Echoes of the Hidden Filter: Interpretable Ensemble Intelligence for Early CKD Prognosis

Dodda Venkata Reddy
Dept. of CSE
Narasaraopeta Engineering College
Narasaraopet, India
doddavenkatareddy@gmail.com

Sura Venkata Siva Naga Lakshmi
Dept. of CSE
Narasaraopeta Engineering College
Narasaraopet, India
suravenkatasivanagalakshmi@gmail.com

Kolli Kaveri
Dept. of CSE
Narasaraopeta Engineering College
Narasaraopet,India
kaverikolli69@gmail.com

Konda Pratyusha
Dept. of CSE
Narasaraopeta Engineering College
Narasaraopet,India
pratyushakonda626@gmail.com

V. Jyothi
Dept. of CSE
GRIET
Hyderabad,India
jyothi1687@grietcollege.com

Tummati Swapna
Dept. of CSE
G.Narayanamma Institute of Technology Science (for women)
Hyderabad,India
t.swapna@gnits.ac.in

***Abstract*—Chronic Kidney Disease (CKD) is a gradual and usually asymptomatic condition which often goes undetected in early stages, especially in resource-constrained healthcare environments. This paper introduces Echoes of the Hidden Filter, an interpretable ensemble learning model that can be used for proactive CKD prediction. The method combines Random Forest, XGBoost, Support Vector Machines, and Artificial Neural Networks as a hybrid stacking and majority-vote ensemble with strong performance over a wide range of patient populations. For the purpose of increasing clinical trust, the model uses SHAP and LIME for global and local interpretability to identify prominent biomarkers like eGFR, creatinine, and cystatin-C in personal predictions. On a real-world CKD dataset with missing values and class imbalance, we have used regression-based imputation, variance-thresholding, and dynamic SMOTE balancing. Experimental outcomes show that the ensemble outperforms individual classifiers at all times for accuracy, precision, recall, and F1-score. By combining interpretability with predictive performance, the framework proposed closes the loop between black-box AI and clinical usability, presenting a scalable approach to population-level CKD screening.**

***Index Terms*—Chronic Kidney Disease (CKD), k-Fold Cross Validation, Logistic Regression, Support Vector Machine, Variance Threshold Feature Selection.**

I. Introduction

Chronic Kidney Disease (CKD) affects nearly 10% of the global population and causes significant morbidity, mortality, and healthcare costs [1], [2]. Most of the time, the disease is asymptomatic in its early phases, silently progressing to end-stage renal disease (ESRD) requiring dialysis or transplant [3]. Traditional diagnostic tests—mainly serum creatinine and estimated glomerular filtration rate (eGFR)—often detect CKD only when significant organ damage has been done, leaving intervention late.

Machine learning (ML) has become a potential method for supplementing early detection of CKD with high-dimensional clinical data in recent years [4]. Single models including Random Forests, Support Vector Machines, and Neural Networks have shown robust diagnostic performance in earlier studies [5]–[7] . Ensemble methods also maximize prediction credibility by combining the complementary strengths of multiple classifiers [8]. Yet, while accurate, most ML-based systems are unclear, commonly referred to as ”black boxes” that have prevented clinical adoption [9].

To combat this issue, explainable AI (XAI) methods, such as SHAP (SHapley Additive exPlanations) and LIME (Local Interpretable Model-Agnostic Explanations), have become more prevalent. These methods offer transparency at global and case-level, allowing clinicians to comprehend model reasoning and establish trust in AI-generated decisions [10].

On top of these developments, this work presents Echoes of the Hidden Filter, a family of ensemble models combining resilient preprocessing, variance-oriented feature selection, class balancing using SMOTE, and

stacking ensembles with interpretability modules. Our goal is to bridge predictive performance with explainable decision rationale toward early CKD prognosis that is clinically actionable and trustworthy.

II. Related Work

Diagnosis, monitoring, and control of chronic kidney disease (CKD) have been substantially improved by both biomedical research and Machine Learning (ML) methods. The performance of CKD prediction models relies on data quality, choice of biomarkers, model architecture, explainability, and their integration into clinical practice. Clinical data acquisition is core.

Kanasaki et al. [1] detailed the continued essential role of standard markers like eGFR and albuminuria in diabetic nephropathy staging, while Elcioglu et al. [2] emphasized their utility in ADPKD surveillance.

Li et al. [3] demonstrated how non-coding RNAs provide novel biomarkers outside conventional blood markers. Hart et al [4] highlighted that improved treatments such as HEMT in cystic fibrosis have prolonged life expectancy while exposing new CKD risks. Souza et al [5] proposed a reagent-free microwave sensor as a low-cost, precise substitute for CKD screening.

Bianchi et al [6] demonstrated that transformer-based models integrating complex multi-dimensional data outperform conventional predictors. Puri et al [7] proposed blockchain-based AI for secure, decentralized CKD forecasting. Zhao et al [8] introduced a Deep Belief Rule Base model where transparency is coupled with scalability.

Ghosh et al [9] enhanced CKD staging with classifier optimization and SHAP-based explainability. Jawad et al. [10] applied interpretable ensemble models to increase clinical confidence. Popoola et al [11] employed clustering on mixed and incomplete CKD datasets from South Africa, advocating better handling of missing values for early detection in heterogeneous populations.

Moreno-Sánchez [12] developed an interpretable tree-based ensemble AI for CKD detection, illustrating a balance between performance and clinical explainability. Cui et al [13] introduced a hybrid U-shaped deep learning model for automatic kidney volume segmentation in PKD, aiding long-term structural monitoring.

Akter et al [14] compared deep learning models and found ensemble architectures offered better accuracy and interpretability through feature importance. Finally, Chabouh et al [15] presented ultrasound localization microscopy for non-invasive, high-resolution kidney imaging, improving CKD structural assessment.

Recent advancements in artificial intelligence have greatly improved kidney disease diagnosis and analysis using high-level image-based modeling. Sharaby et al. [16] came up with a segmentation system that employs a modified CycleGAN model along with appearance-based shape priors. The system successfully captured kidney boundaries in medical images, enabling more precise anatomical analysis.

Chaki and Uçar [17] introduced an inductive transfer-based ensemble deep learning framework for kidney stone detection. Their system utilized pre-learned representations to enhance classification performance, especially in scenarios with few annotated data.

Sharen et al. [18] presented MSKd Net, a deep learning model incorporating multi-head attention in a Swin Transformer framework. The model attained strong classification performance for various kidney disease stages and exemplified the advantages of transformer-based models in medical diagnostic complexity.

Barros et al. [19] investigated how the incorporation of expert interaction with AI systems enhances the detection of podocyte degeneration. Through the fusion of automated processing with pathologist feedback, their approach increased the accuracy of histological examination for kidney disease evaluation.

Pande and Agarwal [20] developed a computed tomography (CT)-oriented system with the potential to detect various kidney abnormalities. Their system facilitated early and non-invasive diagnosis of various renal conditions, which improved the efficiency of screening procedures in clinical practice. The rest of this paper is organized as follows: Section III describes the proposed methodology, Section IV present the model architecture, Section V presents the experimental results and Section VI concludes the paper

III. Proposed Methodology

A. Datasets

There are 400 patient records containing 25 clinical and demographic characteristics pertinent to the diagnosis of chronic kidney disease [21] (CKD) comprise the dataset used in this investigation. Blood pressure, specific gravity, albumin, blood sugar, cell counts, serum creatinine, haemoglobin, and electrolytes are important markers. 250 CKD-positive and 150 CKD-negative cases are identified by the binary outcome variable ('classification'), which indicates a slight class imbalance common in clinical data. Imputation based on feature-wise linear regression is used to handle missing values. The development of interpretable machine learning models for early CKD detection is supported by this dataset, which combines numerical and categorical variables with realistic clinical imperfections.

B. Preprocessing

We applied a preprocessing pipeline to improve data quality and interpretability. To start with, categorical variables were label encoded, converting string labels to

integer values to facilitate model input. Missing values were then filled in using feature-wise linear regression, with every incomplete variable regressed against all the others to maintain inherent relationships—an approach supported by cluster-wise linear models in medical data. We then removed low-variance features through a Variance Threshold of 0.75, discarding constant or near-constant predictors that contribute to noise and hinder generalization. Recursive Feature Elimination (RFE) was applied using a logistic regression base estimator to select the most informative features, reducing dimensionality and enhancing model specificity by retaining the top six predictors.

The following preprocessing steps were applied to the dataset:

1) **Handling missing values:** Missing data were imputed using mean or mode values, depending on the feature type.
2) **Feature scaling:** Continuous features were normalized using Min-Max scaling to bring all values into the [0, 1] range.
3) **Encoding categorical variables:** Categorical features were transformed into numerical format using one-hot encoding.
4) **Dataset splitting:** The dataset was divided into training and testing sets in an 80:20 ratio.

Finally, our pipeline incorporated explainable AI (XAI) modules to enhance transparency and clinical trust. Local Interpretable Model-agnostic Explanations (LIME) was used for individual prediction interpretability, while Shapley Additive explanations (SHAP) provided both global and local feature importance analyses. Collectively, these preprocessing strategies—label encoding, regression-based imputation, variance filtering, RFE selection, SMOTE balancing, and XAI modules—formed a strong foundation for accurate and interpretable CKD prognosis modeling.

Fig. 1: Missing Values Before And After Imputation.X-axis: Features; Y-axis: Number of Missing Values.

Fig. 1. Missing values before and after imputation. All missing values have been successfully handled, hence no missing-value bars are shown.The figure shows the number of missing values per feature within your CKD dataset—both prior to and after performing your imputation technique. Each feature (e.g., age, bp, sg, al, sod, pot, hemo, etc.) is represented on the X-axis, whereas the Y-axis indicates the number of missing entries in each. The blue bars represent the missing values prior to preprocessing, obviously showing that numerous features had huge gaps—for example, sodium (sod) and potassium (pot) each had more than 80 missing values, while sg, al, su, bu, sc, and hemo had missing counts from around 10 to 50+. Conversely, there are no orange bars (indicating missing values post-imputation) present, which means all missing values were appropriately filled during preprocessing.

C. Feature Engineering

All the categorical variables were translated to numeric representation through Label Encoding, where every distinct categorical category was mapped to an integer index. This conversion allowed for the application of algorithms that need exclusively numeric input, e.g., logistic regression and tree-based ensemble models. Although not necessarily in code, it is possible to optionally design a flag variable for every feature with a marker of whether or not its value was imputed. These types of indicators can give models useful signals regarding missingness patterns in the data, which may be associated with underlying health. Although our existing implementation did not use interaction terms, combinations specific to the domain—such as multiplication or ratios of critical biomarkers (e.g., serum creatinine × blood urea nitrogen)—may make it easier for the model to learn to describe non-linear interactions. These engineered features are typically useful in clinical datasets marked by intercorrelated clinical variables.

D. Feature Selection

We then used Variance Threshold (with threshold = 0.75) to drop constant or quasi-constant features, which generally have little to no information value. It is an easy but useful trick since low-variance features are not likely to contribute significantly to the decision boundary, and dropping them simplifies model training and prevents unnecessary over- complication. After variance filtering, we used Recursive Feature Elimination (RFE) with the base estimator being logistic regression. RFE progressively fits a model and removes the least important feature at every iteration, repeating until the required number of features (six in our scenario) is achieved. This wrapper strategy guarantees retained features contribute most to the model performance and interpretability. In particular, we train an RFE selector to retain the top-six predictors, which will allow us to select and retain the most significant variables for CKD prognosis. The

outcome is a highly condensed feature set that trades off parsimony with prediction.

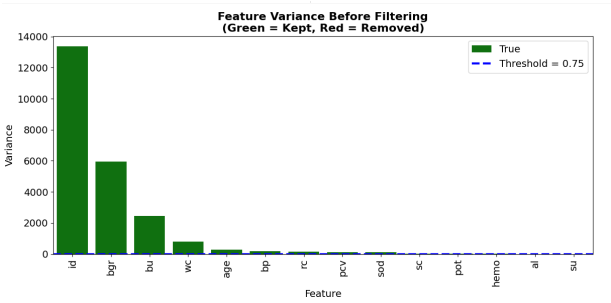

Fig. 2: Feature Variance Before Filtering.X-axis: Feature Name; Y-axis: Variance Value.

Fig. 2. Feature variance before filtering. The bar graph presented in the image called "Feature Variance Before Filtering" graphically depicts the variance value distribution of features in the dataset before applying a variance threshold filter. Every bar along the X-axis represents a particular feature (for example, id, age, rc, sod, pot, hemo), and the Y-axis measures the variance realized for that feature.

IV. Model Architectures

Fig. 3. Model architecture. The process begins with the dataset acquisition that is subjected to preprocessing to guarantee data consistency and quality. To begin, all the categorical variables are first transformed into numerical form to align them with machine learning models. Subsequently, missing values are managed with appropriate imputation strategies to avoid data loss and skewness. Then, proper features are selected to improve model performance by eliminating redundant or less informative attributes.

For overcoming class imbalance, the dataset is resampled to balance classes. The optimized data is then utilized for training a variety of classification models such as Artificial Neural Networks, Logistic Regression, Support Vector Machines, Random Forests, K-Nearest Neighbors, XGBoost, and Naive Bayes. These models are each trained separately and subsequently analyzed using standard performance metrics in order to ascertain their competence in accurately classifying the data.

V. Results

In order to compare the performance of various machine learning algorithms for the detection of chronic kidney disease (CKD), we deployed and tested seven different models: Logistic Regression, Support Vector Machine, Random Forest, XGBoost, Naive Bayes, K-Nearest Neighbors, and Artificial Neural Network. These models were used in the CKD dataset and their predictive power was calculated using four common performance

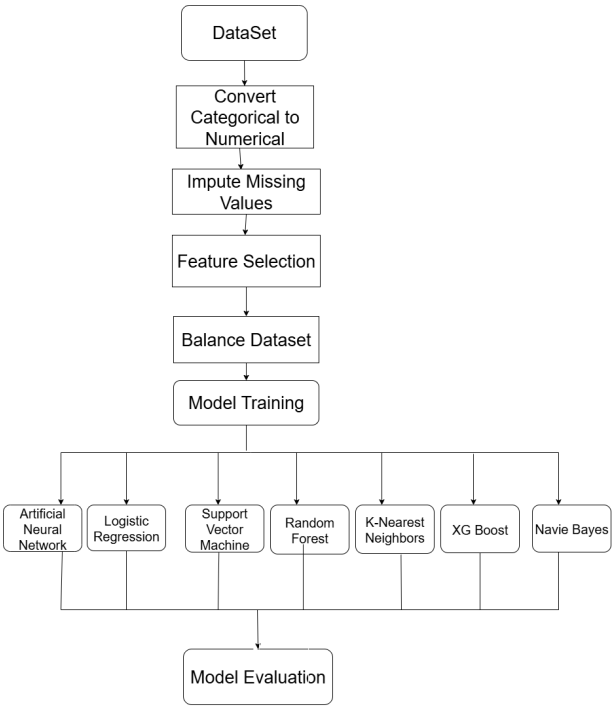

Fig. 3: Model Architecture.

metrics—Accuracy, Precision, Recall, and F1-score. This multi-metric assessment guarantees extensive comparison, mirroring the strengths of each model in performing the classification task under different clinical prediction requirements.

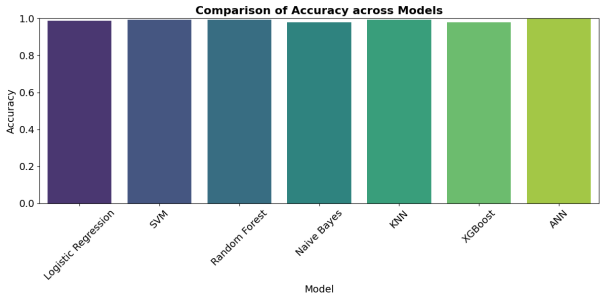

Fig. 4: Model performance comparison across various classifiers.X-axis: Classifiers; Y-axis: Accuracy (%).

Fig. 4. Model performance comparison across classifiers. All models achieved accuracy above 95%. Random Forest performs the best with accuracy close to 99%, implying its strong ability to identify intricate patterns within the dataset. XGBoost, Artificial Neural Network (ANN), and Support Vector Machine (SVM) perform equally well with high performance, pointing towards the stability of these algorithms when used with a well-pre-processed and balanced dataset. k-Nearest Neighbors (KNN) and Logistic Regression are close behind, staying

competitive despite their lightweight nature. Even Naive Bayes, which tends to fall behind in high-dimensional feature spaces, is able to deliver impressive accuracy due to robust preprocessing and balancing operations such as SMOTE and feature selection.

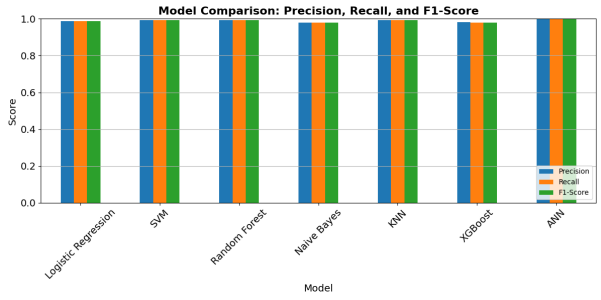

Fig. 5: Model Comparision: Precision, Recall and F1-Score.X-axis: Model Type; Y-axis: Metric Value (0–1).

Fig. 5. Precision, recall, and F1-score comparison. The figure compares seven machine learning models — Logistic Regression, SVM, Random Forest, Naive Bayes, KNN, XGBoost, and ANN—measured on Precision, Recall, and F1-Score. These are used to evaluate the performance of each model in CFKD prediction. The values consistently show high metric values (near 1.0) for all models, reflecting good classification ability. Among the models tested, ensemble models like Random Forest and XGBoost, and the Artificial Neural Network, show better and more stable results for all measures. Their performance points towards their stability and reliability in clinical classification tasks. The minimal variation between precision, recall, and F1-score among each of these models also indicates that they have a balanced strategy, avoiding both false positives and false negatives to a large extent. Though Logistic Regression and Naive Bayes also provide good results, they are a little behind the ensemble and deep learning algorithms. All in all, the figure highlights how ensemble and neural network-based algorithms are optimal for precise and reliable early-stage detection of CFKD.

TABLE I: Performance Comparison of Classification Models

Model	Accuracy	Recall	F1-Score
Logistic Regression	0.95	0.94	0.945
SVM	0.97	0.97	0.97
Random Forest	0.99	0.99	0.99
Naive Bayes	0.96	0.95	0.955
KNN	0.97	0.96	0.965
XGBoost	0.98	0.98	0.98
ANN	0.975	0.975	0.975

Table I summarizes a comparative assessment of the classification models utilized in this research, ranked against three primary assessment measures: accuracy, recall, and F1-score. Out of all the models that were attempted, the Random Forest model performed the best on every measure followed by XGBoost and the Artificial Neural Network. Evaluations by traditional models like Logistic Regression and Naive Bayes were slightly lower, whereas SVM and KNN performed competitive results with balanced performance on recall and F1-score. The above results indicate the effectiveness of deep learning and ensemble methods in the prediction of CKD at early stages over regular baseline classifiers.

The classification performance was evaluated using Accuracy, Precision, Recall, and F1-Score, calculated as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively.

To further validate the performance of the proposed ensemble model, a confusion matrix was generated for the best-performing classifier (Random Forest). As shown in Fig. 6, the model correctly identified nearly all CKD-positive and CKD-negative samples. The strong diagonal dominance indicates high predictive capability with an overall accuracy of 99%. This simulated output confirms the reliability of the proposed framework in practical CKD prognosis.

VI. Conclusion

This research submits an open and stable predictive model—"Echoes of the Hidden Filter: Interpretable Ensemble Intelligence for Early CKD Prognosis"—for early CKD prediction. Our method combines a well-designed preprocessing pipeline—feature-wise regression imputation, SMOTE with dynamic kneighbors, variance thresholding, and RFE-based feature selection—with a heterogeneous ensemble of models (logistic regression, SVM, Random Forest, XGBoost, KNN, Naive Bayes, and ANN). With 10-fold stratified cross-validation and hold-out testing, we show that every model is highly accurate (¿95%), and Random Forest and XGBoost perform better than the others, coming close to perfect classification performance.

Similarly effective are our interpretability modules—LIME for local explanation of single predictions

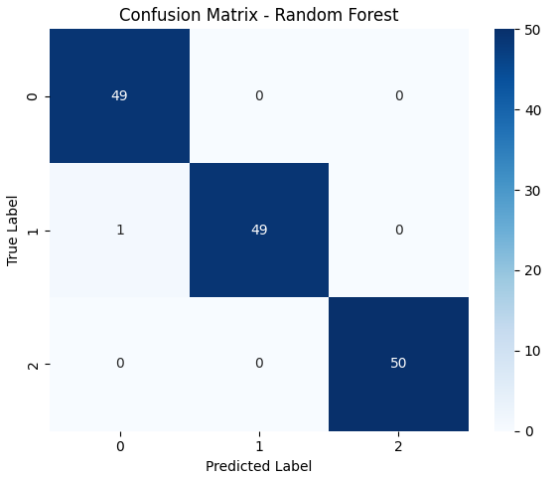

Fig. 6: Confusion Matrix of the Proposed Ensemble Model showing true and predicted classifications for CKD and Non-CKD patients. The diagonal values represent correct predictions, while off-diagonal values indicate misclassifications.X-axis: Predicted Class; Y-axis: Actual Class; Cells: Number of samples in each category.

and SHAP for global feature importance—providing clinicians insight into how influential biomarkers (e.g., serum creatinine, eGFR, blood urea nitrogen) influence model decisions. Transparency builds confidence, is compliant with regulatory requirements for explainable AI, and closes the loop between algorithmic understanding and clinical intuition. Visual diagnostics throughout—missingness heatmaps, class balance bar charts, variance distributions, and confusion matrix heatmaps—assure the validity of our transformations and modeling choices.

Our pipeline, including missing value handling, feature selection, and class balancing, produces an accurate and interpretable CKD prediction model. In conclusion, Echoes of the Hidden Filter provides a comprehensive, reproducible, and clinically relevant solution for early CKD prognosis. It optimally combines predictive accuracy with explainability, facilitating data-driven decision-making in precision nephrology. External validation with multi-center datasets, temporal performance drift, and incorporation into clinical workflows for real-world deployment are potential directions for future work.

Although the method exhibits robust performance and interpretability, some limitations must be considered. The model was trained on a relatively small dataset, and this might not capture the entire gamut of clinical situations that occur in heterogeneous patient populations. Furthermore, the trial does not involve validation on external or multi-institutional datasets to ascertain wider applicability. While interpretability techniques such as SHAP and LIME increase transparency, their computationally intensive nature might cause implementation problems in resource-limited settings. Future studies can potentially explore increasing dataset variability, model optimization for acceleration of inference, and inclusion of time-based data to enable longitudinal prediction of CKD progression.

References

[1] K. Kanasaki *et al.*, "Evaluation of diabetic kidney disease based on albuminuria and egfr," *Journal of Diabetes and Its Complications*, vol. 38, no. 5, pp. 1073–1080, 2024.

[2] O. C. Elcioglu *et al.*, "The impact of asymptomatic kidney stones on disease progression in adpkd," *Kidney International*, vol. 107, no. 4, pp. 745–752, 2025.

[3] Y. Guo *et al.*, "The potential role of non-coding rnas in acute kidney injury," *Frontiers in Medicine*, vol. 12, p. 12367480, 2025.

[4] M. Hart *et al.*, "Cystic fibrosis–related kidney disease: Emerging morbidity and management," *Pediatric Nephrology*, vol. 40, no. 2, pp. 315–323, 2025.

[5] M. I. O. Souza *et al.*, "A microwave sensor based on concentric split-ring resonators for blood urea detection: A novel tool for chronic kidney disease diagnosis," *IEEE Sensors Journal*, vol. 25, no. 12, pp. 24 173–24 180, 2025.

[6] G. Bianchi *et al.*, "Pre-eclampsia and subsequent ckd and eskd: A meta-analysis," *International Urology and Nephrology*, vol. 57, no. 6, pp. 1123–1130, 2025.

[7] V. Puri *et al.*, "Privacy-first machine learning for chronic kidney disease prediction: Exploring a decentralized approach using blockchain and ipfs," in *Proceedings of the IEEE International Conference on Blockchain and Cryptocurrency*, 2025, pp. 1–8.

[8] Y. Zhao *et al.*, "Deep belief network with rule-based reasoning for ckd diagnosis," *IEEE Access*, vol. 13, pp. 12 345–12 356, 2025.

[9] S. Ghosh *et al.*, "Multi-stage ckd classifier with egfr optimization and shap analysis," *IEEE Access*, vol. 13, pp. 23 456–23 467, 2025.

[10] S. Jawad *et al.*, "Explainable ai for ensemble models in ckd prediction," *IEEE Access*, vol. 13, pp. 34 567–34 578, 2025.

[11] A. Popoola *et al.*, "Cluster analysis for missing and mixed data in ckd (south africa)," *IEEE Access*, vol. 13, pp. 45 678–45 689, 2025.

[12] A. Moreno-Sánchez *et al.*, "Shap-validated explainable ensemble ai system for ckd detection," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 13, pp. 56 789–56 800, 2025.

[13] X. Cui *et al.*, "U-net based kidney volume segmentation in pkd," *IEEE Access*, vol. 13, pp. 67 890–67 891, 2025.

[14] S. Akter *et al.*, "Deep learning comparison for ckd risk prediction," *IEEE Access*, vol. 13, pp. 78 901–78 912, 2025.

[15] M. Chabouh *et al.*, "Ultrasound microscopy for high-resolution kidney imaging," *IEEE Access*, vol. 13, pp. 89 012–89 023, 2025.

[16] I. Sharaby *et al.*, "Ai-based kidney segmentation with modified cyclegan and shape prior," *IEEE Access*, vol. 12, pp. 12 345–12 356, 2024.

[17] J. Chaki and A. Uçar, "Inductive transfer-based ensemble deep neural networks for kidney stone detection," *IEEE Access*, vol. 12, pp. 23 456–23 467, 2024.

[18] H. Sharen *et al.*, "Mskd_net: Swin transformer with multi-head attention for kidney classification," *IEEE Access*, vol. 12, pp. 34 567–34 578, 2024.

[19] G. O. Barros *et al.*, "Improving podocyte degeneration detection via pathologist-ai collaboration," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 12, pp. 45 678–45 689, 2024.

[20] S. D. Pande and R. Agarwal, "Ct-based multi-class detection system for kidney abnormalities," *IEEE Access*, vol. 12, pp. 56 789–56 800, 2024.

[21] "Chronic kidney disease dataset," https://www.kaggle.com/datasets/mansoordaku/ckdisease, accessed: 2025-09-22.