



Retinal Image-Based Prediction of Chronic Kidney Disease Using a Hybrid Multimodal Deep Learning Framework

Kalyani Chapa^{1,2*}, Bhramaramba Ravi¹

¹ Department of Computer Science and Engineering, GITAM (Deemed to be University), Visakhapatnam 530045, India

² Department of Computer Science and Engineering, Anil Neerukonda Institute of Technology and Sciences (ANITS), Visakhapatnam 531162, India

Corresponding Author Email: kchapa@gitam.in

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

<https://doi.org/10.18280/mmep.130310>

ABSTRACT

Received: 12 November 2025

Revised: 20 January 2026

Accepted: 2 February 2026

Available online: 10 April 2026

Keywords:

chronic kidney disease, retinal imaging, deep learning, multimodal learning, Convolutional Neural Network, transfer learning, attention mechanism, early screening

Chronic kidney disease (CKD) is a major global health burden among many severe diseases. Many existing and conventional approaches suffer from poor generalization in predicting microvascular features and often show limited accuracy. The prediction task becomes more complex in the presence of diabetes. Retinal imaging provides a non-invasive alternative method to assess changes in retinal dysfunction. To address these challenges, a hybrid architectural model is required that consists of a Convolutional Neural Network (CNN), transfer learning based on ResNet50, and an attention-based mechanism. In this model, CNN is used for the extraction of spatial features, transfer learning is used for the representation of retinal biomarkers, and a two-level attention mechanism is employed to enhance the interpretability by highlighting critical morphological and vascular patterns of CKD risk. The novelty of the proposed framework lies in the integration of transfer learning, attention, and CNN. The use of data preprocessing reduces variability and mitigates class imbalance in retinal images and supports multimodal fusion. The hybrid architecture captures local abnormalities and global context by reducing false positives. This hybrid framework offers a non-invasive screening tool for the early prediction of CKD. For improved detection, the study processes publicly available source on Kaggle EyePACS is considered, which provides high-resolution retinal fundus images.

1. INTRODUCTION

There are severe diseases in the world that cause deaths due to a lack of methodologies and technologies to detect them at an early stage. Prevention of further progression of the disease is possible by early identification. The deadliest disease noticed is kidney disease, which is not detected until severe damage occurs in the retinal fundus. The critical health concern lies in its uncontrolled progression, and late detection might cause severe implications. Processing retinal images with respect to diabetes is a challenge. The retinal images involve many attributes, such as simple attributes by standard tests like estimated glomerular filtration rate (eGFR), serum creatinine, and albuminuria, which allows limited insights at the early stage but emerging vascular alterations like vessel calibre changes, hemorrhages, microaneurysms, tortuosity, etc. are considered noninvasive biomarkers of vascular dysfunction to detect chronic kidney disease (CKD) risks.

Many traditional approaches, such as clinical support methods, imaging-based models, and machine learning methods, lack of detecting subtle retinal biomarkers, lack accuracy assurance, resulting in high false positives. For patients with diabetes, the implications are complex. The performance is degraded if patient databases are large and complex.

The challenge identified in detecting kidney disease is a lack of scalability in managing large and diverse database types. Most methods do not process multimodal data, such as Electronic Health Records (EHRs), images, and clinical data simultaneously, and suffer from high false positive rates, which reduces accuracy. The main issue found is the absence of an effective, interpretable deep learning model that analyzes retinal fundus images by capturing both global retinal context and local vascular abnormalities.

Hence, a hybrid model is demanded that consists of Convolutional Neural Network (CNN), the attention mechanism, and transfer learning, but ordering is applied as transfer learning, Attention, and a custom CNN for effective CKD prediction. Among these, transfer learning uses a pre-trained model, ResNet50, that reduces training time, provides robust feature representation, and increases performance. The attention mechanism identifies critical regions, such as vascular abnormalities, that enhance interpretability through attention maps. Then, a custom CNN is applied for specific region spatial feature extraction. Then, multimodal fusion works by integrating significant features from these three models to enhance prediction with better accuracy. This integration supports cross-modal data and ensures robustness across diverse datasets.

Table 1. Methods involved in chronic kidney disease (CKD) prediction

Methods	Purposes	Key Drawbacks
Logistic regression	Applicable to linear relationships and predicts whether a person has kidney disease	Not suitable for complex, nonlinear types; cannot process spatial features
Random Forest	The best decision trees in a set of trees are chosen for disease classification	Not suitable for raw, unprocessed data; weak in scalability for complex databases.
Support Vector Machine (SVM) (RBF Kernel)	A kernel is used to classify nonlinear databases	Costly for big databases; native feature extraction is unsupported
Basic Convolutional Neural Network (CNN)	Suitable for working on spatial features, especially retinopathy lesions	Requires large labeled datasets; may over-fit without regularization and augmentation.
Transfer learning (ResNet/VGG)	Pre-trained CNNs are used for medical image tasks	Involves complex operations for small datasets; it needs fine-tuning for domain-specific retinal biomarkers.
Attention mechanism	Focuses on critical regions by assigning more weights to relevant CKD retinal significant features	Requires careful parameter tuning. Otherwise yields noise and irrelevance.

Note: RBF: Radial Basis Function; ResNet: Residual Network; VGG: Visual Geometry Group network.

Table 2. Specific hybrid models for chronic kidney disease (CKD) detection

Features	Convolutional Neural Network (CNN), Attention, and Transfer Learning Hybrid Model	CNN, Random Forest Hybrid Model	Recurrent Neural Network (RNN), Support Vector Machine (SVM) Hybrid Model	Autoencoder, Clustering Hybrid Model
Strength	High accuracy and training efficiency, with fewer false positives	Handles nonlinear relationships and is robust to overfitting	Suitable for sequential/time-series data and capturing temporal patterns	Suitable for anomaly detection, uncovering hidden patterns, and unsupervised learning
Supporting datatypes	Images and tabular data	Image data initially, then fused with tabular data	Sequential data	Generally suitable for single-modality data
Interpretability	Highlights specific regions associated with affected kidney disease	Focuses on feature importance but lacks spatial context	RNN is often a black box; SVM provides limited interpretability (especially with nonlinear kernels)	Difficult to interpret
Training efficiency	Fast training due to pre-trained weights	Random Forest trains faster than CNN	SVM scales poorly, RNN is computationally expensive, and training is slow	Autoencoder training is typically slow
Multimodal support	Fuses different data types before classification	Feature fusion is possible	Handles sequential and static data	Combining different modalities often results in poor performance
False positives reduction	Focuses on the most important features	Errors may arise in non-discriminative contexts	Requires sequential data and careful model tuning	Primarily for anomaly detection

Table 1 demonstrates the existing methods with their drawbacks to derive a possible hybrid model and get multiple benefits. From Table 1, the methods demonstrated are logistic regression, which is not suitable for nonlinear types and complex spatial features. The Random Forest, which is also used for kidney disease prediction, lacks a supporting raw, unprocessed dataset and is weak in processing large datasets. The Support Vector Machine (SVM) is not suitable for big datasets, and won't support native feature extraction. The basic CNN lacks in performance when a large labelled dataset is given, and may suffer from an overfitting issue without regularization and augmentation. In these, although transfer learning captures high-level representation but may miss a few local context parameters, which would be captured by the attention mechanism, but may pose noise due to imbalanced data.

Table 2 demonstrates various specific hybrid models with their descriptions. The noted parameters considered are strength, data types supported, interpretability, training efficiency, multimodal support, and false positives reduction. The considered hybrid methods are such as CNN + attention mechanism + transfer learning, CNN + Random Forest, Recurrent Neural Network (RNN) + SVM, and autoencoder + clustering approaches. The former model, CNN + attention

mechanism + transfer learning, would have better and more positive results than other hybrid models.

2. LITERATURE REVIEW

The challenges observed in chronic disease detection are a lack of scalability over large datasets and complications raised due to diabetes in the datasets. The existing methods suffer from false positives and inefficient extraction of features from multimodal databases. Hence, a novel approach is required to overcome these challenges and ensure accuracy in detection and quick implementation. The studies based on their domain were decomposed into deep learning-based, specific CNN, transfer learning-based, biomarkers-based, hybrid, and explainable artificial intelligence (XAI) based, and cloud-related and other detection mechanism-based.

2.1 Deep learning-based models for chronic kidney disease prediction

These deep learning models evolved in capturing robust features, ensemble learning, and optimizations over handling of imbalanced data and heterogeneous clinical attributes. Zhou

et al. [1] addressed the mortality rate issues due to chronic disease detection. It defined a proposed model of CNN, Random Forest, and AdaBoost used in place of the Softmax layer of CNN to ensure optimization. This model is compared against CNN, Random Forest, k-nearest neighbors (KNN), and others regarding accuracy and misclassification. Saif et al. [2] addressed kidney disease prediction, which has a life expectancy of 6 months to a year. It uses deep learning models, such as CNN, long short-term memory (LSTM), and bidirectional long short-term memory (BLSTM). An ensemble of these using majority voting for 6 months is better than a year of life. The ensemble gets the benefit of all, but the performance of the best model is chosen to decide the optimized output. Byeon et al. [3] addressed the issue of picking robust features that would be ensured using a combination of broad learning and an autoencoder. The broad learning avoids gradient descent and suffers from extracting complex features, whereas the Autoencoder reduces the noise, providing robust features. Saif et al. [4] addressed deep learning models and ensemble strategy over CNN, LSTM, BLSTM, and an ensemble of these. The optimization is achieved using CNN-AdaMax, LSTM-Adam, and BLSTM-AdaMax, achieving better accuracy in chronic disease detection. Yuan et al. [5] addressed two problems in chronic disease detection: feature hiding and class distribution imbalance. To overcome the network-limited polynomial neural network, it is designed for high-level feature selection, which avoids overfitting, and NLPNN-based attention improves serious sick cases using data augmentation. The proposed algorithm ensures high performance. Abd Ulsada and Ramaha [6] addressed the consequences of chronic diseases like kidney disease, heart disease, diabetes, and others. The deep learning and machine learning models are applied, such as CNN in the first phase, and machine learning models, in which Linear regression and stochastic gradient descent would yield detection at a very high accuracy rate. Rashid et al. [7] addressed chronic diseases such as kidney disease, heart attack, breast cancer, etc., using augmented AI using Artificial Neural Network (ANN) with Particle Swarm Optimization (PSO). This model is compared against other machine learning and deep learning models, such as Random Forest, SVM, etc. Kulkarni et al. [8] addressed the extraction of robust features using broad learning that builds the model quickly using incremental learning but suffers from extracting complex features, and a denoising autoencoder helps to bring robust features, reducing noise. This combination achieves better accuracy than other models. Yu et al. [9] addressed the conversion of biomedical to bio-informational and developing deep learning models using omics such as Ribonucleic Acid (RNA), Deoxyribonucleic Acid (DNA), and multi-omics databases. All studies of these domains are helpful for the detection of diseases. Islam et al. [10] addressed strategies like minimum redundancy, maximum relevance, Least Absolute Shrinkage and Selection Operator (LASSO), and Relief techniques for significant feature selection. It also demonstrated various machine learning models like Naive Bayes, SVM, Random Forest, etc., as well as specific deep learning models for the detection of chronic diseases like liver disease, neurological disease, kidney disease, etc. These methods are compared for the best results. Although these models demonstrate ensemble operators, temporal modeling improves accuracy but still suffers from limited multimodal data.

2.2 Transfer learning and Convolutional Neural Network in retinal modeling

Transfer learning has emerged as a key approach in processing retinal imaging due to its pretrained visual encoding functionality. Canbay et al. [11] demonstrated the various transfer learning methods in which MobileNet performs better than others, in which feature extraction is done, and ensures better privacy along with accuracy. According to Badawy et al. [12], the two datasets were first discussed with 4 classes, and the second, with 5 classes, for handling kidney stones and renal cancer types. There were many specific transfer learning models used, in which DenseNet is best in both datasets. However, in the first case, it would ensure 99.98%, and in the second, it would ensure 100%. Ganie et al. [13] discussed the various transfer learning models such as VGG, ResNet, Efficient, MobileNet, etc., in which ResNet is identified as the best model compared with others, with the best accuracy. This study detects early, so that the survival rate is increased. Sharon and Anbarasi [14] demonstrated dual bottleneck attentions and two-fold convolution layers in order to fuse the multiple important features, reduce the complexity, and produce an accuracy of 98.5%. This model increases robustness and increases the solution accuracy in classifying into four types. Akhtar et al. [15] discussed two models, such as CNN and attention mechanism, in which the aggregated impact is taken using the attention mechanism that selects important features, for classifying the kidney disease level.

2.3 Biomarker-based chronic kidney disease detection

The emphasis on biological markers covers features such as pathological, biochemical, and physiological markers. Pethó et al. [16] addressed the severity of kidney disease when patients have diabetes and hypertension. Early prediction of the disease would help manage it efficiently. This scenario requires advancements and practical diagnostic tools for better prediction. Vaidya and Aeddula [17] addressed chronic diseases like kidney function and its damage. In such a case, treatment should be provided to slow down the progress, or transplantation should be used. Kidney damage affects other body parts, like cardiovascular disease, bone metabolism, blood pressure, and other health indicators. Kaur et al. [18] addressed that kidney diseases influence 15% of the world population. This kidney disease is not irreversible; it is damaged slowly over time. To detect diseases early and support health professionals, specific machine learning methods are applied. Al-Momani et al. [19] addressed the seriousness of kidney damage, and 2 million people worldwide were affected. Late identification leads to death. The methods used are machine learning, ANN, SVM, and KNN, of which ANN was found to be more accurate than the other models. Rajeashwari and Arunesh [20] addressed chronic diseases like heart disease, kidney disease, breast cancer, and diabetes. The various machine learning and deep learning methods are used and compared to detect chronic diseases. Halder et al. [21] addressed seven machine learning models, in which AdaBoost and Random Forest performed better in accuracy than other models. It involves two steps: data preprocessing for disease classification and the development of an application that auto-predicts the disease. This model involves data preprocessing, feature selection, and machine learning detection. These models rely on biomarker-driven approaches but still depend on clinical aspect processing and lack multimodal data.

2.4 Interpretability-based hybrid and explainable artificial intelligence models

The interpretability-based CKD integrates XAI with a machine learning/deep learning pipeline to improve clinical decisions. Illiyas et al. [22] demonstrated a few methods, such as deep learning, AI, ensemble, and XAI techniques, to improve the detection rate of kidney injuries. The integration of XAI over specific methods would enhance interpretability. The goal of this would be to increase accuracy and reduce human error. Liu et al. [23] addressed decision-making issues in chronic disease. Hence, kernel-based rough fuzzy sets with multi-granularity are derived to detect chronic disease with robustness and better accuracy. Navita et al. [24] addressed the issues and overcome them using four stages: the first stage uses Synthetic Minority Over-sampling Technique combined with Edited Nearest Neighbors (SMOTE-ENN) for data balancing, the second stage uses feature selection using CHI-Square, the third stage uses a hybrid model of AdaBoost, Random Forest, KNN, and optimizer for accurately classifying, and the fourth stage is evaluation of performance measures. Zhang et al. [25] discussed a random forest method with a multivariate model against different types of sources, like multi-omics. This model uses inverse minimal depth for classifying the subtypes of the kidney disease with respect to clinical and biological data. The studies show a shift toward transparent, cum multi-granular decision models, but focused on limited end-to-end deep multimodal architecture integration.

2.5 Miscellaneous studies on cloud and other detections

The studies under this category explore cloud-based analytics and non-renal disease models, relevant to CKD processing. Dey and Sangaraju [26] demonstrated applying global and local load balancing strategies in task distribution using particle optimization technique, and ensured quick responses as well as increased resource optimization. Dey and Sangaraju [27] discussed increasing the performance of the data center in the cloud environment using a novel framework that ensures reduced latency and scalability. Allamudi and Raju [28] explored a hybrid model that assures better accuracy in the detection of fraudulent activities, over considering existing models along with robustness. Raju et al. [29] depicted defect identification in flyovers using integrated deep learning models, so that prevention strategies are applied to extend the life of the bridges. Firdaus et al. [30] discussed the clinical data against the logistic regression model, and ensured 79% performance and 80% accuracy. This study is used to classify the kidney disease type, such as normal or CKD. Mendapara [31] discussed the use of Random Forest against patients' genes and gene sequencing, and classifies into CKD or non-CKD. The model performed the training and validation tests with respect to 10-fold cross-validation, ensuring an accuracy of 94.5%. Li et al. [32] demonstrated the SVM model for predicting the sudden loss of kidney functioning, especially in patients in the ICU. The SVM is compared against the other predictive models in this category. Table 3 highlights the significant studies and their gaps.

Table 3. Significant studies and their gaps

Studies	Contributions	Research Gaps
Zhou et al. [1]	A hybrid model replacing CNN's Softmax layer with ensemble classifiers (Random Forest, AdaBoost) for accuracy, and reducing misclassification	Lack of interpretability and clinical utility in terms of not accurately guiding
Yuan et al. [5]	Two challenges, such as feature hiding and class imbalance, are addressed	Lack of technical focus, not mentioning of external datasets and multimodal data
Saif et al. [2]	Uses CNN, LSTM, and BLSTM to improve prediction accuracy and survival rate	Compared to clinical, the model is more complex and lacks a cost-benefit analysis
Byeon et al. [3]	Uses broad learning and contributions, extracts robust features, auto encoders for noise reduction	Lack of clear clinical description from centric and non-centric methods, weak methodology in prediction
Kulkarni et al. [8]	Data preprocessing and AdaBoost for prediction	Not supporting imaging data
Halder et al. [21]	Handles with and without comorbidities	Lack of robustness
Firdaus et al. [30]	Handles genetic data and uses 10-fold cross-validation	Lack of genetic validation
Mendapara [31]	Handles scenarios that involve acute kidney injury identification	Lack of biomarker scarcity and face challenges in acute kidney injury heterogeneity

Note: CNN: Convolutional Neural Network; LSTM: long short-term memory; BLSTM: bidirectional long short-term memory.

Table 3 demonstrates limitations of models addressed, such as the absence of retinal biomarkers in kidney-disease models, lack of models capturing the global retinal structure and microvascular abnormalities, limited interpretability in kidney-focused deep learning models, lack of multimodal on retinal imaging and CKD clinical indicators, and weak generalization across comorbid populations.

3. PROPOSED METHODOLOGY

The objective of chronic kidney disease detection includes the demonstration of modules involved in Figure 1, the flow of activities for the successful deployment of the hybrid model in Figure 2, and a pseudo-procedure of the hybrid model in PS1. Figure 1 demonstrates the modules, such as data

acquisition, that consider retinal images and then quantify the data using normalization and data augmentation. Reducing training time with transfer learning using a pre-trained model, ResNet50, critical region identification of kidney malfunction using attention-guided mechanisms, Spatial feature handling with respect to region selection using CNN, then fusing the significant features, yields better evaluation of accuracy, performance, and interpretability.

Figure 2 illustrates the order of actions that aggregate the important features that ensure the goal of the hybrid model. This includes detecting whether a kidney malfunction is present in an image and classifying it so that health professionals can guide the patient on the next steps to recover from the disease. It involves activities such as acquisition of data, preprocessing of images and clinical data, transfer learning, attention mechanism, and a custom CNN for regions

of kidney dysfunction. Then, fuse significant features of retinal images and clinical indicators to ensure better classification accuracy. Then evaluate the model's effectiveness against the base models.

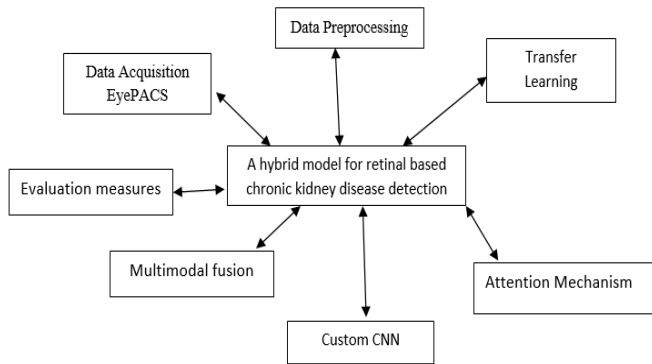


Figure 1. Modules' interaction in the hybrid model for kidney detection

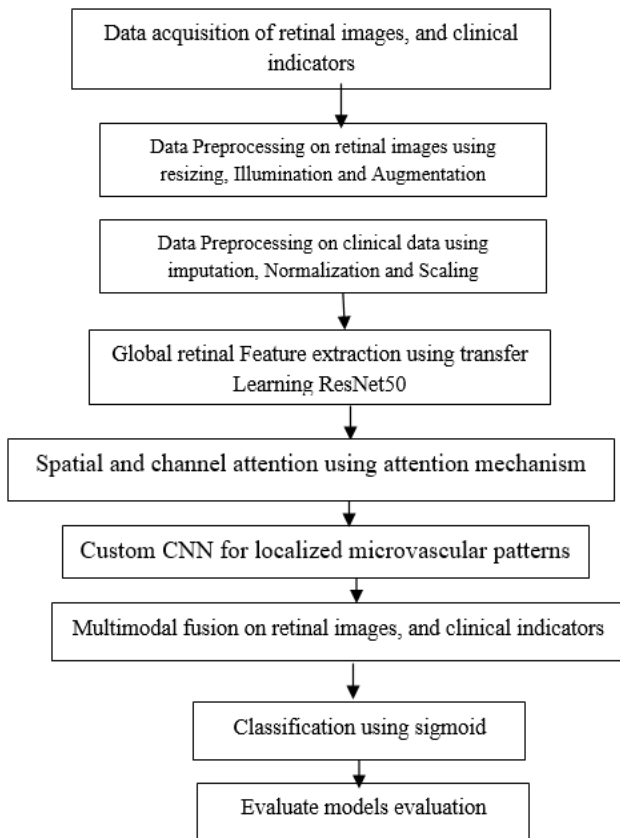


Figure 2. Flowchart of a hybrid model (transfer learning, attention, and custom Convolutional Neural Network (CNN))

The following pseudo-procedure demonstrates the working procedure of the proposed hybrid model to justify the effectiveness of the model.

PS1: Pseudo ProcedureRetinalbased_Chronic disease Detection_TL_Attention_CNN (Database[[]]):

Input: Retinal images and tabular data

Output: CKD or non-CKD with stages 0 to 4, where 0 denotes no DR, 1 denotes mild, 2 denotes moderate, 3 denotes severe, and 4 denotes proliferative DR

Step 1: Load the multimodal database using the link

“<https://www.kaggle.com/c/diabetic-retinopathy-detection/data>” for retinal, and use “<https://physionet.org/content/mimiciv/3.1/>” for clinical data

Step 2: Do data preprocessing

2.1 If data is images:

Apply normalization and resize to standard 224×224 , then do augmentation using rotation and flipping.

2.2 If data is EHR or clinical data:

- Handle missing and imbalanced data using data augmentation and sampling techniques.
- Use a Gaussian filter to remove noise and enhance quality
- Divide the dataset into training, validation, and testing tests

Step 3: Apply transfer learning for high-level feature representation

3.1 Construct a pre-trained CNN backbone

- Load weights from the pre-trained model ResNet50
- Freeze early layers
- Fine-tune deeper layers for kidney lesion detection
- Add a custom classification head using a global average pooling layer, fully connected layers, and output with the final number of classes.

Step 4: Apply attention-guided lesion detection

4.1 Define attention maps that accept input $F \in \mathbb{R}^{H \times W \times C}$, where H, W are spatial dimensions, and C is the number of channels.

4.2 Apply a convolution layer with a 1×1 kernel and ReLU to reduce channels and get spatial dependencies

4.3 Apply a second convolution layer with a 1×1 kernel, use sigmoid activation to generate attention weights $A \in \mathbb{R}^{H \times W \times 1}$ in the range

4.4 Multiply feature map F against attention weights A for kidney abnormalities

4.5 Apply global average pooling that reduces to a 1D vector

4.6 Upsample A to the size of the original image, and a heatmap is overlaid for the critical regions

4.7 The feature map passes into dense layers and Softmax to produce output CKD levels

Step 5: CNN architecture

5.1 Input layer takes preprocessed image $24 \times 224 \times 3$

5.2 Define convolution blocks

- Use multiple convolution layers with multiple filters
- Use kernel size 3×3
- Use the activation function ReLU
- To stabilize learning, use batch normalization
- Use max pooling to downsample

5.3 Use dropout layers to prevent overfitting

5.4 Flatten output feature map to 1D

5.5 Apply fully connected layers to learn complex feature patterns

5.6 Display final output with levels

Step 6: Construct hybrid modal training

6.1 Use pre-trained model ResNet50, and freeze initial layers

6.2 Refine attention maps on critical regions using attention maps

6.3 Extract medical features from images using CNN

- 6.3.1 Add custom classification for output that produces CKD or Non-CKD and stage
- 6.3.2 Use Conv2D(), MaxPooling(), Conv2D(), and MaxAveragePooling() in CNN design

6.4 Define a fusion layer for combining image features

and clinical features

6.5 Use a loss function weighted cross-entropy to handle class imbalance, like CKD or non-CKD, to produce the final output as levels

Step 7: Evaluate measures such as Accuracy, F1-Score, Precision, and Recall:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{Total Number of Cases}} \quad (1)$$

Compute performance measures:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{Recall (Sensitivity)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (4)$$

where, true positives (TP): The number of correctly identified relevant data points, true negatives (TN): The number of correctly identified irrelevant data points, and the total number of cases: The sum of TP, TN, false positives (FP), and false negatives (FN).

From PS1, the part of EyePACS dataset is loaded based on hardware configuration, loads retinal images and tabular data for processing, then data preprocessing is applied on images and tabular data using specific techniques, and then split into 3 categories: training (70%), validation (15%), and testing (15%). Firstly, the transfer learning used is ResNet50, which is applied for enhancing efficiency by freezing initial layers to preserve learned patterns, fine-tuning deep layers for domain

adaptation, and representation at a high level. Then, the attention mechanism generates attention weights from feature maps, highlighting the regions of lesions or blood vessels. Then, a CNN architecture is built to eliminate overfitting using batch normalization and dropouts, and complex representations are done using flattened feature maps into a vector. Then, the hybrid model fuses the image features and clinical features through fusion layers, uses a weighted cross-entropy loss that reduces class imbalance. Hence, the modules involved include data preprocessing, transfer learning, attention mechanism for critical regions identification, CNN architecture for region-selected feature extraction, and then multimodal fusion to ensure more accurate prediction. The effectiveness is evaluated using Eqs. (1)-(4).

The methodology is not a simple sequential pipeline but a hierarchical feature-aggregation architecture, which ensures different representations before fusion. In this, the pre-trained ResNet50 is working like a global feature encoder. It extracts high-level retinal representations like macular texture, optic disc structure, and vascular topology. The initial layers were frozen, but the final layers fine-tune to retinal patterns relevant to CKD. The output of transfer-learning is propagated to the attention module, which produces a map that focuses on discriminative retinal regions relevant to vessel calibre changes, microaneurysms, or hemorrhages. The output of attention is multiplied element-wise with ResNet50, producing an attention-refined embedding. The custom CNN is applied to capture localized textural and microvascular cues that are not captured by ResNet50. The hybrid feature-level fusion captures a global feature vector from ResNet50, an attention-weighted feature vector, and a refined local spatial feature vector, and concatenates these three into a unified feature tensor, which propagates to fully connected layers and a Softmax classifier, for CKD stage prediction.

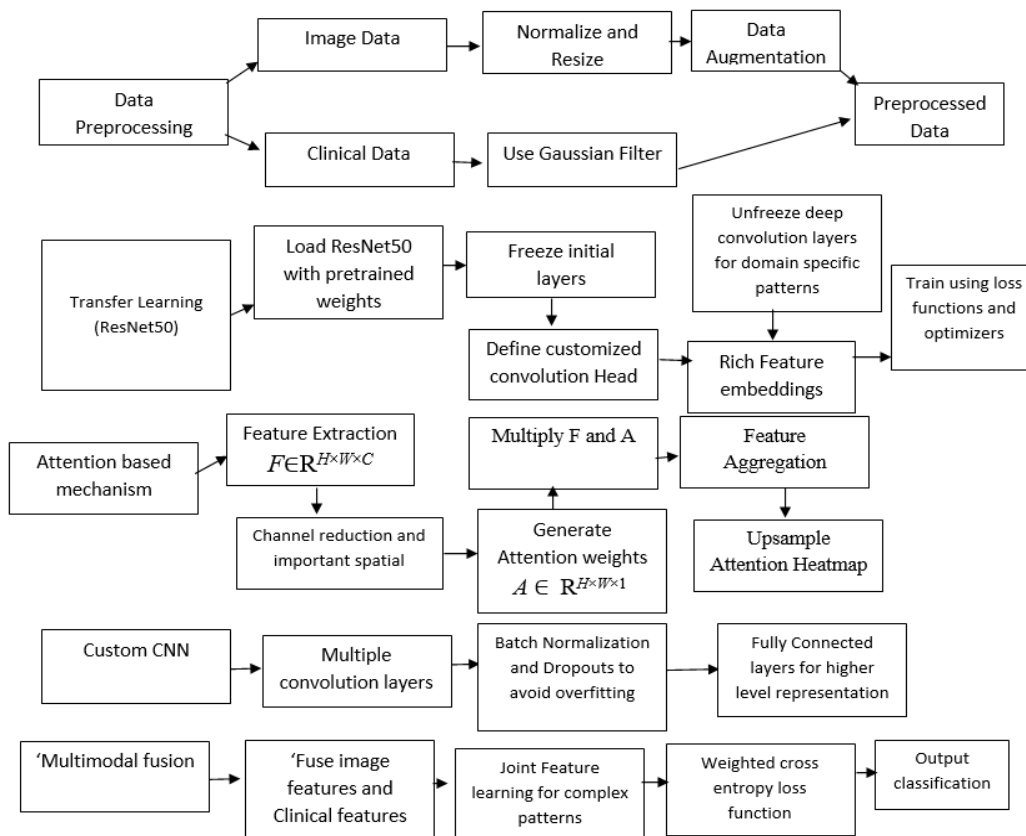


Figure 3. Layer-wise processing of kidney classification

Figure 3 shows 5 layers: the first layer performs data preprocessing, the second uses transfer learning, the third employs an attention-based mechanism, the fourth uses a CNN, and the fifth performs multimodal fusion, producing a classification output. This proposed hybrid model considers multimodal data, such as retinal image data, as well as clinical data considered, such as serum creatinine, eGFR, blood pressure, Urine albumin–creatinine ratio (ACR), age, sex, and comorbidity indicators (hypertension, diabetes), available on Kaggle. The ethical compliance norms are adhered to on secondary, de-identified data. The CKD and non-CKD with stage are determined from retinal images and from clinical data.

4. RESULTS

Based on studies on kidney disease detection, the accuracy and performance are noted and compared against the proposed hybrid model to justify the effectiveness. Table 4 demonstrates the specific studies along with their evaluated measures to assess the effectiveness of the model, and the model's effectiveness is demonstrated in Figure 4. The methods considered are demonstrated in Table 5, which depicts the specific measure of multimodal database support. Table 6 demonstrates on false negative rate in % dictates the

robustness of the models. Table 5 demonstrates an ablation study of methods involved in the proposed hybrid model with cross-modal fusion. The evaluated details of accuracy, performance, and multimodal support are demonstrated in Tables 4 and 5. It is observed that our hybrid model with multimodal support provides better accuracy and performance compared to other models. Table 4 demonstrates that the basic machine learning approaches logistic regression, SVM, Random Forest wont interpret retinal vascular, and pixel abnormalities, Core CNN lacks global representation although provide performance, ResNet50 as transfer learning approach acts as pretrained filter whose performance is increase than earlier approaches, then transfer learning + CNN performs better than individual CNN, and alone transfer learning in emphasizing CKD relevant vascular regions, the proposed hybrid approach without multimodal fusion achieves better accuracy than earlier models in which transfer learning for global representation, attention for local abnormalities, and custom CNN for features specific to critical regions. Then, the proposed model with multimodal fusion yields the highest performance values using integration of retinal biomarkers and clinical indicators. This emphasis is demonstrated in Figure 4. Table 4 visualizes the measures like accuracy, performance, and k-fold cross-validation to evaluate the effectiveness of the model and its robustness.

Table 4. Evaluated measures of chronic kidney disease (CKD) prediction used models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	Area Under the Receiver Operating Characteristic Curve (AUC)
Logistic Regression	79.3	78.5	76.2	77.3	0.76
Support Vector Machine (SVM, RBF kernel)	81.0	80.4	78.1	79.2	0.78
Random Forest	85.2	84.6	82.3	83.4	0.81
Basic CNN (No transfer learning)	82.9	81.0	78.5	79.7	0.86
Transfer Learning (ResNet50)	88.7	87.5	85.9	86.7	0.92
CNN + Attention (No transfer learning)	90.3	89.2	87.0	88.1	0.94
Proposed Hybrid Model (transfer learning + Attention + CNN Refinement)	95.4	94.6	93.8	94.2	0.97
Proposed Multimodal Fusion (Retina + Clinical Data)	97.1	96.4	95.8	96.1	0.98

Note: RBF: Radial Basis Function; CNN: Convolutional Neural Network.

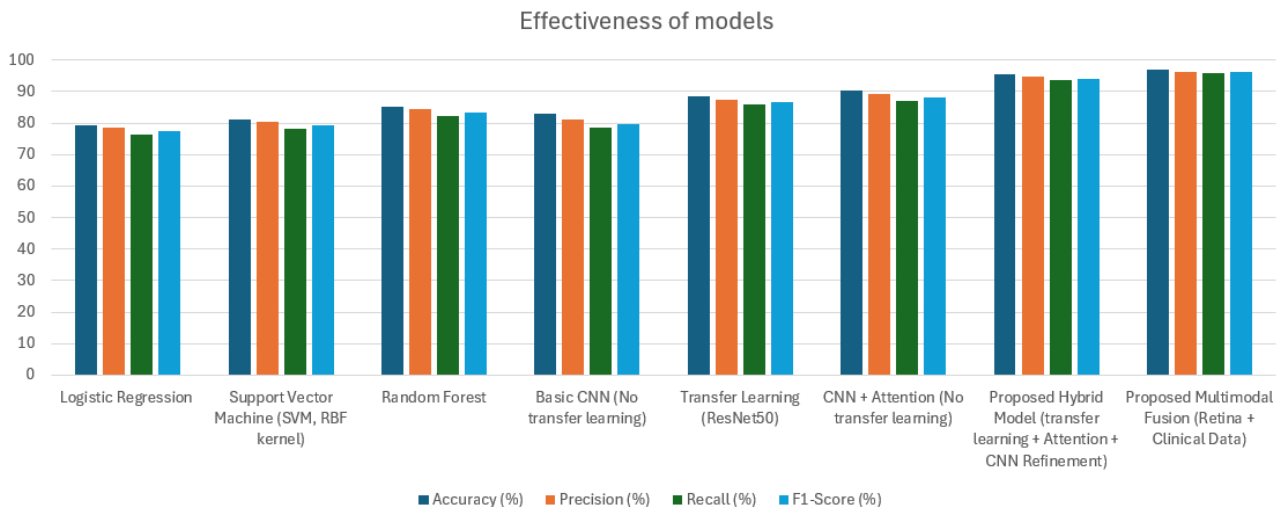


Figure 4. Effectiveness of the considered models

Table 5. Multimodal support of the considered models

Models	Retina-Only Accuracy (%)	Retina + Clinical Accuracy (%)
Logistic Regression	N/A	79.3
Support Vector Machine (SVM, RBF kernel)	N/A	81.0
Random Forest	N/A	85.2
Basic CNN	82.9	89.0
Transfer Learning (ResNet50)	88.7	93.5
CNN + Attention	90.3	94.2
Proposed Hybrid Model (Transfer Learning + Attention + CNN Refinement)	95.4	97.1
Proposed Hybrid Model (Transfer Learning + Attention + CNN + Cross-Fusion)	N/A	97.8

Note: RBF: Radial Basis Function; CNN: Convolutional Neural Network.

Table 6. Evaluated the Area Under the Receiver Operating Characteristic Curves (AUCs) and the false negatives of the models

Models	AUCs	False Negatives (%)
Convolutional Neural Network (CNN)	0.86	12.40
Transfer Learning	0.92	9.80
Attention	0.94	8.60
CNN + Transfer Learning	0.93	8.90
CNN + Attention Mechanism	0.94	7.90
Hybrid Model (Proposed System)	0.97	6.10
Proposed Hybrid Model with Fusion	0.98	4.80

This study ensures unbiased model evaluation that includes data preprocessing, class imbalance is handled by augmentation and sampling, and splitting the dataset into training (70%), validation (15%), and testing (15%). Transfer learning with ResNet50 is used for freezing all initial layers and fine-tuning deep layers for lesion features. The technical details are demonstrated in Table 5. In addition, the Area Under the Receiver Operating Characteristic Curve (AUC) demonstrates a plot drawn on the true positive rate against the False Positive Rate (FPR), defining the effectiveness of the model.

Table 6 demonstrates that CNN achieves an AUC of 0.86, having limitations in capturing global retinal context. Transfer learning increases the AUC to 0.92 and reduces false negatives to 9.80%. Attention-based model handles microvascular abnormalities, achieving an AUC of 0.94 and 8.60% false negatives. CNN + attention configuration observed as a reduction in false negatives to 7.90%. Transfer learning, attention, and CNN refinement increase an AUC of 0.97, and reduce false negatives to 6.10%. The multimodal fusion model uses a combination of retinal features and clinical indicators to ensure an AUC of 0.98 and reduce false negatives to 4.80%. Based on Table 6, the AUC evaluated values are specified for the considered models. The false negative rates are also listed, which declares that the hybrid model is best in effectiveness of the model.

The proposed hybrid model integrating transfer learning, Attention, and CNN refinement achieves better performance against other base models. The architectural strengths contribute to the effectiveness of the model. In this process, the role of transfer learning forms a strong base to stabilize

training and improve generalization by capturing global retinal attributes such as optic disc boundaries, vessel orientation, and macular patterns. Custom CNN refinement layers extract microvascular cues like vessel calibre changes, tortuosity, microaneurysms, and hemorrhages, which reveal CKD risks. Attention mechanisms would enhance discriminative capability by localizing retinal regions and suppressing background noise, illumination artifacts, and non-informative areas. The attention maps produced by the model would focus on retinal regions and suppress background noise, illumination artifacts, and non-informative areas, which induce structural and functional changes in retinal vasculature. The interpretability is enhanced with trust and supports retinal imaging to infer kidney health.

The dataset describes high-resolution retinal images with left and right views, along with subject ID. Key observations as limitations include:

(i) EyePACS is a large retinal dataset of 88670 retinal images are retrospectively paired with clinical data, which introduces bias. The dataset overrepresents diabetes, which results in compounding effects of inadvertent retinal changes.

(ii) Although it supports multimodal support, its effectiveness varies based on varying details of clinical variables.

(iii) Although the model performs a reduction in false negatives, external validation is required on diverse populations and devices used.

(iv) The observed cost of the derived model is higher than that of the base models, which limits deployment in low-resource environments.

5. CONCLUSION

The challenges that traditional and other models posed were reduced and eliminated using a hybrid method that incorporates models such as transfer learning, ResNet50, Attention mechanism, custom CNN, and multimodal fusion support. The purpose of each model depicts that transfer learning for using the pre-trained model ResNet50 to reduce training time, focus on high level representation with faster performance, Attention mechanism for the critically identified regions of vascular abnormalities, the integration of CNN for spatial feature extraction of selected regions, and multimodal fusion combines significant features of retinal imaging and provide clinical indicators. From evaluations, each model's individual performance and the hybrid model are compared. The AUC is one performance assessment measure for judging the model's effectiveness, along with a low false-negative rate. The evaluation measures, especially accuracy, performance, and multimodal database support, are significant strengths of the proposed hybrid model with multimodal fusion. The key advantages ensured by this model are improved accuracy, reduced false positives, and improved interpretability. The model to build must ensure cost-aware model optimization in low-resource constraints, and would be scalable by integrating with tele-ophthalmology pipelines.

ACKNOWLEDGMENT

The author would like to express sincere gratitude to the supervisor for continuous guidance and support. Appreciation is also extended to the university for providing the facilities and resources necessary for this research.

REFERENCES

- [1] Zhou, H., Zhang, P.Y., Zou, X., Liu, J., Wang, W.J. (2023). Chronic disease diagnosis model based on convolutional neural network and ensemble learning method. *Digital Health*, 9. <https://doi.org/10.1177/20552076231198643>
- [2] Saif, D., Sarhan, A.M., Elshennawy, N.M. (2024). Deep-kidney: An effective deep learning framework for chronic kidney disease prediction. *Health Information Science and Systems*, 12: 3. <https://doi.org/10.1007/s13755-023-00261-8>
- [3] Byeon, H., Prashant, G.C., Hannan, S.A., Alghayadh, F.Y., Soomar, A.M., Soni, M., Bhatt, M.W. (2024). Deep neural network model for enhancing disease prediction using auto encoder based broad learning. *SLAS Technology*, 29(3): 100145. <https://doi.org/10.1016/j.slast.2024.100145>
- [4] Saif, D., Sarhan, A.M., Elshennawy, N.M. (2024). Early prediction of chronic kidney disease based on an ensemble of deep learning models and optimizers. *Journal of Electrical Systems and Information Technology*, 11: 17. <https://doi.org/10.1186/s43067-024-00142-4>
- [5] Yuan, X.H., Chen, S.Y., Sun, C., Lu, Y.W. (2022). A novel early diagnostic framework for chronic diseases with class imbalance. *Scientific Reports*, 12: 8614. <https://doi.org/10.1038/s41598-022-12574-x>
- [6] Abd Ulsada, A.A., Ramaha, N.T.A. (2023). Detection of chronic diseases based on the principles of deep and machine learning. *AIP Conference Proceedings*, 2977: 020039. <https://doi.org/10.1063/5.0183661>
- [7] Rashid, J., Batool, S., Kim, J., Wasif Nisar, M., Hussain, A., Juneja, S., Kushwaha, R. (2022). An augmented artificial intelligence approach for chronic diseases prediction. *Frontiers in Public Health*, 10: 860396. <https://doi.org/10.3389/fpubh.2022.860396>
- [8] Kulkarni, C., Quraishi, A., Raparthi, M., Shabaz, M., Khan, M.A., Varma, R.A., Keshta, I., Soni, M., Byeon, H. (2024). Hybrid disease prediction approach leveraging digital twin and metaverse technologies for health consumer. *BMC Medical Informatics and Decision Making*, 24: 92. <https://doi.org/10.1186/s12911-024-02495-2>
- [9] Yu, X.D., Zhou, S.S., Zou, H.L., Wang, Q.J., Liu, C.J., Zang, M.J., Liu, T. (2023). Survey of deep learning techniques for disease prediction based on omics data. *Human Gene*, 35: 201140. <https://doi.org/10.1016/j.humgen.2022.201140>
- [10] Islam, R., Sultana, A., Islam, M.R. (2024). A comprehensive review for chronic disease prediction using machine learning algorithms. *Journal of Electrical Systems and Information Technology*, 11: 27. <https://doi.org/10.1186/s43067-024-00150-4>
- [11] Canbay, Y., Adsiz, S., Canbay, P. (2024). A privacy-preserving transfer learning framework for kidney disease detection. *Applied Sciences*, 14(19): 8629. <https://doi.org/10.3390/app14198629>
- [12] Badawy, M., Almars, A.M., Balaha, H.M., Shehata, M., Qaraad, M., Elhosseini, M. (2023). A two-stage renal disease classification based on transfer learning with hyperparameter optimization. *Frontiers in Medicine*, 10: 1106717. <https://doi.org/10.3389/fmed.2023.1106717>
- [13] Ganie, S.M., Pramanik, P.K.D., Zhao, Z.M. (2024). Deep transfer learning for kidney disease detection using CT scan images. In 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Lisbon, Portugal, pp. 5624-5629. <https://doi.org/10.1109/BIBM62325.2024.10822760>
- [14] Sharon, J.J., Anbarasi, L.J. (2025). An attention-enhanced dilated bottleneck network for kidney disease classification. *Scientific Reports*, 15: 9865. <https://doi.org/10.1038/s41598-025-90519-w>
- [15] Akhtar, M.N., Iqbal, K., Abid, M.S. (2025). Renal vision advanced kidney disease detection using attention-powered ensemble CNNs. *Physical Education, Health and Social Sciences*, 3(1): 31-42. <https://doi.org/10.63163/jpehss.v3i1.114>
- [16] Pethő, Á.G., Tapolyai, M., Csongrádi, É., Orosz, P. (2024). Management of chronic kidney disease: The current novel and forgotten therapies. *Journal of Clinical & Translational Endocrinology*, 36: 100354. <https://doi.org/10.1016/j.jcte.2024.100354>
- [17] Vaidya, S.R., Aeddula, N.R. (2024). Chronic kidney disease. In *StatPearls, Treasure Island (FL)*. <https://pubmed.ncbi.nlm.nih.gov/30571025/>.
- [18] Kaur, B., Goyal, B., Dogra, A., Ramshankar, S., Singh, D., Alkhayyat, A. (2024). Chronic kidney disease detection using machine learning: From analysis to framework development. *Biomedical and Pharmacology Journal*, 17(3): 1739-1747. <https://doi.org/10.13005/bpj/2979>
- [19] Al-Momani, R., Al-Mustafa, G., Zeidan, R., Alquran, H., Mustafa, W.A., Alkhayyat, A. (2022). Chronic kidney disease detection using machine learning technique. In 2022 5th International Conference on Engineering Technology and its Applications (IICETA), Al-Najaf, Iraq, pp. 153-158. <https://doi.org/10.1109/IICETA54559.2022.9888564>
- [20] Rajeashwari, S., Arunesh, K. (2022). Performance analysis for chronic disease prediction using various data mining techniques. In 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, pp. 721-727. <https://doi.org/10.1109/ICESC54411.2022.9885383>
- [21] Halder, R.K., Uddin, M.N., Uddin, M.A., Aryal, S., Saha, S., Hossen, R., Ahmed, S., Rony, M.A.T., Akter, M.F. (2024). ML-CKDP: Machine learning-based chronic kidney disease prediction with smart web application. *Journal of Pathology Informatics*, 15: 100371. <https://doi.org/10.1016/j.jpi.2024.100371>
- [22] Iliyas, I.I., Baukary, S., Gital, A.Y. (2025). Recent trends in prediction of chronic kidney disease using different learning approaches: A systematic literature review. *Journal of Medical Artificial Intelligence*, 8: 62. <https://doi.org/10.21037/jmai-24-256>
- [23] Liu, J.Q., Sun, B.Z., Ye, J., Zhao, X.X., Chu, X.L. (2025). Hybrid deep-learning prediction model based on kernel multi-granularity fuzzy rough sets and its application in the diagnosis and treatment of chronic kidney disease. *Engineering Applications of Artificial Intelligence*, 147: 110297. <https://doi.org/10.1016/j.engappai.2025.110297>
- [24] Navita, Mittal, P., Sharma, Y.K., Lilhore, U.K., Simaiya, S., Saleem, K., Ghith, E.S. (2025). Advanced hybrid machine learning model for accurate detection of cardiovascular disease. *International Journal of Computational Intelligence Systems*, 18: 51.

- <https://doi.org/10.1007/s44196-025-00771-1>
- [25] Zhang, W., Huang, H.C., Wang, L., Lehmann, B.D., Chen, S.X. (2025). An integrative multi-omics random forest framework for robust biomarker discovery. *bioRxiv*. <https://doi.org/10.1101/2025.03.05.641533>
- [26] Dey, N.S., Sangaraju, H.K.R. (2024). A particle swarm optimization inspired global and local stability driven predictive load balancing strategy. *Indonesian Journal of Electrical Engineering and Computer Science*, 35(3): 1688-1701. <https://doi.org/10.11591/ijeecs.v35.i3.pp1688-1701>
- [27] Dey, N.S., Sangaraju, H.K.R. (2023). Hybrid load balancing strategy for cloud data centers with novel performance evaluation strategy. *International Journal of Intelligent Systems and Applications in Engineering*, 11(3): 883-908. <https://ijisae.org/index.php/IJISAE/article/view/3345>.
- [28] Allamudi, A.K., Raju, S.H. (2025). Hybrid machine learning for fraud detection: Balancing accuracy and security in digital transactions. *International Journal of Safety and Security Engineering*, 15(2): 331-337. <https://doi.org/10.18280/ijss.150214>
- [29] Raju, S.H., Adinarayna, S., Jadala, V.C., Rao, K.Y., Sesadri, U., Sreeraman, Y. (2025). An optimized approach for defect detection in the flyovers using faster R-CNN, LSTM, and Transfer Learning. *Mathematical Modelling of Engineering Problems*, 12(8): 2874-2882. <https://doi.org/10.18280/mmep.120828>
- [30] Firdaus, T.F., Fa'rifah, R.Y., Pramesti, D. (2025). Implementation of logistic regression for classification of chronic kidney disease with and without comorbidities: A performance evaluation. In *2025 4th International Conference on Electronics Representation and Algorithm (ICERA)*, Yogyakarta, Indonesia, pp. 443-448. <https://doi.org/10.1109/ICERA66156.2025.11087353>
- [31] Mendapara, K. (2024). Development and evaluation of a chronic kidney disease risk prediction model using random forest. *Frontiers in Genetics*, 15: 1409755. <https://doi.org/10.3389/fgene.2024.1409755>
- [32] Li, X.S., Chen, Y., Sun, A.L., Wang, Y., Liu, Y., Lei, H.K. (2023). Development and validation of prediction model for overall survival in patients with lymphoma: A prospective cohort study in China. *BMC Medical Informatics and Decision Making*, 23: 125. <https://doi.org/10.1186/s12911-023-02198-0>