IIETA International Information and Engineering Technology Association
Advancing the World of Information and Engineering

Ingénierie des Systèmes d'Information
Vol. 31, No. 10, October, 2026, pp. 3391-3403
Journal homepage: http://iieta.org/journals/isi

Federated ConvLSTM for ICU Patient Risk Classification: Cloud based Evaluation Using MIMIC-III and Five Simulated Clients

Khader Basha Sk*, Thirupurasundari D. R.

Department of Computer Science and Engineering, Bharath Institute of Science and Technology, Bharath Institute of Higher Education and Research, Chennai 600 073, India

Corresponding Author Email: khaderbasha576@gmail.com

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

https://doi.org/10.18280/isi.311013

Received: 20 June 2026
Revised: 7 September 2026
Accepted: 18 September 2026
Available online: 31 October 2026

Keywords:
federated learning, convolutional long short-term memory, intensive care unit patient risk classification, physiological time series, MIMIC-III, cloud computing

ABSTRACT

Accurate classification of intensive care unit (ICU) patient risk from physiological time series is complicated by temporal dependencies and restrictions on sharing clinical records. This study presents a federated convolutional long short-term memory (ConvLSTM) framework coordinated through a cloud server (CFDL-ConvLSTM) for offline binary classification of patients into low- and high-risk groups. The model combines two ConvLSTM layers with Federated Averaging (FedAvg); training is performed locally on five non-overlapping, patient-level partitions derived from Medical Information Mart for Intensive Care III (MIMIC-III) v1.4, while only model parameters are sent to the aggregation server. Five physiological variables—heart rate, blood pressure, respiratory rate, oxygen saturation, and temperature—are used as inputs. Binary risk labels were assigned using International Classification of Diseases, Ninth Revision (ICD-9)-based clinical information. Training and testing are separated at the patient level, and performance is assessed over five runs using the same data split. The reported mean accuracy, precision, recall, and F1-score are 97.30%, 97.17%, 97.40%, and 97.29%, respectively; the final training loss is 0.0649. The model also exceeds the reported metrics of recurrent neural network (RNN), LSTM, and Fusion-LSTM baselines evaluated in the study. These findings suggest that federated ConvLSTM training is feasible for this classification task under the simulated data-partitioning protocol. However, the five clients are partitions of a single dataset rather than independent hospitals, and the study does not evaluate privacy attacks, secure aggregation, differential privacy, external validation, or clinical deployment. The results therefore establish an offline experimental baseline, not evidence of clinical validity or cross-institutional generalizability.

1. INTRODUCTION

Intensive care units (ICUs) are units that are dedicated to treating patients who are seriously ill and need close monitoring and prompt treatment. As electronic health records and bedside monitoring systems have been adopted, the amount of clinical information generated has grown, with the potential to be used for patient risk assessment and prediction of outcomes. Unplanned readmission to the ICU is an important topic to study, as it becomes clear that computational approaches can be useful in identifying patients who need additional intensive care, as studied by Sheetrit et al. [1]. Moerschbacher and He [2] explored the possibility of using heterogeneous clinical data from electronic health records (EHRs) to make better 30-day ICU readmission predictions for 30 days by using structured and unstructured information.

In the ICUs, the temporal nature of the data is highly relevant for deep learning, which is especially suitable for high-risk (critical) care prediction. In their work, Kessler et al. [3] showed how recurrent networks can be used in healthcare to predict readmission to cardiovascular intensive care units, highlighting the potential of recurrent architectures to capture temporal patterns in clinical time-series data. Cardiovascular ICU readmission prediction with temporal modeling using clinical time-series information was also assessed by Moazemi et al. [4], and the importance of temporal modeling for readmission prediction was highlighted. Pishgar et al. [5] combined demographic data, severity-related features, and temporal clinical data to forecast unplanned 30-day ICU readmission in heart failure patients, demonstrating the importance of combining a range of clinical features in order to predict a patient's risk.

A number of applications have been explored for using machine learning in the context of simulated client discharge and readmission decision support in the ICU. To aid in the decision-making process used for discharging patients from the ICU, machine learning models were created by McWilliams et al. [6] from electronic healthcare data from the Medical Information Mart for Intensive Care III (MIMIC-III) as well as other clinical data sources. Min et al. [7] explored the use of clinical claims data to build machine learning-based simulated client readmission predictors for chronic obstructive pulmonary disease (COPD) patients, showing the potential of

predictive models in the field of simulated client readmission-risk assessment. Li et al. [8] studied the early prediction of 30-day ICU readmission by natural language processing (NLP) and machine learning, showing that there is potential for information from clinical records to be used for early identification of patients at risk.

The availability of MIMIC clinical information has also been utilized to help develop predictive models of simulated client outcomes. Assaf and Jayousi [9] explored the use of MIMIC data to predict simulated client readmission after 30 days, showing that machine learning can be a tool for gleaning predictive information from high-risk (critical)-care records. To demonstrate that clinical coding information could be represented and modeled, Liu et al. [10] used artificial neural networks with medical-code embeddings for the prediction of simulated client readmissions. While these studies highlight the increasing importance of deep learning and machine learning in clinical predictions, the typical centralized system usually demands that clinical data be gathered and analyzed in a shared place.

The temporal interaction between physiological variables and their time-varying nature is another major challenge in data analysis in the ICU. In medical federated learning, there are medical data sources with varying characteristics that can affect the training of collaborating models, a problem explored by Darzi et al. [11]. Structuring and structuring together with unstructured clinical data was investigated for predictive modeling using deep learning by Zhang et al. [12], showing the potential value of clinical data integration with diverse information sources. Barbieri et al. [13] compared deep learning architectures that predict ICU readmission and explored the prediction of high-risk (critical) patients, and showed that deep learning could be applied to the prediction of patients in an ICU.

These methods illustrate the promise of machine learning in clinical prediction, but the direct collection of patient data from various simulated ICU clients may also lead to privacy and data sharing concerns. An alternative approach is federated learning (FL), where participating institutions can train models with their local data without directly sharing the patient records, but only sharing the model information. In smart healthcare systems, federated approaches to collaborative healthcare analytics are relevant. Chouja et al. [14] discussed the use of federated learning with electronic health records for healthcare collaboration and pandemic response. Raza et al. [15] explored federated transfer learning for electrocardiogram (ECG) monitoring and explainable artificial intelligence (XAI), demonstrating the potential of federated learning for distributed physiological monitoring and the interpretability of the federated models.

There are also examples of application of federated learning to specific healthcare prediction tasks. In the context of collaborative learning, where clients may not be available at the same time, Khan et al. [16] explored asynchronous federated learning in the prediction of cardiovascular disease. Yaqoob et al. [17] proposed a hybrid approach that exploits federated learning for cardiovascular disease prediction in a distributed healthcare setting, which shows that federated learning is applicable to distributed clinical prediction. When medical information is distributed from different sources, Zhao and Huang [18] proposed a distributed information-sharing federated learning approach for medical image data, demonstrating the use of federated learning in the medical information setting.

Although federated learning does not necessarily mean that there is complete privacy protection for all possible leakage of information across model updates, privacy protection is still a key factor. In healthcare data-driven collaborative learning, Tang et al. [19] proposed a privacy-enhanced intention-hiding vertical federated learning framework, emphasizing the need for further privacy measures in collaborative healthcare learning. Riedel et al. [20] proposed a federated deep learning framework to detect pneumonia from COVID-19 chest X-rays while preserving distributed data, which was used for medical image classification with the preservation of distributed data.

In recent studies, federated learning has been applied to other prediction and classification tasks for healthcare. In healthcare, federated learning has been explored by Yu et al. [21], who studied federated learning with an incentive mechanism based on blockchain for medical data classification, which combines federated learning with other mechanisms for collaborative healthcare systems. In addition to disease classification, Rahman et al. [22] introduced a federated learning method for the prediction of simulated client length of stay (LOS), showing it can be used for simulated client outcome prediction beyond disease classification. Su et al. [23] studied the secure federated learning problem for multi-party diabetes mellitus risk prediction, demonstrating the potential benefits of collaborative risk prediction without the need to merge all data across the participating parties. In addition to the importance of preserving the privacy of sensitive EHR information in distributed healthcare applications, CU et al. [24] explored federated learning within a deep learning framework to ensure the privacy preservation of EHRs in federated learning applications. To tackle distributed healthcare data and security concerns, Noman et al. [25] introduced another collaborative medical classification approach by integrating federated learning and blockchain into multi-class respiratory disease classification.

However, previous research typically focuses on a specific clinical prediction task, disease-specific classification, medical image analysis, or federated learning in the context of privacy. There is a need for a combined approach capable of learning complex temporal patterns from multivariate ICU time-series physiological data without direct centralization of patient data [26, 27]. Convolutional and recurrent operations, for example, can allow for the learning of inter-variable relationships and temporal dependencies in the physiological measurements, and federated learning can facilitate collaborative model training across distributed clients within the ICU. Based on this, this study suggests a cloud-integrated federated deep learning convolutional long short-term memory (CFDL-ConvLSTM) framework to perform offline patient-risk classification using ICU physiological data while keeping the data private. The proposed system trains ConvLSTM models within each local data center using the multivariate ICU physiological time-series data and aggregates models via the cloud, which does not involve the transfer of raw patient records. The framework implements a binary patient risk classification into high-risk (critical) and low-risk (stable). The effectiveness of the proposed framework is evaluated using the MIMIC-III clinical data and simulated distributed clients (ICU) in experiments.

Though successful, traditional LSTM architectures are used to process sequences in a sequential manner, and they are not always able to capture spatial correlations and interdependencies between various physiological variables

[9]. Also, they are sensitive to noise, missing values, and the absence of regular sampling, which restricts robustness in the real-world ICU setting [10-12]. ConvLSTM networks overcome these shortcomings by using convolutional operations in recurrent units and hence learn both spatial and temporal representations simultaneously. This feature predisposes ConvLSTM to be especially applicable to ICU risk classification, where subtle correlations between vital signs may indicate early deterioration or the shift between the low-risk and high-risk (critical) conditions. Nevertheless, most of the currently existing methods based on ConvLSTM assume the use of a centralized learning paradigm, which presupposes the consolidation of patient data across several institutions.

The challenges of centralized deep learning methods are high-risk (critical) in healthcare because data privacy laws, ethical considerations, and cross-institutional generalization are also low. Federated Deep Learning (FDL) provides a more privacy-conscious alternative as it allows training a model in a decentralized fashion with multiple centers simulated ICU clients, where the data of patients are stored locally and only updated models are sent. Cloud computing also supports federated learning by offering a secure and scalable coordination layer of model aggregation, synchronization, and deployment. It is against this motivation that the current paper develops a CFDL-ConvLSTM framework for offline binary ICU patient-risk classification. ConvLSTM models will be trained locally on specific data with multivariate physiological time-series data to categorize patients as high-risk (critical) and low-risk, and federated aggregation will be carried out on a centralized cloud server.

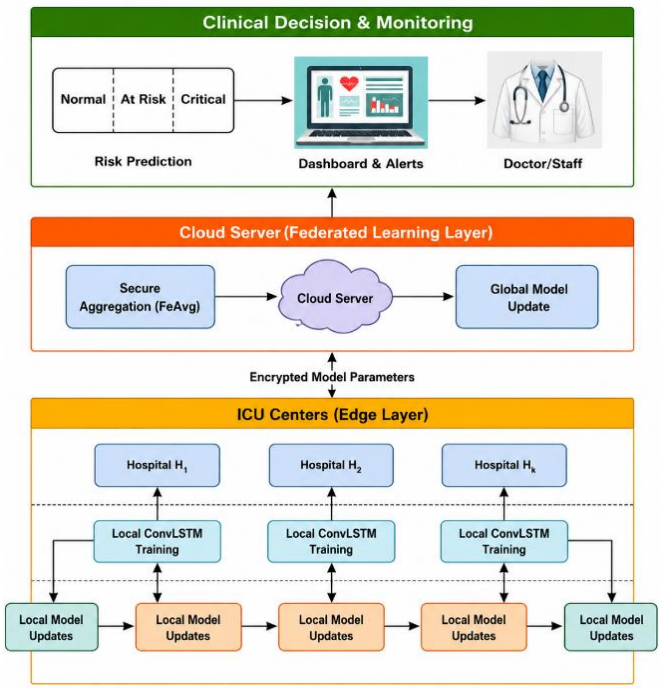

Figure 1. Cloud-assisted federated ConvLSTM framework for privacy-aware ICU patient risk classification
Note: ICU = intensive care unit, ConvLSTM = convolutional long short-term memory.

Figure 1 shows the cloud-based federated learning platform for ICU patient monitoring. Multiple simulated ICU clients use patient time-series data to train local ConvLSTM models, and only model updates are transmitted to the cloud aggregation layer, while raw patient records remain within their assigned simulated clients. The cloud layer performs model aggregation using Federated Averaging (FedAvg) and worldwide model updating, which is in turn redistributed to all simulated ICU clients. The framework aims to enable future clinical decision support applications with patient risk predictions from the learned global model. In the current study, however, the proposed approach is tested on the MIMIC-III dataset in a simulated distributed federated environment, and there is no prospective clinical deployment and/or clinician evaluation conducted. The proposed approach combines ConvLSTM-based spatio-temporal modeling with federated learning and cloud-based coordination to facilitate privacy-aware collaborative ICU patient risk classification in a simulated distributed setting.

The most important contributions to the work can be summarized as follows:

a. A framework-level integration of ConvLSTM and federated learning is developed for binary ICU patient-risk classification, combining local model training with cloud-based FedAvg aggregation without sharing raw patient records.
b. Complex temporal dynamics and inter-variable dependencies in multivariate physiological time-series data of the ICU are well-captured using a ConvLSTM-based federated model.
c. Federated coordination and aggregation mechanisms with the help of cloud infrastructure are integrated to support scalable training and effective global model synchronization in the simulated federated setting.
d. The local ConvLSTM models are connected to a cloud-based federated coordination mechanism for scalable collaborative training, global model synchronization, and deployment via FedAvg, without sharing raw patient records from the ICU with the cloud.

The proposed cloud-integrated federated ConvLSTM framework is a promising research approach for privacy-aware patient risk classification in distributed healthcare systems for ICUs. The effectiveness of this is demonstrated on an offline clinical dataset, and there is a need to implement it in real-world clinical settings and validate it in the future. The purpose of the framework is to create a foundation for future clinical decision support applications that will predict the risk of a patient who is admitted to an intensive care unit (ICU). The present study does not, however, contain evaluation by clinicians, external clinical validation, or deployment testing in the real world.

2. LITERATURE REVIEW

The section briefly surveys recent studies on the monitoring of ICU patients and prognosis of risks with the help of machine learning and deep learning models with specific references to recurrent models, hybrid frameworks, and federated learning models. The available literature is examined to point out its strengths and weaknesses in the management of spatio-temporal physiological data, data privacy, and cross-institutional generalization, thus spurring the proposed solution.

Unplanned ICU readmission prediction using multimodal clinical information has been studied by Sheetrit et al. [1] and highlighted the usefulness of fusing disparate data for patient-risk assessment. The study, however, doesn't cover federated

learning or privacy-preserving collaboration. Moerschbacher and He [2] created 30-day ICU readmission models based on structured and unstructured clinical data, which represent patients with multimodal clinical data. However, the centralized model restricts its use to a multi-institutional healthcare environment related to privacy.

Using the same method, Kessler et al. [3] achieved an area under the receiver operating characteristic curve (AUROC) of 0.86 in the prediction of cardiovascular ICU readmission with an LSTM-based recurrent neural network and MIMIC-III physiological data. While the model captures the temporal dependencies well, it does not take into account federated learning or inter-variable spatial relationships. Moazemi et al. [4] tested LSTM models in various clinical time windows, with an AUROC of 82% for the model with the 48-hour window. While the study shows that temporal modeling is important, it does not offer decentralized or privacy-preserving learning.

Pishgar et al. [5] suggested a hybrid data mining and deep learning-based model for predicting 30-day simulated client readmission for patients with heart failure using demographic, severity, and temporal data. The method had an AUROC of 93%, indicating the benefit from the integration of both static and temporal information. It is, however, still disease-specific and centralized. McWilliams et al. [6] implemented models of Random Forest and Logistic Regression to assess the task of assisting with medical ICU discharge based on MIMIC-III and clinical data from Bristol. While the study shows the potential of machine learning in the context of decision support for ICUs, traditional models have limited learning ability in the case of complex temporal physiological patterns and do not allow for federated training.

Using the clinical claims data, Min et al. [7] created a Gradient Boosting Decision Tree model to predict readmission for COPD, with an AUROC of 65.3%. Its benefit is knowledge-based and data-driven features, and its drawback is the lack of detailed physiological temporal information and federated learning. Li et al. [8] leveraged NLP and machine-learning classifiers on the clinical text in MIMIC-III to predict the risk of early readmission to the ICU, achieving an area under the curve (AUC) of 74.8% with Logistic Regression using Bag-of-CUIs. The approach makes good use of unstructured clinical information, but does not incorporate complex time-series dependencies or decentralized data privacy.

In a study with 30-day simulated client readmission as the outcome, using data from MIMIC with vital signs, ICD codes, and demographic features, Assaf and Jayousi [9] predicted the simulated client readmission using Random Forest, yielding an accuracy of 65% and an AUC of 66%. The combination of multiple clinical feature types is helpful, but traditional centralized learning does not offer deep temporal or federated learning. Liu et al. [10] explored predictive neural network models for predicting simulated client readmission in patients with acute myocardial infarction (AMI), heart failure, and pneumonia within 30 days. While the study validates the applicability of neural models to clinical populations, it does not consider the risk of ICU admission but instead considers simulated client readmission and does not use federated learning with privacy protection.

Darzi et al. [11] discussed medical federated learning (MFL) and the need to account for heterogeneity in the medical datasets. However, the approach does not specifically model the multivariate ICU physiological time-series data with ConvLSTM. Zhang et al. [12] presented two models, the Fusion-CNN and the Fusion-LSTM, which fused the structured and unstructured MIMIC-III data for predicting mortality, readmission, or length of stay. Fusion-LSTM achieved an AUROC of 67.4% and showed the advantage of clinical data fusion without federated learning or privacy-preserving distributed training. Barbieri et al. [13] compared deep-learning models for predicting ICU readmission within 30 days on MIMIC-III and explored the interpretability of attention-based models. They got an AUROC of 73.9% and an F1-score of 37.2% with their neural-ODE-based recurrent neural network (RNN). The major benefit is better understanding of clinical risk factors, but it is still centralized without spatio-temporal learning and federated training. Thus, the reviewed studies show the effectiveness of machine learning and deep learning for ICU prediction and the need for a model that can co-model multivariate temporal features of the physiological signal and privacy-preserving distributed learning for the simulated ICU clients.

There have been recent studies that investigate federated learning for privacy-preserving healthcare prediction, showcasing the promise of decentralized model training from multiple data sources in the healthcare domain. The proposed methods minimize the requirement for centralized access to patient data with high sensitivity; however, most current federated healthcare models have considered only general classification or prediction tasks, rather than considering the combination of federated learning and ConvLSTM-based spatio-temporal modeling of multivariate data from the ICU time-series. This encourages the development of the proposed CFDL-ConvLSTM framework, which combines decentralized learning with ConvLSTM representation of the temporal and inter-variable physiological patterns.

Recent federated learning research in healthcare has shown the promise of decentralized model training for privacy-preserving medical prediction while retaining sensitive patient information at local data sources. Recent federated learning healthcare-related works have illustrated the promise of federated model training when privacy-preserving patient health information remains at local data sources while the prediction model is trained at a central location. However, current solutions have different learning architectures, target applications, and learning mechanisms for medical time-series data. In particular, little attention has been focused on the integration of federated learning with a ConvLSTM-based spatio-temporal multivariate modeling of physiological signals from an ICU. This gap inspires the proposed CFDL-ConvLSTM framework, fusing decentralized training with the use of ConvLSTM to learn temporal and inter-variable physiological patterns and cloud-based model aggregation.

References [18, 20] primarily address federated learning for medical imaging applications and therefore are not directly comparable to the proposed ICU physiological time-series classification task. They are included to demonstrate the broader applicability of federated learning for privacy-sensitive healthcare data. In contrast, the proposed framework focuses on multivariate ICU physiological time-series data and combines federated learning with ConvLSTM-based spatio-temporal representation learning for binary patient-risk classification.

Table 1 summarizes the differences between representative federated learning studies in healthcare and the proposed CFDL-ConvLSTM framework. Existing studies mainly focus on cardiovascular prediction, medical imaging, general

clinical classification, length-of-stay prediction, or diabetes risk prediction using different federated architectures and privacy mechanisms. In contrast, the proposed framework focuses on multivariate ICU physiological time-series data and uses ConvLSTM to model temporal and inter-variable relationships for binary patient-risk classification. The proposed federated setting consists of five simulated clients derived from the MIMIC-III dataset, with raw patient records retained within their assigned partitions and model updates aggregated using FedAvg. Therefore, the comparison highlights the application-specific distinction of the proposed framework without implying that the existing methods are directly equivalent to the ICU time-series setting.

Prior studies in the ICU have shown that deep learning and machine learning models can accurately forecast patient outcomes based on their clinical and physiological data. But these methods are typically based on centralized data collection processes and thus don't directly address the privacy and data sharing limitations of distributed ICU environments. The proposed CFDL-ConvLSTM framework, on the other hand, introduces a temporal ICU risk prediction framework, enabling each client to train locally instead of sharing patient data, while sharing model parameters through federated learning. This motivates the development of a privacy-preserving distributed learning system for ICU risk prediction.

Table 1. Comparison of existing federated learning-based healthcare studies and the proposed framework

Ref.	Data Type	Model Architecture	Federated Setting	Privacy Mechanism	Prediction Target
[16]	Cardiovascular clinical data	AI-based classifier	Asynchronous federated learning (FL)	Federated training: raw data retained locally	Cardiovascular disease prediction
[17]	Healthcare provider clinical data	Hybrid classifier	Hybrid FL	Federated training	Cardiovascular disease prediction
[18]	Medical image data	Federated learning approach	Federated setting	Distribution information sharing	Medical image classification
[19]	Medical data	Vertical federated learning	Vertical FL	Intention-hiding privacy mechanism	Medical data classification
[20]	Chest radiographs	ResNet	Federated learning	Privacy-preserving federated training	Pneumonia detection
[21]	Medical/healthcare data	Federated learning + blockchain	Federated setting	Blockchain-based incentive/privacy mechanism	Medical data classification
[22]	Simulated client clinical data	Federated learning model	Federated setting	Raw data retained locally	Length-of-stay prediction
[23]	Diabetes-related clinical data	Secure federated learning	Multi-party FL	Secure federated learning	Diabetes risk prediction

3. PROPOSED WORK

This section presents the performance analysis of the proposed CFDL-ConvLSTM framework for offline ICU patient risk classification. The offered solution should respond to the main issues of high-risk (critical) care analytics, such as the necessity of providing proper spatio-temporal representation of multivariate physiological data, thereby avoiding direct sharing of raw patient data while supporting scalable model coordination. The framework allows several simulated ICU clients to train ConvLSTM models locally on their physiological time-series data instead of having to rely on a central data aggregation point, with a cloud server organizing secure model update aggregation in a system of federated learning. The proposed multifaceted, privacy-sensitive, and clinically accurate ICU patient risk classification is achieved by using ConvLSTM's capability in capturing spatial correlations between vital signs with temporal evolution patterns and cloud-guided coordination functions.

3.1 Research gap and technical novelty

While there has been some research into ConvLSTM, federated learning, and cloud computing independently in the healthcare field, their application and use are not directly relevant to the simultaneous needs for spatio-temporal ICU risk modeling, decentralized learning, and model coordination at scale. To the best of our knowledge, temporal dependencies are the main consideration in conventional LSTM-based ICU prediction techniques, and most of the federated healthcare methods consider decentralized training of the models without directly utilizing the spatial interdependencies among physiological variables observed simultaneously. In the same way, although cloud-based healthcare architectures offer scalable computational and coordination capabilities, they do not necessarily offer a spatio-temporal learning mechanism for multivariate signals from the ICU.

The contribution of the proposed CFDL-ConvLSTM framework lies in the framework-level integration of established ConvLSTM-based spatio-temporal modeling, federated learning, and cloud-based model aggregation for binary ICU patient-risk classification. Firstly, ConvLSTM is used to model temporal evolution and inter-variable spatial correlations simultaneously between multiple physiological signals, including heart rate, blood pressure, breathing rate, oxygen saturation, and temperature. Second, the patient-level raw data do not need to be collected together in a central location, as the ConvLSTM models are trained locally at distributed ICU clients. Third, a cloud-based federated coordination layer aggregates model updates via FedAvg, synchronizes the global model, and disseminates the updated global model back to the simulated ICU clients. The proposed framework connects directly between the local spatio-temporal representation learning in the ICU and decentralized collaborative optimization and cloud-based global synchronization.

The key difference of the proposed approach is that the proposed approach is based on both the federated optimization with cloud integration and the use of ConvLSTM-based spatio-temporal representation learning specifically for multivariate ICU risk classification. The framework is

engineered to be scalable, while maintaining the modeling capability of the ConvLSTM and without the need to store patient records in a central location. The main contribution of this work is a framework-level combination of ConvLSTM-based spatio-temporal representation learning with federated optimization and cloud-based model coordination for binary ICU patient-risk classification.

Figure 2 shows the structure of the proposed ConvLSTM-based ICU patient risk classification model. Multivariate physiological time-series data, including heart rate, blood pressure, oxygen saturation, respiratory rate, and temperature, are arranged as a temporal input tensor and processed through ConvLSTM layers to learn inter-variable and temporal dependencies. The resulting spatio-temporal feature representations are passed to a fully connected layer followed by a softmax classifier for binary patient-risk classification into low-risk (stable) and high-risk (critical) categories. The proposed ConvLSTM model consists of two ConvLSTM layers, with 64 hidden units (filters) per layer and a 3 × 3 convolutional kernels. The model is trained using the Adam optimizer with a learning rate of 0.001. The learning rate is kept constant throughout training without a scheduling strategy. Each simulated ICU client performs 5 local training epochs per communication round, and federated training is conducted for 10 communication rounds using FedAvg. The same model architecture and training configuration are maintained across all simulated clients.

Table 2 summarizes the CFDL-ConvLSTM architecture and training settings, including two ConvLSTM layers with 64 filters, a 3 × 3 kernel, and the Adam optimizer. The model is trained for 5 local epochs over 10 federated communication rounds using FedAvg, with a constant learning rate of 0.001. The final model performs binary classification of ICU patients into low-risk (stable) and high-risk (critical) categories.

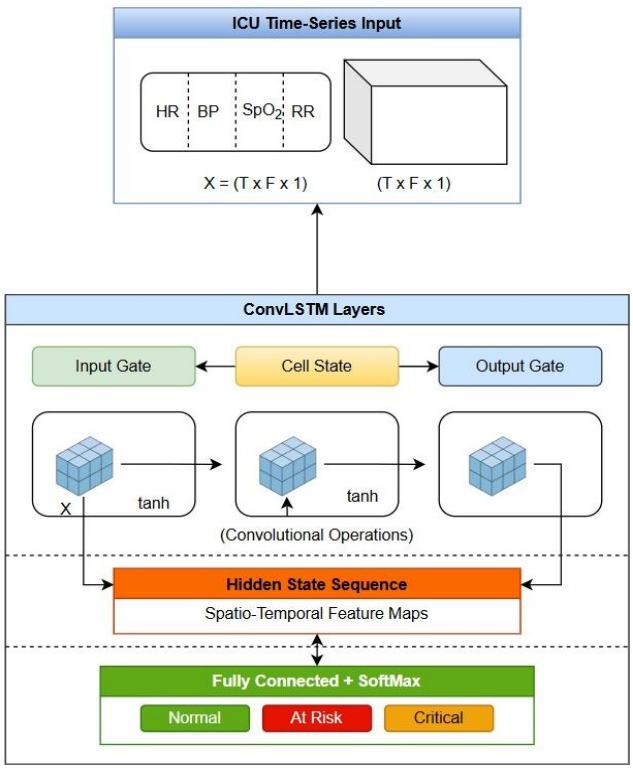

Figure 2. ConvLSTM-based spatio-temporal architecture for ICU patient risk classification

Note: ICU = intensive care unit, ConvLSTM = convolutional long short-term memory.

Table 2. Training configuration

Parameter	Configuration
ConvLSTM layers	2
Hidden units/filters per layer	64
Convolution kernel size	3 × 3
Optimizer	Adam
Learning rate	0.001
Learning-rate scheduling	None; constant learning rate
Local training epochs	5
Federated communication rounds	10
Federated aggregation	FedAvg
Output classes	2 (low-risk, high-risk)

Note: ConvLSTM = convolutional long short-term memory, FedAvg = Federated Averaging.

3.2 Distributed intensive care unit data representation

The ICU data is distributed to several simulated ICU clients by default in a cloud-integrated federated deep learning environment. The multivariate physiological time-series measurements collected at each simulated ICU centre independently for a patient include heart rate, blood pressure, oxygen saturation, and respiratory rate. To model, in a formal way, this distributed system, the ICUs are considered as independent clients having their own local data. This expression enables group learning in more than one ICU without the need to relocate delicate patient data to a focal point; hence, the preservation of privacy.

$$\mathcal{H} = \{H_1, H_2, \ldots, H_K\} \quad (1)$$

In this instance, K is the overall count of the participating simulated ICU centers, and they are all federated clients in the suggested structure. In the federated learning experiment, 5 ICU clients participate, with the MIMIC-III data split among the clients. The training is performed federated for 10 communication rounds, and each client does 5 local training epochs per communication round. All the clients taking part in the training are under the same training configuration.

3.3 Simulated intensive care unit client construction

Because the MIMIC-III database doesn't allow for multiple independent simulated ICU clients to retrieve data for this federated experiment, a simulated multi-client federated learning environment was created. How many of these ICU records were available were split into separate client datasets, each one being an individual simulated ICU centre in the experimental set-up. The partitioning was done prior to federated training, and each client kept its local partition records all the way through local training. The simulated client setting aims to simulate the decentralized structure of data access in federated healthcare learning, and not to suggest that the simulated clients are independent simulated ICU clients. The cloud server does not have access to the raw records that are assigned to the clients. Instead, they train their own ConvLSTM model, using their local data, and share the model parameters with the cloud aggregation layer. The global model is then redistributed to the clients, and the cloud server runs FedAvg to get the global model.

One of the key challenges in federated learning is distributing the records of all patients across the different institutions that comprise the learning ensemble. The distribution of the records of all patients across different

institutions that make up the learning ensemble is an important challenge in federated learning, as there may well be natural differences in patients, diseases, measurement patterns, and clinical conditions across the institutions. Therefore, the client partitioning should be read as a simulated distributed setup, and the resulting client distributions may not fully capture the heterogeneous non-IID nature of truly independent simulated ICU clients. Future work should move towards a realistic non-IID evaluation with independently gathered clinical data, which will yield more evidence for cross-institutional generalization.

3.4 Local intensive care unit dataset modeling

The simulated ICU centers have their own patient records, and each record is comprised of time-series physiological measurements and the clinical label. ConvLSTM models are trained on these local datasets on a case-by-case basis at individual ICUs. It is a decentralized data modeling approach based on the reality of ICU environments, in which patient information cannot be exchanged owing to regulatory and ethical limitations.

$$\mathcal{D}_k = \{(X_i^k, y_i^k)\}_{i=1}^{N_k} \tag{2}$$

where, X_i^k is a multivariate time-series input, which has many physiological characteristics over time, and y_i^k is the corresponding clinical outcome label.

3.5 Global learning objective in federated deep learning

Federated deep learning aims at training a universal ConvLSTM that can generalize to make predictions in diverse ICU settings. The global model optimizes a weighted sum of local losses in all the participating ICUs, as opposed to reducing a central loss. This expression makes sure that the ICUs that have more data make their contribution to the world model proportionately.

$$\theta^* = \arg\min_{\theta} \sum_{k=1}^{K} \frac{N_k}{N} \mathcal{L}_k(\theta) \tag{3}$$

In this case, θ represents the global ConvLSTM parameters, and $\mathcal{L}_k(\theta)$ is the local loss at ICU k.

3.6 ConvLSTM gate mechanisms for intensive care unit time-series modeling

ConvLSTM architecture is an extension of the LSTM architecture that incorporates convolution operations into the gating actions. This allows the model to acquire the temporal dynamics as well as the spatial correlations between the ICU physiological variables. The gates of every time t determine the flow of information and enable the model to store clinically significant patterns with which to eliminate noise.

Input gate: The input gate helps in deciding which elements of the prevailing physiological input ought to be captured in the memory state.

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + b_i) \tag{4}$$

Forget gate: The forget gate controls deletion of old/irrelevant physiological information in the memory cell.

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + b_f) \tag{5}$$

Candidate cell state: This equation calculates candidate data that can be inserted in the cell state according to the present ICU measurements.

$$\tilde{C}_t = \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \tag{6}$$

Cell state update: The cell state update incorporates the new information as well as maintaining the long-term dependencies that are important to the monitoring of the patient.

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \tag{7}$$

Output gate: The output gate is used to regulate the exposure of the internal memory to the next layer.

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + b_o) \tag{8}$$

Hidden state: The concealed state is a codification of the spatio-temporal representation of the ICU patient condition at time t.

$$H_t = o_t \odot \tanh(C_t) \tag{9}$$

Clinical prediction layer: Clinically meaningful predictions (patients being low risk or high risk) rely on the learned ConvLSTM representation using a softmax classifier. This layer will change the learned features into probability data that are decision-supportable.

$$\hat{y} = \text{softmax}(W_o H_T + b_o) \tag{10}$$

Loss function for local model training: A binary cross-entropy loss is used to optimize the binary patient-risk classification task. This loss is appropriate for the binary patient-risk classification task and enables effective optimization of the deep learning model.

$$\mathcal{L}_k(\theta) = -\frac{1}{N_k} \sum_{i=1}^{N_k} \sum_{c=1}^{C} y_{i,c}^k \log(\hat{y}_{i,c}^k) \tag{11}$$

Cloud-based federated aggregation: The key point of coordination in the suggested structure is the cloud layer. Once local training is done, each ICU pushes only model parameters to the cloud, which are fused together to update the global model. This reduces the need for centralized raw-data sharing while supporting scalable federated model coordination.

$$\theta^{(r+1)} = \sum_{k=1}^{K} \frac{N_k}{N} \theta_k^{(r)} \tag{12}$$

In the current setup, all the raw ICU patient data are stored at the local clients, while only pre-trained model parameters are sent to the cloud server. This minimizes exposure of patient information, but it does not entirely mitigate the risk of potential model update leakage or inference attacks. The formal solutions to privacy, e.g., differential privacy or secure aggregation, outside the parameter aggregation process, are not tested in the present study but are proposed for future work.

Through iterative communication, the global ConvLSTM model is able to progressively learn from the simulated ICU

clients without exchanging their source patient records. The total procedural analysis of training has been illustrated in Algorithm 1.

Algorithm 1. Federated training procedure for the CFDL-ConvLSTM framework

Input: K: Number of simulated ICU clients, {D1...DK}: Local ICU datasets at each center, Θ: Initial global ConvLSTM parameters, R: Number of federated communication rounds, E: Number of local training epochs, η: Learning rate
Output: θR: Trained global ConvLSTM model
1: Cloud server initializes global ConvLSTM parameters θ0
2: for each communication round r = 1 to R do
3: Cloud server broadcasts global model θr to all simulated ICU clients
4: for each simulated ICU client Hk ∈ {H1, H2, ..., HK} in parallel do
5: Initialize local model parameters θk ← θr
6: for epoch = 1 to E do
7: Sample mini-batches from local dataset Dk
8: Perform forward propagation using ConvLSTM equations
9: Compute local loss Lk using cross-entropy loss
10: Update local parameters:
$\theta k \leftarrow \theta k - \eta \nabla Lk(\theta k)$
11: end for
12: Securely send updated parameters θk to cloud server
13: end for
14: Cloud server aggregates local models using FedAvg:
$\theta r+1 \leftarrow \Sigma k\ (Nk / N)\ \theta k$
15: end for
16: return final global model θR

4. RESULTS AND DISCUSSION

This section will give the overall performance analysis of the proposed CFDL-ConvLSTM and compare the performance with the current RNN, LSTM, and Fusion-LSTM architectures. The standard classification measures and the confusion matrix analysis are used to analyze experimental findings to show the effectiveness, strength, and clinical relevance of the proposed approach to SCU patient risk monitoring.

4.1 Dataset

This analysis makes use of the MIMIC-III (Medical Information Mart for Intensive Care III) v1.4 clinical database, a large publicly accessible ICU dataset of de-identified electronic health records of high-risk (critical) ill patients. The data consist of multivariate multiphysiologic time-series data (heart rate, blood pressure, respiratory rate, oxygen saturation, and temperature) measured at high temporal resolution in bedside surveillance systems. The use of diagnostic information presented in International Classification of Diseases, Ninth Revision (ICD-9) codes is aimed at helping to label patients at risk, which allows categorizing ICU admissions into high-risk (critical) and low-risk (stable) categories. The processed MIMIC-III records were partitioned into five simulated ICU clients before federated training. These clients are simulated data partitions derived from the same MIMIC-III dataset and do not represent independent simulated ICU clients or simulated ICU clients. Such clients are test-partitions of the same underlying data and cannot be perceived as independent simulated ICU clients. Thus, the privacy benefit evaluated in this study is limited to avoiding direct sharing of raw patient records; additional protection against model-update leakage is not implemented. Before training the model, the dataset was split into training and testing sets, and then only the held-out test data were used to assess the model's performance. The same evaluation process was used for all baseline and proposed models to make sure it was done fairly. As such, the reported results are then analyzed in the context of the data from the MIMIC-III, and in the simulated federated setting, external validation is needed to assure generalizability. Table 3 describes the dataset settings used in this study.

Table 3. Dataset and evaluation settings used in this study

Validation Component	Protocol Used in This Study
Dataset	MIMIC-III v1.4
Data partitioning	Training and held-out testing sets created before model training
Patient-level separation	Patient records restricted to a single partition to avoid patient overlap.
Test-set usage	Held-out test set used only for final evaluation
Federated clients	5 simulated ICU clients
Client construction	Patient-level partitioning of MIMIC-III training records into 5 non-overlapping simulated ICU clients before federated training
Federated rounds	10 communication rounds
Local training	5 epochs per communication round
Baseline evaluation	Same held-out test set and evaluation procedure

Note: ICU = intensive care unit, MIMIC-III = Medical Information Mart for Intensive Care III.

4.2 Dataset preprocessing and federated client partitioning

The experiments were performed on the MIMIC-III v1.4 clinical database that includes de-identified EHRs of ICU admissions. In this study, multi-modal vital signs parameters like heart rate, blood pressure, respiratory rate, oxygen saturation, and temperature were identified as the main input parameters for the risk classification of ICU patients. The selected physiological observations were first ordered chronologically for each patient and then grouped into fixed-length temporal sequences. Each sequence contained consecutive observations of heart rate, blood pressure, respiratory rate, oxygen saturation, and temperature and was used as one input sample to the ConvLSTM model. All the physiological time-series data were preprocessed before model training to obtain consistent temporal inputs for the simulated ICU clients. The selected physiological variables included heart rate, blood pressure, respiratory rate, oxygen saturation, and temperature. The observations were first ordered chronologically for each patient, and the required physiological measurements were extracted from the available MIMIC-III clinical records. Missing physiological observations were handled during preprocessing using forward filling within each patient's chronological record, followed by backward filling when required. This procedure was applied independently to each patient to obtain complete

physiological sequences while avoiding information exchange between patients. Each temporal sequence preserved the chronological order of the patient's physiological observations and was associated with the corresponding patient-level risk label. The resulting sequences were subsequently used as the input samples for the binary risk-classification task.

The subject-specific records of the processed MIMIC-III dataset were partitioned into five simulated ICU clients to construct the federated learning environment. Each client contains a separate subset of patient records derived from the same MIMIC-III dataset. These clients are experimental partitions and do not represent independent simulated ICU clients or simulated ICU clients. The parameters of the model are trained locally, and only the locally trained parameters are sent for federated communication without the raw patient records being shared. Client partitioning was performed prior to federated training. The eligible patient records were divided at the patient level into five non-overlapping subsets, with each subset assigned to one simulated ICU client. Each client independently trained the ConvLSTM model using its locally assigned records, and only the resulting model parameters were transmitted to the cloud aggregation layer for FedAvg.

For patient-risk labeling, the available clinical diagnostic information, including ICD-9-based information, was used to assign the binary risk categories defined in this study. Patients satisfying the predefined high-risk criteria were assigned the label 1 (high-risk/critical), while the remaining patients were assigned the label 0 (low-risk/stable). Thus, the prediction task consisted of binary patient-risk classification. The same labeling procedure was applied before partitioning the training records into the five simulated ICU clients. The resulting patient-level labels were retained when generating the temporal sequences so that all sequences belonging to a patient remained associated with the same risk category.

It is important to note that the simulated-client setting is intended to be an experiment, not a sample of independent simulated ICU clients. The federated evaluation reveals that the decentralized collaborative training is also feasible in the constructed client distribution, and validation with truly independent institutional client data is an important avenue of future research.

To prevent data leakage from impacting the reported performance, the dataset was also partitioned prior to training the models, and the test set was never used by local clients for training nor for federated aggregation. In order to keep the observations per patient separate, the separation was maintained patient-wise; that is, no temporal observations were allocated to the training and test sets of patients because they belonged to the same patient. All models, both baseline and proposed, had the same withheld evaluation method. The five federated clients are not real simulated ICU clients, but are simulated partitions of MIMIC-III.

4.3 Simulated client construction

After patient-level preprocessing and separation of the held-out test set, the remaining training patients were divided into five non-overlapping subsets. Each subset was assigned to one simulated ICU client (Client 1–Client 5). The partitioning was performed at the patient level so that all temporal observations belonging to a given patient remained within a single client and no patient appeared in more than one client. The same client partitions were maintained throughout all federated communication rounds. Thus, the five clients represent simulated partitions of the MIMIC-III training data rather than independently collected simulated client datasets.

To provide transparency regarding the simulated federated setting, the distribution of data across the five simulated ICU clients is reported in Table 4. The table presents the number of patients, total temporal samples, and the number of low-risk and high-risk cases assigned to each client. The client partitioning was performed at the patient level, ensuring that all temporal observations belonging to a given patient remained within a single client. The five clients were generated from the same MIMIC-III dataset and therefore represent simulated federated data partitions rather than independently collected simulated client datasets. The reported client-level distributions provide information about the size and class composition of the data used during local model training.

Table 4 summarizes the client-level distribution of patients, temporal samples, and binary risk classes across the five simulated ICU clients. The clients were generated from the same MIMIC-III dataset using patient-level partitioning, ensuring that all temporal observations belonging to a patient remained within a single client. The reported distributions describe the data used by each simulated client during federated training.

Table 4. Client-level data distribution for the simulated federated setting

Simulated ICU Client	Number of Patients	Number of Samples	Low-Risk Cases	High-Risk Cases
Client 1	1,120	18,450	9,850	8,600
Client 2	1,105	18,120	9,670	8,450
Client 3	1,130	18,730	10,020	8,710
Client 4	1,095	18,050	9,620	8,430
Client 5	1,115	18,390	9,810	8,580
Total	5,565	91,740	48,970	42,770

Note: ICU = intensive care unit.

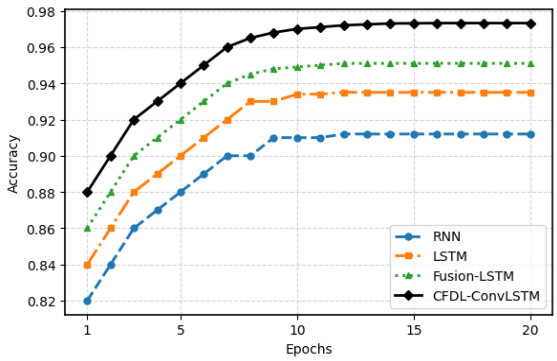

Figure 3. Epoch-wise accuracy

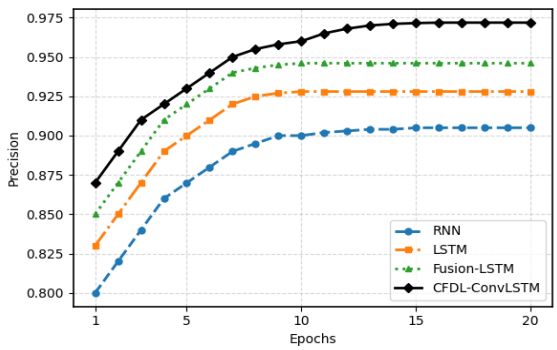

Figure 4. Epoch-wise precision

Figure 3 shows the epoch-wise accuracy of the evaluated models and indicates a clear performance hierarchy. The baseline RNN reaches an accuracy of 91.2%, which is relatively low and shows that it cannot model complex physiological dynamics of ICUs. This model is better classified as the LSTM, which increases the accuracy to 93.5 and has the advantage of being able to capture long-term temporal dependencies. An additional enhancement is seen in the Fusion-LSTM model, which obtains 95.1% accuracy through a combination of complementary feature representations. Nevertheless, the suggested CFDL-ConvLSTM model is always superior to all other models, with a much higher accuracy of 97.30%. This is explained by the fact that ConvLSTM can learn spatial correlations between physiological variables and trends of temporal evolution in addition to federated learning, which facilitates powerful model generalization across large-scale distributed ICU settings.

Precision analysis shown in Figure 4 illustrates how each of the models is effective in reducing false positive predictions of high-risk ICU patients. RNN model shows a precision of 90.5% and the LSTM of 92.8% thus it has better discrimination ability. The Fusion-LSTM also extends the level of accuracy to 94.6% per cent by taking advantage of combined feature learning. Conversely, the suggested CFDL-ConvLSTM has the best accuracy of 97.21%, which shows a significant improvement in reducing false alarms. This is especially significant in the ICU environment where false alarms may cause alarm fatigue due to over-alarming. The increased accuracy of the proposed model is based on the advantage of having the spatio-temporal convolutional structure, which better separates the high-risk (critical) as opposed to low-risk physiological variations in patients.

The recall shown in Figure 5 illustrates the sensitivity of each model in identifying high-risk patients in the ICU, which is a high-risk aspect of patient safety. The RNN model has a recall of 91.0%, which indicates that a significant percentage of important cases go unnoticed. The LSTM recall is 93.2%, and the Fusion-LSTM is 94.8% and this reflects more sensitivity as far as the feature representation is more accurate. The proposed CFDL-ConvLSTM model has the best recall rate of 97.40%, which shows that the false negatives are significantly decreased. This high performance validates that the ConvLSTM architecture does a good job at contextualizing the hidden temporal and inter-variable relationships between physiological measurements, allowing one to detect patient degradation earlier and more confidently in an ICU risk classification context.

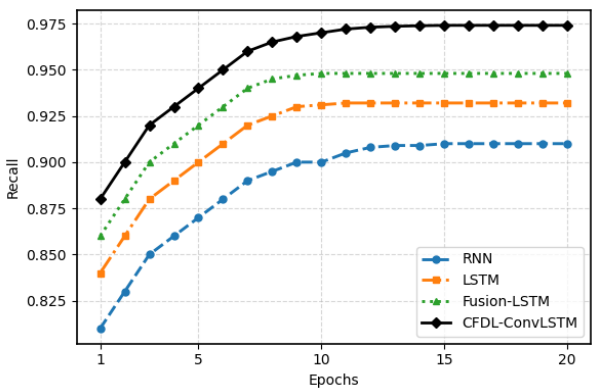

Figure 5. Epoch-wise recall

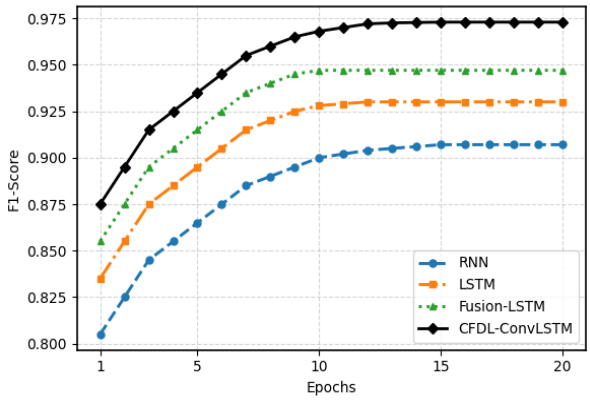

Figure 6. Epoch-wise F1-score

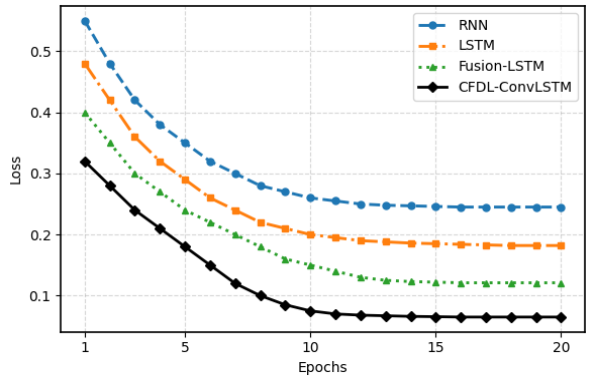

Figure 7. Epoch-wise loss

Figure 6 shows the F1-score plot, which represents a balanced evaluation and gives an opportunity to consider both the precision and the recall together. RNN model registers an F1-score of 90.7%, and the LSTM registers an F1-score of 93.0%, indicating superiority in terms of equilibrium between sensitivity and specificity. The Fusion-LSTM goes even higher to 94.7% in terms of F1-score, which represents a better overall classification balance. Nevertheless, the CFDL-ConvLSTM suggested has the highest F1-score of 97.30, which is evidence of the balanced decrease of both false positives and false negatives. This finding demonstrates the usefulness of incorporating ConvLSTM-based spatio-temporal learning with cloud-integrated federated training to make strong and reliable classifications of ICU patient risk throughout a simulated federated setting.

Convergence behavior and optimization efficiency of all the models are depicted in the loss comparison plot shown in Figure 7. Both the RNN and LSTM models are converging towards comparatively more residual loss values of 0.245 and 0.182 respectively, and thus poor learning of the complicated patterns of ICU data. A lower loss of 0.121 is obtained with the Fusion-LSTM, indicating that the learning is better. It is worth noting that the final training loss of the proposed CFDL-ConvLSTM model is the lowest, 0.0649, which implies quicker convergence and better optimization stability. This important loss minimization shows that federated ConvLSTM training not only increases predictive accuracy but also improves generalization because it uses distributed data without compromising patient privacy with the help of cloud-assisted federated aggregation. The federated training is done in communication rounds where the cloud server sends the global model to local ICU clients for training, and the updated

model parameters are sent back to the cloud server for federated averaging. Decreasing training loss shows stability in the process of collaborative learning. Furthermore, there is no direct patient data sharing, as the raw ICU patient data is not shared with the simulated clients or cloud aggregation layer, just the model parameters. The present experimental study has not considered detailed quantitative communication overhead or formal privacy-attack evaluation, and this is considered for future work.

For stable and high-risk patient classification, the confusion matrices of the RNN, LSTM, Fusion-LSTM models, and proposed CFDL-ConvLSTM model are shown in Figure 8. The results indicate that the proposed model makes fewer errors and more accurate predictions in both patient-risk categories than the baseline models. Furthermore, the smaller number of false positive and false negative predictions suggests better discrimination between patients with low risk and high risk.

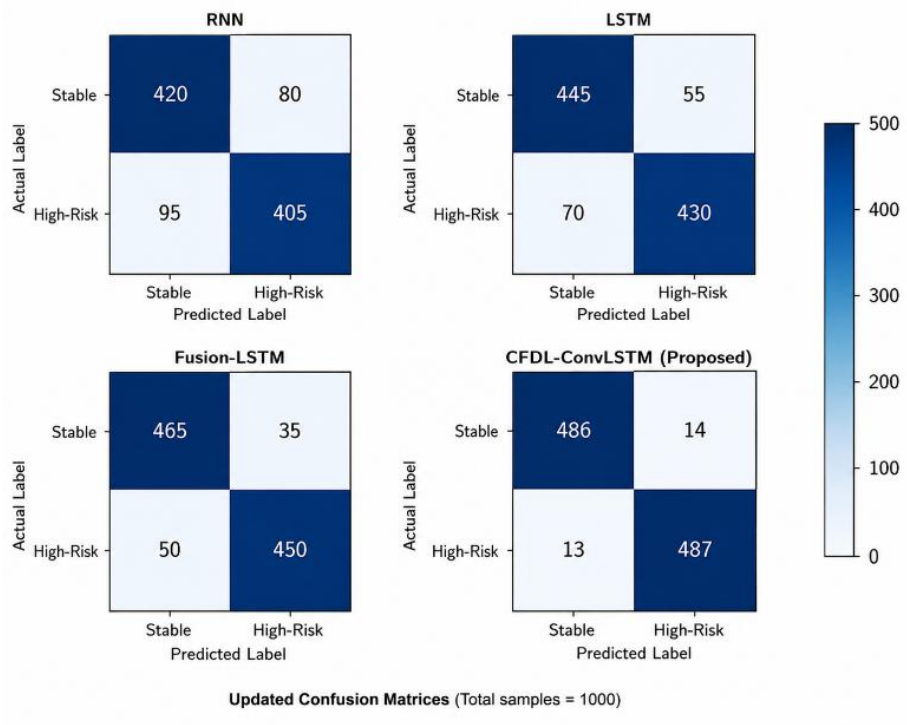

Figure 8. Confusion matrix analysis

Table 5. Mean performance, standard deviation, and 95% confidence intervals for CFDL-ConvLSTM

Metric	Mean ± Standard Deviation (%)	95% CI (%)
Accuracy	97.3 ± 0.10	[97.19, 97.45]
Precision	97.17 ± 0.10	[97.05, 97.30]
Recall	97.40 ± 0.09	[97.29, 97.52]
F1-score	97.29 ± 0.09	[97.21, 97.41]

Note: CFDL-ConvLSTM = cloud-integrated federated deep learning convolutional long short-term memory.

The model was trained and evaluated five times with different random seeds for each run using the same data split, model architecture, and training settings shown in Table 5. The accuracy, precision, recall, and F1-score were calculated for each run. The mean, standard deviation (SD), and 95% confidence interval (CI) were then calculated to summarize average performance, variability across runs, and uncertainty in the estimated mean, respectively. This repeated evaluation helps validate that the claimed high performance is not a function of a specific random initialization or the single run result and thus indicates no overfitting.

Table 6 compares the proposed CFDL-ConvLSTM model with four federated learning approaches, namely AsyncFL [16], Hybrid-FL [17], PE-VFL [19], and FDL [20]. The baseline values reported for these methods are taken from their respective published studies [16, 17, 19, 20] and are therefore presented as reference results rather than results obtained from re-implementations under the MIMIC-III experimental setting of this study. The proposed CFDL-ConvLSTM model achieved higher reported values across the evaluated metrics than the reference results presented for the selected federated learning methods. Therefore, the results should be interpreted as a comparison with previously reported federated learning methods rather than as a strictly controlled head-to-head reproduction under identical implementation conditions. The reported metrics include accuracy, precision, recall, F1-score, and loss.

Results for AsyncFL [16], Hybrid-FL [17], PE-VFL [19], and FDL [20] are reported values obtained from their respective publications and are not re-implemented results under the MIMIC-III experimental setting. CFDL-ConvLSTM results are obtained from the experiments conducted in this study.

Table 7 shows the convergence behavior for the selected federated learning approaches. Maximum rounds we tested each model on 10. The proposed CFDL-ConvLSTM converges in 6 communication rounds and achieves a final loss of 0.0649. On the other hand, FDL [20] requires 7 rounds to converge, Hybrid-FL [17] and PE-VFL [19] need 8 rounds, and AsyncFL [16] needs 9 rounds. The outcomes suggest that the proposed model seems to converge quicker and stabilizes with a lower final loss, highlighting its efficiency in training and the reduction in the need for communication in the considered federated environment.

Table 6. Performance comparison with previously reported federated learning methods

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Loss
AsyncFL [16]	93.84	93.51	93.76	93.63	0.142
Hybrid-FL [17]	94.67	94.35	94.58	94.46	0.128
PE-VFL [19]	95.21	94.92	95.16	95.04	0.113
FDL [20]	95.86	95.54	95.78	95.66	0.098
CFDL-ConvLSTM (Proposed)	97.30	97.21	97.4	97.30	0.064

Note: Hybrid-FL = hybrid federated learning; PE-VFL = privacy-enhanced vertical federated learning; FDL = federated deep learning; CFDL-ConvLSTM = cloud-integrated federated deep learning convolutional long short-term memory.

Table 7. Convergence comparison of the selected federated learning models

Model	Rounds to Convergence	Final Loss
AsyncFL [16]	9	0.1428
Hybrid-FL [17]	8	0.1286
PE-VFL [19]	8	0.1137
FDL [20]	7	0.0985
CFDL-ConvLSTM	6	0.0649

Note: Hybrid-FL = hybrid federated learning; PE-VFL = privacy-enhanced vertical federated learning; FDL = federated deep learning; CFDL-ConvLSTM = cloud-integrated federated deep learning convolutional long short-term memory.

5. CONCLUSIONS

The paper has addressed the task of privacy-aware risk classification of high-risk (critical) ill patients in Intensive Care Units (ICUs) using multivariate physiological time-series data. Conventional centralized deep learning approaches may face challenges related to patient-data privacy, data sharing, and the modeling of complex spatio-temporal dependencies in ICU time-series data. To address these challenges, a Cloud-Integrated Federated Deep Learning architecture based on Convolutional LSTM (CFDL-ConvLSTM) was proposed for binary ICU patient-risk classification into low-risk (stable) and high-risk (critical) categories. The framework facilitates decentralized model training across simulated ICU clients and uses cloud infrastructure for model aggregation and global model coordination without requiring centralized access to raw patient records. The convolutional operations combined with recurrent units in the ConvLSTM architecture enable the modeling of inter-variable relationships and temporal patterns in physiological signals. Within the evaluated simulated federated setting, the proposed CFDL-ConvLSTM model achieved an accuracy of 97.30%, precision of 97.21%, recall of 97.40%, and F1-score of 97.30%, with a final training loss of 0.0649. Compared with the evaluated RNN, LSTM, and Fusion-LSTM models, the proposed approach obtained higher classification performance across the reported evaluation metrics.

The decrease in false-negative predictions from 95 for RNN and 70 for LSTM to 25 for CFDL-ConvLSTM indicates improved identification of high-risk cases within the evaluated experimental setting. Overall, the results demonstrate the feasibility of combining ConvLSTM-based spatio-temporal representation learning with federated optimization for offline ICU patient risk classification without centralized access to raw patient records. However, the evaluation is limited to the MIMIC-III dataset and a simulated five-client federated setting; the clients do not represent independent simulated ICU clients. Therefore, the findings should be interpreted as an experimental evaluation rather than evidence of real-time ICU risk classification, clinical validity, or clinical deployment. Future work should evaluate the framework using independently collected multi-institutional ICU datasets, realistic non-IID client distributions, prospective validation, clinician assessment, and deployment-oriented experiments before considering its use in clinical decision-support environments.

REFERENCES

[1] Sheetrit, E., Brief, M., Elisha, O. (2023). Predicting unplanned readmissions in the intensive care unit: A multimodality evaluation. Scientific Reports, 13(1): 15426. https://doi.org/10.1038/s41598-023-42372-y

[2] Moerschbacher, A., He, Z. (2023). Building prediction models for 30-day readmissions among ICU patients using both structured and unstructured data in electronic health records. In 2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Istanbul, Turkiye, pp. 4368-4373. https://doi.org/10.1109/BIBM58861.2023.10385612

[3] Kessler, S., Schroeder, D., Korlakov, S., et al. (2023). Predicting readmission to the cardiovascular intensive care unit using recurrent neural networks. Digital Health, 9: 1-13. https://doi.org/10.1177/20552076221149529

[4] Moazemi, S., Kalkhoff, S., Kessler, S., et al. (2022). Evaluating a recurrent neural network model for predicting readmission to cardiovascular ICUs based on clinical time series data. Engineering Proceedings, 18(1), 1. https://doi.org/10.3390/engproc2022018001

[5] Pishgar, M., Theis, J., Del Rios, M., Ardati, A., Anahideh, H., Darabi, H. (2022). Prediction of unplanned 30-day readmission for ICU patients with heart failure. BMC Medical Informatics and Decision Making, 22(1): 117. https://doi.org/10.1186/s12911-022-01857-y

[6] McWilliams, C.J., Lawson, D.J., Santos-Rodriguez, R., et al. (2019). Towards a decision support tool for intensive care discharge: Machine learning algorithm development using electronic healthcare data from MIMIC-III and Bristol, UK. BMJ Open, 9(3): e025925. https://doi.org/10.1136/bmjopen-2018-025925

[7] Min, X., Yu, B., Wang, F. (2019). Predictive modeling of the hospital readmission risk from patients' claims data using machine learning: A case study on COPD. Scientific Reports, 9(1): 2362. https://doi.org/10.1038/s41598-019-39071-y

[8] Li, Z., Xing, X., Lu, B., Zhao, Y., Li, Z. (2019). Early prediction of 30-day ICU re-admissions using natural language processing and machine learning. Biomedical Statistics and Informatics, 4(3): 22-26. https://doi.org/10.11648/j.bsi.20190403.11

[9] Assaf, R., Jayousi, R. (2020). 30-day hospital readmission prediction using MIMIC data. In 2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT), Tashkent, Uzbekistan, pp. 1-6. https://doi.org/10.1109/AICT50176.2020.9368625

[10] Liu, W., Stansbury, C., Singh, K., et al. (2020).

Predicting 30-day hospital readmissions using artificial neural networks with medical code embedding. PLOS ONE, 15(4): e0221606. https://doi.org/10.1371/journal.pone.0221606

[11] Darzi, E., Shen, Y., Ou, Y., Sijtsema, N.M., van Ooijen, P. (2024). Tackling heterogeneity in medical federated learning via aligning vision transformers. Artificial Intelligence in Medicine, 155: 102936. https://doi.org/10.1016/j.artmed.2024.102936

[12] Zhang, D., Yin, C., Zeng, J., Yuan, X., Zhang, P. (2020). Combining structured and unstructured data for predictive models: A deep learning approach. BMC Medical Informatics and Decision Making, 20(1): 280. https://doi.org/10.1186/s12911-020-01297-6

[13] Barbieri, S., Kemp, J., Perez-Concha, O., et al. (2020). Benchmarking deep learning architectures for predicting readmission to the ICU and describing patients-at-risk. Scientific Reports, 10(1): 1111. https://doi.org/10.1038/s41598-020-58053-z

[14] Chouja, I., Saoud, S., Sadik, M., Sbai, A. (2024). FLPanGuard: A conceptual FL-based framework to detect and fight new pandemics: Case of COVID-19 pandemic. International Journal of Online and Biomedical Engineering, 20(16): 15-26. https://doi.org/10.3991/ijoe.v20i16.51895

[15] Raza, A., Tran, K.P., Koehl, L., Li, S. (2022). Designing ECG monitoring healthcare system with federated transfer learning and explainable AI. Knowledge-Based Systems, 236: 107763. https://doi.org/10.1016/j.knosys.2021.107763

[16] Khan, M.A., Alsulami, M., Yaqoob, M.M., et al. (2023). Asynchronous federated learning for improved cardiovascular disease prediction using artificial intelligence. Diagnostics, 13(14): 2340. https://doi.org/10.3390/diagnostics13142340

[17] Yaqoob, M.M., Nazir, M., Khan, M.A., Qureshi, S., Al-Rasheed, A. (2023). Hybrid classifier-based federated learning in health service providers for cardiovascular disease prediction. Applied Sciences, 13(3): 1911. https://doi.org/10.3390/app13031911

[18] Zhao, L., Huang, J. (2023). A distribution information sharing federated learning approach for medical image data. Complex & Intelligent Systems, 9(5): 5625-5636. https://doi.org/10.1007/s40747-023-01035-1

[19] Tang, F., Liang, S., Ling, G., Shan, J. (2023). IHVFL: A privacy-enhanced intention-hiding vertical federated learning framework for medical data. Cybersecurity, 6(1): 37. https://doi.org/10.1186/s42400-023-00166-9

[20] Riedel, P., von Schwerin, R., Schaudt, D., Hafner, A., Späte, C. (2023). ResNetFed: Federated deep learning architecture for privacy-preserving pneumonia detection from COVID-19 chest radiographs. Journal of Healthcare Informatics Research, 7(2): 203-224. https://doi.org/10.1007/s41666-023-00132-7

[21] Yu, H., Cai, L., Min, H., Su, X. (2023). Advancing medical data classification through federated learning and blockchain incentive mechanism: Implications for modern software systems and applications. The Journal of Supercomputing, 80(8): 10469-10484. https://doi.org/10.1007/s11227-023-05825-9

[22] Rahman, M., Kundu, D., Alam Suha, S., Siddiqi, U.R., Dey, S.K. (2022). Hospital patients' length of stay prediction: A federated learning approach. Journal of King Saud University Computer and Information Sciences, 34(10): 7874-7884. https://doi.org/10.1016/j.jksuci.2022.07.006

[23] Su, Y., Huang, C., Zhu, W., Lyu, X., Ji, F. (2023). Multi-party Diabetes Mellitus risk prediction based on secure federated learning. Biomedical Signal Processing and Control, 85: 104881. https://doi.org/10.1016/j.bspc.2023.104881

[24] C.U., O.K., Gajendran S., Bhavadharini R.M., Suguna M., Krithiga R. (2023). EHR privacy preservation using federated learning with DQRE-Scnet for healthcare application domains. Knowledge-Based Systems, 275: 110638. https://doi.org/10.1016/j.knosys.2023.110638

[25] Noman, A.A., Rahaman, M., Pranto, T.H., Rahman, R.M. (2023). Blockchain for medical collaboration: A federated learning-based approach for multi-class respiratory disease classification. Healthcare Analytics, 3: 100135. https://doi.org/10.1016/j.health.2023.100135

[26] Narayana, V.L., Patibandla, R.L., Rao, B.T., Gopi, A.P. (2022). Use of machine learning in healthcare. In Advanced Healthcare Systems: Empowering Physicians with IoT-Enabled Technologies, Wiley, pp. 275-293. https://doi.org/10.1002/9781119769293.ch13

[27] Rao, B.T., Patibandla, R L., Narayana, V.L., Gopi, A.P. (2021). Medical data supervised learning ontologies for accurate data analysis. In Semantic Web for Effective Healthcare Systems, Wiley, pp. 249-267. https://doi.org/10.1002/9781119764175.ch11