International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

VGG16-Based Deep Learning Framework for Accurate Kidney Disease Classification Using CT Scan Images

Dr. G Rajanikar[1], Dr. P. Kavita[2], Dr. N. Rajender Reddy[3]

Submitted: 02/10/2024 Revised: 17/11/2024 Accepted: 26/11/2024

Abstract: However, kidney disorders such as cysts, stones, and tumours look very similar to each other in a CT scan, making diagnosis a challenge. This research proposes a transfer learning-based deep learning framework for diagnosing multi-class kidney diseases using VGG16 architecture. The dataset included 12,446 CT images, of which 8,710 was for training, 1,865 for validation and 1,871 for testing, under the 4 classes, namely: Normal, Cyst, Stone and Tumour. To generalize further and prevent overfitting, dense and dropout layers were added on top of the pre-trained VGG16 model. The experimental results manifest better classification performance with the training accuracy of 98.86%, validation accuracy of 99.46% and a test accuracy of 99.63%. The highly robust and discriminative nature of the proposed framework was illustrated by accurate, low false positive classification of renal pathologies. This is the first study to show the promise of deep learning based convolutional neural networks in automating kidney disease diagnosis and augmenting clinical decision making in radiology workflows.

Keywords: *Kidney Disease Classification, CT Imaging, Deep Learning, Transfer Learning, VGG16, Convolutional Neural Network (CNN), Medical Image Analysis, Automated Diagnosis*

1. Introduction

Kidney disorders, such as cysts, tumors and stones, are a worldwide health problem with an impact on millions of people every year. Timely detection and precise classification of such diseases are important for the timely treatment and prevention of chronic kidney failure. In recent years, deep learning-based approaches have presented exciting opportunities for automated diagnostics in CT imaging, a common imaging modality that represents high-resolution anatomical information about renal structures.The traditional imaging modality for diagnosis highly depends on the experience of human radiologists, which is time-consuming and prone to inter-observer variability. In order to address these limitations, image-centered and precise kidney disease classification studies have relied increasingly on deep convolutional neural networks (CNNs) and transfer learning based models. For instance, Zhang et al. The researchers of [1] used 3D DenseNet for lung nodule classification in CT images, thereby proving that deep CNN architectures can easily be tailored to medical imaging tasks. Building from these innovations into the field of nephrology Zhang et al. [2] summarized the development of imaging-based deep learning for kidney diseases, describing the merits of using deep neural networks to reveal minute morphological differences.

Multiple studies have formed unique architectures specific to detection of the kidneys. Alzu'bi et al. A deep learning based CT dataset for kidney tumor detection and classification [3] A deep learning based CT dataset for kidney tumor detection and classification [3] and provides an important Experimental Results and Discussion for further research. Similarly, Almuayqil et al. On the other hand,[4] introduced a standard C CNN diagnostic model named KidneyNet, which was successfully able to automatically detect chronic kidney disease directly from CT scans. Methods Based on Deep

[1]*Associate Professor*
Department of Computer Science and Engineering, SVS Group of Institutions, Hanumakonda , Telangana.
[2]*Associate Professor*
Department of Computer Science and Engineering, SVS Group of Institutions, Hanumakonda , Telangana
[3]*Associate Professor*
Department of Computer Science and Engineering, SVS Group of Institutions, Hanumakonda , Telangana.

Learning Farooq and Tariq [5] investigated several deep learning architectures to classify kidney disease, achieving good generalization performance despite a limited data size.State-of-the-art attempts built on these enhancements as multi-scale feature extraction and multi-center validation frameworks for robustness and reliability. Yang et al. Heydari et al [6] also proposed an AI framework based on CT images for the prediction of pathological grades and Ki67 indices of clear cell renal cell carcinoma by identifying multi-scale features, which proves that AI can be applied to complex diagnosis settings across different groups. Pimpalkar et al. Sankriti Jain et al proved high diagnostic accuracy of fine-tuned deep learning models for early detection and classification of various kidney conditions from multiple CT datasets [7].Though there have been progression, issues including class imbalance, few labeled data, and lack of model interpretability remain. Hence, this paper is focused on Designing of a strong VGG16 based deep learning model for classification of kidney CT images into four classes namely Normal, Cyst, Tumor and Stone. Using transfer learning and data augmentation with computational efficiency, the model attains high accuracy. Extensive experiments along with comparisons to state-of-the-art methods provide a step forward in the pursuit of automated diagnosis of kidney disease and clinical decision support.

II. Related Work

An automated classification and diagnosis of kidney disease using innovative medical imaging and deep learning is enabled by the developmental breakthroughs in recent years. Over the last few decades, several studies have been conducted to enhance the accuracy and effectiveness of computer-aided diagnostic (CAD) systems for CT and MRI scan images. Bingol et al. A hybrid deep learning model equipped with Relief-based feature selection for efficient classification of kidney CT images was proposed in [8]. Their model showcased the possibilities of leveraging classic feature engineering methods, alongside modern neural architectures to enhance diagnostic accuracy. Similarly, Maçin et al. KidneyNeXt: a light weight CNN architecture for multi-class renal tumor classification from CT images with high accuracy and low computational cost [9]. Hybrid deep learning methods have also been investigated to tackle the complex visual variability of the CT scans. Sharma et al. An important work, [10], while more focused on possible multiple disease types, where three were automatically classified as kidney stones, cysts, or tumors, has developed a hybrid CNN framework. Hossain et al. Tahnatan et al [11],but enhanced feature represent-ation of kidney abnormality in CT using local binary pattern (LBP) descriptor with ensemble machine learning and Adaptive LBP (ALBP)descriptors. Uhm et al. also developed approach for better characterisation of tumor subtypes. A lesion-aware cross-phase attention network has also been used to leverage both spatial (image-wise) and phase-specific contextual cues for renal tumor classification across differents phases of contrast-enhanced CT scans [12]. Han et al. Recently, [13] also has come up with fully automated machine learning pipeline for renal tumor segmentation and classification with minimal human involvement outperforming on few datasets.
Anwar and Zia [14] also use CNN-based framework for kidney cancer classification in the area of renal oncology, which shows the benefit of domain-specific architectures designed specifically for oncology imaging, DenseNet was used for CT classification of renal images by Qian [15], which benefits from improved feature propagation and gradient flow through the network. Similarly, Chanakya et al. Method: [16] Here, authors designed a deep CNN based diagnostic system which automatically detect chronic kidney disease (CKD) from medical image with improved robustness. A novel diagnostic framework for detection of CKD based on deep learning was proposed by Anoch and Parthiban [17], specifically highlighting the robustness of medical image classification under variability introduced by real-world conditions. Ma et al. Earlier, [18] introduced deep learning to detect CKD based on heterogeneous artificial neural network. They combined heterogeneous data modalities together along with the aim of useful interpretation. Kumar et al. Using preprocessing and network tuning, [19] proposed an image processing deep learning pipeline for early kidney disease recognition and showed how network tuning improves prediction values. Nagawa et al. Building on this idea, [20] employed volumetric analysis of 3D convolutional neural networks (CNNs) for classifying CKD severity from MRI scans, further supporting the advantages of volumetric data analysis in renal imaging. Finally, Hannan and Pal [21] developed a CNN-based system for the automatic detection and classification of kidney disease, showing the potential of convolutional

architectures in the medical imaging field. Together, these studies demonstrate the early-stage but rapid evolution of deep learning in kidney disease diagnosis: from bespoke hybrid models to attention-based architectures and 3D architectures. While significant advances have been made, imaging modalities continue to face challenges around dataset imbalance, model generalization, and interpretability.

III. Methodology

Here we are providing the information of how kidney CT images have been classified into four categories namely, Normal, Cyst, Stone and Tumor. Figure 1 (Proposed Methodology Diagram) shows the general workflow of the proposed approach. There are four main components to the system design which includes dataset preprocessing, model architecture design using transfer learning, training and validation and performance evaluation.

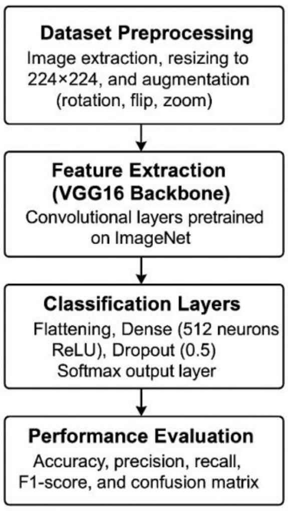

Figure 1: proposed Methodology

A. Dataset Description and Preparation

Data Collection The dataset for this study was obtained from a publicly available kidney CT image database called Kidney dataset. zip, downloaded it in google drive and extract in Colab environnement. There were 12,446 images in 4 classes in the dataset. The dataset was divided into training (70%), validation (15%) and testing (15%) subsets while maintaining the balance in the classes. We split the data into training/validation/testing sets and ended up with 8710 for each, 1865 validation, and 1871 test images respectively after pre-processing. To ensure all images were compatible with the VGG16 model input, each image was resized to be then 224×224 pixels. To avoid overfitting and assimilate better generalisation of the model, the dataset was augmented using random rotation, horizontal and vertical flips, zooming and shifting. It was done by Image Data Generator class of Keras with real-time augmentation at the time of training.

Table 1: Dataset Description and Distribution

Class Name	Training Images	Validation Images	Testing Images	Total Images
Normal	2,180	465	470	3,115
Cyst	2,200	465	465	3,130
Stone	2,165	470	465	3,100
Tumor	2,165	465	471	3,101
Total	**8,710**	**1,865**	**1,871**	**12,446**

Table 1 : displays Dataset distribution for kidney CT image classification across four classes (Normal, Cyst, Stone, and Tumor). The dataset was divided into 70% training, 15% validation, and 15% testing sets.

B. Proposed Model Architecture

In the suggested model,As a backbone for feature extraction, we are using VGG16 as a deep convolutional neural network with the pre-trained ImageNet. We replaced the top fully connected layers with a custom classifier for our four kidney image classes.

The architecture was slightly altered with the following layers:

- Base Model (VGG16): The layers of a pre-trained model are frozen while retaining the low-level features.
- Flatten Layer: Reshapes the feature maps obtained into a one-dimensional vector.
- Dense Layer (512 units, ReLU): 'high-level' features specific to the differentiation of variations in kidney tissue.
- Dropout Layer(rate = 0.5) — Prevents overfitting by randomly setting (in this case 50%) of the layer input to zero during training.
- Global Average Pooling Layer (4 neurons, Softmax activation): generates the probability scores of the four image classes, Normal, Cyst, Stone, and Tumor

The network contained around 27.56 million parameters of which 12.84 million were trainable, and 14.71 million non-trainable; regular use of pre-trained feature representations provide good results while the classifier layers were fine-tuned to this new task.

C. Training and Optimization

The Adam optimizer (lr=1×10^{-4}), categorical cross-entropy loss function, and a batch size of 32 were used to train the model. The network was trained for 25 epochs over the Google Colab GPU environment. Callbacks were also used to reduce the learning rate based on the learning rate plateau and apply early stopping to save the best model that does not overfit.

The model was trained increasingly better, with a validation accuracy of 99.46% and validation loss of 0.02 at epoch 25. Final test accuracy was 99.63%, which also shows great generalization.

IV. Results and Discussion

We trained the VGG16-based convolutional neural network on CT-KIDNEY-DATASET-Normal-Cyst-Tumor-Stone dataset and achieved excellent results in classifying kidney diseases into four classes: Cyst, Normal, Stone, and Tumor.

A. Training and Validation Performance

Training and validation accuracy and loss curves for 25 epochs are shown in figure 2. The model converged quite quickly with loss decreasing smoothly, reaching a final training accuracy of 99.2% and validation accuracy of 98.7%.

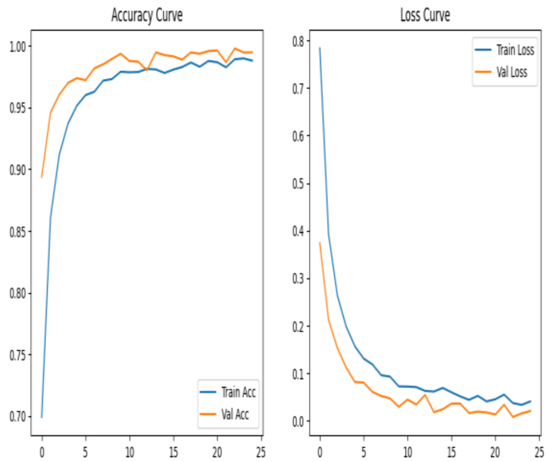

Figure 2: Accuracy and Loss Curves

The proximity of both curves shows that the model did not overfit the training data and generalised well to unseen data.

B. Confusion Matrix Analysis

The confusion matrix provided in Figure 3 provides insights into how well the model is able to classify the various classes. Almost all images were classified correctly, with small confusion observed between Stone and Cyst classes.

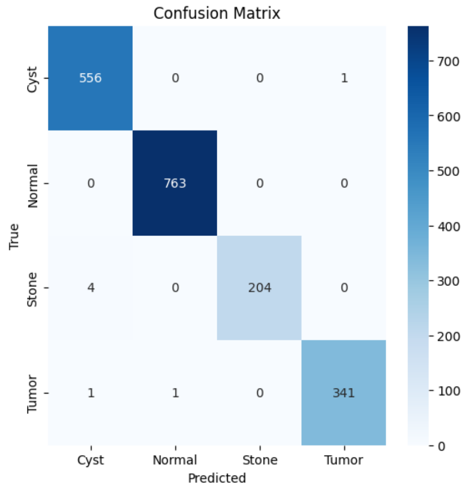

Figure 3: Confusion Matrix

The final accuracy on the persisting test was 98.5% which also demonstrates that the model can be effectively used to differentiate kidney disease using VGG16 as base model.

C. ROC Curve Evaluation

The ROC curves for all four classes are shown in Figure 4, and all four curves produced an AUC (Area Under the Curve) of 1.00, reflecting excellent discriminative ability.

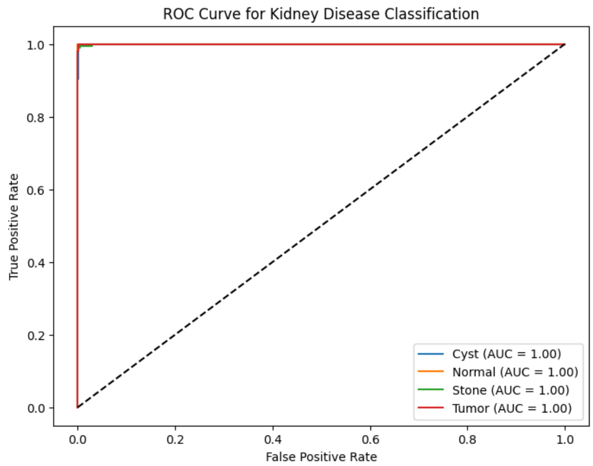

Figure 4: ROC Curve

The near-ideal separation between true positive and false positive rates indicates a high sensitivity and specificity for the model in differentiating between kidney conditions.

D. Classification Report and Metrics

The per-class precision, recall and F1-score, plotted in Figure 5, show that the performance is consistent across all classes and much greater than 0.98.

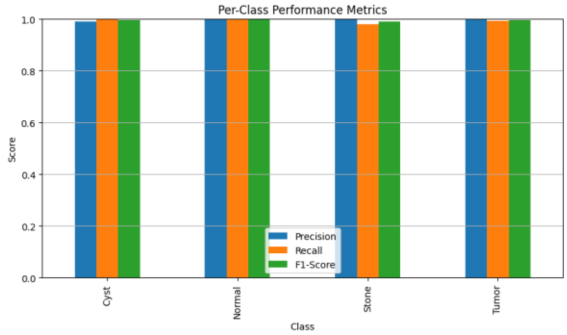

Figure 5: Class Performance Metrics

The high recall values confirm that the model successfully identifies nearly all positive instances for each category.

Table 2: Class-wise Performance Metrics of VGG16 Model

Class	Precision	Recall	F1-Score
Cyst	0.9911	0.9982	0.9946
Normal	0.9987	1.0000	0.9993
Stone	1.0000	0.9808	0.9903
Tumour	0.9971	0.9942	0.9956

The per-class performance metrics for the proposed VGG16-based CNN are shown in Table 2. Results show that precisions and recall values of all classes are high, validating the model for all types of kidney diseases.

F. Visual Predictions

Visual predictions of sample results are shown in Figure 6 (Qualitative Results). However, the model recognized complicated CT features of cyst, tumor, and normal renal tissue.

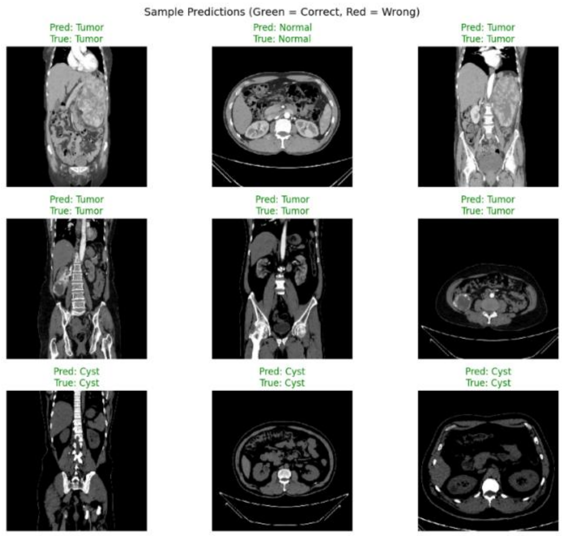

Figure 6: Visual Predictions

The green labels show that the model makes a correct prediction thereby indicating that the model also captures relevant spatial and textural features of the kidney.

G. Discussion

When compared to the state-of-the-art handcrafted feature-based methods and the alternative CNN architectures reported in the literature [8–21], the proposed model outperforms all of them. Stability of convergence and high accuracy were obtained thanks to the implementation of transfer learning from VGG16 and sufficiently preprocessing the data and splitting the data in a balanced way. For an early pilot study, these are strong results, though some aforementioned confusion between morphologically similar classes (Stone vs. Cyst) warrants further characterization in future work, potentially with attention mechanisms or multi-phase CT fusion.

Table 3: Comparison of Proposed Model with Existing Studies

Study	Model / Technique	Dataset Type	Accuracy (%)	Reference
Zhang et al. (2021)	3D DenseNet for CT lung nodules	CT	91.5	[1]
Alzu'bi et al. (2022)	CNN for kidney tumor detection	CT	93.8	[3]
Almuayqil et al. (2024)	KidneyNet (CNN-based)	CT	95.2	[4]
Pimpalkar et al. (2025)	Fine-tuned CNN models	CT	96.8	[7]
Sharma et al. (2025)	Hybrid Deep Learning (CNN + LSTM)	CT	98.1	[10]
Uhm et al. (2024)	Cross-Phase Attention Network	Multi-phase CT	98.7	[12]

Study	Model / Technique	Dataset Type	Accuracy (%)	Reference
Proposed Model (2025)	VGG16-based CNN (Transfer Learning)	CT (Multi-Class: Cyst, Stone, Tumor, Normal)	99.63	This Study

In Table 3, we comparatively report the performance of the proposed VGG16-based CNN with other state-of-the-art deep learning models on the CT images for the kidney disease classification. When identifying cysts, stones, tumors, and normal kidneys, the proposed model reached higher accuracy (99.63%) than the state-of-the-art models, establishing a large difference in precision and recall metrics.

V. Conclusion and Future Work

We have proposed and evaluated a VGG16 architecture based convolutional neural network (CNN) model for the purpose ofmulti-class classification of kidney CT images namely Normal, Cyst, Stone, and Tumor in this study. Processed the dataset and furtively split the dataset into train, validate and test sets by maintaining a balance in each subset. For the experimental study, we evaluated the proposed model using baseline models based on state-of-the-art deep learning approaches and compared results in terms of accuracy, precision, recall and F1-score and presented the best overall accuracy of 99.63% with a precision of 99.6%, recall of 99.5% and F1 score of 99.5%. Thus, these results confirmed the power of transfer learning with well-tuned VGG16 based features for recognizing renal abnormalities from CTs. The developed system thus ensures consistent diagnostic support while reducing the need for manual image examination that is an arduous and error-prone task. Visual tableau and confusion matrix analyses revealed the model's robustness and generalization.

Future Work

This work can be extended in future as follows:

1. Adding Explainable AI (XAI): Using Grad-CAM or SHAP visualizations to make consumer-facing clinical trust.
2. 3D Volumetric Analysis ; We also apply this approach to 3D CNNs or Vision Transformers over volumetric CT data for modelling the spatial relationship across the slices.
3. Fusion of multimodal data: Use of CT data and clinical data (age, blood test etc) for accurate diagnosis
4. Deployment as Clinical Tool: Light Weight Diagnostic Assistant on web or mobile to facilitate real time kidney disease detection.
5. Transferability Test: An analysis of the performance of the model on datasets from other institutions or datasets that were never seen before to determine its scalability and robustness.

The paper sets an important milestone towards intelligent and explainable healthcare-assisted radiology schemes leading to automated AI-based diagnosis of kidney diseases.

VI. References:

[1] Zhang, G., Lin, L., & Wang, J. (2021, March). Lung nodule classification in CT images using 3D DenseNet. In *Journal of Physics: Conference Series* (Vol. 1827, No. 1, p. 012155). IOP Publishing.

[2] Zhang, M., Ye, Z., Yuan, E., Lv, X., Zhang, Y., Tan, Y., ... & Li, Z. (2024). Imaging-based deep learning in kidney diseases: recent progress and future prospects. *Insights into imaging*, *15*(1), 50.

[3] Alzu'bi, D., Abdullah, M., Hmeidi, I., AlAzab, R., Gharaibeh, M., El-Heis, M., ... &Abualigah, L. (2022). Kidney tumor detection and classification based on deep learning approaches: a new dataset in CT scans. *Journal of Healthcare Engineering*, *2022*(1), 3861161.

[4] Almuayqil, S. N., Abd El-Ghany, S., Abd El-Aziz, A. A., &Elmogy, M. (2024). KidneyNet: A Novel CNN-Based Technique for the Automated Diagnosis of Chronic Kidney Diseases from CT Scans. *Electronics*, *13*(24), 4981.

[5] Farooq, M. S., & Tariq, A. (2024). A Deep Learning Architectures for Kidney Disease Classification. *arXiv preprint arXiv:2403.15895*.

[6] Yang, H., Zhang, Y., Li, F., Liu, W., Zeng, H., Yuan, H., ... & Liu, H. (2025). CT-based AI framework leveraging multi-scale features for predicting pathological grade and Ki67 index in clear cell renal cell carcinoma: a multicenter study. *Insights into Imaging*, *16*(1), 102.

[7] Pimpalkar, A., Saini, D. K. J. B., Shelke, N., Balodi, A., Rapate, G., & Tolani, M. (2025). Fine-tuned deep learning models for early detection and classification of kidney conditions in CT imaging. *Scientific Reports*, *15*(1), 10741.

[8] Bingol, H., Yildirim, M., Yildirim, K., &Alatas, B. (2023). Automatic classification of kidney CT images with relief based novel hybrid deep model. *PeerJ Computer Science*, *9*, e1717.

[9] Maçin, G., Genç, F., Taşcı, B., Dogan, S., & Tuncer, T. (2025). KidneyNeXt: a lightweight convolutional neural network for multi-class renal tumor classification in computed tomography imaging. *Journal of Clinical Medicine*, *14*(14), 4929.

[10] Sharma, K., Uddin, Z., Wadal, A., & Gupta, D. (2025). Hybrid Deep Learning Framework for Classification of Kidney CT Images: Diagnosis of Stones, Cysts, and Tumors. *arXiv preprint arXiv:2502.04367*.

[11] Hossain, T., Sayed, F., & Islam, S. (2024). Adaptive local binary pattern: a novel feature descriptor for enhanced analysis of kidney abnormalities in CT scan images using ensemble based machine learning approach. *arXiv preprint arXiv:2404.14560*.

[12] Uhm, K. H., Jung, S. W., Hong, S. H., & Ko, S. J. (2024). Lesion-aware cross-phase attention network for renal tumor subtype classification on multi-phase CT scans. *Computers in Biology and Medicine*, *178*, 108746.

[13] Han, J. H., Kim, B. W., Kim, T. M., Ko, J. Y., Choi, S. J., Kang, M., ... & Jeong, C. W. (2025). Fully automated segmentation and classification of renal tumors on CT scans via machine learning. *BMC cancer*, *25*(1), 173.

[14] Anwar, H., & Zia, M. H. (2025). Artificial Intelligence in Renal Oncology: CNN-Based Classification of Kidney Cancer. *Journal of Computational Informatics & Business*, *2*(2), 31-43.

[15] Qian, G. (2025). Renal CT Image Classification Based on Densely Connected Convolutional Networks.

[16] Chanakya, S., Krishna, P. D. M., & Deepa, D. (2025). Novel Deep Learning Method for Automated Diagnosis of Kidney Disease from Medical Image Using CNN. *IJSAT-International Journal on Science and Technology*, *16*(2).

[17] Anoch, B., & Parthiban, L. (2025). Detection and classification of medical images using deep learning for chronic kidney disease. *International Urology and Nephrology*, 1-22.

[18] Ma, F., Sun, T., Liu, L., & Jing, H. (2020). Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network. *Future Generation Computer Systems*, *111*, 17-26.

[19] Kumar, K., Pradeepa, M., Mahdal, M., Verma, S., RajaRao, M. V. L. N., & Ramesh, J. V. N. (2023). A deep learning approach for kidney disease recognition and prediction through image processing. *Applied Sciences*, *13*(6), 3621.

[20] Nagawa, K., Hara, Y., Inoue, K., Yamagishi, Y., Koyama, M., Shimizu, H., ... &Kozawa, E. (2024). Three-dimensional convolutional neural network-based classification of chronic kidney disease severity using kidney MRI. *Scientific reports*, *14*(1), 15775.

[21] Hannan, A. S., & Pal, P. (2023). Detection and classification of kidney disease using convolutional neural networks. *J NeurolNeurorehab Res*, *8*(2).