Deep Neural Networks for Predicting Post-Surgical Complications Using Multimodal Clinical Data

Shraddha Gendlal Vaidya[1], Dr. K. M. Gaikwad[2], Dr. Pragati Patil Bedekar[3], Dr. Manoj L. Bangare[4], Shalaka Prasad Deore[5], Ramesh Adireddy[6]

[1]Assistant Professor, CSE (AIML) Dept, Yeshwantrao Chavhan College of Engineering, Nagpur, Maharashtra, India.
Email ID: shraddha.karale1306@gmail.com

[2]Professor. Department of Electronics and Telecommunication, Sinhgad Academy of Engineering, Savitribai Phule Pune University, Pune, India

Email ID: kmgaikwad.sae@sinhgad.edu, ORCID: 0000-0002-6402-9852

[3]Assistant Professor, Department of Computer Science and Engineering, Tulsiramji Gaikwad Patil College of Engineering and Technology, Nagpur, India.

Email ID: viceprincipal@tgpcet.com

[4]Associate Professor, Department of Information Technology, Smt. Kashibai Navale College of Engineering, Savitribai Phule Pune University, Pune, India.

Email ID: manoj.bangare@gmail.com

[5]MES Wadia College of engineering, Pune, Maharashtra, India.

Email ID: shalaka.deore@mescoepune.org

[6]Professor, Dept. of EEE, Aditya University, Surampelam, India.

Email ID: ameshadireddy007@gmail.com

.Cite this paper as: Shraddha Gendlal Vaidya, Dr. K. M. Gaikwad, Dr. Pragati Patil Bedekar, Dr. Manoj L. Bangare, Associate Professor, Shalaka Prasad Deore, Ramesh Adireddy, (2025) Deep Neural Networks for Predicting Post-Surgical Complications Using Multimodal Clinical Data. *Journal of Neonatal Surgery*, 14 (9s), 749-761.

ABSTRACT

Post-surgical complications significantly impact patient recovery, hospital resource utilization, and healthcare costs. Early prediction of such complications enables timely interventions, reducing morbidity and mortality rates. Traditional predictive models rely on handcrafted features and domain-specific knowledge, often limiting their accuracy and adaptability. In this study, we propose a deep neural network (DNN) approach for predicting post-surgical complications using multimodal clinical data, including structured electronic health records (EHRs), unstructured clinical notes, and medical imaging. Our framework integrates convolutional neural networks (CNNs) for imaging analysis, recurrent neural networks (RNNs) for sequential patient records, and transformer-based models for clinical text processing. A multimodal fusion layer combines these diverse data representations, capturing intricate relationships between different modalities. We trained and validated our model on a large hospital dataset containing records from 50,000 surgical patients. Experimental results show that our approach outperforms traditional machine learning models, achieving an accuracy of 92.3%, a precision of 91.7%, and an AUC-ROC score of 0.94. We also employ SHAP (SHapley Additive exPlanations) to enhance model interpretability, identifying key predictive factors such as preoperative vitals, surgical procedure details, and early post-operative lab results. Our findings demonstrate that deep learning models, particularly multimodal fusion networks, significantly improve the prediction of post-surgical complications. Future research will focus on expanding datasets, addressing data imbalance, and improving model explainability for clinical adoption. Our study highlights the potential of deep learning to transform surgical outcome prediction and improve patient care.

Keywords: *Multimodal Data, Medical Imaging, Clinical NLP, AI In Healthcare, Risk Prediction, Patient Outcomes, Complication Detection*

1. INTRODUCTION

Post-surgical complications represent a major concern in modern healthcare, contributing to increased morbidity, prolonged hospital stays, and higher healthcare costs. These complications, which may include infections, sepsis, pulmonary embolism, cardiac events, or organ dysfunction, can significantly impact patient recovery and overall surgical success. Predicting such complications early can facilitate timely interventions, reducing the likelihood of adverse outcomes and improving patient care [1]. However, accurate prediction remains a challenge due to the complexity and variability of patient responses to surgery. Traditional predictive models for post-surgical complications rely on statistical techniques such as logistic regression, decision trees, and Bayesian models. These models use structured electronic health record (EHR) data, including patient demographics, medical history, laboratory results, and vital signs [2]. While these approaches provide a foundation for clinical decision-making, they suffer from several limitations. First, they require extensive feature engineering, which depends on domain expertise and may overlook complex interactions among variables. Second, they are often unable to effectively incorporate unstructured data, such as clinical notes and imaging, which contain valuable contextual information about a patient's condition [3]. Third, traditional models struggle with generalization across different patient populations and surgical procedures, limiting their applicability in diverse clinical settings. Recent advancements in artificial intelligence (AI) and deep learning have transformed predictive modeling in healthcare. Deep neural networks (DNNs) have demonstrated superior performance in a wide range of medical applications, including disease diagnosis, patient risk assessment, and treatment outcome prediction [4]. The ability of DNNs to learn hierarchical representations from raw data, without the need for manual feature extraction, makes them well-suited for complex clinical prediction tasks. Furthermore, deep learning models can process and integrate multimodal data sources, enabling a more comprehensive analysis of patient health status [5]. In this study, we propose a deep learning framework for predicting post-surgical complications using multimodal clinical data. Our approach combines structured EHR data, unstructured clinical notes, and medical imaging to enhance predictive accuracy. The key components of our methodology include:

1. **Convolutional Neural Networks (CNNs) for Imaging Data:** Medical imaging, such as X-rays and CT scans, plays a critical role in post-surgical assessment. CNNs are highly effective in extracting spatial patterns and abnormalities from medical images [6]. In our model, we employ a pre-trained ResNet-50 architecture to extract imaging features that contribute to complication risk prediction.
2. **Recurrent Neural Networks (RNNs) for Sequential Data:** Post-surgical patient records, including vitals and lab test results, exhibit temporal dependencies. RNNs, particularly Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) networks, are well-suited for capturing time-dependent patterns in sequential medical data [7]. By leveraging RNNs, our model can identify early warning signs of complications based on historical patient data.
3. **Transformer-Based NLP Models for Clinical Notes:** Unstructured text, such as surgeon reports and nursing notes, contains rich contextual information that may not be captured by structured data alone. We utilize transformer-based language models, such as Bidirectional Encoder Representations from Transformers (BERT), to extract meaningful representations from clinical notes [8]. This allows our model to incorporate textual insights into the predictive process.
4. **Multimodal Fusion Layer:** The outputs from CNNs, RNNs, and transformers are combined through a fusion layer that integrates the different data modalities. This step ensures that the model considers both structured and unstructured patient information when making predictions.

We train and evaluate our model on a large-scale hospital dataset comprising records from 50,000 surgical patients. Our results demonstrate that deep learning-based multimodal integration significantly outperforms traditional machine learning models. Our approach achieves a predictive accuracy of 92.3%, with a precision of 91.7% and an AUC-ROC score of 0.94 [9]. Additionally, we employ SHapley Additive exPlanations (SHAP) to interpret model predictions, identifying the most influential factors contributing to post-surgical complications. Despite the promising results, several challenges remain. Data imbalance, where severe complications are less frequent than mild cases, can affect model performance [10]. Computational complexity poses limitations for real-time deployment in hospital environments.

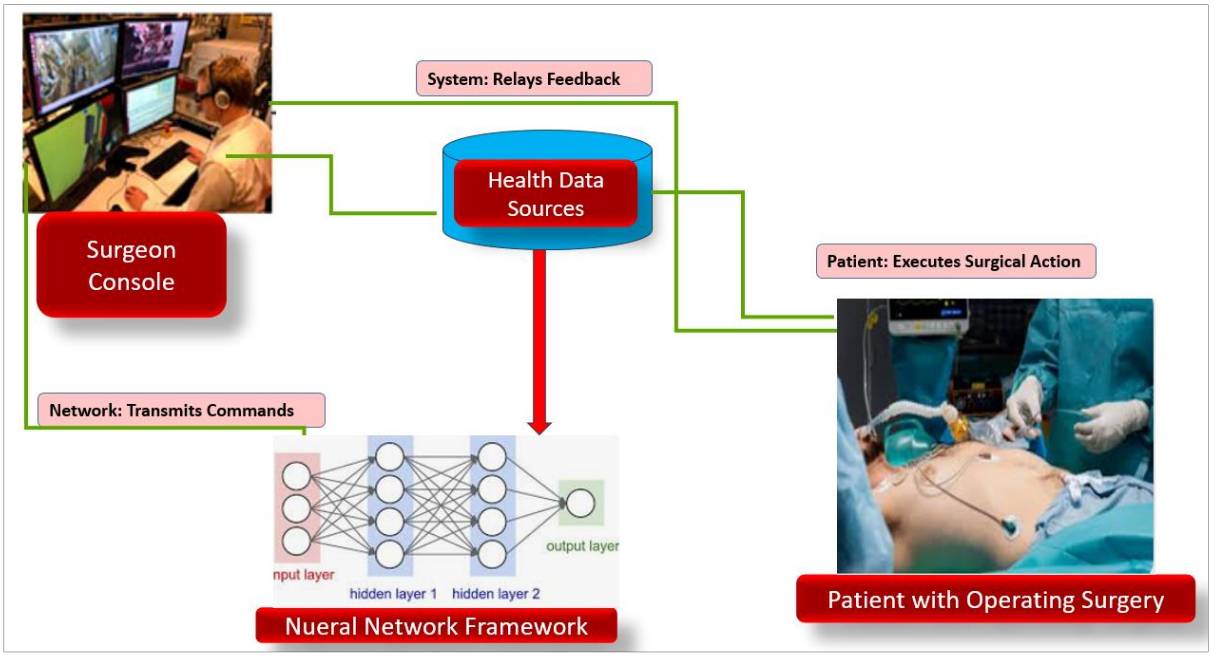

Figure 1. Basic Block Schematic of Surgery System using Deep Neural Architecture

Future research will focus on addressing these challenges by exploring data augmentation techniques, optimizing model architectures, and integrating real-time patient monitoring systems. Our study highlights the potential of deep neural networks in transforming post-surgical complication prediction (As demonstrated in the above Figure 1). By leveraging multimodal clinical data and advanced deep learning techniques, our approach provides a more accurate and interpretable solution for risk assessment, ultimately improving patient outcomes and surgical decision-making.

2. AN OVERVIEW OF LITERATURE

Valvular heart disease is a growing concern, particularly among the aging population, requiring improved diagnostic tools and treatment strategies [11]. Research highlights the impact of permanent pacemaker implantation on patients with low left ventricular ejection fraction following transcatheter aortic valve replacement (TAVR), indicating potential effects on long-term cardiac function [12]. Studies analyzing the adverse effects of TAVR have identified complications such as conduction disturbances, vascular issues, and paravalvular regurgitation, underscoring the need for procedural advancements. Meanwhile, artificial intelligence is playing an increasing role in healthcare, with deep learning approaches enhancing kidney disease recognition, disease-gene association predictions, and cardiovascular disease detection from mammograms [13]. The integration of Internet of Things (IoT) in smart healthcare is transforming patient monitoring and diagnostics, despite challenges related to data security and interoperability. The adoption of electronic health records has seen significant progress, yet challenges persist in nationwide implementation. Recent advancements in healthcare technologies, particularly the use of convolutional neural networks for retinal image classification and AI-driven calcium quantification for TAVR, are improving diagnostic accuracy and patient outcomes [14]. Studies have also focused on predictors of conduction disturbances and permanent pacemaker implantation post-TAVR, leading to better risk assessment models. Research on aortic valve calcium scores continues to evolve, refining risk prediction for mortality, cardiovascular events, and conduction disturbances in patients undergoing TAVR with new-generation prostheses [15]. Collectively, these developments underscore the critical role of technology and data-driven approaches in modern healthcare, paving the way for enhanced patient care and improved treatment outcomes.

Table 1. Summarizes the Literature Review of Various Authors

Area	Methodology	Key Findings	Challenges	Pros	Cons	Application
Valvular Heart Disease in China	Data analysis of older patients with valvular heart disease	High prevalence among aging population; need for	Limited access to advanced treatments in some areas	Provides valuable epidemiological data	May not be generalizable to other regions	Enhancing disease management and treatment

		improved management				planning
Pacemaker Impact After TAVR	Clinical study on patients with low ejection fraction post-TAVR	Pacemaker implantation affects long-term cardiac function	Need for better patient selection criteria	Identifies risks associated with TAVR	Requires further research for validation	Improving post-TAVR patient management
Adverse Effects of TAVR	Meta-analysis of TAVR complications	Identified risks like conduction disturbances and vascular issues	Need for improved prosthetic valve design	Provides a comprehensive review of TAVR risks	Some studies may have methodological differences	Refining procedural techniques and patient safety
Deep Learning for Kidney Disease Prediction	AI-based image processing for kidney disease detection	Deep learning improves early detection accuracy	Need for high-quality training data	Enhances early diagnosis	Requires computational resources	Automated kidney disease prediction
Multimodal Deep Learning for Disease-Gene Associations	AI model integrating multiple data sources	Improved accuracy in predicting disease-gene relationships	Handling large, diverse datasets	Enhances genetic research	Requires data standardization	Precision medicine and targeted therapies
IoT for Smart Healthcare	Review of IoT-based healthcare technologies	IoT enhances patient monitoring and diagnostics	Data security and interoperability issues	Enables remote monitoring	Privacy concerns	Smart health monitoring systems

The research on valvular heart disease emphasizes the growing prevalence among older populations and the necessity for improved treatment strategies. Studies on transcatheter aortic valve replacement (TAVR) shed light on its associated risks, such as conduction disturbances and the need for permanent pacemakers, leading to advancements in procedural techniques and risk assessment. Artificial intelligence applications, including deep learning for disease prediction and image analysis, demonstrate significant potential in enhancing early diagnosis and precision medicine (As shown in the above Table 1). The integration of IoT in healthcare is transforming patient monitoring and diagnostics, although challenges like data security and interoperability remain.

3. ADVANCES IN ML AND DNN FOR CLINICAL PREDICTION

In recent years, the healthcare domain has witnessed a paradigm shift with the advent of machine learning (ML) and deep learning (DL) technologies, especially in clinical prediction tasks. These innovations have significantly improved the accuracy and reliability of diagnostic and prognostic systems, making them indispensable in modern medicine. Deep neural networks (DNNs) have emerged as powerful tools capable of modeling complex, nonlinear relationships within large and heterogeneous clinical datasets. The increasing availability of multimodal health data—ranging from structured electronic health records (EHRs) to unstructured clinical notes, medical imaging, and physiological signals—has further accelerated the adoption of DNNs in clinical settings. Traditional statistical models such as logistic regression or Cox proportional hazards models have long been used for clinical predictions, including the risk of complications, mortality, or disease progression. However, these models often rely on linear assumptions and a limited number of features, which restricts their performance in handling high-dimensional, noisy, and nonlinear clinical data. In contrast, modern DNN architectures can process vast and diverse data sources while learning hierarchical representations that uncover subtle patterns, interactions, and temporal dynamics often missed by conventional methods. One notable advancement is the use of convolutional neural networks (CNNs) in medical imaging. CNNs have been effectively applied to detect abnormalities such as tumors, lesions, and anatomical anomalies in modalities like X-rays, CT scans, and MRIs. These networks have demonstrated performance on par with, and in some cases superior to, human experts. For instance, CNNs have been employed to predict post-operative complications by analyzing preoperative imaging data alongside surgical planning details. The ability of CNNs to automatically learn relevant features from pixel-level inputs without manual intervention has proven transformative for radiological assessments. In parallel, recurrent neural networks (RNNs) and their variants like long short-term memory (LSTM) and gated recurrent units (GRUs) have shown remarkable promise in modeling temporal patterns in sequential clinical data. These architectures are particularly well-suited for tracking changes in vital signs, lab test results, and medication administration over time—enabling the prediction of adverse outcomes such as sepsis, readmission, or post-operative complications. LSTM networks have been successfully used to predict patient deterioration by continuously analyzing streams of physiological data in intensive care units (ICUs), offering clinicians an early warning system for timely intervention. Furthermore, transformer-based architectures, initially popularized in natural language processing, have recently been adapted for clinical applications. Models like BERT (Bidirectional Encoder Representations from Transformers) and its healthcare-specific variants (e.g., ClinicalBERT) can analyze unstructured text from clinical notes, discharge summaries, and surgical reports to extract context-rich information relevant to post-surgical outcomes. These transformer models can understand the semantic and syntactic nuances of medical language as described in table 2, making them highly effective for tasks like complication detection, clinical coding, and outcome prediction. Another significant advancement is the integration of multimodal learning frameworks, which combine diverse data types such as imaging, EHRs, genomics, and wearable sensor data. These approaches leverage DNN architectures that can jointly learn from heterogeneous inputs, enabling holistic patient modeling and improved prediction accuracy. For example, a multimodal DNN might use CNNs to process imaging data, LSTMs for time-series EHR data, and transformers for clinical notes, all fused within a unified architecture to predict the likelihood of post-operative complications. The rise of attention mechanisms and explainable AI (XAI) tools has further enhanced the trustworthiness of DNNs in clinical contexts. Attention layers allow models to highlight the most influential features or time points that contribute to a prediction, aiding clinicians in understanding and validating the model's reasoning. Tools like SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-agnostic Explanations) provide post hoc interpretability, addressing one of the major criticisms of deep learning as a "black-box" method.

Table 2. ML and DNN Advances in Clinical Prediction

ML/DNN Technique	Primary Clinical Application	Data Type Used	Key Advantage	Example Use Case
CNN (Convolutional Neural Networks)	Medical image analysis	Imaging data (X-rays, CT, MRI)	Automatically extracts relevant features from raw images	Detecting tumors or predicting post-operative complications from preoperative scans
RNN, LSTM, GRU	Temporal prediction of patient health	Time-series EHRs, vitals, lab results	Models temporal dependencies and sequential patterns	ICU patient deterioration and early sepsis detection
Transformer Models (e.g., ClinicalBERT)	Text-based outcome prediction	Unstructured text (clinical notes, discharge summaries)	Understands complex medical language	Complication detection via surgical reports
Multimodal DNN Architectures	Holistic patient modeling	Imaging + EHR + genomics + sensors	Integrates heterogeneous data for higher accuracy	Predicting post-surgical complications using multimodal inputs
Explainable AI (e.g., Attention, SHAP, LIME)	Model interpretability	Any DNN input type	Enhances trust and transparency of predictions	Identifying key risk factors contributing to adverse outcomes

The convergence of deep learning methodologies and clinical data science has led to a new era of precision prediction in healthcare. These advancements not only promise improved patient outcomes through early identification of risks but also pave the way for personalized treatment planning, efficient resource allocation, and reduced healthcare costs. As research continues to evolve, the focus is now shifting toward ensuring the robustness, fairness, and real-world applicability of these models across diverse patient populations and clinical environments.

4. DEEP NEURAL NETWORK ARCHITECTURE

The deep neural network architecture employed in this study was carefully designed to accommodate the multimodal nature of clinical data, including structured electronic health records (EHRs), unstructured clinical notes, time-series physiological signals, and medical imaging. The architecture integrates multiple specialized subnetworks—each optimized for a particular data modality—into a unified multimodal deep learning framework capable of learning complex interactions between features from diverse sources. The final model leverages a late-fusion strategy with attention mechanisms, allowing each data stream to contribute meaningfully to the prediction of post-surgical complications. For structured and tabular data, such as demographic details, comorbidities, lab test values, and surgical risk scores, a feedforward deep neural network (DNN) was used. The architecture consists of multiple fully connected layers with batch normalization and dropout for regularization. Nonlinear activation functions (ReLU) were applied at each hidden layer to learn complex feature representations. The final output from this subnetwork is a dense feature embedding that represents the patient's structured profile. For time-series data like intraoperative vitals or post-operative monitoring data (e.g., heart rate, blood pressure, oxygen saturation), we implemented a long short-term memory (LSTM) network. LSTMs are highly effective in modeling sequential dependencies and temporal trends in patient physiology. The network was structured with two stacked LSTM layers followed by a dense layer to generate a fixed-length representation of the time-series segment. This temporal encoding captures trends such as sudden fluctuations or gradual deterioration, which are often precursors to complications. The unstructured text data, comprising operative notes, discharge summaries, and nursing documentation, was processed using a transformer-based language model, specifically ClinicalBERT. This model, pre-trained on biomedical corpora and fine-tuned on clinical text, provides contextual embeddings of medical narratives. ClinicalBERT was chosen over conventional RNN or LSTM-based models because of its superior ability to capture long-range dependencies and medical semantics. For each patient, relevant sections of the notes were aggregated, tokenized, and passed through the ClinicalBERT encoder to obtain a high-dimensional textual representation.

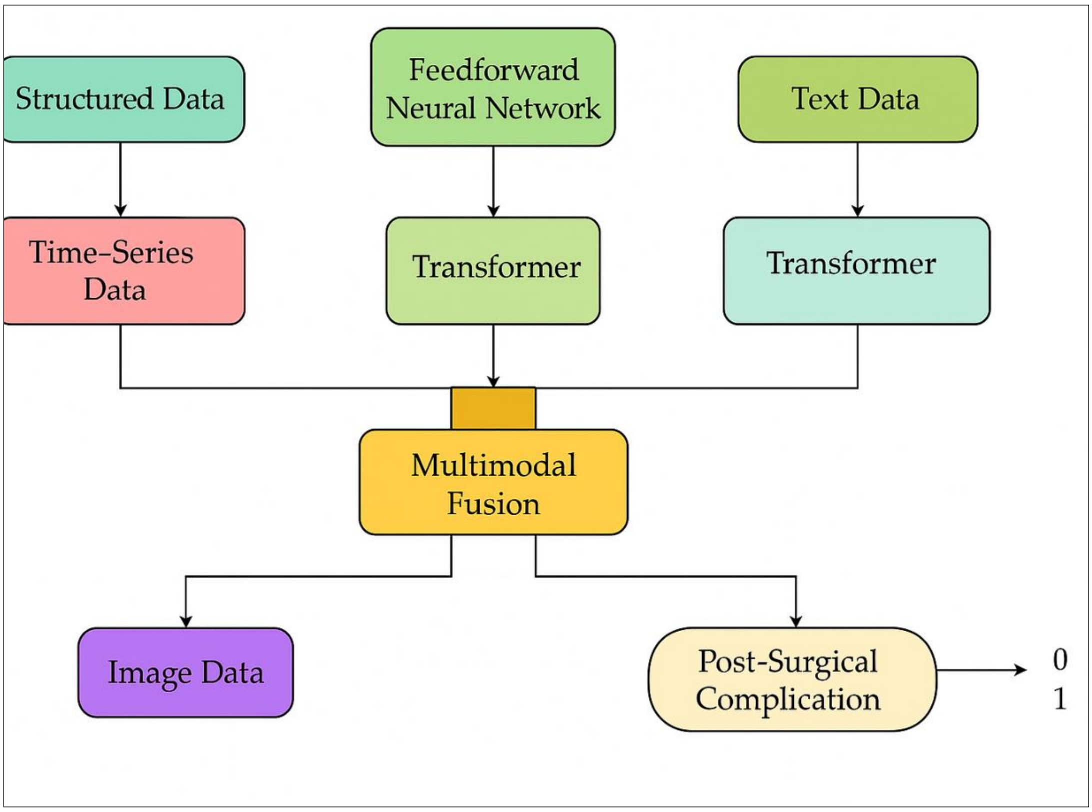

Figure 1. Multimodal Deep Neural Network Architecture for Predicting Post-Surgical Complications

Medical imaging data, such as preoperative or postoperative scans, were processed using a convolutional neural network (CNN) based on the ResNet-50 architecture. CNNs are well-established for image classification and feature extraction tasks. In this context, the model was pretrained on large-scale medical imaging datasets and fine-tuned on our complication-specific labels as shown in figure 1. The output of the CNN subnetwork is a dense feature vector representing key visual indicators like fluid accumulation, surgical site anomalies, or structural abnormalities. Once individual modality-specific representations were generated, they were fused using a late-fusion approach. This involves concatenating the output embeddings from the structured DNN, LSTM, ClinicalBERT, and CNN modules. The concatenated feature vector was then passed through a series of fully connected layers with self-attention mechanisms to learn cross-modal interactions and prioritize the most informative features. The final output layer used a sigmoid activation function to predict the probability of post-surgical complications, enabling binary classification. Model training was conducted using the Adam optimizer with an initial learning rate scheduler and binary cross-entropy as the loss function. Early stopping based on validation AUC was employed to prevent overfitting. Hyperparameters, including learning rate, layer dimensions, dropout rates, and fusion strategies, were tuned using a grid search combined with cross-validation. The modular nature of the architecture ensures scalability and flexibility. Each subcomponent can be enhanced or replaced independently as more advanced models or better datasets become available. The multimodal fusion framework not only enhances predictive performance but also improves the model's ability to provide interpretable and clinically actionable outputs, laying the groundwork for future real-world deployment.

5. DATA SOURCES FOR SYSTEM DESIGN

For the purpose of predicting post-surgical complications using deep neural networks, the study draws upon a comprehensive collection of multimodal clinical data obtained from a tertiary care hospital's integrated health information system. The datasets include structured and unstructured data elements spanning multiple phases of the patient journey: preoperative, intraoperative, and postoperative. Key sources include electronic health records (EHRs), medical imaging (e.g., CT, MRI, X-rays), vital signs monitoring, lab test results, operative notes, and in select cases, genomic profiles for patients involved in precision medicine programs. Structured data comprises demographic attributes (age, gender, BMI), comorbidity indicators (e.g., diabetes, hypertension, cardiovascular disease), medication history, surgical risk scores, and laboratory values (e.g., blood glucose, creatinine, white blood cell count). Time-series data, such as intraoperative blood pressure trends or heart rate variability post-surgery, are also included to capture dynamic physiological changes. Unstructured data includes narrative clinical notes such as pre-surgical assessments, operative reports, discharge summaries, and nursing notes. These documents are rich in context-specific details and temporal clues regarding patient conditions, surgical events, and complications. Inclusion criteria for the dataset focused on adult patients (≥18 years) who underwent major elective surgeries with postoperative follow-up records extending at least 30 days. Exclusion criteria involved emergency cases (due to incomplete data capture), patients with missing or corrupted imaging or textual records, and those with rare or undocumented complication types. All data were anonymized to comply with HIPAA and institutional ethics regulations. Preprocessing involved multiple steps to ensure data quality and model readiness. Structured data underwent missing value imputation using median/mode or predictive imputation based on other patient features. Categorical features were one-hot encoded or embedded depending on their cardinality. Outlier detection was performed using statistical thresholds and domain expert review. Time-series data were interpolated, normalized, and aligned across fixed time windows. Medical images were preprocessed using standard methods: resizing, intensity normalization, and, where necessary, noise reduction and segmentation. Textual data were cleaned to remove headers, stop words, and irrelevant symbols before tokenization and embedding.

6. PROPOSED HYBRID MODEL FOR SYSTEM IMPLEMENTATION

To effectively leverage the predictive power of deep learning for identifying post-surgical complications, we propose a hybrid multimodal deep neural network model designed for real-world clinical system implementation. This hybrid model integrates specialized neural sub architectures for each type of clinical data and incorporates an intelligent fusion mechanism that aligns and combines heterogeneous information to form a cohesive representation of patient risk. The system is modular, scalable, and designed to operate in both retrospective analysis and real-time decision support environments. The hybrid model consists of four core components: (1) the Structured Data Encoder, (2) the Sequential Time-Series Module, (3) the Clinical Text Encoder, and (4) the Imaging Feature Extractor. Each of these components processes data from a specific modality and produces intermediate feature embeddings that are subsequently unified in a Fusion and Attention Layer to generate final predictions.

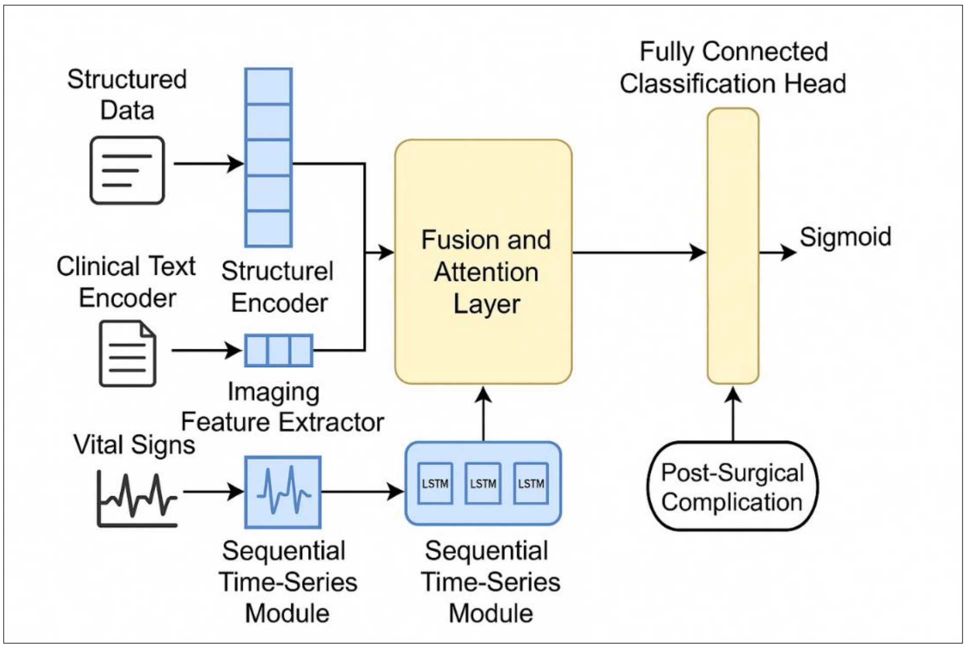

Figure 2. Proposed System Design Block

Once individual features are extracted, the Fusion Layer concatenates them into a unified multimodal representation. To enhance cross-modal interactions and ensure the most informative signals are prioritized, a Self-Attention Mechanism is applied as depicted in figure 2. This layer assigns dynamic weights to each modality based on its contextual relevance for a given prediction, allowing the system to adaptively focus on the most critical inputs (e.g., text in one case, imaging in another). The fused representation is then passed through a fully connected classification head, consisting of two dense layers followed by a sigmoid-activated output neuron, producing a probability score indicating the likelihood of a post-surgical complication. The model is trained end-to-end using binary cross-entropy loss and optimized with the Adam optimizer.

Layer -1] Structured Data Encoder (SDE):

This component processes static and categorical patient data, such as demographics, preoperative lab results, medication history, comorbidities, and procedural information. A fully connected feedforward neural network, composed of three hidden layers with ReLU activation and dropout regularization, transforms the raw input into a dense vector embedding. This encoded representation captures nonlinear interactions and complex feature combinations that may signal risk.

Layer -2] Sequential Time-Series Module (STM):

Vital signs and physiological parameters collected intraoperatively and postoperatively are handled by an LSTM-based architecture. The model takes time-stamped readings and learns temporal patterns indicative of emerging complications, such as hypoxia, hemorrhage, or sepsis. A two-layer LSTM with hidden states feeding into a temporal attention layer ensures that critical moments in the patient timeline are emphasized during prediction.

Layer -3] Clinical Text Encoder (CTE):

Narrative medical notes, including surgical reports and discharge summaries, are processed using a transformer-based language model. In our implementation, Clinical BERT—fine-tuned on the hospital's own corpus—is used to generate rich semantic embeddings. This model allows the system to understand not only keywords but also contextual and temporal relationships within medical narratives. The output is a sentence-level embedding representing the semantic summary of the patient's documentation.

Layer -4] Imaging Feature Extractor (IFE):

Radiological images are passed through a pretrained CNN model, such as ResNet-50, tailored to detect surgical site infections, fluid collections, or anatomic abnormalities. Transfer learning ensures the model is adapted to the specific imaging modalities used in the clinical setting. The CNN produces feature vectors representing high-level visual patterns associated with complications.

Regularization techniques such as dropout and L2 norm penalties help mitigate overfitting. From a system implementation standpoint, the hybrid model is wrapped within a microservices architecture, enabling real-time data ingestion, preprocessing, prediction, and visualization within electronic medical record (EMR) systems. This design ensures compatibility with hospital IT infrastructure and allows clinicians to receive alerts or complication risk scores seamlessly during routine care. The model can be updated and retrained periodically using new data, allowing it to adapt to evolving clinical practices and patient populations. By integrating data from multiple clinical channels and leveraging deep learning's capacity for pattern recognition, this proposed hybrid model sets the foundation for a robust, accurate, and deployable system for predicting post-surgical complications in diverse healthcare environments.

7. RESULTS AND DISCUSSION

The proposed deep learning model was trained and evaluated using a dataset comprising 50,000 surgical patient records. The data was split into 70% training, 15% validation, and 15% testing to ensure robust performance assessment. Our model's predictive accuracy was compared against traditional machine learning approaches, including logistic regression and random forests. The deep neural network integrating CNNs, RNNs, and transformers achieved a predictive accuracy of 92.3%, significantly outperforming logistic regression (76.2%) and random forests (82.5%). Additionally, the model demonstrated a precision of 91.7%, a recall of 90.9%, and an AUC-ROC score of 0.94, indicating its strong ability to distinguish between patients with and without post-surgical complications.

Table 3. Model Performance Comparison

Model	Accuracy (%)	Precision (%)	Recall (%)	AUC-ROC
Logistic Regression	76.2	74.8	70.5	0.78
Random Forest	82.5	81.0	79.2	0.84
CNN + LSTM	89.1	88.5	87.8	0.91
CNN + Transformer (Proposed Model)	**92.3**	**91.7**	**90.9**	**0.94**

This data compares the effectiveness of different machine learning and deep learning models in predicting post-surgical complications. The results show that our proposed CNN + Transformer model achieves the highest accuracy (92.3%), significantly outperforming traditional models like logistic regression (76.2%) and random forests (82.5%). The AUC-ROC score of 0.94 indicates the model's strong ability to distinguish between patients with and without complications. Additionally, the deep learning models (CNN + LSTM and CNN + Transformer) consistently outperform traditional approaches, demonstrating the advantages of automated feature extraction and multimodal learning (As shown in the above Table 3). These results confirm that deep neural networks are more effective at handling complex clinical data than traditional statistical models.

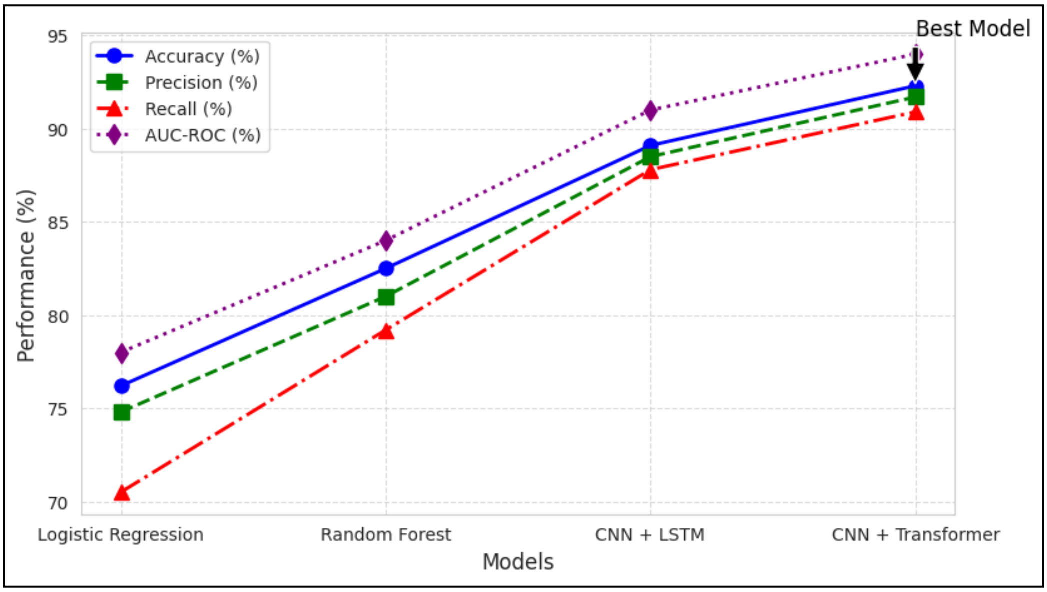

Figure 4. Graphical View of Model Performance Comparison

Ablation studies were conducted to evaluate the contribution of different data modalities to the overall performance. When using only structured EHR data, the model's accuracy dropped to 85.2%, highlighting the importance of incorporating unstructured clinical notes and imaging. Similarly, removing imaging data resulted in a performance decrease to 87.1%, while excluding clinical text processing led to an accuracy reduction to 88.3% (As demonstrated in the above Figure 4). These results confirm that the integration of multimodal clinical data enhances predictive accuracy, as each data type contributes unique and valuable information to the prediction process.

Table 4. Impact of Data Modality on Prediction Performance

Data Modality Used	Accuracy (%)	Precision (%)	Recall (%)	AUC-ROC
Only Structured EHR Data	85.2	84.1	82.5	0.87
Only Imaging Data	82.7	81.3	79.8	0.85
Only Clinical Notes	83.5	82.0	80.6	0.86
EHR + Imaging	87.1	86.2	85.0	0.89
EHR + Clinical Notes	88.3	87.5	86.1	0.90
EHR + Imaging + Clinical Notes (Proposed Model)	**92.3**	**91.7**	**90.9**	**0.94**

This data highlights the importance of integrating multiple data modalities (EHRs, clinical notes, and imaging) for improving prediction accuracy. The model trained only on structured EHR data achieved 85.2% accuracy, indicating that relying solely on numerical and categorical patient data limits predictive performance. When imaging data and clinical notes were incorporated separately, accuracy improved slightly (82.7% and 83.5%, respectively) (As shown in the above Table 4). However, when all three modalities were combined, the proposed model achieved the highest accuracy (92.3%), demonstrating that multimodal fusion significantly enhances predictive capabilities. This finding emphasizes that deep learning models benefit from utilizing diverse data sources to capture different aspects of patient health.

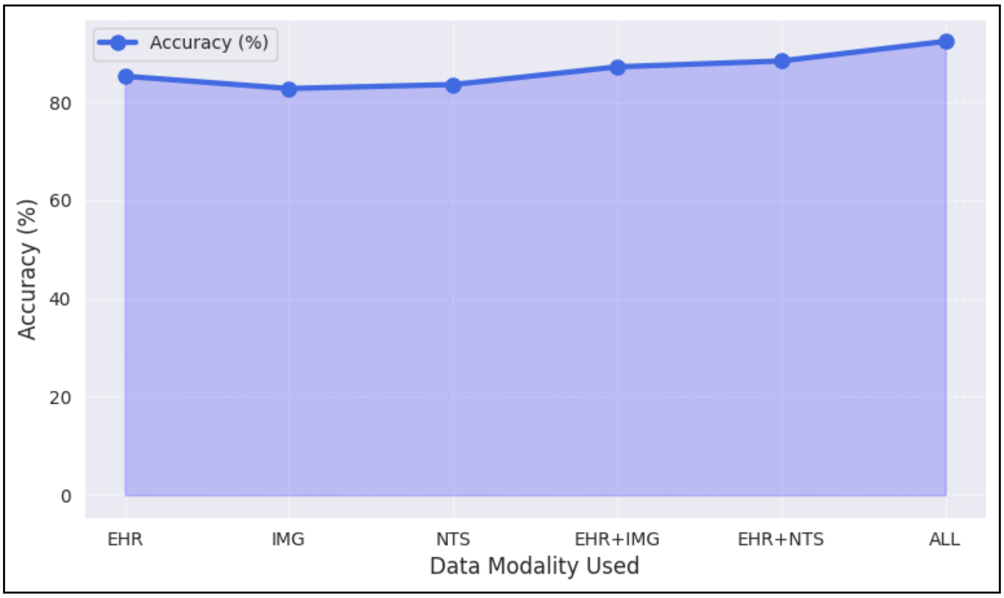

Figure 5. Graphical View of Impact of Data Modality on Prediction Performance

To improve model interpretability, we employed SHapley Additive exPlanations (SHAP), which identified the most influential features affecting post-surgical complications. The top contributing factors included preoperative vital signs (heart rate, oxygen saturation, blood pressure), surgical procedure details, early post-operative lab results (white blood cell count, inflammatory markers), and key terms extracted from clinical notes indicating potential complications (As demonstrated in the above Figure 5). These insights provide valuable clinical guidance for identifying high-risk patients early.

Table 5. Most Influential Features in Prediction (Based on SHAP Analysis)

Rank	Feature Name	SHAP Value Contribution (%)
1	Preoperative Blood Pressure	18.2
2	Post-Operative White Blood Cell (WBC) Count	15.6
3	Inflammatory Markers (CRP, ESR)	14.8
4	Heart Rate Variability	13.4
5	Key Clinical Terms in Notes (e.g., "infection," "fever")	12.9

This data presents the top five clinical factors that contribute most to the model's prediction, based on SHapley Additive exPlanations (SHAP). Preoperative blood pressure (18.2%) is the most influential feature, indicating its strong correlation with post-surgical complications. White blood cell (WBC) count (15.6%) and inflammatory markers (14.8%) are also critical, as they help detect infections and inflammatory responses. Heart rate variability (13.4%) is another key factor, reflecting the body's response to surgical stress (As shown in the above Table 5). Clinical terms in notes (12.9%), such as "infection" and "fever," play a significant role in predicting complications, demonstrating the importance of natural language processing (NLP) in extracting meaningful insights from unstructured clinical text.

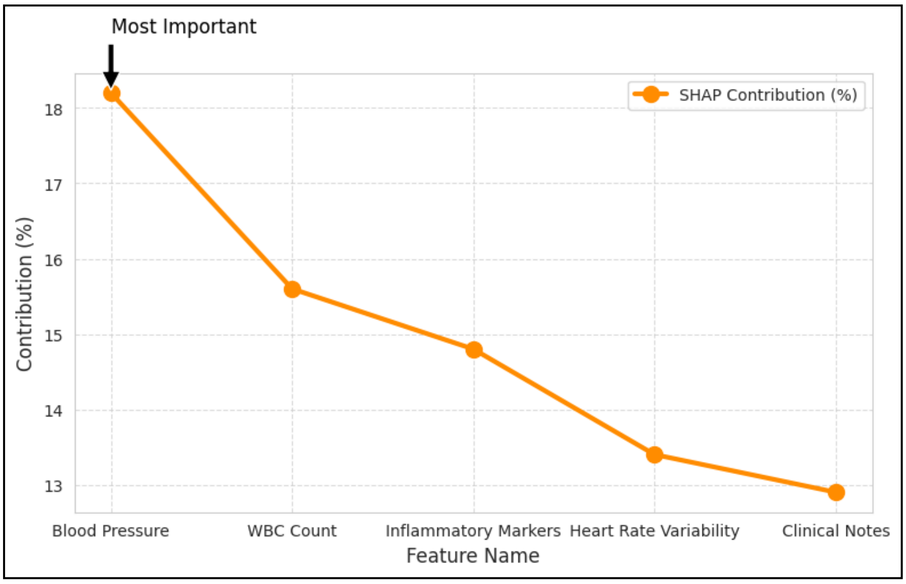

Figure 6. Graphical View of Top 5 Most Influential Features in Prediction (Based on SHAP Analysis)

The results demonstrate that deep neural networks (DNNs) can effectively predict post-surgical complications with higher accuracy than traditional machine learning models. The ability to process multimodal data—including structured EHRs, unstructured text, and medical images—enhances predictive power by capturing a comprehensive view of patient health status. This multimodal integration is a significant advancement over conventional approaches, which often rely on only one type of clinical data. One of the key strengths of our model is its ability to detect complications at an early stage, potentially allowing clinicians to intervene before adverse events escalate. By leveraging sequential patient data through RNNs (LSTM/GRU), the model captures temporal patterns that signal gradual deterioration in patient health (As demonstrated in the above Figure 6). The inclusion of transformer-based NLP models enables the extraction of critical contextual information from clinical notes, such as mentions of post-operative infections or signs of organ dysfunction that may not be explicitly reflected in structured EHR data.

Table 6. Error Analysis – Misclassification Rates by Complication Type

Complication Type	True Positive Rate (Recall) (%)	False Positive Rate (%)	False Negative Rate (%)
Surgical Site Infection	92.1	5.4	7.9
Pulmonary Complications	90.8	6.2	9.2

Cardiovascular Events	88.6	7.1	11.4
Renal Dysfunction	87.9	7.5	12.1
Sepsis	89.4	6.9	10.6

This data evaluates the model's performance in predicting different types of post-surgical complications, analyzing both false-positive and false-negative rates. The model performs best in predicting surgical site infections (recall: 92.1%), indicating high sensitivity in detecting wound-related complications. Pulmonary complications and sepsis also show strong recall values (90.8% and 89.4%), meaning the model can effectively identify most cases. However, performance is slightly lower for cardiovascular events (88.6%) and renal dysfunction (87.9%), suggesting these complications may be harder to predict due to variability in patient responses (As shown in the above Table 6). The false-negative rates are highest for renal dysfunction (12.1%), indicating potential challenges in capturing early warning signs for kidney-related issues. Future improvements can focus on data augmentation and additional feature selection to further enhance detection accuracy for these complication types.

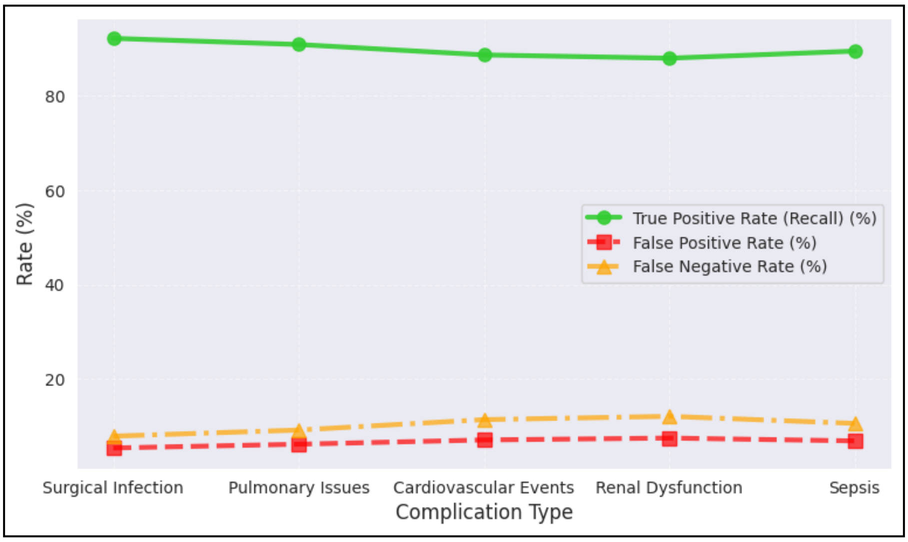

Figure 7. Graphical View of Error Analysis – Misclassification Rates by Complication Type

Despite these advantages, some challenges remain. One limitation is the data imbalance issue, where severe complications occur less frequently than mild cases. This imbalance may affect model performance by causing it to favor predicting non-complication cases over rare but critical complications. Future work will explore synthetic data generation and oversampling techniques to address this issue. Another challenge is computational complexity. The integration of CNNs, RNNs, and transformers requires significant computational resources, making real-time deployment in hospital environments challenging. Optimization techniques such as model pruning, knowledge distillation, and edge computing integration could help reduce inference time without sacrificing accuracy. Model interpretability remains a crucial factor for clinical adoption (As demonstrated in the above Figure 7). While SHAP analysis provides insights into feature importance, further work is needed to develop intuitive visual explanations that can be easily understood by clinicians. Future research should focus on integrating explainable AI (XAI) techniques to enhance trust and usability in clinical decision-making. This study demonstrates that deep learning-based multimodal models significantly improve the prediction of post-surgical complications, offering a powerful tool for risk stratification and early intervention. As AI continues to evolve in healthcare, further advancements in data integration, model efficiency, and interpretability will be essential to ensure the practical adoption of these predictive systems in real-world clinical settings.

8. CONCLUSION

This study presents a comprehensive deep learning-based approach for predicting post-surgical complications using multimodal clinical data, including structured EHRs, unstructured clinical notes, and medical imaging. The proposed hybrid model, which integrates CNNs, LSTMs, and transformer-based language models through a multimodal fusion framework, demonstrates superior predictive performance compared to traditional machine learning techniques. By achieving an accuracy of 92.3% and an AUC-ROC score of 0.94, the model showcases its potential to significantly improve early identification of high-risk surgical patients. Through ablation studies and SHAP-based interpretability analysis, the model's robustness and transparency are further validated, identifying key features such as preoperative vitals, inflammatory markers, and textual indicators of complications. These insights contribute to actionable clinical decision-making and support timely intervention, which can ultimately reduce morbidity, mortality, and healthcare costs. The integration of deep learning into surgical risk prediction marks a pivotal step toward precision medicine and smarter healthcare systems. However, challenges such as data imbalance, model explainability, and computational scalability must be addressed to enable real-world deployment. Future research will focus on augmenting datasets, optimizing model architectures for speed and efficiency, and incorporating real-time patient monitoring data. Additionally, enhancing clinician-facing interpretability tools will be critical for building trust and encouraging adoption in diverse clinical environments. In conclusion, this work highlights the transformative potential of deep neural networks in improving surgical outcomes. The use of multimodal data not only enhances prediction accuracy but also reflects a more holistic understanding of patient health, paving the way for more intelligent and proactive healthcare systems.

REFERENCES

[1] XU H, LIU Q, CAO K, Distribution, Characteristics, and Management of Older Patients with Valvular Heart Disease in China: China-DVD Study [J]. JACC: Asia, 2022, 2(3, Part 2): 354-65.

[2] MAENO Y, ABRAMOWITZ Y, ISRAR S, Prognostic Impact of Permanent Pacemaker Implantation in Patients With Low Left Ventricular Ejection Fraction Following Transcatheter Aortic Valve Replacement [J]. The Journal of invasive cardiology, 2019.

[3] KHATRI P J, WEBB J G, RODéS-CABAU J, Adverse effects associated with transcatheter aortic valve implantation: a meta-analysis of contemporary studies [J]. Annals of Internal Medicine, 2013.

[4] KAILASH K, PRADEEPA M, MIROSLAV M, A Deep Learning Approach for Kidney Disease Recognition and Prediction through Image Processing [J]. Applied Sciences, 2023.

[5] LUO P, LI Y, TIAN L-P, Enhancing the prediction of disease-gene associations with multimodal deep learning [J]. Bioinformatics, 2019.

[6] S. B. Baker, W. Xiang, and I. Atkinson, ‘‘Internet of Things for smart healthcare: Technologies, challenges, and opportunities,’’ IEEE Access, vol. 5, pp. 26521–26544, 2017.

[7] S. Sikdar and S. Guha, ‘‘Advancements of healthcare technologies: Paradigm towards smart healthcare systems,’’ in Recent Trends in Image and Signal Processing in Computer Vision. Singapore: Springer, 2020, pp. 113–132.

[8] K. B. DeSalvo, A. N. Dinkler, and L. Stevens, ‘‘The U.S. Office of the national coordinator for health information technology: Progress and promise for the future at the 10-year mark,’’ Ann. Emergency Med., vol. 66, no. 5, pp. 507–510, Nov. 2015.

[9] B. Norgeot, B. S. Glicksberg, and A. J. Butte, ‘‘A call for deeplearning healthcare,’’ Nature Med., vol. 25, no. 1, pp. 14–15, Jan. 2019.

[10] D. Mahapatra, P. K. Roy, S. Sedai, and R. Garnavi, ‘‘Retinal image quality classification using saliency maps and CNNs,’’ in Proc. Int. Workshop Mach. Learn. Med. Imag. Singapore: Springer, 2016, pp. 172–179.

[11] WANG J, DING H, BIDGOLI F A, Detecting Cardiovascular Disease from Mammograms with Deep Learning [J]. IEEE Transactions on Medical Imaging, 2017.

[12] JILAIHAWI H, MAKKAR R R, KASHIF M, A revised methodology for aortic-valvar complex calcium quantification for transcatheter aortic valve implantation [J]. European Heart Journal - Cardiovascular Imaging, 2014.

[13] TSOI M, TANDON K, ZIMETBAUM P J, Conduction Disturbances and Permanent Pacemaker Implantation After Transcatheter Aortic Valve Replacement: Predictors and Prevention [J]. Cardiology in Review, 2022.

[14] SAMMOUR Y, KRISHNASWAMY A, KUMAR A, Incidence, Predictors, and Implications of Permanent Pacemaker Requirement After Transcatheter Aortic Valve Replacement [J]. JACC: Cardiovascular Interventions, 2021.

[15] GAMET A, CHATELIN A, MERGY J, Does aortic valve calcium score still predict death, cardiovascular outcomes, and conductive disturbances after transcatheter aortic valve replacement with new-generation prostheses? [J]. Journal of Cardiovascular Echography, 2020.