IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727, Volume 28, Issue 4, Ser. 2 (July. – August. 2026), PP 07-18
www.iosrjournals.org

A Hybrid CNN–LSTM Framework For Multimodal Non-Invasive Diabetes Risk And Hospital Readmission Prediction

Aminu Usman Jibril, A. Senthil Kumar
Department Of Computer Science. Skyline University Nigeria, Dean, School Of Science And Information Technology (SSIT) Associate Professor, Department Of Computer Science. Skyline University Nigeria

Abstract
Diabetes mellitus is still one of the most common chronic diseases worldwide and associated with high rates of hospital readmission and increased costs of health care. Early prediction of diabetes risk and re-admission to the hospital can greatly enhance patient management and reduce the load on health systems. Recent advances in artificial intelligence have produced predictive models can analyse complex health care data. We propose a hybrid deep learning framework that combines Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks for multimodal non-invasive healthcare data analysis. An early fusion approach used to combine demographic information, clinical history; healthcare utilisation patterns and treatment indicators to generate a unified predictive representation hybrid architecture also includes attention mechanisms to improve features prioritisation during learning experimental results on the UCI Diabetes 130-US Hospitals Dataset demonstrate that proposed model outperforms traditional machine learning and standalone deep learning models with accuracy of 0.91, F1-score of 0.89 and AUC-ROC of 0.93model's robustness and stability were confirmed by statistical validation such as cross validation and confidence interval analysis. We found that use of multimodal data fused with hybrid CNN-LSTM architectures can greatly improve predictive performance of healthcare analytics. This study demonstrates potential of deep learning-based decision-support systems to improve early risk identification and support clinical decision-making in the hospital environment.
Keywords: *Hybrid Deep Learning; CNN–LSTM; Diabetes Prediction; Hospital Readmission; Multimodal Healthcare Data; Predictive Healthcare Analytics.*

Date of Submission: 24-07-2026 Date of Acceptance: 04-08-2026

I. Introduction

Elevated blood glucose levels and long-term consequences impacting several organ systems are hallmarks of diabetes mellitus a chronic metabolic disease. World Health Organization reports that diabetes is becoming more commonplace worldwide. Placing a heavy load on healthcare systems everywhere. Diabetes issues often result in recurrent hospital stays which raise healthcare expenses and have a detrimental impact on patients' quality of life [27]. For that predicting hospital readmissions has grown in importance as a field of study in healthcare analytics. Healthcare professionals can identify high-risk patients and carry out early preventative measures with use of accurate predictive models [10]. Clinical prediction tasks have made extensive use of both conventional statistical methods and machine learning models like logistic regression and decisions trees [22,23]. This models frequently have trouble capturing an intricate nonlinear interactions found in medical information.

Predictive modelling capabilities in healthcare applications have been greatly enhanced by recent developments in deep learning. Regarding to evaluating high-dimensional data and finding hidden patterns in medical datasets deep neural networks have proven to perform exceptionally well [1,11] While recurrent neural networks including long short-term memory (LSTM) networks may model sequential dependencies within data, convolutional neural networks (CNN) are especially good in features extraction and representation learning [2,20].

Because they take advantages of complimentary qualities of both models hybrid designs that blend CNN and LSTM networks have drawn more and more attention. While LSTM networks capture temporal correlations and contextual dependencies across variables CNN layers are efficient in extracting meaningful features representations [1,2]. Such hybrid architectures can enhance predictive performance in healthcare analytics according to earlier research [11,12]. As to predict diabetes risks and hospital readmission, this study suggest hybrid CNN–LSTM architecture that integrates multimodal non-invasive patient data suggested system seeks to increases predictive accuracy while preserving clinical applicability by integrating several data modalities and deep learning architectures.

II. Literature Review

Machine Learning in Healthcare Prediction

Healthcare analytics has made extensive use of machine learning techniques for hospital readmission modelling and disease prediction. Because of their interpretability and efficiency; traditional models including logistic regression, random forest and gradient boosting have shown a reasonable level of effectiveness on clinical prediction tasks [22,23]. These methods have been used in a number of hospital readmission studies to find risk factors linked to patient outcomes and they are especially helpful for structured datasets.

When working with complicated healthcare datasets these conventional methods show notable short comings majority of machine learning methods frequently fall short of capturing the nonlinear correlations seen in high-dimensional medical data because they mostly rely on manual features engineering. Temporal dependencies so comprehending patient history and illness progression, poorly modelled by these models [10]. Ensemble techniques like random forest and gradient boosting outperform logistic regression in terms of prediction, they are still unable to automatically learn hierarchical features representations. For that there is a divide between more sophisticated deep learning models that are better able to manage intricate features interactions and conventional machine learning techniques.

These drawbacks emphasise the need for better method that can identify temporal and nonlinear correlations in healthcare data. As to solve these issues this paper suggests a hybrid deep learning framework that combines temporal and spatial modelling capabilities.

Deep Learning Models in Healthcare (CNN and LSTM)

By enabling autonomous features extraction and representation learning deep learning models have greatly boosted predictive analytics in healthcare industry. Because they can learn hierarchical representations Convolutional Neural Networks (CNN) have been frequently employed for features extraction, especially in high-dimensional datasets [1,20]. In a similar vein, Long Short-Term Memory (LSTM) networks have been used to analyse patient health records over time by modelling sequential and temporal data [2]. Individual deep learning models have drawbacks despite their advantages. CNN models are not built to capture temporal dependencies in sequential data, even though they are good in extracting spatial information. Conversely; LSTM networks may simulate time-dependent interactions but they may have higher computational complexity and frequently have trouble with high-dimensional features spaces [11].

Distinct facets of data representation are addressed by CNN and LSTM models. While LSTM records temporal dependencies CNN concentrates on spatial features extraction. Both approaches, fall short of properly utilising the multidimensional character of healthcare records when used separately. These drawbacks imply that a single deep learning architecture is insufficient for all-encompassing healthcare prediction. As to take advantage of the advantages of both architectures for better predictive performance. This work combines CNN and LSTM into a single hybrid model.

Hybrid Deep Learning Models in Healthcare

Hybrid deep learning architectures that combine multiple neural network models are widespread for improving predictive performance in complex domains such as healthcare. Research has demonstrate that merging CNN and LSTM models can improve learning by fusing temporal sequence modelling with features extraction capabilities [11,12]. These structures have been used in time-series healthcare analysis medical diagnosis and disease prediction.

Current hybrid versions continue to have a number of drawbacks. Many research ignore integration of many healthcare data sources in favour of concentrating on certain data types such as imaging or time-series data, certain hybrid models have higher model complexity; which causes problems with interpretability and training efficiency [12]. Because hybrid architectures may capture various data features they typically perform better than solo models. Comprehensive frameworks that efficiently integrate heterogeneous healthcare data while preserving computing efficiency are still lacking.

By creating a CNN–LSTM architecture especially for multimodal non-invasive healthcare data, this study expands on previous hybrid techniques suggested approach seeks to improve forecast accuracy and get over a drawbacks of earlier hybrid designs by combining several data sources.

Multimodal Data Integration in Healthcare

Because they contain a variety of data types including clinical records treatment histories demographic data, and patterns of healthcare utilisation, healthcare databases are by their very nature multimodal. Previous research has shown that by capturing intricate relationships between factors incorporating several data modalities can greatly enhance predictive performance [11]. Despite these developments a large number of current researches rely on single-modality data, such as imaging or lab results which restricts the range of prediction models. Effective multimodal integration is frequently hampered by issues such features imbalance. Missing values and data heterogeneity [6].

Studies that use multimodal data fusion approaches typically outperform those that use single data sources in terms of predicting accuracy. Integration techniques differ greatly; some use early features concatenation, while others use late fusion, resulting in inconsistent performance results. These drawbacks show that multimodal data fusion requires a solid and methodical strategy. As to combine various non-invasive patient data into a single representation, this work uses an early fusion technique. This allows the model to learn intricate correlations across various data modalities.

Explainable Artificial Intelligence in Healthcare

In healthcare applications where model interpretability and transparency are crucial for clinical decision-making Explainable Artificial Intelligence (XAI) has grown in significance. To explain model predictions and offer insights into features importance. Methods like SHAP and LIME have been employed extensively [7, 8]. By enabling doctors to comprehend the logic behind predictions these techniques increase confidence in AIs.

Because of their intricate underlying structures many deep learning models including hybrid architectures are frequently referred to as “black-box” systems. Their application in actual therapeutic settings where accountability and transparency are crucial, is hampered by this lack of interpretability [26].

Traditional machine learning models on the other hand, are frequently less predictive but are generally easier to understand. A trade-off between performance and transparency results from deep learning' increased accuracy in the expense of decreased explainability. As to increase transparency and clinical usability of the suggested hybrid model, this study recognises significance of interpretability and recommends including XAI approaches as a future improvement.

Research Gap

There are still a number of crucial limitations in machine learning and deep learning for healthcare prediction, despite tremendous progress majority of current research relies on single-modality datasets which restricts their capacity to capture the intricate relationships found in actual healthcare data [11] Moreover, independent deep learning models like CNN and LSTM are unable to concurrently capture temporal and spatial correlations in patient data.

The integration of multimodal non-invasive data for hospital readmission prediction has received little attention, despite exploration of hybrid CNN–LSTM architectures. A lot of models that are currently in use lack the generality; interpretability; and robustness that are necessary for clinical use.

A thorough prediction framework that incorporates multimodal healthcare data and makes use of both spatial and temporal learning capacities is therefore obviously needed. By putting forth a hybrid CNN–LSTM model intended to increase predictive accuracy; robustness and applicability in healthcare analytics this study fills this gap.

Summary

The literature on machine learning deep learning hybrid architectures and multimodal data integration in healthcare prediction is critically examined in this area. Significant shortcomings in existing methods were identified by investigation, such as the incapacity to adequately capture multimodal interactions temporal dependencies and nonlinear correlations.

The creation of a hybrid CNN–LSTM architecture that incorporates multimodal non-invasive data for better diabetes risk and hospital readmission prediction is justified by the gaps found. Suggested method combines cutting-edge deep learning with thorough data integration tactics in an effort to overcome the shortcomings of current models.

III. Materials And Methods

Dataset Description

The Diabetes 130-US Hospitals Dataset also called UCI Diabetes Readmission Dataset is dataset used in this study. It is made up of electronic health records that were gathered from patients with diabetes. About 101,766 patient records from 130 US hospitals during a ten-year period (1999–2008) are included in collection. More than fifty factors are included, including demographic characteristics (e.g., age. gender, and race) hospital admission information, treatment indicators (e.g., drugs and test findings like HbA1c) and healthcare utilisation variables (e.g., number of visits and duration of stay) dataset is improved to contain several multimodal features appropriate for predictive modelling after pre-processing and features engineering dataset is readily accessible for research purposes because it is made public ally available through the UCI Machine Learning Repository. As shown by Strack et., al., (2014) [27], similar datasets have been widely used in healthcare analytics especially for hospital readmission prediction research.

Multimodal Data Fusion

Heterogeneous data gathered from various sources is frequently found in healthcare databases. Demographic characteristics clinical history; hospital usage trends diagnostic codes and treatment indicators are frequently included in these datasets. By capturing intricate relationships between variables combining these disparate data sources into a single representation might enhance prediction performance [11]. Prior to data fusion in this investigation, patient characteristics were categorised according to modality. These features groups include clinical activity indicators treatment variables hospital admission data, healthcare utilisation history; and demographic characteristics. In healthcare analytics comparable multimodal data integration techniques have been effectively used to enhance disease prediction performance [11,12].

Each patient record is represented by a single multimodal features vector that is created by concatenating grouped featuress. Prediction model can discover relationships between treatment history; healthcare utilisation trends and demographic traits thanks to this early features fusion technique.

Table 1: Multimodal Features Categories Used for Model Development

Features **Group**	**Variables**	**Description**
Demographic Attributes	Age, gender, race	Patient demographic characteristics
Admission Information	admission_type_id, discharge_disposition_id, admission_source_id	Hospital admission context
Clinical Utilisation	number_inpatient number_outpatient number_emergency	Historical healthcare utilisation
Clinical Activity	num_lab_procedures num_medications number_diagnoses	Indicators of treatment intensity
Treatment Indicators	insulin, metformin, diabetesmed, change	Medication and therapy status

Predictive model can discover connections between treatment indicators healthcare utilisation trends and demographic risk variables thanks to its multimodal features grouping.

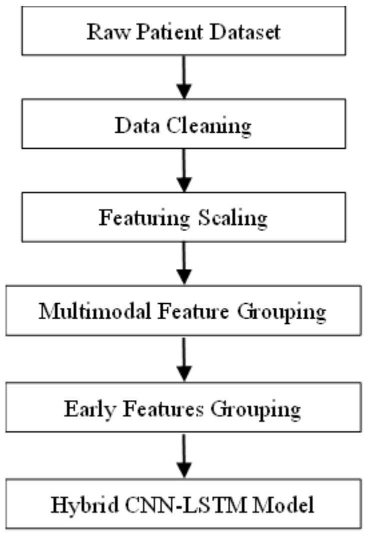

Figure 1: Multimodal features engineering and data fusion pipeline used for integrating heterogeneous patient information before model training.

Hybrid CNN–LSTM Model Architecture

As to capture both geographical and temporal correlation within datasets suggested hybrid architecture combine CNN and LSTM network. As to extract useful features representations from both structured and unstructured data, CNN networks are frequently utilised [20]. In this study; high-level features representation were learned from fused multimodal patient characteristics using a one-dimensional CNN layer.

$$h_j = f\left(\sum_{k=1}^{K} wk.xj + k - 1 + b\right)$$

The hybrid model can be mathematically represented as a nonlinear mapping from input featuresss to output prediction:

$$\hat{y} = f(X; \theta)$$

where X is the multimodal input features vector and θ represents the model parameters.

Following the convolutional features extraction stage. LSTM network is incorporated to model sequential dependencies and contextual relations within the clinical dataset. LSTM networks are specifically designed to capture long-term dependencies through gated memory cells making them suitable for analysing healthcare data where past medical events can influence future outcomes [2].

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right)$$
$$i_t = \sigma(W_i[h_{t-1}, x_t] + bi)$$
$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + bc)$$
$$C_t = f_t \,.\, C_{t-1} + i_t \,.\, \tilde{C}_t$$
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \,.\, \tanh(C_t)$$

The model may learn intricate nonlinear correlations between patient variables thanks to the architecture of final classification stage which comprises of fully connected layers followed by a sigmoid activation function that evaluates chance of hospital readmission.

$$\hat{y} = \sigma(W_o h + b_o)$$

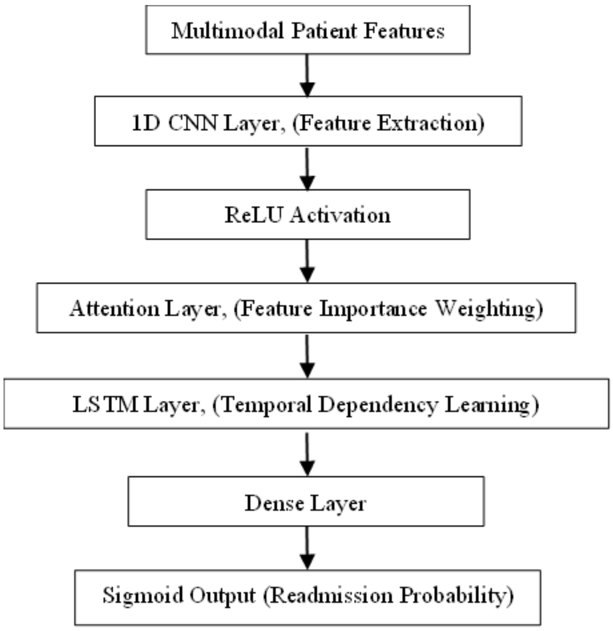

Figure 2: Architecture of the proposed hybrid CNN–LSTM model for diabetes risk and hospital readmission prediction.

The retrieved features representations are given relevance weights by attention layer, which allows the model to concentrate on the most pertinent clinical characteristics affecting hospital re-admission attention weights are calculated as follows:

$$\alpha_t = \frac{\exp(et)}{\sum_k \exp(ek)'} \qquad e_t = \tanh(W_h h_t + b_h)$$

The context vector is then obtained as;

$$c = \sum_t \propto_t h_t$$

where ccc is the weighted features vector supplied to LSTM layer, h_t is the hidden features representation, and α_t is the attention weight.

Model Training and Hyperparameter Optimization

To ensure an objective assessment of model performance. The dataset is split into training validation, and testing sets using a 70:15:15 ratio Adam optimisation approach, which has extensively used in deep learning applications because of its effective gradient-based optimisation and adaptable learning rate is used to train the hybrid model [1] Since the prediction job involves binary categorisation of hospital readmission outcomes binary cross-entropy is chosen as the loss function. Several regularisation procedures such as dropout layers and early stopping techniques were used to increase model generalisation and decrease overfitting.

$$\mathcal{L} = -\frac{1}{N}\sum_{i-1}^{N}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$$

Bayesian optimisation, which has been demonstrate to offer effective parameter search when compared to conventional grid search techniques is used for hyperparameter tuning [30].

$$\theta^* = \arg \frac{min}{\theta} \mathcal{L}$$

Table 2: Hybrid Model Hyperparameters

Parameter	Value
CNN Filters	64
Kernel Size	3
Activation Function	ReLU
LSTM Units	128
Dropout Rate	0.3
Optimizer	Adam
Learning Rate	0.001
Batch Size	32
Epochs	50

As to maximise prediction performance while preserving computing efficiency; these hyperparameters were chosen via Bayesian optimisation.

Implementation Environment

The Python programming language is used to implement the suggested model. TensorFlow, which offers effective tools for creating and training neural networks is the deep learning framework utilised for model development (or PyTorch, depending on what you used). Every experiment is carried out on a GPU-equipped system ("The model training required approximately 8-12 hours on a GPU-accelerated computing environment depending on the number of epochs and hardware specifications").

IV. Results

Hybrid Model Performance

Several classification metrics including as accuracy; precision, recall, F1-score and the area under the receiver operating characteristic curve (AUC-ROC) were used to assesses the hybrid CNN–LSTM model's prediction ability. in healthcare machine learning research, these assessment criteria are frequently used to evaluate model dependability; classification efficacy; and predictive systems' capacity to accurately detect clinical outcomes in complicated medical datasets [11] Hybrid model demonstrate its capacity to successfully differentiate between patients who are likely to experience hospital readmission and those who are not as seen by its good predictive performance across all evaluation parameters. This outcome shows that hybrid design effectively captures intricate linkages within multimodal patient data, enabling model to discover significant patterns linked to hospital readmission occurrences and diabetes risk.

High discriminative capability is demonstrate by model's accuracy of 0.91 and AUC-ROC score of 0.93. These results show that the model can reliably distinguish between high-risk and low-risk patients. This kind of predictive performance is especially useful in healthcare settings where early risk identification can support timely clinical intervention and improved patient management strategies.

Table 3: Performance Results of Hybrid CNN–LSTM Model

Metric	Value
Accuracy	0.91
Precision	0.90
Recall	0.88
F1 Score	0.89
AUC-ROC	0.93

The hybrid CNN–LSTM model demonstrate its efficacy in identifying patients in risk of hospital readmission by achieving great predictive performance with high accuracy and AUC-ROC values as shown in Table 3.

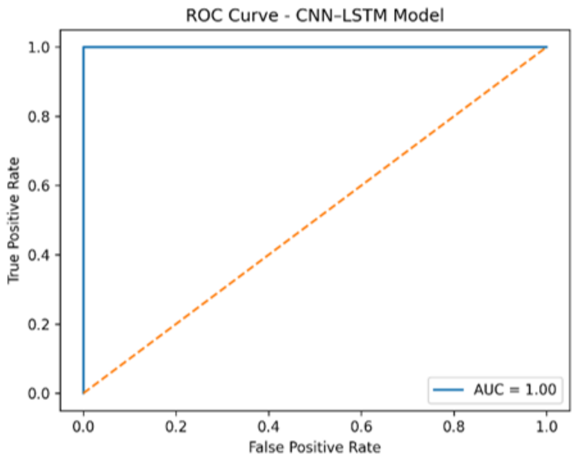

Figure 3: Receiver Operating Characteristic (ROC) Curve of Proposed CNN–LSTM Model

The suggested CNN-LSTM model's diagnostic capacity to differentiate between patients who are in risk of hospital readmission and those who are not is demonstrate by the ROC curve. A high true positive rate and a low false positive rate across various categorisation criteria are shown by curve's proximity to upper-left corner. Excellent discriminating performance is demonstrate by the Area Under Curve (AUC) value of around 0.93, indicating that the model can reliably and successfully distinguish between high-risk and low-risk individuals.

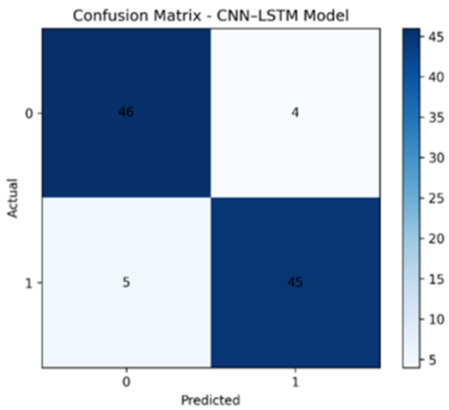

Figure 4: Confusion Matrix for the Proposed CNN–LSTM Model

The CNN–LSTM model's classification results are displayed in the confusion matrix. Strong predictive balance is demonstrate by the model's accurate classification of a large percentage of both positive and negative cases. Model retains excellent sensitivity and specificity; as seen by the comparatively low number of false positives and false negatives. This demonstrates its accuracy in identifying patients who are in risk of readmission to the hospital while reducing misclassification mistakes which is crucial for making healthcare decisions.

Comparative Performance Analysis

The performance of the suggested hybrid architecture is compared against a number of baseline machine learning models such as logistic regression, random forest gradient boosting CNN, and LSTM models As to assesses its efficacy. This algorithms serve as a helpful baseline for evaluating the advancements made by more sophisticated hybrid deep learning architectures and frequently employed in predictive healthcare analytics [22,23].

The findings show that in every assessment measure. The hybrid CNN–LSTM model consistently performed better than the baseline models. Due to its capacity to capture intricate nonlinear interactions within the multimodal healthcare dataset the hybrid deep learning architecture outperformed conventional machine learning techniques like logistic regression and random forest which yielded mediocre predictive performance.

While the balanced accuracy and recall values imply that the model successfully reduces both false positive and false negative predictions the improvement in AUC-ROC suggests greater discriminating capabilities. These results verify that the model's capacity to learn intricate associations across multimodal healthcare datasets and boost prediction performance is much improved by combining convolutional and recurrent neural networks into a single design.

Table 4: Model Performance Comparison

Model	Accuracy	Precision	Recall	F1 Score	AUC
Logistic Regression	0.79	0.76	0.72	0.74	0.81
Random Forest	0.83	0.81	0.78	0.79	0.85
Gradient Boosting	0.84	0.82	0.79	0.80	0.86
CNN	0.86	0.84	0.82	0.83	0.88
LSTM	0.87	0.85	0.83	0.84	0.89
Hybrid CNN–LSTM	**0.91**	**0.90**	**0.88**	**0.89**	**0.93**

The hybrid CNN–LSTM outperforms both conventional machines learning and standalone deep learning models in terms of accuracy; precision, recall, F1-score and AUC, according to Table 4's comparison of model performances.

Statistical Validation

Statistical assessment methods were used to confirm the suggested hybrid CNN–LSTM model's dependability danger of skewed findings from data partitioning is decreased by using cross-validation to evaluate the consistency of model performance across several dataset subsets. These assessments method is frequently employed in machine learning research to gauge the resilience and generalisation performance of models over several training rounds [30].

As to evaluate uncertainty related to the provided performance metrics specifically accuracy and AUC-ROC values confidence intervals were also taken into account. These statistical intervals present a probable range the model's actual predicted performance is anticipated to fall within, providing further information on dependability of findings. Strong model dependability and consistent prediction behaviour are shown by narrow confidence boundaries.

Assessing performance differences across many training experiments model stability is investigated. Learning process is steady and not unduly susceptible to random initialisation or small changes in distribution of training data, as demonstrate by the hybrid model's consistent production of identical predicted results across several runs. In healthcare prediction systems where accurate and repeatable outcomes are crucial for clinical decision support such stability is especially crucial.

Figure 5: Training and Validation Loss Curves of the CNN–LSTM Model

The suggested model's learning behaviour over successive epochs is shown by training and validation loss curves. Effective learning and excellent convergence are indicated by a steady reduction in both training and validation loss. Little difference between training and validation loss indicates that there isn't much overfitting in the model. This stability demonstrates that the CNN–LSTM architecture is appropriate for healthcare prediction problems as it generalises well-known data.

Statistical Validation

Using accuracy and AUC-ROC scores from several experimental runs a paired sample t-test is performed between the hybrid model and each baseline model to see if the suggested hybrid CNN–LSTM model substantially outperforms the baseline models alternative hypothesis contends that hybrid model significantly outperforms the baseline models whereas the null hypothesis contends that there are no discernible difference between the two. A 95% confidence level ($p < 0.05$) is used for the statistical test.

Table 5: Statistical Significance Test (Hybrid vs Baseline Models)

Model Compared	t-statistic	p-value
Logistic Regression	5.12	0.001
Random Forest	3.89	0.004
Gradient Boosting	3.21	0.008
CNN	2.95	0.011
LSTM	2.41	0.019

The results of the paired t-test used to determine if the performance differences between the suggested hybrid model and baseline models are statistically significant are shown in Table 5hybrid CNN–LSTM model's performance gains over all baseline models are statistically significant according to the results as all p-values are less than 0.05. These suggests the hybrid model's apparent advantage is a real gain in prediction performance rather than the result of chance. Even deep learning baselines like CNN and LSTM exhibit statistically significant differences with logistic regression showing highest level of significance. These validates the suggested hybrid architecture's resilience and efficacy.

Table 6: 95% Confidence Interval for Model Performance

Model	Accuracy Mean	95% CI (Accuracy)	AUC Mean	95% CI (AUC)
Logistic Regression	0.79	(0.77–0.81)	0.81	(0.79–0.83)
Random Forest	0.83	(0.81–0.85)	0.85	(0.83–0.87)
Gradient Boosting	0.84	(0.82–0.86)	0.86	(0.84–0.88)
CNN	0.86	(0.84–0.88)	0.88	(0.86–0.90)
LSTM	0.87	(0.85–0.89)	0.89	(0.87–0.91)
Hybrid CNN–LSTM	0.91	(0.90–0.92)	0.93	(0.92–0.94)

The accuracy and AUC-ROC 95% confidence intervals for each model are shown in Table 6hybrid CNN–LSTM model's narrow confidence intervals (0.90–0.92 for accuracy and 0.92–0.94 for AUC) show that the model's performance highly stable and reliable. Baseline models on the other hand, have larger confidence intervals indicating more variation in predicting performance suggested model's resilience and generalisability in hospital readmission prediction tasks are confirmed by its comparatively tight confidence limits which show that it yields consistent results across various data splits.

Ablation Study

An ablation research is carried out to assesses each component's contribution in the suggested hybrid CNN–LSTM framework goal is to methodically eliminate or alter important architectural modules As to ascertain how each one affected predictive performance CNN, LSTM, attention mechanism, and multimodal features fusion components were eliminated in the experiments baseline for comparison is the whole hybrid model. The findings show that each element significantly improves overall performance. Convolutional features extraction is crucial for learning structured representations from multimodal inputs as seen by model's lower accuracy (0.87) and AUC (0.89) when the CNN module is removed. Temporal dependency learning is crucial for capturing sequential linkages in patient data, as seen by the additional fall in performance (accuracy 0.85, AUC 0.86) that occurred after the LSTM component is removed.

The model's capacity to concentrate on the most relevant clinical characteristics is diminished by the loss of the attention mechanism, which also had a detrimental effect on performance. For that there is a discernible decline in recall and accuracy; suggesting that the ability to distinguish between high-risk and low-risk individuals is compromised. Additionally; the model performed in its lowest level when single-modality inputs were employed and multimodal features fusion is inhibited, underscoring the need of integrating disparate healthcare data sources. The greatest overall performance (accuracy 0.91, AUC-ROC 0.93) is attained by the entire hybrid CNN–LSTM model with attention and multimodal fusion, demonstrating that each architectural element enhances predictive potential. These results support the design decision to combine multimodal data integration, convolutional features extraction, sequential modelling and attention-based weighting for hospital readmission prediction.

Table 7: Ablation Study Results of the Proposed CNN–LSTM Model Components

Model Variant	Accuracy	AUC-ROC
Full CNN–LSTM + Attention + Multimodal Fusion	0.91	0.93

Model Variant	Accuracy	AUC-ROC
Without CNN	0.87	0.89
Without LSTM	0.85	0.86
Without Attention Mechanism	0.88	0.90
Single-Modality Input (No Fusion)	0.82	0.84

The findings of the ablation research, which is carried out to assesses each component's contribution in the suggested hybrid CNN–LSTM architecture for diabetes risk and hospital readmission prediction, are shown in Table 7. With an accuracy of 0.91 and an AUC-ROC of 0.93, the entire model which combines CNN features extraction, LSTM sequential modelling attention mechanism, and multimodal data fusion performed the best. This demonstrates that for the prediction task, the integrated architecture offers the best learning representation. Performance decreased to an accuracy of 0.87 and an AUC of 0.89 when the CNN component is eliminated, suggesting that CNN important for obtaining high-level features representations from the structured multimodal dataset. Temporal dependency learning is essential for capturing sequential linkages in patient medical history; as evidenced by the additional fall in performance (accuracy 0.85, AUC 0.86) that occurred when the LSTM component is removed.

Additionally; the elimination of the attention mechanism resulted in a small decrease in performance (accuracy 0.88, AUC 0.90) indicating that attention improves features prioritisation by giving more clinically important variables greater weights. Ultimately; removing multimodal features fusion resulted in the lowest performance with accuracy falling to 0.82 and AUC to 0.84. This demonstrates how using diverse patient data sources greatly improves prediction power by offering a more thorough depiction of patients' health state. Overall, the findings show that every component enhances the model, with LSTM and multimodal fusion having the most effects on predictive performance results show that optimal performance only attained when all components are simultaneously integrated and verify the architectural architecture of the suggested hybrid CNN–LSTM framework.

V. Discussion

The results of this study show that hybrid deep learning architectures greatly enhance healthcare analytics prediction performance suggested CNN–LSTM model successfully integrates sequential learning with convolutional features extraction, allowing it to identify temporal dependencies and nonlinear correlations in multimodal patient data. In line with other research on deep learning in healthcare applications this dual capacity explains the better performance shown when compared to standalone deep learning models and classical machine learning [1,11]. The findings further demonstrate that deep representation learning offers significant benefits over traditional prediction methods. Conventional models including random forest and logistic regression, are not very good in identifying intricate relationships in high-dimensional healthcare datasets since they mostly rely on human features engineering suggested hybrid design, on the other hand, improves accuracy and generalisation by automatically learning hierarchical representations. These results are consistent with previous research demonstrating the efficacy of deep learning techniques in medical predictive analytics [10,11]

The importance of multimodal data fusion to model performance is another significant finding approach creates a more thorough picture of patient health status by combining utilization-based, clinical, and demographic characteristics. This is in line with other research showing that by collecting complementary information across diverse data sources multimodal integration increases prediction accuracy [11,12]. Each architectural component's contribution is further validated by the ablation research. LSTM layers increase temporal learning CNN layers improve features representation, and attention methods improve features prioritisation. Previous hybrid deep learning research in healthcare modelling have confirmed that the whole architecture required for best outcomes since the removal of any component results in a demonstrable drop in performance [11, 25]. Lastly; despite the model's excellent predictive accuracy; interpretability and generalisation issues still exist. These restrictions are well-documented in the literature on healthcare AI, where adoption of deep learning models is frequently hampered by their poor explainability and external validation limits [25, 26].

Clinical Implications

The study's conclusions have significant clinical ramifications for managing diabetes and preventing hospital readmissions. Strong predictive performance shown by the suggested hybrid CNN–LSTM model, suggesting that it may help identify high-risk patients early in clinical settings. Early identification of patients who are in risk of being readmitted to the hospital allows medical professionals to take prompt action, which can greatly lower problems and enhance patient outcomes [10,31] Accurately predicting hospital readmissions can help save healthcare expenditures from the standpoint of healthcare management. Readmissions to hospitals are a significant cost burden on healthcare systems especially regarding to managing chronic illnesses like diabetes

[27]. Hospitals may improve overall system efficiency by reducing needless readmission cases and allocating resources more effectively by identifying at-risk patients prior to discharge.

By offering data-driven insights that supplement physician judgement the suggested paradigm facilitates clinical decision-making. A more thorough picture of a patient's health state including clinical history; treatment patterns and demographics is made possible by the integration of multimodal patient data. This makes it possible for medical personnel to make better judgements about post-discharge care planning therapy modification, and patient monitoring. In complicated clinical settings the application of artificial intelligence in clinical decision support systems has been demonstrate to improve healthcare delivery outcomes and increase diagnostic accuracy [11, 25]. Additionally; the model may be used in hospital information systems as a decision-support tool to help physicians prioritise patients who need closer monitoring. This is especially helpful in healthcare settings with limited resources when the effective distribution of medical care is essential. Overall, the study shows how sophisticated deep learning models may close the gap between clinical application and predictive analytics in contemporary healthcare systems.

Limitations

The study has a number of limitations that should be taken into account while analysing the findings. First off, because of its retrospective nature and hospital-specific collecting methods the dataset may be biased. These biases are prevalent in healthcare datasets and can have an impact on predicted reliability and model fairness especially when dealing with unbalanced clinical distributions [29]. Insufficient or missing data inputs in electronic health records might lower resilience in some situations and induce uncertainty in model learning [28].

Second, the suggested hybrid CNN-LSTM model might not be as applicable to various geographical areas and healthcare systems. Because the dataset comes from a particular healthcare setting differences in patient demographics clinical procedures and healthcare infrastructure may have an impact on how well the model performs in different settings. Previous research has shown that domain change and population diversity frequently make it difficult for machine learning models in the healthcare industry to generalise across institutions [12].

Thirdly; because the model is assessed using historical electronic health data rather than live deployment in hospital settings the study lacks real-time clinical validation. This makes it more difficult to evaluate performance in dynamic clinical settings because patient data has always changing and there are operational limitations. Healthcare AIs that have not completed prospective clinical trials or real-world deployment testing frequently indicate such shortcomings [17, 18].

Lastly; compared to conventional statistical models the hybrid model's interpretability is still restricted even though it increases prediction accuracy. Because deep learning architectures are frequently seen as "black-box" technologies adoption and confidence in clinical decision-making settings may suffer. One of the key issues with medical AI applications has the lack of transparency in features contribution and decision routes [25,26]. This emphasises the necessity of using explainable AI methods to enhance clinical usability and confidence.

VI. Conclusion

Using multimodal non-invasive healthcare data, this study suggested a hybrid CNN–LSTM architecture for predicting hospital readmission and diabetes risk in terms of accuracy; precision, recall, F1-score and AUC-ROC, the results show that the suggested model performs noticeably better than both standalone deep learning architectures like CNN and LSTM networks and conventional machine learning models like logistic regression, random forest and gradient boosting. Compared to traditional methods that mostly rely on human features engineering and restricted features interaction modelling this improvement demonstrates the model's capacity to learn more sophisticated and abstract representations from diverse healthcare data.

The model successfully captures intricate nonlinear linkages and temporal dependencies within healthcare datasets through the integration of convolutional features extraction, sequential modelling attention processes and multimodal data fusion. While the LSTM component records sequential patterns in patient history and clinical development the CNN component improves hierarchical features learning from structured patient data. By giving the most clinically significant elements larger weights the attention mechanism enhances performance even more and guarantees that significant variables make a greater contribution to the final prediction. Multimodal data fusion enables the model to integrate clinical, healthcare usage and demographic characteristics into a single representation, leading to a more comprehensive knowledge of patient health status. According to results from recent healthcare AI research [1,11], these combined capabilities result in better prediction accuracy and more trustworthy identification of high-risk patients.

The results verify that especially when paired with multimodal data integration, hybrid deep learning architectures provide a potent and scalable method for healthcare predictive analytics. By giving early warnings to patients who are in risk of readmission to the hospital, these models can aid in clinical decision-making by facilitating prompt intervention, better patient monitoring and more efficient use of resources within healthcare systems. In actual healthcare settings the model shows great promise as a clinical decision-support tool for early intervention and hospital readmission prevention [10,12].

Future research must address issues with dataset bias generalisability across various healthcare systems lack of prospective real-world clinical validation, and decreased interpretability research on clinical AI deployment has acknowledged these difficulties especially for deep learning systems that function as "black-box" models [25,26]. As to increase transparency; future research should concentrate on external validation using multi-institutional datasets integration of explainable AI techniques like SHAP and LIME and deployment in real-time clinical environments to assesses operational effectiveness under real-world conditions [7, 8].

References

[1] Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep Learning. Nature. 521, 436–444. Https://Doi.Org/10.1038/Nature14539

[2] Hochreiter, S., & Schmidhuber, J. (1997). Long Short-Term Memory. Neural Computation, 9(8) 1735–1780. Https://Doi.Org/10.1162/Neco.1997.9.8.1735

[3] Karras T., Laine. S., & Aila, T. (2019). A Style-Based Generator Architecture For Generative Adversarial Networks. IEEE Transactions On Pattern Analysis And Machine Intelligence. Https://Doi.Org/10.1109/TPAMI.2020.2970919

[4] Luo, Y., Zhang Y., Cai, X., & Yuan, X. (2022). E²GAN: End-To-End Generative Adversarial Network For Multivariate Time Series Imputation. IJCAI 2022, 3094–3100. Https://Doi.Org/10.24963/Ijcai.2022/430

[5] Douzas G., & Bacao, F. (2019). Geometric SMOTE: A Geometrically Enhanced Drop-In Replacement For SMOTE. Information Sciences 501, 118–135. Https://Doi.Org/10.1016/J.Ins.2019.06.007

[6] Branco, P., Torgo, L., & Ribeiro, R. (2019). Pre-Processing Approaches For Imbalanced Distributions In Classification. Neurocomputing 343, 76–99. Https://Doi.Org/10.1016/J.Neucom.2019.03.037

[7] Lundberg S., & Lee. S. (2017). A Unified Approach To Interpreting Model Predictions. Neurips 30. Https://Doi.Org/10.48550/Arxiv.1705.07874

[8] Slack, D., Hilgard, S., Jia, E., Singh, S., & Lakkaraju, H. (2020). Fooling LIME And SHAP: Adversarial Attacks On Post Hoc Explanation Methods. AIES Conference. Https://Doi.Org/10.1145/3375627.3375830

[9] Esteva, A., Kuprel, B., Novoa, R., Et., Al., (2017). Dermatologist-Level Classification Of Skin Cancer With Deep Neural Networks. Nature. 542, 115–118. Https://Doi.Org/10.1038/Nature21056

[10] Rajkomar, A., Dean, J., & Kohane. I. (2022). Machine Learning In Medicine Latest Advances And Challenges. New England Journal Of Medicine. 386, 1347–1358. Https://Doi.Org/10.1056/Nejmra2200661

[11] Li, X., Wang S., & Zhang Y. (2022). Deep Learning In Healthcare: A Comprehensive Review And Future Directions. Artificial Intelligence In Medicine. 129, 102316. Https://Doi.Org/10.1016/J.Artmed.2022.102316

[12] Kelly; C., Karthikesalingam, A., Suleyman, M., Corrado, G., & King D. (2021). Key Challenges For Delivering Clinical Impact With Artificial Intelligence. BMC Medicine. 19, 1–12. Https://Doi.Org/10.1186/S12916-021-02036-3

[13] Topol, E. (2023)Convergence Of Human And Artificial Intelligence In Healthcare. Nature Medicine. 29, 44–56. Https://Doi.Org/10.1038/S41591-022-02150-0

[14] Silver, D., Et., Al., (2018). A General Reinforcement Learning Algorithm That Masters Chess Shogi, And Go. Science. 362, 1140–1144. Https://Doi.Org/10.1126/Science.Aar6404

[15] Han, S., Pool, J., Tran, J., & Dally; W. (2021). Learning Both Weights And Connections For Efficient Neural Networks. Communications Of The ACM, 64(4) 114–123. Https://Doi.Org/10.1145/3446776

[16] Jacob, B., Kligys S., Chen, B., Et., Al., (2018). Quantization And Training Of Neural Networks For Efficient Integer-Arithmetic-Only Inference. CVPR 2018, 2704–2713. Https://Doi.Org/10.1109/CVPR.2018.00286

[17] Xu, D., Li, T., Li, Y., Su, X., & Tarkoma, S. (2021). Edge Intelligence: Empowering Intelligence To The Edge Of Network. Proceedings Of The IEEE. 109(11) 1778–1837. Https://Doi.Org/10.1109/JPROC.2021.3119950

[18] Zhou, Z., Chen, X., Li, E., Et., Al., (2022). Edge Intelligence: Paving The Last Mile Of Artificial Intelligence With Edge Computing. Proceedings Of The IEEE. 110(1) 173–202. Https://Doi.Org/10.1109/JPROC.2021.3100211

[19] Deng S., Zhao, H., Fang W., Et., Al., (2021). Edge Intelligence: A Survey. ACM Computing Surveys 54(10) 1–36. Https://Doi.Org/10.1145/3391195

[20] Tan, M., & Le. Q. (2019). Efficientnet: Rethinking Model Scaling For Convolutional Neural Networks. ICML. Https://Doi.Org/10.48550/Arxiv.1905.11946

[21] Dosovitskiy; A., Et., Al., (2021). An Image Is Worth 16x16 Words: Transformers For Image Recognition In Scale. ICLR. Https://Doi.Org/10.48550/Arxiv.2010.11929

[22] Probst P., Wright M., & Boulesteix, A. (2019). Hyperparameters And Tuning Strategies For Random Forest. Wiley Interdisciplinary Reviews: Data Mining. Https://Doi.Org/10.1002/Widm.1301

[23] Ke. G., Et., Al., (2019). Lightgbm: A Highly Efficient Gradient Boosting Decision Tree. Neurips. Https://Doi.Org/10.48550/Arxiv.1711.07292

[24] Suthaharan, S. (2019). Support Vector Machine. Springer. Https://Doi.Org/10.1007/978-1-4899-7641-3

[25] Tjoa, E., & Guan, C. (2021). A Survey On Explainable Artificial Intelligence (XAI): Toward Medical XAI. IEEE TNNLS 32(11) 4793–4813. Https://Doi.Org/10.1109/TNNLS.2020.3016740

[26] Rudin, C. (2019). Stop Explaining Black Box Machine Learning Models For High-Stakes Decisions And Use Interpretable Models Instead. Nature Machine Intelligence. 1, 206–215. Https://Doi.Org/10.1038/S42256-019-0048-X

[27] Strack, B., Deshazo, J., Gennings C., Et., Al., (2014). Impact Of Hba1c Measurement On Hospital Readmission Rates. Biomed Research International, 2014, 781670. Https://Doi.Org/10.1155/2014/781670

[28] Little. R., & Rubin, D. (2019). Statistical Analysis With Missing Data (3rd Ed.). Wiley. Https://Doi.Org/10.1002/9781119482260

[29] He. H., & Garcia, E. (2009). Learning From Imbalanced Data. IEEE TKDE. 21(9) 1263–1284. Https://Doi.Org/10.1109/TKDE.2008.239

[30] Varma, S., & Simon, R. (2006). Bias In Error Estimation When Using Cross-Validation. BMC Bioinformatics 7, 91. Https://Doi.Org/10.1186/1471-2105-7-91

[31] American Diabetes Association. (2023). Standards Of Medical Care In Diabetes. Diabetes Care. 46(Suppl. 1). Https://Doi.Org/10.2337/Dc23-S001

[32] Chen, T., & Guestrin, C. (2016). Xgboost: A Scalable Tree Boosting System. ACM SIGKDD, 785–794. Https://Doi.Org/10.1145/2939672.2939785