing Interests

The authors declare that they have no competing interests.

### Author Contributions

- Ahmed S. Almasoud conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Ghada Moh Samir Elhessewi conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Munya A. Arasi conceived and designed the experiments, performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Abdulsamad Ebrahim Yahya conceived and designed the experiments, analyzed the data, authored or reviewed drafts of the article, and approved the final draft.
- Menwa Alshammeri performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Donia Badawood performed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Faisal Mohammed Nafie performed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Mohammed Assiri performed the experiments, performed the computation work, prepared figures and/or tables, and approved the final draft.

### Data Availability

The following information was supplied regarding data availability:

The Prediction of Sepsis dataset is available at Kaggle:

<https://www.kaggle.com/datasets/salikhussaini49/prediction-of-sepsis>. The dataset and code are available in the [Supplemental Files](#).

### Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj-cs.2958#supplemental-information>.

## REFERENCES

Al-Mualemi BY, Lu L. 2020. A deep learning-based sepsis estimation scheme. *IEEE Access* **9**:5442–5452 DOI [10.1109/access.2020.3043732](https://doi.org/10.1109/access.2020.3043732).

Bone RC, Balk RA, Cerra FB, Dellinger RP, Fein AM, Knaus WA, Schein RM, Sibbald WJ. 1992. Definitions for sepsis and organ failure and guidelines for the use of innovative therapies in sepsis. *Chest* **101**(6):1644–1655 DOI [10.1378/chest.101.6.1644](https://doi.org/10.1378/chest.101.6.1644).

Burdick H, Pino E, Gabel-Comeau D, Gu C, Huang H, Lynn-Palevsky A, Das R. 2017. Evaluating a sepsis prediction machine learning algorithm in the emergency department and intensive care unit: a before and after comparative study. *BioRxiv* **4**(3):224014 DOI [10.1101/224014](https://doi.org/10.1101/224014).

Daothong P, Jampa-ngern S, Senavongse W. 2024. Utilizing machine learning predictive analytics to enhance early sepsis diagnosis in critical care setting. In: *2024 IEEE/ACIS 9th International Conference on Big Data, Cloud Computing, and Data Science (BCD)*. Piscataway: IEEE, 44–47.

de Hond TA, Niemantsverdriet MS, van Solinge WW, Oosterheert JJ, Haitjema S, Kaasjager KA. 2022. Sepsis labels defined by claims-based methods are ill-suited for training machine learning algorithms. *Clinical Microbiology and Infection* **28**(8):1170–1171 DOI [10.1016/j.cmi.2022.03.029](https://doi.org/10.1016/j.cmi.2022.03.029).

Duan Y, Huo J, Chen M, Hou F, Yan G, Li S, Wang H. 2023. Early prediction of sepsis using double fusion of deep features and handcrafted features. *Applied Intelligence* **53**(14):17903–17919 DOI [10.1007/s10489-022-04425-z](https://doi.org/10.1007/s10489-022-04425-z).

Faisal M, Scally A, Richardson D, Beatson K, Howes R, Speed K, Mohammed MA. 2018. Development and external validation of an automated computer-aided risk score for predicting sepsis in emergency medical admissions using the patient's first electronically recorded vital signs and blood test results. *Critical Care Medicine* **46**(4):612–618 DOI [10.1097/ccm.0000000000002967](https://doi.org/10.1097/ccm.0000000000002967).

Gholamzadeh M, Abtahi H, Safdari R. 2023. Comparison of different machine learning algorithms to classify patients suspected of having sepsis infection in the intensive care unit. *Informatics in Medicine Unlocked* **38**:101236 DOI [10.1016/j.imu.2023.101236](https://doi.org/10.1016/j.imu.2023.101236).

Giordano C, Brennan M, Mohamed B, Rashidi P, Modave F, Tighe P. 2021. Accessing artificial intelligence for clinical decision-making. *Frontiers in Digital Health* **3**:645232 DOI [10.3389/fdgth.2021.645232](https://doi.org/10.3389/fdgth.2021.645232).

Gong M, Liu J, Li C, Guo W, Wang R, Chen Z. 2022. Early warning model for death of sepsis via length insensitive temporal convolutional network. *Medical & Biological Engineering & Computing* **60**(3):875–885.

Henry KE, Hager DN, Pronovost PJ, Saria S. 2015. A targeted real-time early warning score (TREWScore) for septic shock. *Science Translational Medicine* **7**(299):299ra122–299ra122 DOI [10.1126/scitranslmed.aab3719](https://doi.org/10.1126/scitranslmed.aab3719).

Hochreiter S, Schmidhuber J. 1997. Long short-term memory. *Neural Computation* **9**(8):1735–1780.

Horng S, Sontag DA, Halpern Y, Jernite Y, Shapiro NI, Nathanson LA. 2017. Creating an automated trigger for sepsis clinical decision support at emergency department triage using machine learning. *PLOS ONE* **12**(4):e0174708 DOI [10.1371/journal.pone.0174708](https://doi.org/10.1371/journal.pone.0174708).

Hu C, Li L, Huang W, Wu T, Xu Q, Liu J, Hu B. 2022. Interpretable machine learning for early prediction of prognosis in sepsis: a discovery and validation study. *Infectious Diseases and Therapy* **11**(3):1117–1132 DOI [10.1007/s40121-022-00628-6](https://doi.org/10.1007/s40121-022-00628-6).

**Hussaini S.** 2021. Prediction of sepsis. Available at <https://www.kaggle.com/datasets/salikhussaini49/prediction-of-sepsis>.

**Iskander KN, Osuchowski MF, Stearns-Kurosawa DJ, Kurosawa S, Stepien D, Valentine C, Remick DG.** 2013. Sepsis: multiple abnormalities, heterogeneous responses, and evolving understanding. *Physiological Reviews* **93**(3):1247–1288 DOI [10.1152/physrev.00037.2012](https://doi.org/10.1152/physrev.00037.2012).

**Iwashyna TJ, Ely EW, Smith DM, Langa KM.** 2010. Long-term cognitive impairment and functional disability among survivors of severe sepsis. *The Journal of the American Medical Association* **304**(16):1787–1794 DOI [10.1001/jama.2010.1553](https://doi.org/10.1001/jama.2010.1553).

**Kang S-H, Kim HK, Ham CK, Lee DS, Cho H-I.** 2008. Comparison of four hematology analyzers, CELL-DYN sapphire, ADVIA 120, coulter lh 750, and sysmex XE-2100, in terms of clinical usefulness. *International Journal of Laboratory Hematology* **30**(6):480–486 DOI [10.1111/j.1751-553x.2007.00981.x](https://doi.org/10.1111/j.1751-553x.2007.00981.x).

**Lauritsen SM, Kalør ME, Kongsgaard EL, Lauritsen KM, Jørgensen MJ, Lange J, Thiesson B.** 2020. Early detection of sepsis utilizing deep learning on electronic health record event sequences. *Artificial Intelligence in Medicine* **104**(11):101820 DOI [10.1016/j.artmed.2020.101820](https://doi.org/10.1016/j.artmed.2020.101820).

**Lukaszewski RA, Yates AM, Jackson MC, Swingler K, Scherer JM, Simpson AJ, Sadler P, McQuillan P, Titball RW, Brooks TJG, Pearce MJ.** 2008. Presymptomatic prediction of sepsis in intensive care unit patients. *Clinical and Vaccine Immunology* **15**(7):1089–1094 DOI [10.1128/cvi.00486-07](https://doi.org/10.1128/cvi.00486-07).

**Marik PE, Farkas JD.** 2018. The changing paradigm of sepsis: early diagnosis, early antibiotics, early pressors, and early adjuvant treatment. *Critical Care Medicine* **46**(10):1690–1692 DOI [10.1097/ccm.0000000000003310](https://doi.org/10.1097/ccm.0000000000003310).

**Nemati S, Holder A, Razmi F, Stanley MD, Clifford GD, Buchman TG.** 2018. An interpretable machine learning model for accurate prediction of sepsis in the ICU. *Critical Care Medicine* **46**(4):547–553 DOI [10.1097/ccm.0000000000002936](https://doi.org/10.1097/ccm.0000000000002936).

**Parlato M, Philippart F, Rouquette A, Moucadel V, Puchois V, Blein S, Bedos J-P, Diehl J-L, Hamzaoui O, Annane D, Journois D, Boutieb MB, Estève L, Fitting C, Treluyer J-M, Pachot A, Adib-Conquy M, Cavaillon J-M, Misset B, The Captain Study Group.** 2018. Circulating biomarkers may be unable to detect infection at the early phase of sepsis in ICU patients: the CAPTAIN prospective multicenter cohort study. *Intensive Care Medicine* **44**(7):1061–1070 DOI [10.1007/s00134-018-5228-3](https://doi.org/10.1007/s00134-018-5228-3).

**Pierrickos C, Velissaris D, Bisdorff M, Marshall JC, Vincent J-L.** 2020. Biomarkers of sepsis: time for a reappraisal. *Critical Care* **24**(1):1–15 DOI [10.1186/s13054-020-02993-5](https://doi.org/10.1186/s13054-020-02993-5).

**Raghu A, Komorowski M, Ahmed I, Celi L, Szolovits P, Ghassemi M.** 2017. Deep reinforcement learning for sepsis treatment. ArXiv DOI [10.48550/arXiv.1711.09602](https://doi.org/10.48550/arXiv.1711.09602).

**Reyna MA, Nsoesie EO, Clifford GD.** 2022. Rethinking algorithm performance metrics for artificial intelligence in diagnostic medicine. *The Journal of the American Medical Association* **328**(4):329–330 DOI [10.1001/jama.2022.10561](https://doi.org/10.1001/jama.2022.10561).

**Rosnati M, Fortuin V.** 2021. MGP-AttTCN: an interpretable machine learning model for the prediction of sepsis. *PLOS ONE* **16**(5):e0251248.

**Rudd KE, Johnson SC, Agesa KM, Shackelford KA, Tsoi D, Kievlan DR, Colombara DV, Ikuta KS, Kissoon N, Finfer S, Fleischmann-Struzek C, Machado FR, Reinhart KK, Rowan K, Seymour CW, Watson RS, West TE, Marinho F, Hay SI, Lozano R, Lopez AD, Angus DC, Murray CJL, Naghavi M.** 2020. Global, regional, and national sepsis incidence and mortality,

1990–2017: analysis for the global burden of disease study. *The Lancet* **395**(10219):200–211  
DOI [10.1016/s0140-6736\(19\)32989-7](https://doi.org/10.1016/s0140-6736(19)32989-7).

**Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, Bellomo R, Bernard GR, Chiche J-D, Coopersmith CM, Hotchkiss RS, Levy MM, Marshall JC, Martin GS, Opal SM, Rubenfeld GD, van der Poll T, Vincent J-L, Angus DC.** 2016. The third international consensus definitions for sepsis and septic shock (sepsis-3). *The Journal of the American Medical Association* **315**(8):801–810 DOI [10.1001/jama.2016.0287](https://doi.org/10.1001/jama.2016.0287).

**Spasic I, Nenadic G.** 2020. Clinical text data in machine learning: systematic review. *JMIR Medical Informatics* **8**(3):e17984 DOI [10.2196/17984](https://doi.org/10.2196/17984).

**Strickler EA, Thomas J, Thomas JP, Benjamin B, Shamsuddin R.** 2023. Exploring a global interpretation mechanism for deep learning networks when predicting sepsis. *Scientific Reports* **13**(1):3067 DOI [10.1038/s41598-023-30091-3](https://doi.org/10.1038/s41598-023-30091-3).

**Sun C, Hong S, Song M, Shang J, Li H.** 2021. Personalized vital signs control based on continuous action-space reinforcement learning with supervised experience. *Biomedical Signal Processing and Control* **69**(1):102847 DOI [10.1016/j.bspc.2021.102847](https://doi.org/10.1016/j.bspc.2021.102847).

**Sun D-Y, Hu Y-J, Li X, Peng J, Dai Z-J, Wang S.** 2025. Unlocking the full potential of memory t cells in adoptive t cell therapy for hematologic malignancies. *International Immunopharmacology* **144**(8):113392 DOI [10.1016/j.intimp.2024.113392](https://doi.org/10.1016/j.intimp.2024.113392).

**Van der Ven WH, Schuurmans J, Schenk J, Roerhorst S, Cherpanath TGV, Lagrand WK, Thoral P, Elbers PWG, Tuinman PR, Scheeren TWL, Bakker J, Geerts BF, Veelo DP, Paulus F, Vlaar APJ.** 2022. Monitoring, management, and outcome of hypotension in intensive care unit patients, an international survey of the european society of intensive care medicine. *Journal of Critical Care* **67**(11):118–125 DOI [10.1016/j.jcrc.2021.10.008](https://doi.org/10.1016/j.jcrc.2021.10.008).

**Vellido A, Ribas V, Morales C, Ruiz Sanmartín A, Ruiz Rodríguez JC.** 2018. Machine learning in critical care: state-of-the-art and a sepsis case study. *Biomedical Engineering Online* **17**(S1):1–18 DOI [10.1186/s12938-018-0569-2](https://doi.org/10.1186/s12938-018-0569-2).

**Wang K, Yin J, Chen J, Ma J, Si H, Xia D.** 2024. Inhibition of inflammation by berberine: molecular mechanism and network pharmacology analysis. *Phytomedicine* **128**:155258 DOI [10.1016/j.phymed.2023.155258](https://doi.org/10.1016/j.phymed.2023.155258).

**Wong A, Otles E, Donnelly JP, Krumm A, McCullough J, DeTroyer-Cooley O, Pestrule J, Phillips M, Konye J, Penoza C, Ghous M, Singh K.** 2021. External validation of a widely implemented proprietary sepsis prediction model in hospitalized patients. *JAMA Internal Medicine* **181**(8):1065–1070 DOI [10.1001/jamainternmed.2021.2626](https://doi.org/10.1001/jamainternmed.2021.2626).

**Zhang Y, Liu X, Luo J, Liu H, Li Y, Liu J, Zhu L, Wang J, Zeng H.** 2025. Dual recombinase polymerase amplification system combined with lateral flow immunoassay for simultaneous detection of staphylococcus aureus and vibrio parahaemolyticus. *Journal of Pharmaceutical and Biomedical Analysis* **255**:116621 DOI [10.1016/j.jpba.2024.116621](https://doi.org/10.1016/j.jpba.2024.116621).

**Zhou A, Beyah R, Kamaleswaran R.** 2021. OnAI-comp: an online AI experts competing framework for early sepsis detection. *IEEE/ACM Transactions on Computational Biology and Bioinformatics* **19**(6):3595–3603 DOI [10.1109/tcbb.2021.3122405](https://doi.org/10.1109/tcbb.2021.3122405).