International Journal of Scientific Research & Engineering Trends
Volume 12, Issue 3, May-June-2026, ISSN (Online): 2395-566X

Kidney Net: An Intelligent Deep Learning Model for Kidney Disease Detection

Parul Tyagi, Dr. Brij Mohan Singh

Abstract- Kidney disease is a growing global health challenge requiring early, accurate, and automated diagnostic solutions. This paper introduces KidneyNet, a deep learning framework designed for automated kidney disease detection and classification from Computed Tomography (CT) scan images. KidneyNet leverages the power of transfer learning through ResNet50, enhanced with custom classification layers and advanced data augmentation strategies, to classify kidney CT images into four categories: cyst, normal, stone, and tumor. The proposed system is compared against two baseline architectures — Artificial Neural Network (ANN) and Convolutional Neural Network (CNN) — using a publicly available dataset of 12,446 kidney CT images. Experimental results demonstrate that KidneyNet (ResNet50) achieves superior performance with an accuracy of 92%, precision of 91.44%, recall of 92%, and an F1-score of 91.72%, outperforming both ANN (86% accuracy) and CNN (89% accuracy). These findings confirm the effectiveness of deep residual transfer learning as a reliable computer-aided diagnostic tool for kidney disease classification.

Keywords— KidneyNet, deep learning, kidney disease detection, ResNet50, transfer learning, CNN, ANN, CT scan, medical image classification, computer-aided diagnosis.

I. INTRODUCTION

Kidney disease is among the leading causes of morbidity and mortality worldwide, affecting over 850 million individuals globally. Conditions such as renal cysts, kidney stones, and tumors require timely and accurate detection to prevent progression toward chronic kidney disease (CKD) or renal failure. Traditional diagnostic pipelines rely heavily on radiologists manually interpreting CT scans — a process that is both time-intensive and susceptible to inter-observer variability [1].

Recent advances in artificial intelligence (AI) and deep learning (DL) have demonstrated transformative potential in medical imaging. Convolutional Neural Networks (CNNs) and transfer learning architectures have achieved remarkable accuracy in detecting diseases from radiological images, reducing diagnostic time and improving reproducibility [2]. Despite these advances, the application of deep learning specifically to multi-class kidney disease detection from CT images remains underexplored with systematic comparative evaluation.

This paper proposes KidneyNet, a deep learning system based on the ResNet50 transfer learning architecture, designed to classify kidney CT images into four clinically relevant categories: cyst, normal, stone, and tumor. We systematically compare KidneyNet against Artificial Neural Network (ANN) and standard Convolutional Neural Network (CNN) baselines to quantify the performance advantage of deep residual transfer learning for this task.

The key contributions of this paper are:

- A novel KidneyNet pipeline integrating ResNet50 transfer learning with custom classification layers tailored for kidney CT image analysis.
- A rigorous comparative evaluation of ANN, CNN, and ResNet50 architectures using identical experimental conditions on a 12,446-image dataset.
- Comprehensive performance analysis using accuracy, precision, recall, F1-score, and confusion matrix.
- Discussion of clinical applicability, limitations, and future directions for AI-assisted kidney disease diagnosis.

II. RELATED WORK

The application of machine learning to kidney disease diagnosis has a rich literature. Early studies by Rouhani et al. [3] and Kayaer et al. [4] demonstrated that Artificial Neural Networks could classify medical conditions from structured clinical data with reasonable accuracy, establishing a foundation for AI-driven diagnosis. However, these methods required manual feature extraction and showed limited performance on complex image data.

The emergence of Convolutional Neural Networks introduced automated spatial feature extraction, enabling researchers to build more robust disease detection systems. CNN-based models were applied to a variety of medical imaging tasks — including lung nodule detection, retinal disease classification, and tumor segmentation — achieving significantly higher accuracy than traditional ML methods. For kidney disease specifically, CNN architectures demonstrated improved classification of cysts and stones from CT images [5].

While prior work has examined individual architectures for kidney disease detection, systematic three-way comparisons of ANN, CNN, and ResNet50 under identical experimental conditions on standardized kidney CT datasets remain limited. KidneyNet addresses this gap by providing a rigorous comparative framework and a production-ready diagnostic pipeline.

Transfer learning further advanced medical image classification by allowing models pre-trained on large-scale datasets (such as ImageNet) to be fine-tuned on smaller medical datasets. ResNet50, introduced by He et al. [6], employs residual (skip) connections to enable training of very deep networks without vanishing gradients. Several studies have reported that ResNet50-based transfer learning outperforms shallower architectures in multi-class medical image classification tasks, including diabetic retinopathy and skin lesion detection [7].

While prior work has examined individual architectures for kidney disease detection, systematic three-way comparisons of ANN, CNN, and ResNet50 under identical experimental conditions on standardized kidney CT datasets remain limited. KidneyNet addresses this gap by providing a rigorous comparative framework and a production-ready diagnostic pipeline.

III. DATASET AND PREPROCESSING

1. Dataset Description

The study utilizes the CT Kidney Dataset [8], sourced from the Picture Archiving and Communication System (PACS) of multiple hospitals in Dhaka, Bangladesh. The dataset comprises 12,446 kidney CT scan images categorized into four classes, collected from both contrast and non-contrast studies using coronal and axial views.

Table 1: Class-wise distribution of the kidney CT scan dataset.

S.No.	Kidney Disease Class	Number of ImagesPercentage (%)	
1	Cyst	3,709	29.8%
2	Normal	5,077	40.8%
3	Stone	1,377	11.1%
4	Tumor	2,283	18.3%
	Total	12,446	100%

2. Preprocessing Pipeline

All images were subjected to a standardized preprocessing pipeline prior to model training:

- Resizing: All CT images were uniformly resized to 224x224 pixels to match ResNet50 input requirements.
- Normalization: Pixel intensity values were normalized to the [0, 1] range to improve convergence.
- Data Augmentation: Random rotation (+/-15 degrees), horizontal flipping, zoom (10%), and brightness adjustment were applied to training images to reduce overfitting.
- Dataset Split: Images were divided into training (70%), validation (15%), and testing (15%) subsets using stratified sampling to maintain class balance.

IV. METHODOLOGY: THE KIDNEYNET FRAMEWORK

1. ANN Baseline Architecture

The ANN baseline consists of a fully connected feed-forward network. CT scan images are first flattened into one-dimensional feature vectors. The architecture includes three hidden layers with 128, 64, and 32 neurons respectively, each activated by ReLU. A Softmax output layer produces probability scores for the four kidney disease classes. The model is trained using the Adam optimizer with categorical cross-entropy loss and a batch size of 32 for 50 epochs.

2. CNN Baseline Architecture

The CNN baseline follows a sequential convolutional architecture: three convolutional blocks (32, 64, and 128 filters, 3x3 kernels), each followed by batch normalization, ReLU activation, and 2x2 max pooling. The feature maps are flattened and passed through a fully connected dense layer (128 neurons, ReLU) and a Softmax output layer. Dropout (rate=0.5) is applied before the dense layer to reduce overfitting. Identical training configuration to the ANN is used for fair comparison.

3. KidneyNet (ResNet50 Transfer Learning)

KidneyNet is built upon ResNet50 pre-trained on ImageNet. The convolutional base of ResNet50 (50 layers with residual blocks) is retained as a frozen feature extractor during the initial training phase. The original classification head is replaced with:

- Global Average Pooling (GAP) layer to reduce spatial dimensions.
- Fully connected dense layer (2048 units, ReLU activation).
- Dropout layer (rate=0.5) for regularization.
- Softmax output layer (4 units) for classification into cyst, normal, stone, and tumor.

In a second fine-tuning phase, the top residual blocks of ResNet50 are unfrozen and the entire network is retrained with a reduced learning rate (1e-5) to adapt the pre-trained features to kidney CT imaging characteristics.

Table 2: Experimental environment and training configuration

4. Training Configuration	
Parameter	Specification
Programming Language	Python 3.9
Deep Learning Framework	TensorFlow 2.x / Keras
Development Platform	Google Colab (GPU T4)
Image Size	224 x 224 x 3
Batch Size	32
Initial Learning Rate	1e-4 (Adam optimizer)
Fine-tuning Learning Rate	1e-5
Epochs (Phase 1 / Phase 2)	30 / 20
Loss Function	Categorical Cross-Entropy
Dataset Split	70% Train / 15% Val / 15% Test

V. Evaluation Metrics

Model performance is evaluated using four standard classification metrics:

Accuracy measures the proportion of correctly classified samples over all predictions:

Accuracy = (TP + TN) / (TP + TN + FP + FN)

Precision quantifies how many predicted positive cases are truly positive:

Precision = TP / (TP + FP)

Recall (Sensitivity) measures the model's ability to detect all actual disease cases:

Recall = TP / (TP + FN)

- F1-Score provides the harmonic mean of Precision and Recall:
- F1-Score = 2 x (Precision x Recall) / (Precision + Recall)

VI. RESULTS AND DISCUSSION

1. Comparative Performance

Table 3 summarizes the overall performance of all three models evaluated on the held-out test set. KidneyNet (ResNet50) consistently outperforms both ANN and CNN baselines across all four metrics.

Table 3: Comparative performance of ANN, CNN, and KidneyNe

Table 3: Comparative performance of ANN, CNN, and KidneyNe

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
ANN	86.00	85.32	86.00	85.65
CNN	89.00	88.47	89.00	88.73
KidneyNet (ResNet50)	92.00	91.44	92.00	91.72

VI. CONFUSION MATRIX ANALYSIS

The confusion matrix for KidneyNet reveals the per-class classification quality across all four kidney conditions. Of 250 cyst samples, 230 (92.0%) were correctly identified. Among 250 normal cases, 229 (91.6%) were correctly classified. Stone detection, the most challenging class due to visual similarity with other conditions, achieved 90.4% recall (226/250). Tumor classification matched cyst performance at 92.0% (230/250). The overall test set contained 1,000 samples, and KidneyNet achieved an aggregate accuracy of 92%.

Misclassification analysis shows that stones are most frequently confused with tumors (4.8%) due to overlapping radiodensity patterns in CT imaging — a finding consistent with clinical radiology literature [9]. This highlights an important area for future improvement through enhanced feature engineering or multi-modal data integration.

Analysis of Model Comparisons

The ANN model, operating on flattened pixel features, achieved 86% accuracy but demonstrated inherent limitations in capturing spatial relationships within CT images. The absence of spatial convolution constrains its ability to detect subtle morphological features distinguishing cysts from tumors.

The CNN model improved performance to 89% accuracy by automatically learning spatial hierarchies through convolution and pooling. However, training from random weight initialization on a relatively small dataset (approximately 8,700 training images) limited its generalization capability.

KidneyNet achieved the highest performance across all metrics (92% accuracy, 91.72% F1-score). The performance advantage derives from two key factors: (1) ResNet50's residual connections enable deeper feature extraction without gradient degradation; and (2) ImageNet pre-training provides robust low-level feature representations (edge detectors, texture filters) that generalize effectively to CT image patterns, even across domain differences. The two-phase fine-tuning strategy further adapted these representations to kidney-specific pathological features.

Clinical Significance

A 92% classification accuracy, combined with a recall of 92% for disease categories, positions KidneyNet as a viable decision-support tool for radiologists. In clinical deployment, high recall is particularly critical for pathological classes (cyst, stone, tumor) to minimize missed diagnoses. KidneyNet's balanced precision-recall profile (F1 = 91.72%) suggests it can reduce false negatives without generating an excessive false positive burden.

Furthermore, the automated processing capability of KidneyNet could reduce per-image diagnostic time from several minutes (manual analysis) to sub-second inference, enabling high-throughput screening in resource-constrained healthcare environments.

VIII. CONCLUSION AND FUTURE WORK

1 Conclusion

This paper presented KidneyNet, a deep learning framework for automated kidney disease detection from CT scan images. KidneyNet employs ResNet50 transfer learning with a two-phase fine-tuning strategy and was systematically compared against ANN and CNN baselines using 12,446 kidney CT images across four disease categories.

The experimental results conclusively demonstrate that KidneyNet outperforms both baselines, achieving 92% accuracy, 91.44% precision, 92% recall, and a 91.72% F1-score. The residual learning architecture and pre-trained feature representations from ImageNet are identified as the primary contributors to this performance advantage. These results establish KidneyNet as an effective, reliable computer-aided diagnostic solution with clear clinical applicability.

Future Work

Several directions are identified for extending this research:

- Larger and more diverse datasets: Expanding training data across multiple geographic populations and imaging systems to improve model generalizability.
- Advanced architectures: Exploration of EfficientNet, DenseNet, and Vision Transformer (ViT) models for potential accuracy gains.
- Explainability: Integration of Gradient-weighted Class Activation Mapping (Grad-CAM) and other XAI techniques to provide radiologists with interpretable visual explanations.
- Multi-modal fusion: Combining CT imaging with clinical laboratory markers (eGFR, creatinine levels) for improved diagnostic confidence.
- Clinical deployment: Development of a web-based or mobile diagnostic application for real-time kidney disease screening in clinical environments.

REFERENCES

1. Litjens, G., et al. (2017). A survey on deep learning in medical image analysis. Medical Image Analysis, 42, 60-88.

2. Shen, D., Wu, G., & Suk, H. I. (2017). Deep learning in medical image analysis. Annual Review of Biomedical Engineering, 19, 221-248.
3. Rouhani, M., & Haghighi, M. (2009). The diagnosis of hepatitis diseases by support vector machines and artificial neural networks. Islamic Azad University, Gonabad.
4. Kayaer, K., & Yildirim, T. (2003). Medical diagnosis on Pima Indian diabetes using general regression neural networks. International Conference on Artificial Neural Networks.
5. Kumar, K., & Abhishek. (2012). Artificial neural network for diagnosis of kidney stones disease. International Journal of Information Technology and Computer Science, 4(7), 20-25.
6. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. IEEE CVPR, 770-778.
7. Esteva, A., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115-118.
8. Islam, M. N. (2021). CT Kidney Dataset: Normal-Cyst-Tumor and Stone. Kaggle. https://www.kaggle.com/datasets/nazmul0087/ct-kidney-dataset-normal-cyst-tumor-and-stone
9. Mitri, F. G. (2011). Vibro-acoustography imaging of kidney stones in vitro. IEEE Transactions on Biomedical Engineering.