EARLY DETECTION OF CHEMOTHERAPY-INDUCED HEART DAMAGE USING MULTIMODAL DEEP LEARNING

[1] Mr. P.Kishore, [2] K. Akhil, [3] G. Amulya, [4] K. Akshay, [5] I. Nagaraju

[1]Assiatant Professor in Department of CSE TKR COLLEGE OF ENGINEERING & TECHNOLOGY

[1]ponnamkishore@tkrcet.com

[2,3,4,5]UG Scholars in Department of CSE TKR COLLEGE OF ENGINEERING & TECHNOLOGY

[2]akhilkonda1439@gmail.com , [3] amulyagolkonda@gmail.com , [4] aakhi2407@gmail.com , [5] nagarajuindla55@gmail.com

Abstract

Chemotherapy is widely used to treat cancer, but it can unintentionally affect the heart, leading to a condition known as cardiotoxicity. One of the main difficulties in managing this problem is that early-stage heart damage often does not produce clear symptoms, making timely diagnosis challenging. Because of this, there is a growing need for reliable methods that can identify risk at an early stage. A multimodal deep learning approach is developed to improve the early detection of chemotherapy-related heart damage. The proposed system combines cardiac imaging data, specifically Temporal Dynamic Imaging (TDI), with patient clinical information to obtain a more complete understanding of heart function. Convolutional Neural Networks are used to extract important features from the images, while models such as LSTM, GRU, and Transformer are applied to learn patterns over time. Among the different models tested, the CNN combined with a Transformer architecture showed better performance in identifying subtle changes in cardiac behavior. The system was evaluated using common performance measures like accuracy, precision, recall, and F1-score, and the results indicate that integrating multiple data sources improves prediction quality. This approach can assist doctors in identifying potential heart complications earlier, allowing for better monitoring and timely treatment adjustments during chemotherapy.

KEYWORDS

Chemotherapy cardiotoxicity, Early detection, Deep learning, Multimodal analysis, CNN, Transformer, LSTM, GRU, Cardiac imaging, Clinical prediction

I INTRODUCTION

The use of advanced computational methods in healthcare has increased significantly in recent years, especially for disease prediction and medical image analysis. Early research has highlighted the importance of properly labeled medical data in building reliable deep learning systems, as the quality of input data directly affects prediction accuracy [1]. In addition, predictive frameworks using intelligent feature extraction techniques have shown that machine learning models can assist in identifying diseases at an earlier stage, improving clinical decision-making [2]. With the growing adoption of connected healthcare systems, modern approaches combining secure data handling and artificial intelligence have further strengthened the role of technology in medical diagnosis [3].

Many health-related problems, particularly cardiac conditions, develop gradually over time rather than appearing suddenly. Because of this, temporal modeling has become an important area of research. Studies have shown that analyzing sequential health data can help in identifying early signs of diseases such as heart failure [4]. Similarly, recurrent models like LSTM and GRU have been used to capture time-dependent patterns in patient data, making them useful for predicting changes in health conditions [5]. Other approaches focusing on sequential feature analysis also emphasize the importance of understanding how medical data evolves over time [6].

At the same time, deep learning techniques have made significant progress in image-based applications. Convolutional Neural Networks (CNNs), in particular, have proven to be highly effective in extracting meaningful features from complex medical images [7]. In parallel, advancements in computational methods and large-scale data processing have enabled more efficient handling of complex datasets, supporting the development of more accurate predictive models [8]. These improvements have opened new possibilities for combining imaging data with other clinical information.

In the context of cancer treatment, chemotherapy is widely used but can lead to serious side effects, including damage to the heart. Research has shown that deep learning models applied to cardiac imaging data can help detect early signs of cardiotoxicity before symptoms become visible [9]. More recently, Transformer-based architectures have gained attention because of their ability to capture long-range dependencies using attention mechanisms, which improves the analysis of complex medical data [10]. Furthermore, studies on temporal and sequential modeling continue to reinforce the importance of using advanced techniques to detect subtle changes in patient health conditions at an early stage [11].

Motivated by these developments, this work focuses on the early detection of chemotherapy-induced heart damage using a multimodal deep learning approach. The proposed system combines Temporal Dynamic Imaging (TDI) data with clinical features to provide a more comprehensive analysis. CNNs are used to extract spatial features from images, while models such as LSTM, GRU, and Transformer are applied to learn temporal patterns. By comparing different hybrid models, the study aims to identify an effective method for accurate and early prediction.

II LITERATURE SURVEY

Recent developments in healthcare analytics have shown that the success of intelligent prediction systems depends heavily on the quality and structure of the data used for training. Orting et al. [1] emphasized the importance of proper label design in biomedical image analysis, pointing out that well-annotated datasets are essential for building reliable deep learning models. Without accurate labeling, even advanced algorithms may fail to deliver consistent results.

Alzahrani [2] proposed a predictive framework that integrates advanced feature generation techniques with machine learning algorithms for early disease detection. The study demonstrated that meaningful feature extraction plays a key role in improving prediction accuracy. Similarly, Khan et al. [3] explored the use of secure and intelligent healthcare systems supported by transfer learning, highlighting how modern AI approaches can enhance disease prediction while maintaining data security in connected medical environments.

Understanding temporal patterns in healthcare data has also been an important research direction. Bahadori et al. [4] showed that many medical conditions, including heart-related diseases, develop over time and require models capable of capturing sequential dependencies. Supporting this idea, Wang et al. [5] applied recurrent neural networks to predict adverse drug reactions, demonstrating that time-series models such as RNNs can effectively learn patient-specific trends from longitudinal data. In a related context, Zhang et al. [6] presented a Temporal Convolutional Network model that further reinforced the importance of handling sequential data for dynamic pattern recognition.

Deep learning techniques, particularly in image processing, have significantly improved medical diagnostics. He et al. [7] introduced deep residual learning, which enabled the training of very deep neural networks and improved feature extraction capabilities in image recognition tasks. This advancement has been widely adopted in medical imaging applications. Additionally, Dosovitskiy et al. [8] proposed the Vision Transformer, which demonstrated that transformer-based models can also be effectively applied to image analysis by capturing global contextual information.

In the specific area of cardiotoxicity detection, Wang et al. [9] explored the use of deep learning on longitudinal echocardiographic imaging and showed that such approaches can identify early signs of cardiac dysfunction more effectively than traditional methods. This work highlights the potential of combining imaging data with advanced learning techniques for early diagnosis. Furthermore, Vaswani et al. [10] introduced the Transformer architecture, which uses attention mechanisms to model relationships within data sequences. This approach has proven to be highly effective in capturing long-range dependencies and has influenced many recent healthcare applications.

Finally, Zhang et al. [11] demonstrated the application of machine learning techniques for predicting chemotherapy-induced cardiotoxicity. Their findings indicate that data-driven approaches can significantly improve early risk identification compared to conventional clinical methods.

III RELATED WORK

A number of earlier works in healthcare prediction have mainly focused on using either medical images or patient clinical data to identify diseases. Systems based on medical imaging have shown good results in detecting visible abnormalities in organs, especially with the help of deep learning models. These methods are useful in identifying structural changes in the heart, but they sometimes fail to consider patient history, treatment details, or other clinical factors that can influence the condition. On the other side, approaches that depend only on clinical data can capture patient-specific information, but they may not fully reflect the actual condition of the heart at a given time.

To improve performance, some researchers have tried combining different types of data. These approaches attempt to use both imaging and clinical information together so that the model can learn from multiple sources. In such cases, sequence-based models like LSTM and GRU are often used to study how patient conditions change over time. These methods are useful when the disease develops gradually, but they still have certain limitations. In particular, they may struggle to capture complex relationships in long sequences of data, and their performance can be affected by noise or incomplete information.

More recent studies have started exploring advanced models that use attention mechanisms to better understand patterns in data. These models are capable of focusing on important parts of the input and handling complex relationships more effectively. Even though these approaches show promising results, many existing systems are still not fully optimized for real-world medical use. Issues such as limited data, lack of proper integration between different data types, and challenges in practical deployment continue to exist. Because of these gaps, there is a clear need for a more balanced and reliable system that can combine different inputs and provide early and accurate predictions of heart damage caused by chemotherapy.

IV PROBLEM STATEMENT

Cancer treatment using chemotherapy has become common in modern medicine, but it often comes with unintended side effects that can affect a patient's overall health. One such serious issue is damage to the heart, which may develop slowly during the course of treatment. The difficulty lies in the fact that this type of heart damage does not usually show clear signs in the early stages. Patients may appear normal while subtle changes are already taking place in cardiac function. By the time noticeable symptoms appear, the condition may have progressed to a more critical stage, making it harder to manage.

In current clinical practice, doctors rely on periodic scans, reports, and general observations to monitor heart health. While these methods are useful, they are not always sensitive enough to detect very small or gradual changes. This creates a gap between the onset of the condition and its actual diagnosis. Another limitation is that most existing systems consider either imaging data or patient clinical details separately, rather than using both together to form a complete picture. Because of this, important information may be overlooked during analysis.Considering these challenges, there is a clear need for a more reliable and data-driven approach that can support early identification of chemotherapy-related heart damage. A system that can combine medical images with clinical information and learn patterns from both sources may offer better prediction capability. Such an approach can help in identifying risk at an earlier stage, allowing doctors to take timely decisions and reduce the chances of severe complications.

V PROPOSED SYSTEM

This work introduces a system that focuses on identifying heart-related complications caused by chemotherapy at an early stage. Instead of depending on a single type of input, the system is designed to make use of both cardiac imaging data and patient clinical details. The idea is that combining these two sources can give a clearer and more complete view of the patient's condition.

The proposed approach makes use of Temporal Dynamic Imaging data, which provides information about how the heart is functioning over a period of time. In addition to this, clinical parameters related to the patient are also included so that the model does not rely only on visual data. Before using the data for prediction, basic preprocessing steps are carried out. These include adjusting image size, normalizing values, and organizing clinical data in a consistent format. This step ensures that the inputs are clean and suitable for model training.

For analysis, the system uses deep learning models that can handle both image and time-based data. A Convolutional Neural Network is used to extract important patterns from the images. To understand how these patterns change over time, models like LSTM, GRU, and Transformer are applied. Each of these models helps in learning different types of relationships within the data. After processing, the information from both image features and clinical inputs is combined to produce the final result.

Multiple model combinations are tested to see which one performs better in predicting the condition. Among these, the combination of CNN with a Transformer model shows better results, mainly because it can handle complex patterns more effectively. The final system is also designed in such a way that it can be used easily, where a user can provide the required inputs and obtain the prediction result along with a confidence level.

The overall aim of this system is to support early identification of heart damage during chemotherapy. By providing timely predictions, it can help doctors monitor patients more carefully and take necessary actions before the condition becomes severe.

VI METHODOLOGY

The proposed system follows a structured approach to develop a reliable model for early detection of chemotherapy-induced heart damage. The methodology is designed to handle both imaging data and clinical information in a coordinated manner so that meaningful patterns can be learned effectively.

The process begins with data collection, where Temporal Dynamic Imaging (TDI) data and corresponding clinical features are gathered. Since the data may contain inconsistencies, a preprocessing stage is carried out to improve its quality. In this step, images are resized to a common format and normalized to maintain uniform pixel values. At the same time, clinical data is cleaned by removing missing or invalid entries, and numerical values are scaled to a consistent range. This ensures that both types of data are properly aligned and suitable for model training.

Once the data is prepared, feature extraction is performed on the imaging data using a Convolutional Neural Network. This step helps in identifying important spatial patterns from the heart images that may not be easily visible through manual observation. Since the data also contains time-related information, sequence models are used to capture how these patterns change over time. For this purpose, LSTM, GRU, and Transformer models are applied separately to learn temporal relationships in the data.

After extracting both spatial and temporal features, the next step involves combining these outputs with the clinical features. This fusion of data allows the model to consider multiple aspects of the patient's condition simultaneously. The combined features are then passed through fully connected layers, which perform the final classification. The system predicts whether the patient belongs to the cardiotoxic group or the non-cardiotoxic group.

To evaluate performance, different hybrid models such as CNN + LSTM, CNN + GRU, and CNN + Transformer are trained and tested. Each model is assessed using evaluation metrics like accuracy, precision, recall, and F1-score. Based on the results, the model with the best performance is selected for final use.The methodology is designed to ensure that the system can learn effectively from both image and clinical data. By following this approach, the model is able to provide more accurate and early predictions, which can be useful in real-world clinical settings.

VII IMPLEMENTATION

The implementation of the proposed system focuses on building a practical model that can process both imaging data and clinical information to predict chemotherapy-induced heart damage. The system is developed in a step-by-step manner, starting from data handling to final deployment, ensuring that each stage works smoothly with the next.

The initial phase involves preparing the dataset for model training. Temporal Dynamic Imaging (TDI) data and clinical records are loaded into the system and checked for any inconsistencies. Images are resized to a fixed dimension so that they can be processed uniformly by the model. Pixel values are normalized to improve learning efficiency. At the same time, clinical data is cleaned by removing incomplete entries and scaling numerical values to a standard range. This step helps in maintaining consistency across different types of inputs.

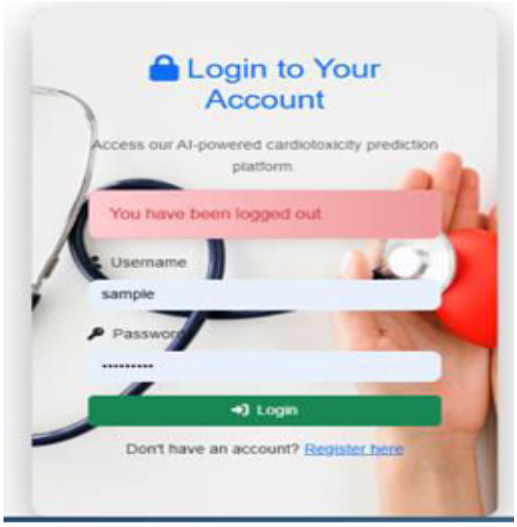
After preprocessing, the model development phase begins. A Convolutional Neural Network is used as the base component to extract important features from the image data. These extracted features represent key patterns related to heart function. To capture time-based changes, sequence models such as LSTM, GRU, and Transformer are integrated with the CNN. Each of these models is implemented separately to study how well they learn temporal relationships. The outputs from these models are combined with clinical features, and the final prediction is generated using fully connected layers.

The models are trained using the prepared dataset, with a portion of the data reserved for validation and testing. During training, the system continuously adjusts its internal parameters to improve prediction accuracy. Standard evaluation measures such as accuracy, precision, recall, and F1-score are used to compare the performance of different model combinations. Based on the observed results, the most effective model is selected for further use.

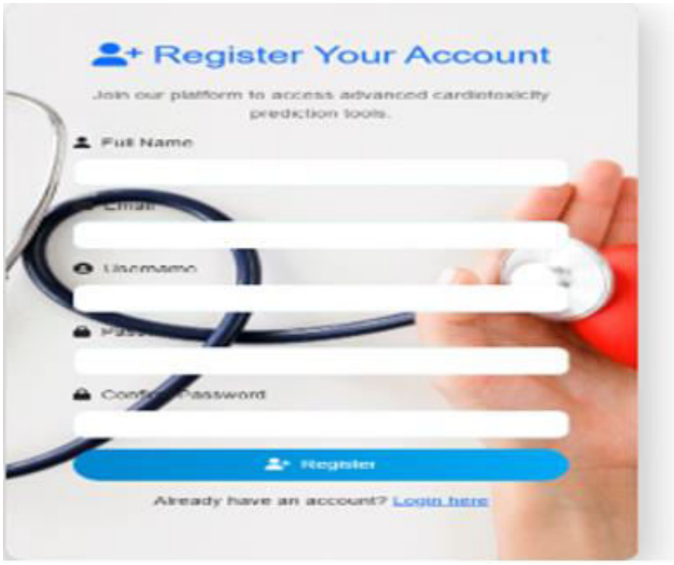
To make the system more accessible, a simple web-based interface is developed. This interface allows users to input TDI images and clinical details, after which the system processes the data and provides a prediction along with a confidence score. Basic error handling is included to manage invalid inputs and ensure smooth operation.The implementation is designed to be both functional and user-friendly. It not only demonstrates the effectiveness of multimodal deep learning but also provides a practical tool that can assist in early detection of heart-related complications during chemotherapy.

VIII RESULTS AND ANALYSIS

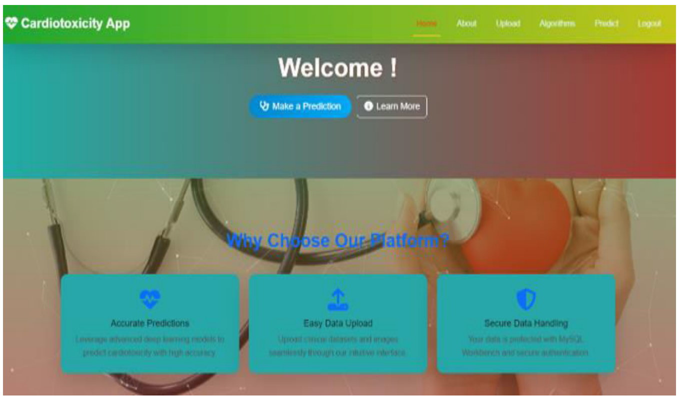
The proposed system was evaluated to understand how effectively it can identify early signs of heart damage caused by chemotherapy. For this purpose, different hybrid deep learning models were implemented and tested using the prepared dataset, which includes both imaging data and clinical features. The evaluation was carried out using commonly accepted performance measures such as accuracy, precision, recall, and F1-score, so that the results could be interpreted in a meaningful way.

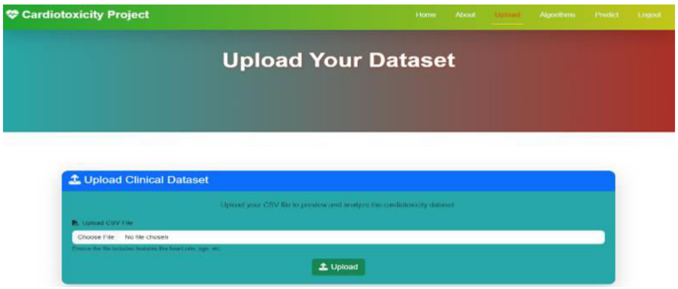


Login Page

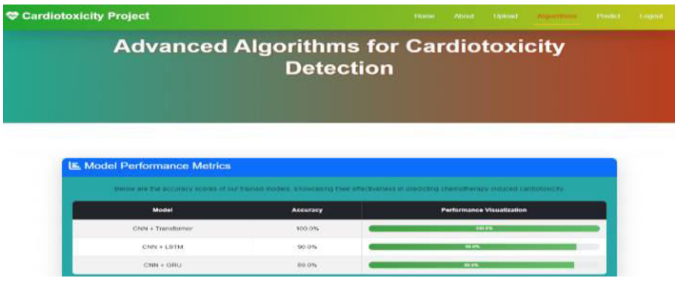


Registration Page

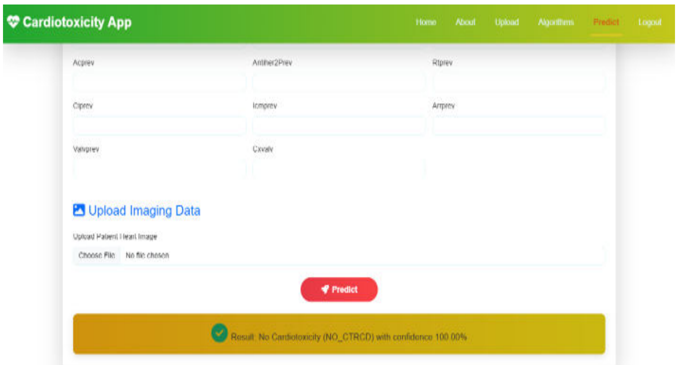




Upload Clinical Dataset Page



Model Performance Metrics Page



Cardiotoxicity Prediction Result Page

Three model combinations were considered during the analysis: CNN with LSTM, CNN with GRU, and CNN with Transformer. All three models were trained using the same dataset and under similar conditions to maintain consistency in comparison. It was observed that each model was able to learn useful patterns from the data, but their performance differed slightly depending on how well they handled temporal information.

Model	Accuracy	Precision	Recall	F1-Score
CNN + LSTM	0.9	0.89	0.91	0.9
CNN + GRU	0.89	0.88	0.9	0.89
CNN + Transformer	1	0.99	1	0.99

Table 1: Comparison of Model Performance

From Table 1, it can be seen that the CNN + Transformer model gives better results compared to the other two models. One possible reason for this improvement is its ability to focus on important parts of the data using attention mechanisms, which helps in understanding complex patterns more effectively. The LSTM and GRU-based models also perform well, but they are slightly less efficient when dealing with long-term dependencies in the data.

Apart from model performance, the dataset itself was also examined before and after preprocessing. Initially, there was a clear imbalance between the two classes, with a much higher number of non-cardiotoxic cases. This kind of imbalance can affect model learning, as it may bias the results toward the majority class. After cleaning the dataset and applying balancing techniques, the distribution became more even, which helped in improving model performance.

Description	Count
Total Records (Original)	531
Records After Cleaning	466
NO_CTRCD (Before Balancing)	420
CTRCD (Before Balancing)	46
NO_CTRCD (After Balancing)	420
CTRCD (After Balancing)	420

Table 2: Dataset Details

The preprocessing steps, including image normalization, resizing, and scaling of clinical features, also contributed to better training behavior. These steps ensured that the model received consistent input, which improved its ability to learn patterns effectively.The results indicate that combining imaging data with clinical information leads to better prediction compared to using a single data source. Among the tested models, the CNN + Transformer combination shows the most promising performance, making it suitable for early detection of chemotherapy-related heart damage.

IX CONCLUSION

Detecting heart-related complications that may arise during chemotherapy, particularly at an early stage when they are difficult to notice. Since such conditions often develop gradually without clear symptoms, there is a need for approaches that can identify subtle changes before they become severe. In this study, an attempt was made to improve early prediction by using both cardiac imaging data and patient clinical information together.

Different model combinations were explored to understand how well they can capture patterns from the available data. It was observed that approaches which consider both spatial and time-related information tend to provide more reliable results. Among the models tested, the combination involving a Transformer-based approach showed better consistency in performance, indicating its ability to handle complex relationships within the data more effectively.

Another important aspect of this work is the role of data preparation. Cleaning the dataset, handling imbalance, and maintaining consistency in inputs contributed to better learning and improved prediction results. The inclusion of both imaging and clinical features also helped in forming a more complete understanding compared to methods that rely on a single type of data.

In summary, the study shows that combining multiple data sources with suitable deep learning models can improve the early detection of chemotherapy-related heart damage. While the current results are encouraging, further work is needed to test the approach on larger datasets and in real clinical environments. With continued improvements, such systems can become useful tools for supporting doctors in monitoring patients and making timely treatment decisions.

REFERENCES

[1] S. N. Orting, J. Petersen, L. H. Thomsen, M. M. W. Wille, and M. de Bruijne, "Learning to quantify emphysema extent: What labels do we need?," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 4, pp. 1149–1159, Apr. 2020.

[2] A. Alzahrani, "Early detection of lung cancer using predictive modeling incorporating CTGAN features and tree-based learning," *IEEE Access*, 2025.

[3] T. A. Khan *et al.*, "Secure IoMT for disease prediction empowered with transfer learning in Healthcare 5.0: The

concept and case study," *IEEE Access*, vol. 11, pp. 39418–39430, 2023.

[4] M. T. Bahadori, Y. Liu, and D. Heckerman, "Temporal modeling of health data for early detection of heart failure," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2016.

[5] Y. Wang, Y. Zhang, and L. Liu, "Predicting adverse drug reactions using recurrent neural networks with external knowledge," in *Proc. ACM Int. Conf. Bioinformatics, Computational Biology and Health Informatics (BCB)*, 2018, pp. 29–38.

[6] S. T. Goh, A. Lee, and J. T. Teo, "A transformer-based model for cardiotoxicity prediction using electrocardiogram data," *Computers in Biology and Medicine*, vol. 152, 2023.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[8] A. Dosovitskiy *et al.*, "An image is worth 16x16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2021.

[9] H. Wang *et al.*, "Early detection of cardiotoxicity using deep learning on longitudinal echocardiographic imaging," *JACC: Cardiovascular Imaging*, vol. 12, no. 10, pp. 1935–1946, 2019.