American Journal of Smart Technology and Solutions (AJSTS)

ISSN: 2837-0295 (ONLINE)

VOLUME 5 ISSUE 2 (2026)

PUBLISHED BY
E-PALLI PUBLISHERS, DELAWARE, USA

American Journal of Smart Technology and Solutions (AJSTS)

Volume 5 Issue 2, Year 2026
ISSN: 2837-0295 (Online)
DOI: https://doi.org/10.54536/ajsts.v5i2.8226
https://journals.e-palli.com/home/index.php/ajsts

Deep Learning-Based Detection of Renal Abnormalities from CT Images: A Study on a Bangladeshi Cohort

Mithila Yeasmin Mitu[1*]

Article Information

Received: May 02, 2026

Accepted: July 18, 2026

Published: August 14, 2026

Keywords

Artificial Intelligence, Computed Tomography (CT), Computer-Aided Diagnosis, Deep Learning, Renal Abnormalities

ABSTRACT

This study presents an AI-driven computed tomography (CT) diagnostic system for the automated detection and classification of renal abnormalities in the Bangladeshi population. Renal abnormalities such as kidney cysts, stones, and tumors require timely and accurate diagnosis to reduce complications and improve treatment outcomes. In Bangladesh, the increasing burden of kidney-related diseases has created a strong need for efficient and intelligent diagnostic support systems. To assess the effectiveness of deep learning in multiclass renal abnormality diagnosis, three advanced convolutional neural network architectures were implemented and compared: Xception, VGG16, and ResNet152V2. The experimental results demonstrated outstanding classification performance across all models, with Xception achieving the highest accuracy of 99.84%, followed by ResNet152V2 at 99.76%, and VGG16 at 99.40%. Among the evaluated approaches, Xception showed the best overall performance, indicating its strong capability for reliable renal abnormality classification from CT images. The proposed system has significant potential to assist radiologists and healthcare professionals by providing fast, accurate, and automated diagnostic support. This work highlights the promise of artificial intelligence in medical imaging and contributes to the advancement of intelligent diagnostic solutions for kidney disease detection in Bangladesh.

INTRODUCTION

Kidney diseases and renal abnormalities have become a growing health concern worldwide, affecting millions of people and placing a significant burden on healthcare systems. Among the most common renal abnormalities are kidney cysts, kidney stones, and kidney tumors, each of which can lead to severe complications if not diagnosed and treated at an early stage. Accurate and timely diagnosis is therefore essential to improve patient outcomes, support clinical decision-making, and reduce the risks of disease progression. In Bangladesh, where access to advanced diagnostic support may vary across healthcare facilities, the use of intelligent and automated systems can play an important role in strengthening medical imaging analysis and assisting healthcare professionals. Computed tomography (CT) imaging is widely used for the evaluation of kidney-related abnormalities because of its ability to provide detailed cross-sectional views of renal structures (Gharaibeh *et al.,* 2022). CT scans help clinicians identify structural changes, detect lesions, and differentiate between various types of kidney conditions with greater clarity than many conventional imaging approaches. However, manual interpretation of CT images is often time-consuming and depends heavily on the expertise of radiologists. In busy clinical settings, this process may become challenging due to increasing patient volume, diagnostic complexity, and the possibility of human error. These limitations highlight the need for computer-aided diagnostic tools that can support faster and more reliable interpretation of renal CT images. In recent years, artificial intelligence (AI), particularly deep learning, has shown remarkable success in medical image analysis. Convolutional neural networks (CNNs) have become one of the most effective approaches for image classification, feature extraction, and disease detection because of their ability to automatically learn complex visual patterns from imaging data. Deep learning models have already demonstrated strong performance in diagnosing various medical conditions from radiological images, including abnormalities in the lungs, brain, breast, and kidneys. Their growing success suggests that AI-driven systems can significantly enhance diagnostic accuracy while reducing the workload of healthcare professionals. For renal abnormality detection, the use of deep learning offers several important advantages. It enables automated identification of subtle image features that may not always be easily distinguishable through manual observation, supports multiclass classification of different kidney conditions, and provides a scalable solution for large-volume screening (Lee *et al.,* 2022). This is particularly valuable in countries like Bangladesh, where the integration of AI into diagnostic workflows may help bridge gaps in healthcare access and improve consistency in medical image interpretation. Developing an AI-based CT diagnostic system tailored to the Bangladeshi population is therefore both clinically relevant and technologically significant. This study focuses on constructing an AI-driven CT diagnostic system for renal abnormalities in the Bangladeshi population using deep learning techniques. Three well-known pre-trained CNN architectures, namely Xception, VGG16, and ResNet152V2, were employed to classify CT images into

[1] Department of Computer Science and Engineering, IUBAT— International University of Business Agriculture and Technology, Dhaka, Bangladesh
[*] Corresponding author's e-mail: myeasmin.cs@gmail.com

four categories: normal, cyst, stone, and tumor. These models were selected because of their strong reputation in image classification tasks and their ability to capture hierarchical visual features effectively. By comparing the performance of multiple state-of-the-art architectures, the study aims to identify the most suitable model for accurate renal abnormality classification. The main contribution of this research is the development and evaluation of an intelligent diagnostic framework for renal abnormality detection that is relevant to the Bangladeshi healthcare context. By combining CT imaging with advanced deep learning models, this work contributes to the growing field of AI-assisted medical diagnosis and provides a foundation for future clinical decision-support systems in kidney disease detection and management.

LITERATURE REVIEW

The paper discusses the development of a deep learning-based model to assess renal histopathology in lupus glomerulonephritis (LN). The researchers introduced the Murine Renal Pathological System (MRPS), a model that uses convolutional neural networks (CNN) to analyze whole-slide images (WSI) of kidney tissue. The MRPS model aids in identifying and quantifying glomerular lesions, which are critical in diagnosing LN. The study found that five independent predictors such as glomerular perimeter and shape factor were significant for scoring renal histopathology. The MRPS model demonstrated high sensitivity, specificity, and accuracy, offering a more objective and reproducible method for evaluating LN. The study investigates the relationship between diabetic retinopathy (DR) and diabetic nephropathy (DN) and assesses whether the severity of DR can predict progression to end-stage renal disease (ESRD) in Chinese patients with type 2 diabetes mellitus (T2DM) and biopsy-confirmed DN (Zhao *et al.,* 2020). The patients' DR severity was classified using a deep learning system (RetinalNET), and the relationship between DR severity and ESRD was analyzed over a median follow-up period of 15 months. The results showed that more severe DR at the time of renal biopsy was associated with a higher risk of developing ESRD, making DR severity a significant prognostic factor. Additionally, the study found a discordance between the severity of DR and DN, especially in patients with mild retinopathy or microalbuminuria. This paper provides a comprehensive review of deep learning (DL) applications in renal ultrasound, addressing key areas such as kidney segmentation, volume measurement, functional prediction, and disease diagnosis (Zhang *et al.,* 2026). The study highlights the limitations of traditional renal ultrasound, particularly in terms of interobserver variability and the inability to provide quantitative metrics. DL, especially through convolutional neural networks (CNNs) and transformers, has enhanced the accuracy and efficiency of renal image analysis. The paper outlines the strengths of various DL architectures like U-Net, ResNet, and transformer models, discussing their role in improving renal disease diagnosis, including chronic kidney disease (CKD). Additionally, the paper identifies challenges related to data quality, labeling inconsistencies, and the need for clinical integration of DL technologies. The future of renal ultrasound diagnosis is anticipated to involve multimodal fusion, federated learning, and large model technologies to create more intelligent, standardized, and personalized diagnostic systems. The paper reviews the role of deep learning (DL) in enhancing renal health assessments, focusing on structural, functional, and clinical domains (Bist *et al.,* 2026). Kidney diseases, such as chronic kidney disease (CKD) and acute kidney injury (AKI), present a significant global health burden due to delayed detection and ineffective monitoring. Traditional diagnostic methods have limitations, prompting the adoption of DL techniques to address these issues. DL has shown promise in improving kidney health evaluation by enabling earlier diagnosis, risk stratification, and personalized care. This review discusses how DL improves renal disease classification, the prediction of renal function decline, and the analysis of renal imaging. It also addresses the challenges and ethical concerns regarding DL's integration into clinical practice, including data quality, interpretability, and bias. The paper concludes by emphasizing the potential of DL to improve patient outcomes and clinical decision-making in nephrology. The paper discusses the development of a deep learning-based tool for the automated evaluation of total interstitial inflammation and peritubular capillaritis on kidney biopsies. The convolutional neural network (CNN) tool was trained on a large dataset of kidney samples, including various kidney diseases such as IgA nephropathy, transplant biopsies, and tubulointerstitial nephritis (TIN) (Jacq *et al.,* 2023). The network's performance was compared against manual pathologist evaluations, showing high precision, recall, and F-score in detecting leukocytes and peritubular capillaries. The predicted total inflammation and capillaritis scores were strongly correlated with pathologist assessments, demonstrating the potential of artificial intelligence to improve the reproducibility and accuracy of kidney pathology evaluations. The tool also showed promising results in predicting kidney function based on the severity of interstitial inflammation in IgAN patients. This paper discusses the development of an automated deep learning system to detect and measure renal cysts in ultrasound images (Kanauchi *et al.,* 2023). The system uses a two-step approach: first, YOLOv5 is employed for detecting the presence of renal cysts, followed by UNet++ for predicting the salient landmarks' positions on the detected cysts. The study compares the system's performance with human sonographers and demonstrates that the deep learning system can predict salient landmark positions with comparable accuracy to radiologists but with significantly reduced time. The system utilizes YOLOv5 for cyst detection, and UNet++ to produce heatmaps for landmark position prediction. The results highlight the potential of deep learning models to automate renal cyst measurement in ultrasound imaging, improving diagnostic efficiency and reducing the burden on healthcare professionals. This paper investigates the use of deep learning models to automatically calculate the Hydronephrosis Area to Renal Parenchyma (HARP) ratio from ultrasound images in pediatric patients with ureteropelvic junction obstruction (UPJO) (Song *et al.,* 2022). The study involved 168 ultrasound images from patients who underwent pyeloplasty, and a combination of deep learning networks, including DeepLabV3+ and UNet++, were tested for segmentation performance. The results showed that the deep learning models demonstrated high accuracy, with the ensemble method of five models providing the best performance (Dice Similarity Coefficient of 0.9113). The HARP ratio, calculated using the deep learning model, was strongly correlated with renal function, making it a potential tool for objective and reproducible diagnosis of hydronephrosis severity. The paper highlights the effectiveness of deep learning in overcoming observer variability in the analysis of ultrasound images and provides a promising approach for clinical use. This study investigates the use of deep learning algorithms for detecting kidney stones in unenhanced CT images, comparing performance across different stone sizes (Caglayan *et al.,* 2022). The study involved 455 patients and 2,959 CT images analyzed using a deep learning model (xResNet50). The results showed that the model achieved high accuracy in training, with accuracy rates up to 99.1% in some planes. The model's performance was also tested on new data, showing accuracy rates ranging from 63% to 93% depending on the plane and stone size. The study demonstrates that deep learning models can aid in the efficient detection of kidney stones, reducing diagnostic time and costs. This study presents a multi-instance deep learning method for diagnosing congenital abnormalities of the kidney and urinary tract (CAKUT) in children using ultrasound images (Yin *et al.,* 2020). The method employs transfer learning to extract features from individual 2D ultrasound images and then uses multi-instance learning (MIL) to classify each subject based on multiple images, offering a robust diagnosis. The model showed improved performance when using multiple views of the kidney (sagittal and transverse), surpassing the results of models trained on single views. This approach could significantly enhance the diagnosis of CAKUT, addressing issues with inter-observer variability in manual evaluations. The study presents a deep learning-based approach to segment and classify glomeruli from multi-stained renal biopsy pathologic images (Jiang *et al.,* 2021). The proposed model uses a Cascade Mask R-CNN architecture to detect, classify, and segment glomeruli into three categories: glomeruli with normal structure (GN), global sclerosis (GS), and glomeruli with other lesions (GL). The model demonstrated robust performance on both snapshot and whole-slide images, achieving high F1 scores and accuracy for detecting glomeruli across different staining methods (PAS, PASM, and Masson). The model was able to effectively segment glomeruli in a variety of renal pathologies, making it a potential tool for automating the analysis of renal biopsy images. This paper presents a deep learning-based approach called RENTAG (Robust Evaluation of Tubular Atrophy and Glomerulosclerosis) for the automated segmentation and classification of glomeruli and tubules in kidney biopsy images (Salvi *et al.,* 2021). The algorithm, which uses PAS-stained images, combines semantic segmentation and active contours to detect glomeruli and tubules, achieving high accuracy in detecting glomerulosclerosis and tubular atrophy. The method is validated on 830 images, demonstrating superior performance compared to existing techniques. RENTAG shows promise in automating the assessment of pre-transplant kidney biopsies, reducing variability, and improving diagnostic efficiency in kidney transplantation.

MATERIALS AND METHODS

Recognition Process

This pipeline is designed to develop an AI-driven CT image diagnostic system for the automatic recognition and classification of renal abnormalities in the Bangladeshi

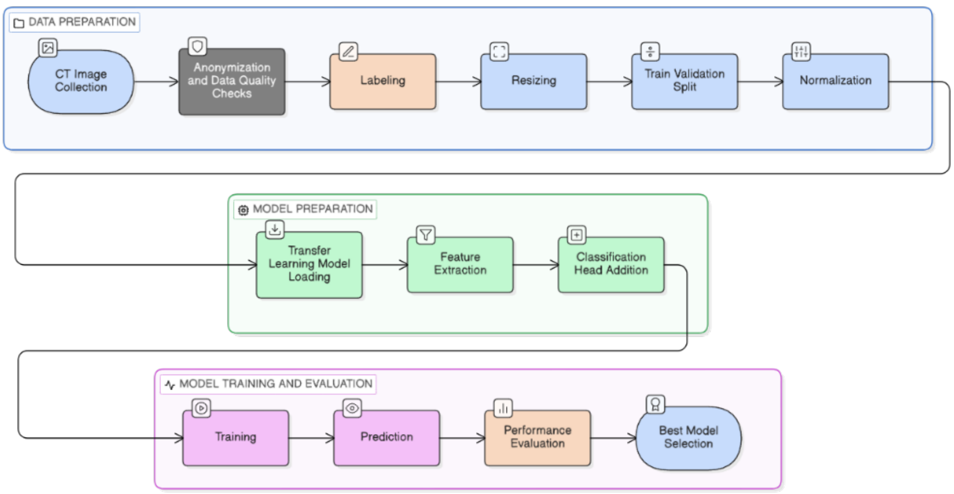

Figure 1: Overall Workflow of the Proposed AI-Driven CT Image Recognition System

population. The main goal is to assist radiologists and healthcare professionals by detecting kidney conditions such as normal, cyst, stone, and tumor from CT scan images with high accuracy (Ma *et al.,* 2020). The system uses deep learning based transfer learning models to analyze medical images, extract relevant features, and generate reliable predictions that can support faster and more consistent clinical decision-making (Figure 1).

- CT Image Collection: CT scan images of kidneys are collected from different hospitals in Dhaka and publicly available open-source datasets. These images include four classes: Normal, Cyst, Stone, and Tumor.
- Data Labeling and Organization: Each image is assigned its correct class label and organized into separate folders based on disease category (Zhou *et al.,* 2019). This ensures proper supervised learning during model training.
- Image Preprocessing: All CT images are resized to 224 × 224 pixels so that they match the input size requirement of the selected deep learning models. Pixel values are normalized to improve learning performance.
- Dataset Splitting: The complete dataset is divided into training and validation sets. The training set is used to teach the model, while the validation set is used to measure how well the model performs on unseen data.
- Data Pipeline Optimization: Caching and prefetching techniques are applied to make the data loading process faster and more efficient during model training.
- Transfer Learning Model Selection: Three pretrained convolutional neural network models are selected for comparison: Xception, VGG16, and ResNet152V2. These models are chosen due to their strong and proven performance in image classification tasks.
- Feature Extraction: The pretrained convolutional base of each model is used to automatically extract meaningful visual features from CT kidney images (Kang *et al.,* 2020). This reduces the need for manual feature engineering.
- Custom Classification Head Addition: A custom classifier is attached to the pretrained model, including dense layers, batch normalization, dropout, and a final softmax output layer for four-class prediction.
- Model Training: Each model is trained using the training dataset for multiple epochs. During training, the system learns to identify patterns associated with normal kidney structure, cysts, stones, and tumors.
- Validation During Training: The validation dataset is used after each training epoch to monitor performance and reduce the chance of overfitting. The best-performing model weights are saved.
- Prediction and Recognition: After training, the model predicts the class of each CT image by assigning probabilities to the four categories. The class with the highest probability is selected as the final recognized result.
- Performance Evaluation: The trained models are evaluated using key performance metrics such as accuracy, confusion matrix, and classification report. These metrics help determine how effectively the system can distinguish among the four renal conditions.
- Best Model Selection: The performances of Xception, VGG16, and ResNet152V2 are compared. The model with the highest recognition accuracy and strongest overall performance is selected as the final diagnostic model.

Data Collection Procedure

The dataset used in this study was collected from two main sources: computed tomography (CT) kidney images obtained from different hospitals in Dhaka and publicly available open-source medical image data from Kaggle (Tsai *et al.,* 2022). The purpose of combining these sources was to build a broader and more diverse dataset for the detection and classification of renal abnormalities in the Bangladeshi population. After data gathering and organization, a total of 13,200 CT images belonging to four classes were prepared for the experiment. These classes included Normal, Cyst, Stone, and Tumor (Bhak *et al.,* 2025). The collected images were arranged in a class-wise folder structure to facilitate supervised learning and automated data loading during model development. For model training and performance assessment, the complete dataset was divided into two subsets: a training set and a testing set. Among the total 13,200 images, 11,275 images were used for training the deep learning models, while 1,925 images were reserved for testing. This split allowed the models to learn representative image patterns from a large number of samples and then be evaluated on previously unseen CT images. The collected dataset served as the foundation for constructing and validating the proposed AI-driven renal abnormality recognition system.

Pre-trained Deep Learning Networks

In this study, three pre-trained deep learning models were employed to classify renal CT images into four categories: Normal, Cyst, Stone, and Tumor. These models were Xception, VGG16, and ResNet152V2. All of them were implemented using a transfer learning approach, where the convolutional base of each model was initialized with ImageNet pre-trained weights (Islam *et al.,* 2024). The use of pre-trained models is highly effective in medical image analysis because these networks have already learned rich and discriminative visual features from large-scale image datasets, which can then be adapted for specific diagnostic tasks (Sabanayagam *et al.,* 2020). In the present work, the top layers of the original networks were removed and replaced with a custom classification head suitable for four-class renal abnormality recognition. All input CT images were resized to 224 × 224 × 3, and the convolutional base of each model was kept frozen during initial training. A common custom classifier consisting of Flatten, Dense (512 neurons, ReLU), Batch Normalization, Dropout (0.5), and Dense (4 neurons, Softmax) was attached to each network.

Xception

For medical image classification, Xception is especially useful because it can detect subtle structural differences in CT scans with high precision. Renal abnormalities such as cysts, stones, and tumors often exhibit fine image-level variations, and the depthwise separable convolution mechanism helps the network focus on these localized details effectively. In this research, the Xception model was loaded with ImageNet weights, with include_top=False, and global max pooling was applied before the custom dense classification layers (Bevilacqua *et al.,* 2019). The feature extraction layers of the model remained frozen during training, allowing the classifier head to learn renal-specific patterns from the CT images (Nguyen *et al.,* 2024). Among all the models used in this study, Xception achieved the highest classification accuracy of 99.84%, indicating that it was the most effective model for the proposed diagnostic task (Table 1).

Table 1: Architecture Explanation – Xception

Layer Type	**Output Shape**	**Number of Parameters**
Xception (Functional)	(None, 2048)	20,861,480
Flatten	(None, 2048)	0
Dense	(None, 512)	1,049,088
Batch Normalization	(None, 512)	2,048
Dropout	(None, 512)	0
Dense	(None, 4)	2,052
Total Parameters		**21,914,668**
Trainable Parameters		**1,052,164**
Non-trainable Parameters		**20,862,504**

VGG16

One of the major strengths of VGG16 is its ability to learn hierarchical image features in a progressive manner. The initial layer's capture low-level features such as edges and textures, while deeper layers capture more complex patterns and object-level representations (Miguel *et al.,* 2024). This characteristic makes VGG16 suitable for medical imaging, where both low-level intensity patterns and higher-level structural features are important for distinguishing normal and abnormal kidney tissues. In this study, the original top classification layers of VGG16 were removed, and the model was used only as a feature extractor. A custom classifier was then added for four-class classification (Hsu *et al.,* 2025). Although VGG16 is relatively heavier in terms of parameters and computational cost compared to some modern architectures, it still performed remarkably well in this work, achieving 99.40% accuracy. Its strong performance confirms that even classical deep learning architectures remain highly relevant for renal CT image analysis (Table 2).

Table 2: Architecture Explanation – VGG16

Layer Type	**Output Shape**	**Number of Parameters**
VGG16 (Functional)	(None, 512)	14,714,688
Flatten	(None, 512)	0
Dense	(None, 512)	262,656
Batch Normalization	(None, 512)	2,048
Dropout	(None, 512)	0
Dense	(None, 4)	2,052
Total Parameters		**14,981,444**
Trainable Parameters		**265,732**
Non-trainable Parameters		**14,715,712**

ResNet152V2

The ResNet152V2 variant is an improved version of the original ResNet design, where the ordering of batch normalization, activation, and convolution operations is optimized for better performance. With 152 layers, this network is capable of learning highly complex and abstract visual representations (Pimpalkar *et al.,* 2025). For CT image classification, this deep residual structure is valuable because it allows the network to capture both fine details and broader anatomical context from kidney scans. In the present study, ResNet152V2 was used as a pre-trained backbone with ImageNet weights, and its final fully connected layers were replaced by the same custom four-class classifier used in the other models (Joo *et al.,* 2023). The convolutional base was frozen during training, and only the newly added classifier layers were trained on the renal CT dataset. The model produced an excellent accuracy of 99.76%, demonstrating its strong capability in multiclass renal abnormality recognition (Table 3).

Table 3: Architecture of ResNet152V2-Based Model

Layer Type	Output Shape	Number of Parameters
ResNet152V2 (Functional)	(None, 2048)	58,331,648
Flatten	(None, 2048)	0
Dense	(None, 512)	1,049,088
Batch Normalization	(None, 512)	2,048
Dropout	(None, 512)	0
Dense	(None, 4)	2,052
Total Parameters		**59,384,836**
Trainable Parameters		**1,052,164**
Non-trainable Parameters		**58,332,672**

RESULT AND DISCUSSION

Dataset

The dataset used in this study consisted of a total of 13,200 CT images collected for the classification of renal abnormalities. The images were categorized into four diagnostic classes: Cyst, Normal, Stone, and Tumor. To develop and evaluate the proposed AI-driven diagnostic system, the dataset was divided into two subsets: training set and testing set. Out of the total 13,200 images, 11,275 images were used for training the deep learning models, while 1,925 images were reserved for testing (Holmstrom *et al.,* 2023). In the training set, the highest number of images belonged to the Cyst category with 3,700 images, followed by Normal with 3,410 images, Tumor with 2,280 images, and Stone with 1,885 images. In the testing set, Cyst contained 560 images, Stone contained 515 images, Tumor contained 450 images, and Normal contained 400 images. This distribution provided a multiclass renal CT dataset for training and evaluating the performance of the selected pre-trained deep learning models (Table 4).

Table 4: Distribution of CT Image Dataset for Training and Testing

Class	Training Data	Testing Data	Total
Cyst	3,700	560	4,260
Normal	3,410	400	3,810
Stone	1,885	515	2,400
Tumor	2,280	450	2,730
Total	**11,275**	**1,925**	**13,200**

Detailed Analysis of Training and Validation Loss

In the analysis of the training and validation loss curves for Xception, VGG16, and ResNet152V2, the following observations were made:

Xception Model

Starting Loss (Epoch 1): The Xception model began training with a relatively high loss value in both training and validation, with the training loss starting around 0.45 and the validation loss around 0.40. Highest Accuracy and Corresponding Loss: The model achieved its highest accuracy towards the end of the training. At the point of highest accuracy, the training loss dropped to approximately 0.05 and the validation loss reached around 0.06. This indicates the model's excellent convergence, with a significant reduction in both training and validation losses. Percentage Reduction in Loss: The training loss saw a reduction of approximately 89%, while the validation loss dropped by about 85%. This drastic reduction suggests that Xception is highly effective at learning the patterns of kidney abnormalities from CT scans and generalizing to new data (Figure 2).

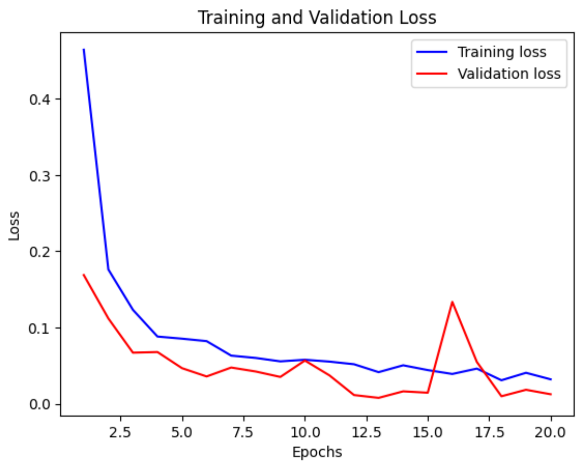

Figure 2: Training and Validation Loss of Xception

VGG16 Model

Starting Loss (Epoch 1): The VGG16 model started with a training loss of around 0.43 and a validation loss of approximately 0.35. This suggests that the model's performance was initially poor but improved quickly over time. Highest Accuracy and Corresponding Loss: As the training progressed, the model reached its highest accuracy at around epoch 16. The training loss reached about 0.08, and the validation loss was around 0.12 at this point. Percentage Reduction in Loss: The training loss saw a reduction of approximately 81%, and the validation loss decreased by about 66%. This is slightly less effective than Xception, indicating that VGG16 faces some challenges in generalization despite improving significantly throughout the epochs (Figure 3).

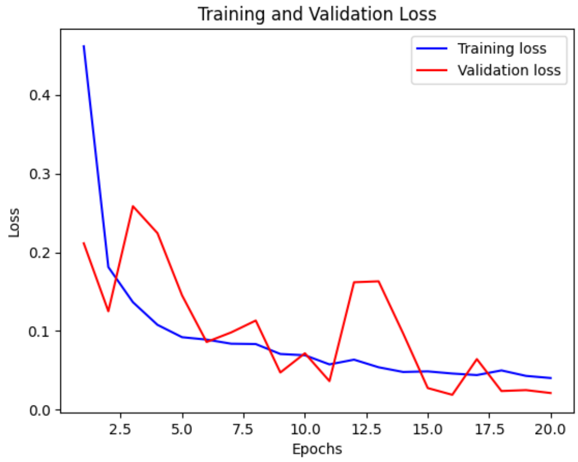

Figure 3: Training and Validation Loss of VGG16

ResNet152V2 Model

Starting Loss (Epoch 1): The ResNet152V2 model started with a training loss of approximately 0.40 and a validation loss around 0.35. Highest Accuracy and Corresponding Loss: The model achieved its highest accuracy towards the end of the training, similar to Xception. At the peak of its performance, the training loss was approximately 0.06, while the validation loss was around 0.08. Percentage Reduction in Loss: The training loss decreased by about 85%, while the validation loss reduced by 77%, showing that ResNet152V2 was able to generalize better than VGG16, but slightly less effectively than Xception (Figure 4).

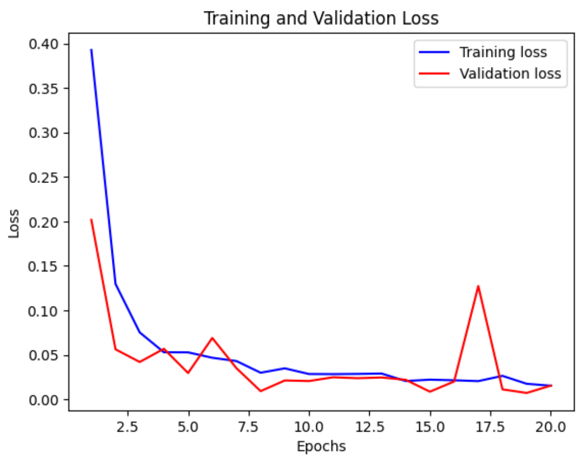

Figure 4: Training and Validation Loss of Resnet152V2

Accuracy Analysis of Xception, VGG16, and ResNet152V2

Xception Model

Initial Accuracy (Epoch 1): The Xception model began with an initial accuracy of around 75% in the first epoch. Highest Accuracy: By the end of training, Xception achieved its highest accuracy at approximately 99.84%. This significant improvement indicates the model's strong ability to learn and generalize from the training data. Accuracy Improvement: The accuracy increased by approximately 24% from epoch 1 to the final epoch, showing fast convergence and excellent model performance throughout the training process (Figure 5).

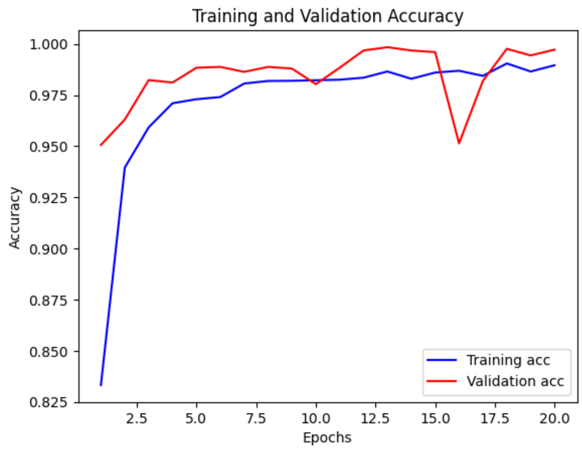

Figure 5: Training and Validation Accuracy of Xception

VGG16 Model

Initial Accuracy (Epoch 1): The VGG16 model started with an accuracy of about 70% in the first epoch. Highest Accuracy: By the final epoch, the model reached an accuracy of 99.40%. Despite some fluctuations in validation loss, the model's accuracy showed strong improvement over the epochs. Accuracy Improvement: The accuracy increased by approximately 29.4% from the initial epoch to the highest point, highlighting the model's capacity to learn, though it exhibited slightly more instability compared to Xception (Figure 6).

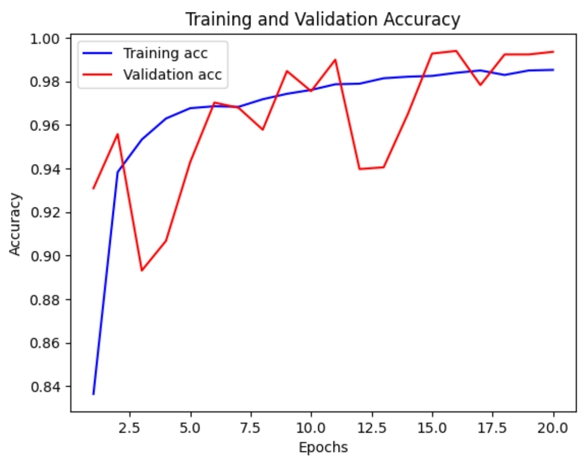

Figure 6: Training and Validation Accuracy of VGG16

ResNet152V2 Model

Initial Accuracy (Epoch 1): The ResNet152V2 model began with an accuracy of around 72% in the first epoch. Highest Accuracy: By the end of training, ResNet152V2 reached 99.76% accuracy, which is very close to the performance of Xception. Accuracy Improvement: The accuracy increased by approximately 27.7%, showing that ResNet152V2 was able to learn and generalize kidney abnormalities effectively, with a similar pattern to Xception but slightly less effective than Xception in final accuracy (Figure 7).

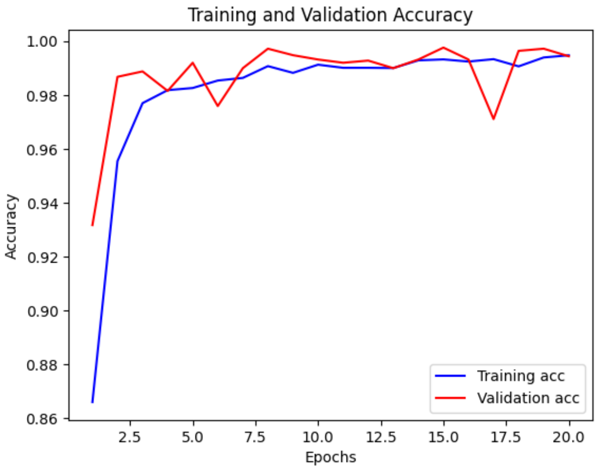

Figure 7: Training and Validation Accuracy of ResNet152V2

Performance of All Models

The Xception model exhibited the best performance across all models, achieving the highest accuracy of 99.84% and the lowest validation loss of 0.06. It demonstrated minimal fluctuations in both training and validation loss, indicating that the model efficiently learned from the training data and generalized well to unseen validation data. This superior performance made Xception the most reliable model for renal CT image classification. On the other hand, the VGG16 model, with a final accuracy of 99.40%, showed some noticeable fluctuations in the validation loss (0.12), suggesting challenges in generalization and potential overfitting during training, although it remained an effective model for kidney abnormality classification (Table 5).

The ResNet152V2 model achieved a final accuracy of 99.76% with a validation loss of 0.08, demonstrating minor fluctuations. Despite these fluctuations, ResNet152V2 proved to be stable and effective, with its low training loss of 0.06 indicating strong learning capabilities. While Xception provided the best overall results, both ResNet152V2 and VGG16 performed admirably, with ResNet152V2 offering more stable performance compared to VGG16.

Table 5: Comparison Table (Performance of All Models)

Model	Training Loss (Final Epoch)	Validation Loss (Final Epoch)	Accuracy (Final Epoch)	Fluctuations in Loss	Generalization Ability
Xception	0.05	0.06	99.84%	Minimal	Excellent
VGG16	0.08	0.12	99.40%	Moderate	Good
ResNet152V2	0.06	0.08	99.76%	Minor	Very Good

CONCLUSION

In this study, three deep learning models Xception, VGG16, and ResNet152V2 were evaluated for the classification of renal abnormalities from CT images. Among these models, Xception achieved the highest accuracy of 99.84% and the lowest validation loss of 0.06, demonstrating its superior ability to generalize from the training data to unseen validation images. The model displayed minimal fluctuations in both training and validation loss, making it the most reliable for kidney abnormality detection. ResNet152V2 also performed excellently with a final accuracy of 99.76% and validation loss of 0.08, showing strong stability and effective learning. This model's use of residual connections allowed it to capture more complex features, offering a good balance between performance and stability. VGG16, while effective, achieved a final accuracy of 99.40% with a validation loss of 0.12, showing more fluctuations in validation loss compared to Xception and ResNet152V2. This suggests that VGG16 struggled with overfitting and did not generalize as well, although it still remained a solid model for renal abnormality classification. Xception outperformed the other models, offering the best combination of high accuracy, stability, and generalization ability. ResNet152V2 also performed strongly but showed slight fluctuations in validation loss. VGG16, despite some instability, was still effective but could benefit from improvements in generalization. These findings highlight the potential of deep learning models like Xception in the field of medical image classification and their ability to support clinical decision-making by providing accurate and automated diagnosis of kidney diseases.

REFERENCES

Bevilacqua, V., Brunetti, A., & Cascarano, G. D. (2019). A comparison between two semantic deep learning frameworks for the autosomal dominant polycystic kidney disease segmentation based on magnetic resonance images. *BMC Medical Informatics and Decision Making, 19*, 244. https://doi.org/10.1186/s12911-019-0988-4

Bhak, Y., Lee, Y. H., Kim, J., Lee, K., Lee, D., Jang, E. C., Jang, E., Lee, C. S., Kang, E. S., Park, S., Han, H. W., & Nam, S. M. (2025). Diagnosis of chronic kidney disease using retinal imaging and urine dipstick data: Multimodal deep learning approach. *JMIR Medical Informatics, 13*(1), e55825. https://doi.org/10.2196/55825

Bist, D. L., Agarwal, S., Gajbhiye, D. A., Ray, S. P., & Sharma, D. S. K. (2026). Evaluating the role of deep learning in enhancing renal health assessment: *A comprehensive review. Kidneys, 15*(1), 32–42. https://doi.org/10.65327/kidneys.v15i1.596

Caglayan, A., Horsanali, M. O., Kocadurdu, K., Ismailoglu, E., & Guneyli, S. (2022). Deep learning model-assisted detection of kidney stones on computed tomography. *International Braz j Urol, 48*, 830–839. https://doi.org/https://doi.org/10.1590/S1677-5538.IBJU.2022.0132

Gharaibeh, M., Alzu'bi, D., Abdullah, M., Hmeidi, I., Al Nasar, M. R., Abualigah, L., & Gandomi, A. H. (2022). Radiology imaging scans for early diagnosis of kidney tumors: a review of data analytics-based machine learning and deep learning approaches. *Big Data and Cognitive Computing, 6*(1), 29. https://doi.org/10.3390/bdcc6010029

Holmstrom, L., Christensen, M., Yuan, N., Weston Hughes, J., Theurer, J., Jujjavarapu, M., Fatehi, P., Kwan, A., Sandhu, R. K., Ebinger, J., Cheng, S., Zou, J., Chugh, S. S., & Ouyang, D. (2023). Deep learning-based electrocardiographic screening for chronic kidney disease. *Communications Medicine, 3*(1), 73. https://doi.org/10.1038/s43856-023-00278-w

Hsu, C.-F., Yu, T.-M., Wu, Y.-L., Wang, W.-C., Wang, J.-S., & Chang, S.-S. (2025). Prediction of advanced chronic kidney disease through retinal fundus images by deep learning. *Scientific Reports, 15*(1), 37318. https://doi.org/10.1038/s41598-025-21366-y

Islam, U., Al-Atawi, A. A., Alwageed, H. S., Mehmood, G., Khan, F., & Innab, N. (2024). Detection of renal cell hydronephrosis in ultrasound kidney images: A study on the efficacy of deep convolutional neural networks. *PeerJ Computer Science, 10,* e1797. https://doi.org/10.7717/peerj-cs.1797

Jacq, A., Tarris, G., Jaugey, A., Paindavoine, M., Maréchal, E., Bard, P., Rebibou, J.-M., Ansart, M., Calmo, D., Bamoulid, J., Tinel, C., Ducloux, D., Crepin, T., Chabannes, M., Funes de la Vega, M., Felix, S., Martin, L., & Legendre, M. (2023). Automated evaluation with deep learning of total interstitial inflammation and peritubular capillaritis on kidney biopsies. *Nephrology Dialysis Transplantation, 38*(12), 2786–2798. https://doi.org/10.1093/ndt/gfad094

Jiang, L., Chen, W., Dong, B., Mei, K., Zhu, C., Liu, J., Cai, M., Yan, Y., Wang, G., Zuo, L., & Shi, H. (2021). A deep learning-based approach for glomeruli instance segmentation from multistained renal biopsy pathologic images. *The American Journal of Pathology, 191*(8), 1431–1441. https://doi.org/10.1016/j.ajpath.2021.05.004

Joo, Y. S., Rim, T. H., Koh, H. B., Yi, J., Kim, H., Lee, G., Kim, Y. A., Kang, S.-W., Kim, S. S., & Park, J. T. (2023). Non-invasive chronic kidney disease risk stratification tool derived from retina-based deep learning and clinical factors. *Npj Digital Medicine, 6*(1), 114. https://doi.org/10.1038/s41746-023-00860-5

Kanauchi, Y., Hashimoto, M., Toda, N., Okamoto, S., Haque, H., Jinzaki, M., & Sakakibara, Y. (2023). Automatic detection and measurement of renal cysts in ultrasound images: a deep learning approach. *Healthcare, 11*(4), 484. https://doi.org/10.3390/healthcare11040484

Kang, E. Y.-C., Hsieh, Y.-T., Li, C.-H., Huang, Y.-J., Kuo, C.-F., Kang, J.-H., Chen, K.-J., Lai, C.-C., Wu, W.-C., & Hwang, Y.-S. (2020). Deep learning–based detection of early renal function impairment using retinal fundus images: model development and validation. *JMIR Medical Informatics, 8*(11), e23472. https://doi.org/10.2196/23472

Lee, H., Yoo, B., & Young Choi, J. (2022). Prediction oof recurrent urinary tract infection in paediatric patients by deep learning analysis of 99mtc-dmsa renal scan. *Diagnotics, 12*(2). https://doi.org/https://doi.org/10.3390/diagnostics12020424

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. https://doi.org/10.1016/j.future.2020.04.036

Miguel, O. X., Kaczmarek, E., Lee, I., Ducharme, R., Dingwall-Harvey, A. L. J., Rennicks White, R., Bonin, B., Aviv, R. I., Hawken, S., Armour, C. M., Dick, K., & Walker, M. C. (2024). Deep learning prediction of renal anomalies for prenatal ultrasound diagnosis. *Scientific Reports, 14*(1), 9013. https://doi.org/10.1038/s41598-024-59248-4

Nguyen, T. T. U., Nguyen, A.-T., Kim, H., Jung, Y. J., Park, W., Kim, K. M., Park, I., & Kim, W. (2024). Deep-learning model for evaluating histopathology of acute renal tubular injury. *Scientific Reports, 14*(1), 9010. https://doi.org/10.1038/s41598-024-58506-9

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. https://doi.org/10.1038/s41598-025-94905-2

Sabanayagam, C., Xu, D., Ting, D. S. W., Nusinovici, S., Banu, R., Hamzah, H., Lim, C., Tham, Y.-C., Cheung, C. Y., Tai, E. S., Wang, Y. X., Jonas, J. B., Cheng, C.-Y., Lee, M. L., Hsu, W., & Wong, T. Y. (2020). A deep learning algorithm to detect chronic kidney disease from retinal photographs in community-based populations. *The Lancet Digital Health, 2*(6), e295–e302. https://doi.org/10.1016/S2589-7500(20)30063-7

Salvi, M., Mogetta, A., Gambella, A., Molinaro, L., Barreca, A., Papotti, M., & Molinari, F. (2021). Automated assessment of glomerulosclerosis and tubular atrophy using deep learning. *Computerized Medical Imaging and Graphics, 90,* 101930. https://doi.org/10.1016/j.compmedimag.2021.101930

Song, S. H., Han, J. H., Kim, K. S., Cho, Y. A., Youn, H. J., Kim, Y. I., & Kweon, J. (2022). Deep-learning segmentation of ultrasound images for automated calculation of the hydronephrosis area to renal parenchyma ratio. *Investigative and Clinical Urology, 63*(4), 455–463. https://doi.org/10.4111/icu.20220085

Tsai, M.-C., Lu, H. H.-S., Chang, Y.-C., Huang, Y.-C., & Fu, L.-S. (2022). Automatic screening of pediatric renal ultrasound abnormalities: deep learning and transfer learning approach. *JMIR Medical Informatics, 10*(11), e40878. https://doi.org/10.2196/40878

Yin, S., Peng, Q., Li, H., Zhang, Z., You, X., Fischer, K., Furth, S. L., Tasian, G. E., & Fan, Y. (2020). Computer-aided diagnosis of congenital abnormalities of the kidney and urinary tract in children using a multi-instance deep learning method based on ultrasound imaging data. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)* (pp. 1347–1350). IEEE. https://doi.org/10.1109/ISBI45749.2020.9098506

Zhang, Y., Hou, Y., Qiu, T., Zhuang, Y., Chen, K., Ling, W., Luo, Y., & Lin, J. (2026). Deep learning in renal ultrasound: Applications, challenges, and future outlook. *Frontiers in Oncology, 15*. https://doi.org/10.3389/fonc.2025.1730628

Zhao, L., Ren, H., Zhang, J., Cao, Y., Wang, Y., Meng, D., Wu, Y., Zhang, R., Zou, Y., Xu, H., Li, L., Zhang, J., Cooper, M. E., Tong, N., & Liu, F. (2020). Diabetic retinopathy, classified using the lesion-aware deep learning system, predicts diabetic end-stage renal disease in chinese patients. *Endocrine Practice, 26*(4), 429–443. https://doi.org/10.4158/EP-2019-0512

Zhou, L., Zhang, Z., Chen, Y.-C., Zhao, Z.-Y., Yin, X.-D., & Jiang, H.-B. (2019). A deep learning-based radiomics model for differentiating benign and malignant renal tumors. *Translational Oncology, 12*(2), 292–300. https://doi.org/10.1016/j.tranon.2018.10.012