Volume 4 Issue VII, August 2025, No. 70, pp. 910-926
Submitted 22/7/2025; Final peer review 15/8/2025
Online Publication 17/8/2025
Available Online at http://www.ijortacs.com

# TRANSFER LEARNING MULTI-MODAL APPROACH FOR CARDIOVASCULAR DISEASE PROGNOSIS

[1]Kamalu Maureenwendy C., [2]Okeke Ogochukwu C., [3]Aghaizu Herman Chijioke
[1,2]Department of Computer Science, Chukwuemeka Odumegwu Ojukwu,
[3]Department of Computer Engineering, Enugu State University of Science and Technology
Authors Email:Kamalumaureen@gmail.com[1],co.okeke@coou.edu.ng[2],
chijiokeaghaizu@gmail.com[3]

**Abstract**

Despite the development of modern medicine being in progress, CVD is still the leading cause of death worldwide (responsible for 32% mortality). Although great achievements have been made by predictive models already, these current available systems seem to be based on single-modality data only that may not calculate a high accuracy when identifying the comprehensive risk in clinical cases. This study fills a crucial unmet need for multi-modal prediction of CVD by creating an integrated model that harmonizes clinical, imaging, genetic, lifestyle and environmental data within the framework of sophisticated transfer learning methods to improve both predictive capabilities in early detection as well as risk stratification accuracy. The research employed a hybrid architecture combining pre-trained ResNet models for medical imaging analysis with BERT transformers to process genetic and textual data, all integrated in conjunction with CNN and fully connected layers. We created and structured a dataset of 100 patient records. The model was trained with 10-fold stratified cross validation technique to treat class imbalance problems efficiently. Built in Python using TensorFlow and PyTorch, the system produced good outcomes whereby it performed better and reached an accuracy that was slightly above those previously obtained with a 93.06% accuracy rate in predicting 10-year cardiovascular disease risk, yielding a Mean Absolute Error (MAE) of 0.0694 These results significantly outperform traditional single-modality methods, demonstrating the value of combining multiple data sets for increased prediction and suggesting exciting prospects for early intervention and personalized patient care.

**Keywords:** Cardiovascular Diseases, Multi-modal Data, Transfer Learning, Deep Learning, Prognosis Prediction.

## 1. INTRODUCTION

Cardiovascular diseases (CVDs) are conditions that involve blocked arteries, heart attacks and strokes. Heart disease is the top world-wide killer, responsible for some 17.9 million deaths each year, 85% of which are due to heart attacks and strokes. Typically associated with lifestyle choices, such as a poor diet, physical inactivity, and smoking; these CVDs progress slowly. It is essential to identify these conditions early and act on them to reduce the death and illness that they cause.

Figure 1 is adapted from a conceptual model of cardiovascular diseases and their risk factors (e.g., including coronary artery disease, stroke, heart failure; risk factors such as hypertension, diabetes, obesity, smoking, and genetics) (Xu &Sathyapalan, 2024).

Advances in analytics and machine-learning techniques have transformed disease detection in healthcare, enabling providers to shift towards prevention from reaction.

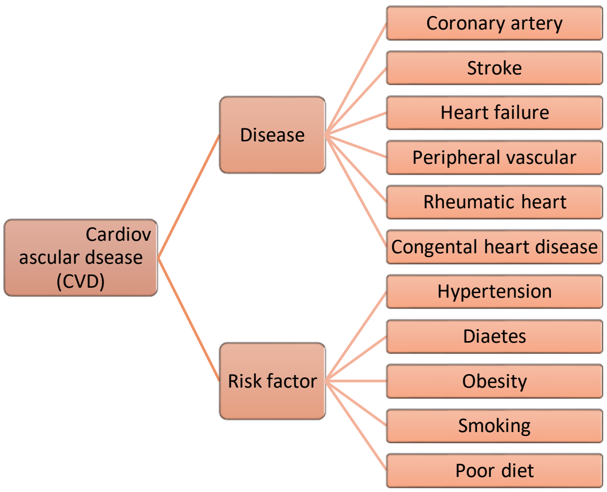


**Figure 1 Overview of Cardiovascular Diseases and Risk Factors (Xu &Sathyapalan, 2024).**

Recent years have witnessed a large increase in the amount and variety of health data available to biomedical researchers; however, traditional models are almost exclusively built on single-modal data (e.g., clinical records or imaging), which may provide an incomplete view but do not exploit all these diverse data sources. While improving diagnostic accuracy for CVDs, this integration plays an equally important role in preventive care by developing predicting models to identify at-risk populations who may benefit from interventions.

Advances in analytics and machine-learning techniques have transformed disease detection in healthcare, enabling providers to shift towards prevention from reaction. Recent years have witnessed a large increase in the amount and variety of health data available to biomedical researchers; however, traditional models are almost exclusively built on single-modal data (e.g., clinical records or imaging), which may provide an incomplete view but do not exploit all these diverse data sources. While improving diagnostic accuracy for CVDs, this integration plays an equally important role in preventive care by developing predicting models to identify at-risk populations who may benefit from interventions.

But as far as employing machine learning in HealthCare is concerned, there are still major hurdles. The major dangers of accessing unauthorized by unauthorized access, data interception and accidental data seize (Okeke & Ezenwegbu, 2024). Challenges with respect to data fragmentation (in institutions, absence of comprehensive datasets) combined with privacy concerns in data (sensitive information which does not pass through firewalls) makes it difficult to develop AI tools that are trustworthy. The advent of Electronic Health Records (EHR) has radically changed the preservation and administration of patient information in healthcare (Okeke &Ezenwegbu, 2024).

Although useful, unstructured data in Electronic Health Records (EHRs) makes analysis more difficult, and data silos further limit access. Transfer learning techniques, which use the knowledge from previously trained models on large datasets to enhance performance on new tasks with little data, are being investigated as a solution to these problems. Biases in training data and the high expense of creating precise forecasting models from scratch are also addressed by this method. Transfer learning has shown promise in improving the robustness, adaptability, and dependability of machine learning systems, especially in the healthcare industry where it can be difficult to obtain sizable, annotated medical datasets.

This paper pioneers a multi-modal predictive solution for cardiovascular diseases utilizing transfer learning, integrating clinical, imaging, and genetic data to improve prediction accuracy and reliability, ultimately supporting early diagnosis and personalized treatment strategies.

## 2. Literature Review

Single-modal analyses have given way to more integrated, multi-modal approaches in the development of predictive models for cardiovascular diseases.

### 2.1 Predictive Models for Cardiovascular Diseases

Several studies have concentrated on creating cardiovascular disease prediction models, frequently using clinical information like vital signs, lab results, and patient history. For example, Chen et al. (2023) showed significant improvements over traditional statistical methods by developing a machine learning model that uses electronic health records to predict the risk of heart failure. Figure 2, titled "Multi-Modal Data Sources for CVD Prognosis," is a conceptual diagram that illustrates how various types of data are collected and consolidated into a central repository for the purpose of cardiovascular disease prognosis.

Yang et al. (2020) used information from 29,930 high-risk individuals to develop a CVD prediction model. Nearly 30 CVD indicators were found using their logistic regression analysis, and Random Forest outperformed the benchmark multivariate regression model (AUC 0.7143) with the highest AUC of 0.787 among the different approaches tested. Although it requires large and balanced datasets for optimal training, this robust model successfully evaluates 3-year CVD risk in a high-risk population in eastern China. Shorewala, V. (2021) used the 'Cardiovascular Disease Dataset' with 70,000 records to assess different coronary heart disease (CHD) prediction methods.

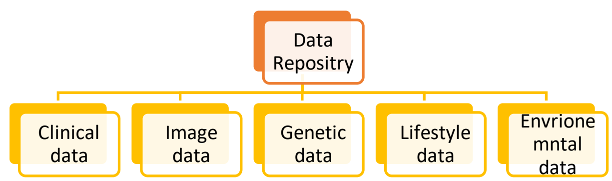


**Figure 2 Multi-Modal Data sources for CVD Prognosis**

K-Nearest Neighbours, Binary Logistic Classification, Naive Bayes, and ensemble techniques (bagging, boosting, and stacking) were among the techniques examined in the study. Results showed that combining several machine learning models produced better outcomes than using just one model. In particular, boosted models attained roughly 73% accuracy in predictions, whereas the bagged approach improved prediction accuracy by almost 2%. The most effective method, which achieved over 75% accuracy, involved stacking three different models: KNN, Random Forest, and SVM.A deep learning system for the automated measurement of calcium accumulation in heart arteries from CT scan images was created by

Zeleznik and his research team in 2021. The automated score, which was tested on 20,084 people from various cohorts, had hazard ratios of up to 4.3 and was a strong predictor of cardiovascular events even in the absence of conventional risk factors. The system's measurements showed good test-retest reliability and were in close agreement with manual quantification. The potential of automated heart disease prediction tools in clinical settings to support public health is highlighted by this study.

A deep learning model for heart disease prediction was also presented by Sharma and Parmar (2020), who emphasised the use of deep neural networks to increase classification accuracy. The study took into account a number of classification techniques, such as Random Forest, KNN, SVM, and Naïve Bayes. Their study proved the effectiveness of Talos Hyper-parameter optimisation using the Heart Disease UCI dataset. One drawback was that successful implementation necessitated a sizable dataset. CardioHelp, a convolutional neural network (CNN)-based technique for predicting the likelihood of cardiovascular disease, was presented by Mehmood et al. (2021). It focuses on temporal data modelling for the prediction of early-stage heart failure. Although the suggested method's application was restricted to a single-modal dataset, it demonstrated an impressive 97% accuracy in early heart disease prediction. Kavitha et al. (2021) addressed the significant mortality rates associated with heart disease by proposing a novel machine learning approach for early prediction. Using the Cleveland heart disease dataset, they employed a data mining technique which involves extraction of hidden predictive information from large databases (Okeke & Okonkwo, 2017).

Gárate-Escamila et al. (2020) proposed a method to enhance cardiac disease prediction using machine learning. Analyzing the heart disease dataset from the UCI Machine Learning Repository with six ML classifiers, they found that combining Chi-square and principal component analysis (CHI-PCA) with random forests achieved the highest accuracy: 98.7% for the Cleveland dataset, 99.0% for the Hungarian dataset, and 99.4% for the combined dataset. A machine learning system for detecting heart disease was created by Chang and associates in 2022. Data processing, categorical variable management, and dataset attribute evaluation were all done by their Python application. Improved diagnostic accuracy was demonstrated by a random forest classifier that obtained 83% accuracy on training data. Dependency on data quality, possible overfitting, and high data processing demands were among the drawbacks.

Ekwonwune et al. (2023) used patient data from Umezuruike Hospital in Owerri to create a data model for identifying cardiovascular diseases using a clustering algorithm. Their study, "Development of a Data Mining Model to Detect Cardiovascular Disease," emphasised the significance of cholesterol levels in the identification of cardiovascular disease. Limitations that were identified included difficulties with data organisation and management, multi-modal data integration, real-time data processing, and guaranteeing data consistency and quality. In evaluating cardiac function, Ouyang et al.

(2020) introduced EchoNet-Dynamic, a deep learning algorithm that outperformed human experts. Trained on echocardiogram videos, it effectively segmented the left ventricle (Dice coefficient 0.92), predicted ejection fraction (mean absolute error 4.1%), and classified heart failure (AUC 0.97). It maintained high accuracy (mean absolute error 6.0%, AUC 0.96) in an external dataset. Its dependence on a single data modality and possible dataset biases were among its drawbacks. To promote more creativity, a dataset of 10,030 annotated echocardiogram videos was made publicly accessible.

### 2.2 Use of Imaging Data

Echocardiograms and MRI scans are two examples of imaging data that offer vital information about the anatomy and physiology of the heart. Research has repeatedly shown that adding imaging data to prediction models can greatly increase accuracy. An example of how ML research has advanced significantly in image analysis using deep learning and convolutional neural networks is the transformer-based architecture developed by Alaa et al. (2022) that fused retinal fundus images, echocardiograms, and electronic health records. This makes Machine learning tools more effective for imaging the heart and blood vessels. Sermesant et al. (2021) discussed various clinical questions that machine learning can address, key AI methodologies, and representative examples of both data-driven and model-integrating approaches. Their work also highlighted challenges like generalizability and explainability, proposing potential solutions.

### 2.3 Genetic Data in CVD Prediction

Despite being a powerful predictor of cardiovascular events, routine quantification of coronary artery calcium is frequently hampered by the time and expertise needed. A deep learning system for automating calcium quantification on CT scans was presented by Sermesant et al. (2021) and tested on 20,084 people. The system showed dependability, correlated well with manual quantification, and made accurate event predictions. Potential biases in the data and the requirement for wider validation were among the limitations. By automating imaging biomarkers, this approach has the potential to greatly improve clinical management and population health. Predispositions to heart problems or other cardiovascular issues can be revealed by genetic information, which is derived from an individual's DNA. Many genetic variations have been linked to the risk of CVD by genome-wide association studies (GWAS). Personalised risk assessment can be enhanced by integrating genetic data into predictive models. For instance, Kullo et al. (2022) showed enhanced predictive performance when they combined clinical data and genetic risk scores to forecast coronary artery disease. Although imaging methods are essential for identifying and treating medical abnormalities, their accessibility and the problem of professional fatigue restrict the efficacy of qualified practitioners. In image understanding, convolutional neural networks (CNNs) usually perform better than humans. With an emphasis on the brain, breast, lung, and other organs, Sarvamangala and Kulkarni (2021) reviewed CNN applications in medical imaging for tasks like classification, segmentation, and detection. In an effort to stimulate more research, they presented CNN frameworks and talked about difficulties. Potential model biases and the requirement for sizable, high-quality datasets were among the limitations.

A thorough medical imaging model that overcomes the drawbacks of existing segmentation techniques was created by Ma et al. in 2024. More than 1.5 million medical photos with matching masks, spanning ten distinct imaging modalities and more than 30 cancer types, were used to train their system. It outperformed modality-specific models in terms of accuracy and robustness when tested on 86 internal and 60 external tasks. Potential biases in training data, the requirement for clinical validation, and the computational demands of large-scale deployment are some of the difficulties. However, MedSAM is a major breakthrough in medical image segmentation for a variety of clinical uses.

**2.4 Multi-Modal Approaches**

It has been demonstrated that accuracy is greatly increased when clinical, imaging, and genetic data are combined into a single predictive model. In order to forecast patient outcomes in a hospital setting, Zhang et al. (2020) and Arnaud et al. (2020) created deep-learning models that combined various data types. Their methods outperformed models that used only one type of data. Large amounts of multimodal big data have been produced as a result of the growth of heterogeneous networks, which presents difficulties for conventional data fusion techniques. Gao et al. (2020) summarised popular architectures and top models while examining innovative deep learning models for multimodal data fusion. Their work also highlighted ongoing challenges and future research directions, aiming to inspire new techniques in multimodal deep learning. Limitations included the need for further advancements to address existing challenges and enhance model performance. Applications including digital clinical trials, digital twins, remote monitoring, pandemic surveillance, personalised medicine, and virtual health assistants were examined by Acosta et al. in 2022.

To optimise the effects of multimodal AI in healthcare, they also talked about important issues in data integration, modelling, and privacy that need to be resolved. In order to improve the diagnosis of Parkinson's disease (PD) through deep learning, Pahuja and Prasad (2022) looked into multi-modal features. They introduced two frameworks: a modal-level framework that reduces MRI features prior to integration, and a feature-level framework that integrates MRI, SPECT, and CSF data. CNN models demonstrated up to 93.33% accuracy when tested on an unbalanced dataset (73 PD, 59 healthy). Multi-modal features improved the accuracy of PD classification despite their increased complexity. Validating these techniques in actual medical settings and ensuring data reliability were the biggest obstacles.

The advanced machine learning system that Ehtisham Khan Jadoon and his research team created in 2023 uses a variety of patient data, such as DNA variations, medical history, and genetic activity patterns, to better predict the course of breast cancer. The model builds a stacked feature set, employs CNNs and DNNs to extract features, and then makes predictions using a random forest. This method performs better than homogeneous and unimodal models. High computational demands, complicated multi-modal integration, and the requirement for additional validation are some of the limitations. Using machine learning to integrate electrocardiogram (ECG) and phonocardiogram (PCG) data is essential for the detection of cardiovascular disease (CVD). A novel approach that combines heart sound information from PCG recordings with heart rhythm data from ECGs was presented by Li et al. in 2021.Deep-coding features are extracted via neural networks for both modalities, and a genetic algorithm optimizes feature selection. Support vector machine classification achieved superior performance over single-modal methods, with an AUC of 0.936 demonstrating efficacy in CVD prediction and diagnosis. However, challenges include the need for larger and more diverse datasets for robust model generalization across different patient demographics and clinical settings.

Cardiovascular disease (CVD) remains a leading global cause of death, challenging early detection due to its complex nature involving genetic, environmental, and lifestyle factors. Through a specialised transfer learning system, Prakash et al. (2024) created a novel method that combines various forms of medical information, such as patient records, medical images, and genetic information, using attention mechanisms. To improve the model's ability to identify relationships between data, ABCM makes use of attention mechanisms. Extensively validated, ABCM outperforms conventional and other cutting-edge techniques with 93.5% accuracy, 92.0% precision, 94.5% recall, and an AUC of 97.2%. Data heterogeneity and the requirement for strong validation across various patient cohorts present difficulties, though.

**2.5 Transfer Learning Applications in Healthcare**

In order to address data scarcity and enhance model performance, transfer learning has been used more and more in the healthcare industry. For instance, Rocheteau& Kim (2020) used transfer learning in dermatology to develop models that could diagnose skin cancer as accurately as dermatologists. In order to help overworked medical professionals, Horry et al. (2020) used transfer learning from deep learning models to detect COVID-19 from X-ray, ultrasound, and CT scan images. The VGG19 model was found to be the most successful for these particular data types after a number of different strategies were tested. Small dataset sizes and inconsistent quality were obstacles that affected the trainability of the model. To lower noise and enhance dataset quality, a stage for pre-processing images was suggested. Results indicated that the VGG19 model achieved detection accuracy rates of 86% for X-rays, 100% for ultrasound, and 84% for CT scans. Limitations included dataset size, quality, and the model not being tested for cardiovascular disease prediction. Chouhan et al. (2020) built on transfer learning knowledge to create a new deep-learning framework system that makes it easier to detect pneumonia in medical images. The image features were extracted using pre-trained neural network models from ImageNet and fed into a classifier for prediction. After testing five models, an ensemble model that combined the outputs of all the pre-trained models performed at a state-of-the-art level, achieving 99.62% recall and 96.4% accuracy on unseen data from the Guangzhou Women and Children's Medical Centre dataset. Nevertheless, this framework only used single-modal image data and was not used for CVD prediction.

A multi-source adversarial transfer learning framework was presented by De Bois et al. (2021) to deal with the lack of data in deep learning applications for healthcare. This framework improves generalisation by facilitating the transfer of feature representations across various data sources. When used for glucose forecasting in diabetic patients, it outperformed state-of-the-art techniques, particularly when dealing with limited or heterogeneous data, and showed improved statistical and clinical accuracies across three datasets. By learning general feature representations free from patient or dataset bias, the framework improves the effectiveness of data sharing between health actors. Potential biases in training data and the requirement for additional clinical validation, particularly for cardiovascular diseases, were among the limitations.

Pathak et al. (2022) investigated the use of chest CT images for COVID-19 diagnosis in order to address the sharp rise in COVID-19 infections and the scarcity of testing kits. They used a deep transfer learning technique for classification because infected patients' chest CT images showed bilateral changes. To deal with noisy and unbalanced datasets, they included a top-2 smooth loss function with cost-sensitive features. This deep transfer learning-based method performs better than supervised learning models, according to experimental results. The difficulty of forecasting bilateral changes and the requirement for additional validation in various clinical contexts, such as cardiovascular disorders, are obstacles, though. Significant memory loss is a symptom of Alzheimer's disease, which is common in people over 65. Its progression can be controlled with early detection. Ghazal et al. (2022) created a technique to classify Alzheimer's disease into four different levels: patients with mild symptoms, those with very slight cognitive decline, those without dementia, and patients exhibiting moderate impairment. They did this by applying deep learning and transfer learning on MRI images. It outperformed earlier techniques with an accuracy of 91.70%, but more clinical validation is required.

Deep learning and medical imaging-based computer-aided diagnosis systems are becoming more and more common. Spatial information is often overlooked by standard CNN models, which impairs their overall performance. The residual structure prevents overfitting and degradation, dilated convolution maintains spatial information, and transfer learning manages limited data by initialising the model with parameters from large-scale datasets. Liang and Zheng (2020) addressed this by proposing a deep learning framework that combines residual learning, dilated convolution, and transfer learning for the diagnosis of paediatric pneumonia. Their method outperformed previous methods and reliably classifies paediatric pneumonia even with low-resolution and partially occluded images, earning them a 96.7% recall and 92.7% F1-score.

## 3. Research Methodology

The shortcomings of conventional single-modal predictive models for CVDs are addressed by the suggested system. These models frequently have poor generalisation across a variety of populations, limited data integration, and a lack of large datasets for training. Accurate prediction is made more difficult by the fact that CVDs are multifactorial, involving lifestyle, environmental, and genetic factors.

### 3.1 Data Collection

A comprehensive dataset comprising 100 patient records was systematically compiled. This dataset integrates diverse modalities:

A. **Clinical Data:** Patient demographics, medical history, vital signs (e.g., systolic and diastolic blood pressure), cholesterol levels.
B. **Imaging Data:** Medical images (e.g., echocardiograms, CT scans) providing visual insights into cardiovascular health.
C. **Genetic Data:** Genetic markers and sequences relevant to CVD predisposition.
D. **Lifestyle Data:** Information on smoking status, dietary habits, physical activity.
E. **Environmental Data:** Factors such as air quality index.

Rigorous pre-processing was applied to ensure data quality and uniformity, including data cleaning, normalization, and handling of missing values.

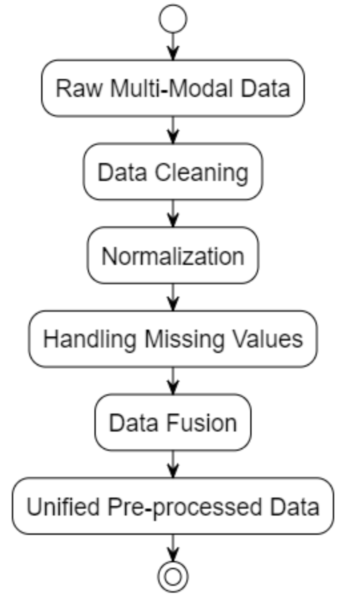


**Figure 3 Data Pre-processing Pipeline (Zhang et al., 2024)**

Figure 3 is Adapted from the standard multimodal preprocessing pipelines outlined by Zhang et al (2024) illustrates the diverse data modalities collected for the study: Clinical Data, Imaging Data, Genetic Data, Lifestyle Data, and Environmental Data.

### 3.2 Feature Extraction with Transfer Learning

Transfer learning techniques are critical for extracting meaningful features from the multi-modal data.

1. **Imaging Data:** Pre-trained Convolutional Neural Networks (CNNs), specifically ResNet models, are utilized to extract high-level features from medical images. ResNet's architecture is effective in mitigating overfitting and degradation, particularly useful for medical imaging where data may be limited.

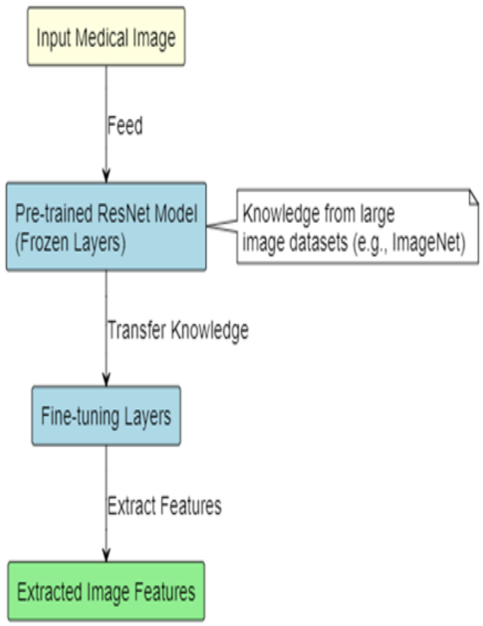


**Figure 4 Transfer Learning for Imaging Data (ResNet) (Kim et al., 2022)**

Figure 4 diagram illustrating how a pre-trained ResNet model is used for feature extraction from medical images **Genetic and Textual Data:** Transformer-based architectures that have already been trained, like BERT, are refined to extract significant patterns and characteristics from textual and genetic data. After each modality's features are extracted, they are carefully examined and aligned in order to be combined into a single predictive model. This improves accuracy and efficiency by guaranteeing that domain-specific knowledge from previously trained models is successfully transferred

The figure above shows the application of a pre-trained BERT transformer for extracting features from genetic sequences or textual clinical notes. Illustrate input genetic/textual data, the BERT model, and the extracted semantic/pattern features.

### 3.3 Model Architecture and Training

Utilising new hybrid architecture, the feature extractors are integrated into a single CNN and fully connected layers framework. In this advanced hybrid framework, fully connected layers optimised for numerical and categorical data processing are combined with CNNs for image feature extraction. In comparison to traditional single-modality systems, this multi-layered approach improves predictive accuracy and permits thorough feature integration. Subsets of the processed dataset are used for testing, validation, and training. In order to avoid overfitting and guarantee generalisation to unobserved data, the model is trained using 10-fold stratified cross-validation, which specifically addresses class imbalance issues. The validation set is used for hyperparameter tuning, and the test set is saved for the last performance assessment..

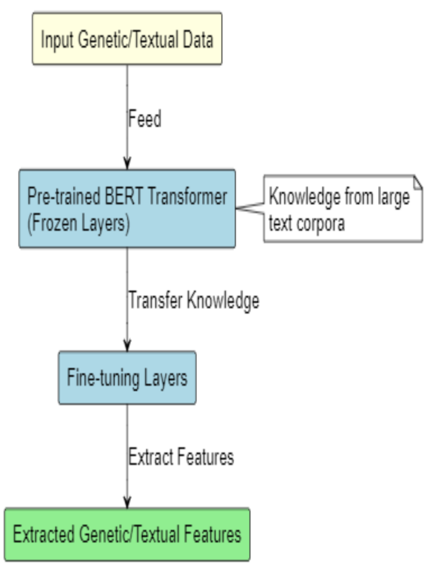


**Figure 5: Transfer Learning for Genetic/Textual Data (Rogers et al., 2020)**

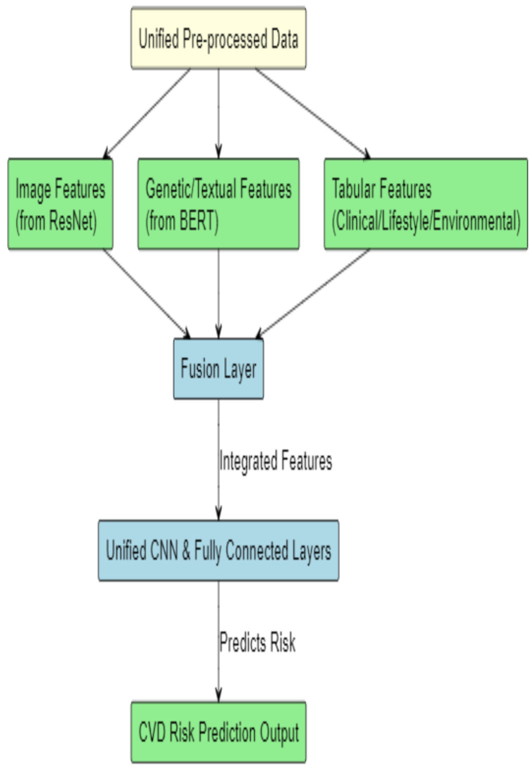


**Figure 6 Hybrid Multi-Modal Deep Learning Architecture (Lee et al., 2023)**

Figure 6 is a core architectural diagram. It shows the parallel processing of features from different modalities (e.g., Image Features from ResNet, Genetic/Textual Features from BERT, Tabular Features from clinical data). These individual feature sets then feed into a "Fusion Layer," followed by a unified Convolutional Neural Network (CNN) and fully connected layers, culminating in the CVD risk prediction output(Lee et al., 2023).

3.4 Model Evaluation

The model's performance is rigorously evaluated using standard metrics, including accuracy, precision, recall, F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC-ROC). A comparative analysis against single-modal models demonstrates the superior predictive capability of the multi-modal approach.

## 4. SYSTEM DESIGN

The system is implemented as a user-friendly application primarily for healthcare professionals at Vaden Specialist Hospital, Owerri, using Python.

### 4.1 System Specifications

1. **Hardware Requirements:** Intel i7 processor or equivalent, 16 GB RAM or higher, 500 GB SSD, NVIDIA GTX 1080 GPU or higher for deep learning tasks.
2. **Software Requirements:** Windows 10/11 or Linux operating system, Python with libraries such as TensorFlow, PyTorch, Scikit-learn, Pandas, NumPy, and a database like PostgreSQL or MongoDB.
3. **Programming Language:** Python was selected because of its vast ecosystem of data processing and machine learning libraries, which make intricate data manipulation and model building easier.
4. **Integrated Development Environment (IDE):**Visual Studio Code (VSCode) was used because of its powerful debugging and version control capabilities, interactive development support, and lightweight yet extensible environment.

### 4.2 System Modules

Several essential modules make up the system:

1. **Data Integration Module**: In charge of combining and pre-processing various data formats into a single dataset.

2. **Feature Extraction Module**: Uses transfer learning techniques (BERT for genetic/textual data, ResNet for images) to extract meaningful features from genetic and imaging data.
3. **Risk Prediction Module**: Provides confidence intervals for predictions by combining extracted features and using machine learning algorithms to determine cardiovascular disease risk scores.
4. **User Interface Module:** Offers medical practitioners a user-friendly interface for entering patient data, visualising outcomes, and displaying prediction details and reports.

Figure 7 showcases the detailed patient data input sections within the **User Interface Module** named CARDIO APP. It includes input forms for clinical, imaging, genetic, lifestyle, and environmental data.

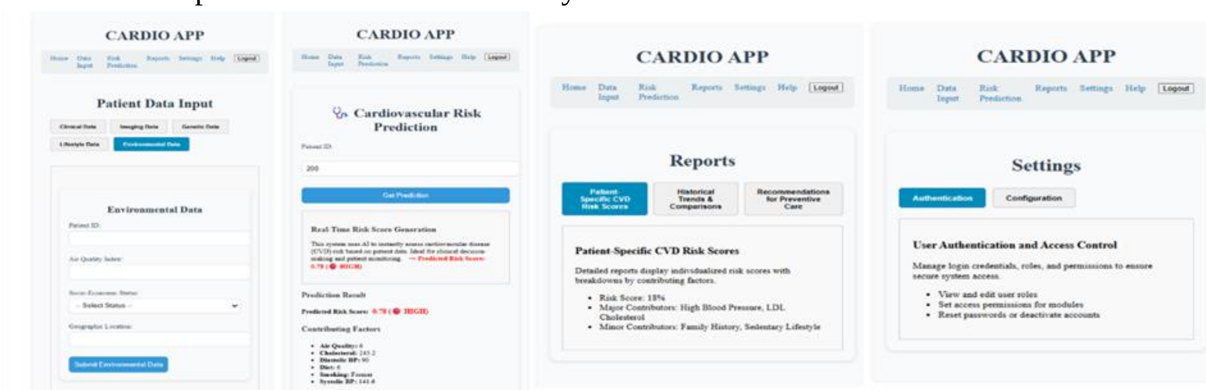


**Figure 7: Patient Data Input Modules User Interface**

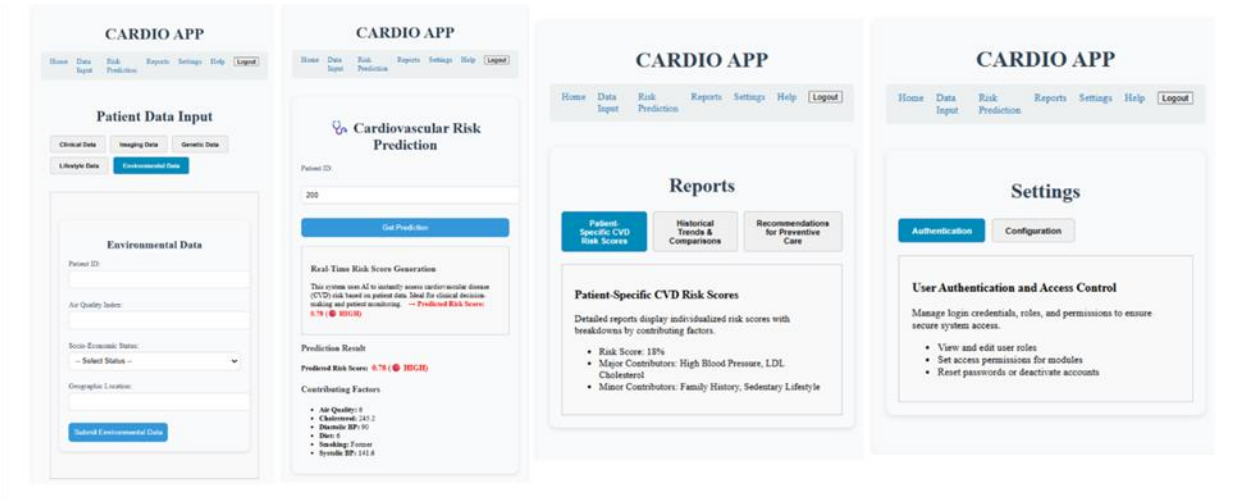


**Figure 8 Mukti-Modal Data Collection and Prediction Interface**

Figure 8 illustrates the integrated workflow of the CARDIO APP, designed for cardiovascular disease (CVD) prognosis using a transfer learning approach. The interface includes modules for environmental data input, real-time cardiovascular risk prediction, detailed reporting, and secure user settings.

### 4.3 Security Protocols

To guarantee data privacy and adherence to healthcare regulations, advanced security protocols are integrated, such as encryption and secure login. Using werkzeug.security for secure hashing, password protection enforces stringent password policies. To protect health information from unwanted access, AES-256 encryption is used on all sensitive data while it is being transmitted (Data in Transit) and stored (Data at Rest).

## 5. Results and Discussions

The outcomes of the multi-modal predictive model for assessing the risk of cardiovascular disease are shown in this section. The model's efficacy and dependability were determined through a thorough evaluation utilising a variety of metrics.

### 5.1 Performance Evaluation

The model performed exceptionally well, predicting the risk of cardiovascular disease over a 10-year period with 93.06% accuracy and a Mean Absolute Error (MAE) of 0.0694. A thorough 10-fold stratified cross-validation was used to achieve these results, guaranteeing the model's capacity for generalisation and resilience to class imbalance. Compared to the conventional single-modality approach, the model's accuracy in predicting risk factors and outcomes represents a major improvement.

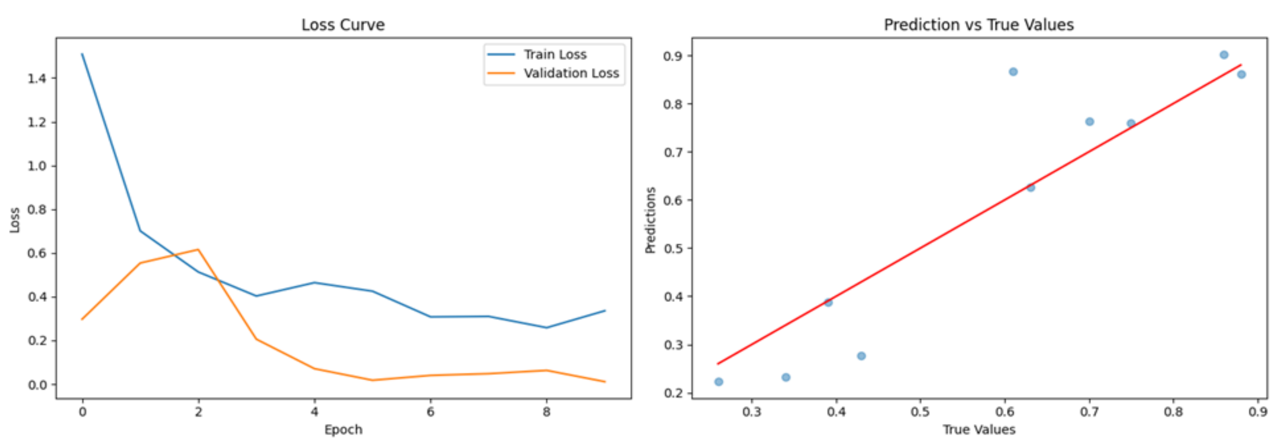


**Figure 9: Model Performance: Loss Curve and Prediction vs True Values**

Two important facets of the model's functionality are shown in Figure 9. The model's learning progress and possible overfitting or underfitting are indicated by the "Loss Curve" on the left, which shows the training and validation loss over epochs. The model's predictions and the actual true values are visually compared in the "Prediction vs. True Values" scatter plot on the right, where a regression line shows how well the fit was linear.

### 5.2 Comparative Analysis

A comparative analysis with existing single-modal models in Table 1 below highlighted the substantial benefits of integrating diverse data sources. The multi-modal approach consistently outperformed models trained on individual data types, demonstrating its capacity to capture the complex interdependencies among various health indicators. This integration provides a more holistic view of patient health, leading to more precise and actionable risk assessments.

**Table 1: Comparative Analysis of Proposed Model vs. Other Approaches**

| **Author(s)** | **Data Modality(s)** | **Key Technique(s)** | **Performance Metric(s)** |
|---|---|---|---|
| Shorewala, (2021) | Clinical Dataset (70,000 records) | KNN, Naive Bayes, SVM, RF, Stacking | Stacking accuracy: ~75% |
| Chang et al. (2022) | Clinical Data | RF, App Framework | Accuracy: 83% |
| Li et al. (2021) | ECG + PCG | Deep Coding + SVM + GA | AUC: 0.936 |
| Prakash et al. (2024) | Clinical + Imaging + Genetic | Attention + Transfer Learning | Accuracy: 93.5%, |
| Horry et al. (2020) | X-ray, Ultrasound, CT | Transfer Learning (VGG19) | Accuracy: 86% |
| Chouhan et al. (2020) | X-ray Images | TL + Ensemble CNNs | Accuracy: 96.4% |
| Liang & Zheng (2019) | Chest Imaging | TL + Residual + Dilated Conv | Recall: 96.7%, |
| New System | Clinical, imaging. environmental, lifestyle and genetic | Transfer learning | Accuracy: 93.06% |
| Ghazal et al. (2022) | MRI Imaging | Deep Learning + TL | Accuracy: 91.7% |

### 6. Conclusion and Recommendations

Using cutting-edge transfer learning techniques, this study effectively created and put into use a comprehensive multi-modal predictive model for cardiovascular disease risk assessment. By developing an integrated system that can process various data modalities, such as clinical parameters, medical imaging, genetic markers, lifestyle factors, and environmental variables, to produce accurate and actionable CVD risk scores, the research successfully addressed significant gaps in conventional cardiovascular risk prediction.

Predictive accuracy was greatly increased by the novel method, which used pre-trained deep learning architectures such as ResNet for medical image analysis and BERT for textual and tabular modalities in conjunction with a hybrid Convolutional Neural Network and fully connected layers framework. With an accuracy of 93.06%, the model's proven superior performance emphasises the significant advantages of combining various data sources for a comprehensive understanding of patient health.

## REFERENCE


Acosta, J. N., Falcone, G. J., Rajpurkar, P., & Topol, E. J. (2022). Multimodal biomedical AI. Nature Medicine, 28(9), 1773–1784. https://doi.org/10.1038/s41591-022-01981-2

Alaa, A., van der Schaar, M., & Weller, A. (2022). *Multimodal learning for cardiovascular risk prediction using retinal images, echocardiograms, and clinical data.* Nature Digital Medicine, 5(1), 1-13. https://doi.org/10.1038/s41746-022-00629-2

Chang, V., Bhavani, V. R., Xu, A. Q., & Hossain, M. (2022a and 2022b). An artificial intelligence model for heart disease detection using machine learning algorithms. *Healthcare Analytics*, *2*, 100016. https://doi.org/10.1016/j.health.2022.100016

Chouhan, V., Singh, S. K., Khamparia, A., Gupta, D., Tiwari, P., Moreira, C. P., Damasevicius, R., & De Albuquerque, V. H. C. (2020). A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images. *Applied Sciences*, *10*(2), 559. https://doi.org/10.3390/app10020559

De Bois, M., Yacoubi, M. a. E., & Ammi, M. (2021). Adversarial multi-source transfer learning in healthcare: Application to glucose prediction for diabetic people. *Computer Methods and Programs in Biomedicine*, *199*, 105874. https://doi.org/10.1016/j.cmpb.2020.105874

Ekwonwune, E. N., Ubochi, C. I., Duruoha, A. E., & Duru, N. A. (2023). Development of a Data Mining Model to Detect Cardiovascular Disease. *International Journal of Communications Network and System Sciences*, *16*(05), 77–96. https://doi.org/10.4236/ijcns.2023.165006

Gao, J., Li, P., Chen, Z., & Zhang, J. (2020). A Survey on Deep Learning for Multimodal Data Fusion. *Neural Computation*, *32*(5), 829–864. https://doi.org/10.1162/neco_a_01273

Gárate-Escamila, A. K., Hassani, A. H. E., &Andrès, E. (2020). Classification models for heart disease prediction using feature selection and PCA. *Informatics in Medicine Unlocked*, *19*, 100330. https://doi.org/10.1016/j.imu.2020.100330

Ghazal, T. M., Abbas, S., Munir, S., Khan, M. A., Ahmad, M., Issa, G. F., Zahra, S. B., Khan, M. A., & Hasan, M. K. (2022). Alzheimer Disease Detection Empowered with Transfer Learning. *Computers, Materials & Continua/Computers, Materials & Continua (Print)*, *70*(3), 5005–5019. https://doi.org/10.32604/cmc.2022.020866

Horry, M. J., Chakraborty, S., Paul, M., Ulhaq, A., Pradhan, B., Saha, M., & Shukla, N. (2020). COVID-19 Detection Through Transfer Learning Using Multimodal Imaging Data. *IEEE Access*, *8*, 149808–149824. https://doi.org/10.1109/access.2020.3016780

Jadoon, Ehtisham & Khan, Fiaz & Shah, Sajid & Khan, Ahmad &Elaffendi, Mohammed. (2023). Deep Learning-Based Multi-Modal Ensemble Classification Approach for Human Breast Cancer Prognosis. IEEE Access. PP. 1-1. 10.1109/ACCESS.2023.3304242.

Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021). *Heart Disease Prediction using Hybrid machine Learning Model*. https://doi.org/10.1109/icict50816.2021.9358597

Kim, H. E., Cosa-Linan, A., Santhanam, N., Jannesari, M., Maros, M. E., &Ganslandt, T. (2022). Transfer learning for medical image classification: a literature review. *BMC Medical Imaging*, *22*(1), 1-13.

Kullo, I. J., Lewis, C. M., Inouye, M., Martin, A. R., Ripatti, S., & Chatterjee, N. (2022). Polygenic scores in biomedical research. *Nature Reviews. Genetics*, *23*(9), 524–532. https://doi.org/10.1038/s41576-022-00470-z

Lee, Y. C., Cha, J., Shim, I., Park, W.–Y., Kang, S. W., Lim, D. H., & Won, H.–H. (2023). *Multimodal deep learning of fundus abnormalities and traditional risk factors for cardiovascular risk prediction.*npj Digital Medicine, 6, Article 14.

Li, P., Hu, Y., & Liu, Z. P. (2021). Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods. *Biomedical Signal Processing and Control*, *66*, 102474. https://doi.org/10.1016/j.bspc.2021.102474

Liang, Gaobo& Zheng, Lixin. (2019). A transfer learning method with deep residual network for pediatric pneumonia diagnosis. Computer Methods and Programs in Biomedicine. 187. 10.1016/j.cmpb.2019.06.023.

Ma, J., He, Y., Li, F., Han, L., You, C., & Wang, B. (2024). Segment anything in medical images. *Nature Communications*, *15*(1). https://doi.org/10.1038/s41467-024-44824-z

Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T., & Masood, M. (2021). Prediction of Heart Disease Using Deep Convolutional Neural Networks. *Arabian Journal for Science and Engineering*, *46*(4), 3409–3422. https://doi.org/10.1007/s13369-020-05105-1

Okeke, O. C &Ezenwegbu, N.C (2024). Enhancing Healthcare Data Entry Efficiency and Accuracy through Voice Assistant Systems: A Case Study of Obijackson Hospital Okija. International Journal of Research and Innovation in Applied Science (IJRIAS). Vol IX, Issue X, Pg 277-286.

Okeke, O. C &Ezenwegbu, N.C (2024). Security and Privacy Considerations in Voice Assistant Data Entry Systems for Healthcare Providers. International Journal of Computer Applications Technology and Research. Vol. 13, Issue 11, Pg 18-24.

Okeke, O. C & Okonkwo, E.N (2017). Design and Development of a Data-mining Model for Census Analysis with Geographical Information System (GIS). International Journal of Emerging Technologies and Innovative Research. Vol. 4, Issue 11, pg. 388-397.

Ouyang, D., He, B., Ghorbani, A., Yuan, N., Ebinger, J., Langlotz, C. P., Heidenreich, P. A., Harrington, R. A., Liang, D. H., Ashley, E. A., & Zou, J. Y. (2020). Video-based AI for beat-to-beat assessment of cardiac function. *Nature*, *580*(7802), 252–256. https://doi.org/10.1038/s41586-020-2145-8

Pahuja, G., & Prasad, B. (2022). Deep learning architectures for Parkinson's disease detection by using multi-modal features. *Computers in Biology and Medicine*, *146*, 105610. https://doi.org/10.1016/j.compbiomed.2022.105610

Pathak, Y., Shukla, P., Tiwari, A., Stalin, S., Singh, S., & Shukla, P. (2022). Deep Transfer Learning Based Classification Model for COVID-19 Disease. *IRBM*, *43*(2), 87–92.

https://doi.org/10.1016/j.irbm.2020.05.003

Prakash, J., S, A. a. V., P, G. K., & NK, K. (2024). A novel attention-based cross-modal transfer learning framework for predicting cardiovascular disease. *Computers in Biology and Medicine*, *170*, 107977. https://doi.org/10.1016/j.compbiomed.2024.107977

Rocheteau, E., & Kim, D. (2020). Deep Transfer Learning for Automated Diagnosis of Skin Lesions from Photographs. *ArXiv*, *abs/2011.04475*.

Rogers, A., Kovaleva, O., & Rumshisky, A. (2020). A primer in BERTology: What we know about how BERT works. *Transactions of the Association for Computational Linguistics*, *8*, 451-471.

Sermesant, M., Delingette, H., Cochet, H., Jaïs, P., & Ayache, N. (2021b). Applications of artificial intelligence in cardiovascular imaging. *Nature Reviews. Cardiology*, *18*(8), 600–609. https://doi.org/10.1038/s41569-021-00527-2

Sharma, S., & Parmar, M. (2020). Heart Diseases Prediction using Deep Learning Neural Network Model. *International Journal of Innovative Technology and Exploring Engineering*, *9*(3), 2244–2248. https://doi.org/10.35940/ijitee.c9009.019320

Shorewala, V. (2021). Early detection of coronary heart disease using ensemble techniques. *Informatics in Medicine Unlocked*, *26*, 100655. https://doi.org/10.1016/j.imu.2021.100655

Xu, S., &Sathyapalan, T. (2024). Molecular aspects of cardiovascular risk factors. *Biomolecules*, *14*(8), 1032. https://doi.org/10.3390/biom14081032

Yang, L., Wu, H., Jin, X. *et al.* Study of cardiovascular disease prediction model based on random forest in eastern China. *Sci Rep* **10**, 5245 (2020). https://doi.org/10.1038/s41598-020-62133-5

Zeleznik, R., Foldyna, B., Eslami, P., Weiss, J., Alexander, I., Taron, J., Parmar, C., Alvi, R. M., Banerji, D., Uno, M., Kikuchi, Y., Karady, J., Zhang, L., Scholtz, J. E., Mayrhofer, T., Lyass, A., Mahoney, T. F., Massaro, J. M., Vasan, R. S., . . . Aerts, H. J. (2021). Deep convolutional neural networks to predict cardiovascular risk from computed tomography. *Nature Communications*, *12*(1). https://doi.org/10.1038/s41467-021-20966-2

Zhang, J., & Verma, V. (2020). Discover Discriminatory Bias in High Accuracy Models Embedded in Machine Learning Algorithms. *Advances in Natural Computation, Fuzzy Systems and Knowledge Discovery*.

Zhang, Q., Wei, Y., Han, Z., Fu, H., Peng, X., Deng, C., Hu, Q., Xu, C., Wen, J., Hu, D., & Zhang, C. (2024, April 27). Multimodal fusion on Low-quality Data: A Comprehensive survey. arXiv.org. https://arxiv.org/abs/2404.18947