*Grenze International Journal of Engineering and Technology, June Issue*

# DL based Technique for Detecting Renal Disorder using MobileNetV2

Jeni Angel J[1] and M Bhuvaneshwari[2]
[1-2]Division of Data Science and Cyber Security Karunya Institute of Technology and Sciences Coimbatore, India
Email: jeniangel@karunya.edu.in, bhuvaneshwaril@karunya.edu

***Abstract*— Proper diagnosis of kidney diseases including cysts, stones and tumors continues to be a key clinical issue. The paper will suggest a lightweight deep learning model to classify the images of CT scan images of kidney disease into four categories: Cyst, Stone, Tumor, and Normal. The suggested method aims at resource-aware and high-efficiency classification, based on the use of MobileNetV2 with ImageNet pre-trained weights, and fine-tuned with Adam optimizer. The imbalance in classes is solved by augmenting images, and each category has 1,377 samples. The final accuracy of the model is 95.95% and F1-score of 0.9596, being better than ResNet18, AlexNet and ShuffleNetV2 and having much lower computational cost. A web based Flask application is further extended to allow real-time clinical execution thus the proposed system is feasible in resource constrained medical settings.**

***Keywords*—Deep Learning, Medical Image analysis, MobileNetV2, Transfer Learning, Image Augmentation, Image Classification, Flask Web Application.**

## I. INTRODUCTION

Chronic kidney disease and other related renal diseases like kidney cysts, kidney stones and tumors have risen as a major health problem of the masses in the world and a monumental health burden globally [1]. Among other essential physiological functions, the kidney performs the functions of filtering the metabolic waste products, electrolyte and fluid balance, and blood pressure regulation. Any structural or functioning impairment, left unattended, will proceed to their end-stage renal disease, cardiovascular complications and premature death in a cascade of silent phases until death ensues [2]. One of the most outstanding risk factors of CKD is diabetes mellitus and provokes the acceleration of the development of the nephrons by the oxidative stress generated by persistent hyperglycemia and therefore diabetic kidney disease is one of the primary reasons of renal failure in the world [3]. The traditional diagnosis of kidney diseases has relied on laboratory results that encompass serum creatinine and glomerular filtration rate besides imaging such as the computed tomography scanning [4]. These methods have a clinical application with significant disadvantages of the significant delay before results, expense and is likely to be affected by inter-observer variability - particularly when attempting to detect weak morphological changes between cysts, stones, tumor and normal renal tissue [5], [6]. The recent and rapid growth of the artificial intelligence and deep learning technologies have provided unprecedented opportunities in the sphere of automated and scalable analysis of the medical images. It is established that deep convolutional neural networks possess extraordinary power to detect characteristic of high-dimensional CT imaging data with high diagnostic accuracy in different renal pathology [7].

*Grenze ID: 01.GIJET.12.2.377_24*

The models that integrate classical machine learning classifiers with the hybrid CNN based models have also demonstrated competitive performance as well as reduced complexity in computation [8]. Multi-architectural deep learning models that are optimally fused have been found to be highly effective in detection of individual conditions of renal diseases [9].
Other architectures constructed over vision transformers have also been applied to improve the analysis of renal images to measure the long range spatial dependencies of the CT images [10]. Explainable residual learning transformer models have also improved the clinical degrees of trust in that they provide high accuracy and interpretable visual representations of spatial attention [11]. Focus on U-Net based on vision transformers, two-stage segmentation and classification pipelines have increased sensitivity to fine abnormalities of the kidney [12]. Special data of CT-based kidney disease have been preserved to address the long term data scarcity issues [13] and hybrid deep neural networks on the fuzzy rough set theory has been demonstrated to be effective with noisy and uncertain clinical data [14].
Even though AI-based kidney disease detection has been advanced, AI based kidney disease detection still possesses some critical gaps. The paper will address these gaps by presenting a four class kidney disease classification system trained with lightweight MobileNetV2.

## II. Literature Survey

The recent advances in the area of AI and DL have affected kidney disease recognition and classification on a massive scale. In an effort to improve the diagnostic accuracy and clinical utility, researchers have tested many architectures, modalities and clinical targets. The contributions most relevant in the five key areas are also analyzed in the section and the research gaps have been concluded.

### *A. ML and Deep Learning to determine the risk of CKD*

The suggested TabNet-based deep learning model proposed by Chowdhury et al. [1] uses attention sequentiality to clinical tabular data, which explains the prediction of CKD progression in diabetic patients in different stages. The ensemble deep learning and the Eurygasters Optimisation Algorithm method was employed by Yosef et al. [2] to select features and optimise the weights of the features to increase the performance with unbalanced clinical data. An overview of different AI models developed by Sawhney et al. [3] on the UCI CKD dataset indicates that both the ensemble and deep neural network models are more sensitive and ROC than the traditional classifiers.

### *B. Explainable and interpretable AI in the diagnosis of kidney diseases*

The SHAP-enhanced XGBoost algorithm developed by Lin et al. [4] is used to predict the sepsis-related AKI to CKD progression, with serum lactate levels and vasopressor use shown to be some of the significant ICU variables. SHAP-optimized XGBoost was utilized by To et al. [5], to distinguish non-invasive diabetic kidney disease and cardiac output is a significant biomarker. Rehan and Aamir [6] revealed the application of AI-assisted chronic kidney disease (CKD) diagnosis of nail biomarkers, a non-invasive and inexpensive population screening methodology of laser-induced breakdown spectroscopy (LIBS).

### *C. CNN-Based and Ensemble Deep Learning of CT Image Analysis.*

Ayogu et al. [7] indicated that multi-CNN with soft and hard voting can enhance the detection accuracy of CKD on CT images of visually similar pathological cohorts. Ramu et al. [8] proposed a hybrid CNN-SVM model that was trained on deep features rather than the softmax head. This model better classified and recollected all classes of CKD.
Asif et al. [9] created a multi-CNN fusion network with the best possible weights for finding kidney stones. This was done through the use of switchable inter-network contributions that were ablation tested to create the best CT performance.

### *D. ViT and Hybrid Segmentation Structures.*

Alzughaibi et al. [10] introduce a ViT-CNN hybrid, that is, a combination of global self-attention and local feature extraction to enhance the differentiation between kidney pathologies. Firos et al. [11] used a ResNet-transformer hybrid with Grad-CAM visualisation, RLTNT, which had better benchmarks with CNN and ViT and was able to offer spatial explainability.
Sudha and Senthamarai [12] employed a two-stage pipeline, which integrates the Attention U-Net to come up with the renal segmentation and ViT classification, which removes noise in the background and maximizes sensitivity to small calculi and cysts at an early stage.

*E. Construction of Hybrid Fuzzy DL and Dataset.*

To solve the problem of the lack of data to carry out research on renal AI, Abdalla et al. [13] offered a standardised annotated axial CT kidney stone data set. To combine uncertainty and noise in the incomplete clinical data, Liu et al. [14] suggested a hybrid deep learning system that diagnoses chronic kidney disease (CKD) with the incorporation of the kernel multi- granularity fuzzy rough sets into the deep learning system.

*F. Overview and Research Gaps.*

Significant progress has been made in AI-based kidney disease detection, and they include clinical models, transformer-based models, and imaging models. However, there remain significant weaknesses: studies are focused on individual pathologies, most architectures are resource-hungry, there is an imbalance between classes that limits their realistic application, and they have not been applied to clinical practice. The gaps are filled in this paper by creating a lightweight MobileNetV2-based four- class framework (Cyst, Normal, Stone, Tumor) and comparing it with ResNet18, AlexNet, and ShuffleNetV2 and a web interface based on Flask that will be deployed in the clinical environment.

## III. Proposed Methodology

The proposed kidney disease detection system as shown in Fig.1 starts with a data collection process that will collect kidney medical images (CT scans) which will then be divided into four different groups that include Cyst Normal Stone and Tumor. To correct the system's initial class imbalance, the workflow creates a comprehensive data balancing methodology. System first checks if each class has enough images to meet the required standard of 1000 images per class which needs to be fulfilled before it implements multiple augmentation techniques that include rotation up to 10 degrees and horizontal mirror imaging and brightness adjustment at $\alpha$=1.1 and 20. The augmentation procedure continues until all classes reach their target image count which will create a balanced dataset that contains accurate image transformations needed to duplicate actual clinical imaging conditions while preventing overfitting with the extensive ImageNet dataset. The image sample is undergoes a series of standard transformations which include tensor conversion to a PyTorch compatible tensor as well as resizing the dimensions of the image to 224x224 pixels, which are the specified dimensions for input into pre-trained models as well as the normalization of the images using ImageNet normalization which has given mean values of [ 0.485 0.456 0.406] as well as specified standard deviation values. The preprocessing of the images ensures that the essential features of medical images are secured as required by the models for accurate classification of the conditions of the kidneys as well as the standardization of the input images for all the models. The processing of the images involves the creation of batches of 32-sized groups. The system achieves optimal GPU memory usage.


Fig.1. Proposed Kidney Disease Detection Workflow

The project uses transfer learning with four selected deep learning models which each provide different benefits for medical image classification tasks. ResNet18 uses residual connections which enable effective training of deeper network structures. ShuffleNet V2 applies channel shuffle operations combined with efficient group convolutions to achieve high performance while using less computational resources. MobileNet V2 uses depthwise separable convolutions combined with inverted residual blocks to create a lightweight design. In order to create the new classifier layers that are intended to identify the four kidney conditions, the transfer learning process starts with feature extractions using frozen pre- trained weights.The system uses computational resources to train only the new classifier layers while maintaining the feature representations that models have developed through training on millions of natural images.

All models use MobileNetV2 ResNet18 ShuffleNetV2 and AlexNet as base models to create their fully connected output layers which produce four different results. The system employs the use of CrossEntropyLoss along with a learning rate of 0.001 for the Adam optimiser. The system uses the same parameters to judge how well it works. These parameters involve the same preprocessing techniques, 32 batches and five epochs. Evaluation process of the system employs the use of five different metrics, which involve the confusion matrix, accuracy, precision, recall, and the F1 score. Also, the evaluation process of the system involves the assessment of the time taken to train the system. The system employs the use of the MobileNetV2 model as the base because of the efficiency and how well the model works. Efficiency of an model is improved by engineers to create an kidney mobile net pth model. The model operates based on the creation of a flask web application.

## IV. Implementation and Result

### A. Dataset Description

For this study, the existing medical images of the kidneys were used. The images were divided in four classes cyst, normal, stone, and tumor. Data was not balanced as it was based on the initial data that the researchers used for their study. The data balancing was applied to this data to make all the classes balanced by selecting the images accordingly. This helps the model to learn all four classes of data in an unbiased way.

### B. Preprocessing of Data

Before the training process commences, the system demands that all the image data should be preprocessed. After determining the image size to 224 pixels, the system presupposes that all the images should be transformed into tensors. The standard deviation and mean values must be normalized using the standard mean and standard deviation according to the system. The data should be augmented to increase the data and enhance the generalization of the model.

### C. Model Selection and Design

The study employs several deep learning models including the AlexNet, ShuffleNetV2, ResNet18 and MobileNetV2 models. The models proposed by the researchers are good illustrations of models that perform best in image classification tasks. To conduct the classification of the four categories of kidney conditions, the researchers constructed new layers with the use of the ready models which were developed by other researchers. The existing fully connected layers will be modified by introducing new layers consisting of neurons that are equivalent to the four classes of outputs. The best model was identified as the MobileNetV2 due to its model accuracy and low cost of calculation.

### D. Training Phase

The Adam optimizer, the CrossEntropyLoss function, model is trained using 0.001 as the learning rate. Model was trained with five epochs and a batch size of 32. Thereby, model's accuracy and loss are also calculated. MobileNetV2 has been trained to 10 more epochs to achieve better performance. Finally, the Flask web application has preserved the model to be utilized in prediction. Accuracy, precision, recall, and F1-score are used to assess the models' performance. The prediction of the classes is determined using the confusion matrix.MobileNetV2 proven strong performance and accuracy across all classes. Every class had the same level of precision. The outcomes prove how well the suggested system classifies kidney diseases and how a web application could use for live prediction.

### E. Performance Matrix

The models' performance is assessed using common metrics like accuracy, precision, recall, and F1-score. With input from four kidney disease classes, these metrics provide a breakdown of the classification accuracy. With an accuracy of 95.95, MobileNetV2 outperformed all other models, followed by ResNet18 (88.29%), AlexNet

(81.49%), and ShuffleNetV2 (76.32%). The outcome prove that MobileNetV2 performs well in generalisation as compared with other models. In comparison to ShuffleNetV2, which attained a comparatively lower accuracy of prediction of equal measure of 0.7651 at the same time that ResNet18 and Alexnet recorded 0.8829 and 0.821, respectively, MobileNetV2 achieved high prediction accuracy using the macro average of 0.9601.The macro average of 0.9601 indicates that the value of recall demonstrates that MobileNetV2 has a high sensitivity in all the classes than ShuffleNetV2 whose sensitive value was 0.7627.Since ResNet18 had a higher sensitivity of 0.8825, ShuffleNet had a higher sensitivity of 0.7627, and AlexNet had a higher sensitivity of 0.8151, all true positive cases were less likely to be identified. The models' performance is assessed using common metrics like accuracy, precision, recall, and F1-score. Furthermore, MobileNetV2's F1-score (0.9596) was the highest, then ResNet18 (0.8824), AlexNet (0.8152), and ShuffleNetV2 (0.7601). F1-score is the best metric since it balances recall and precision. By and large, the findings do show that MobileNetV2 is the most appropriate model using the outcomes to classify kidney disease.

TABLE I: PERFORMANCE METRICS

| Algorithm | Macro Avg Precision | Macro Avg Recall | Macro Avg F1-Score | Final Accuracy |
|---|---|---|---|---|
| MobileNetV2 | 0.9601 | 0.9595 | 0.9596 | 95.95% |
| ResNet18 | 0.8829 | 0.8825 | 0.8824 | 88.29% |
| Alex Net | 0.8210 | 0.8151 | 0.8152 | 81.49% |
| ShuffleNetV2 | 0.7651 | 0.7627 | 0.7601 | 76.32% |

Table 2 lists the classifications. Performance metrics show that the model has high precision, recall, and F1-score and performs well across all classes. The models' performance is assessed using common metrics like accuracy, precision, recall, and F1- score. Model's implementation for tumor class is found to be the best among all disease types. This is due to the model's excellent tumor type classification. By contrast, the performance of the model for stone type is slightly lower since there is a smaller value of recall for this type. This arises from the fact that some of images under this class might be mistaken with images under other classes depending on visual features of images. The models' performance is assessed using standard metrics like accuracy, precision, recall, and F1-score. The model consistently performs well in terms of precision and recall for both the normal and cyst classes. All classes have high values (more than 0.95) and hence it signifies that correct classification of instances from all respective classes as there no biasing towards a specific class.

TABLE II: CLASS-WISE TABLE

| Label | Precision | Recall | F1-Score |
|---|---|---|---|
| Cyst | 0. 9552 | 0. 9593 | 0. 9572 |
| Normal | 0. 9733 | 0. 9521 | 0. 9626 |
| Stone | 0. 9286 | 0. 9731 | 0. 9504 |
| Tumor | 0. 9835 | 0. 9535 | 0. 9683 |

## V. EXPERIMENTAL RESULT ANALYSIS

The researchers also used an experimental evaluation method, which helped them understand how the different deep learning models performed while carrying out the task for which the model was designed. In the study, the researcher identified four different models for the evaluation, which included MobileNetV2, ResNet18, AlexNet, and ShuffleNetV2. In order to establish unbiased evaluation criteria, the researcher ensured that the same training and evaluation procedures were followed for all the models. From the model comparison graph Fig. 2, it can be identified that the highest accuracy, i.e., 95.95%, was achieved by the MobileNetV2 model, followed by the ShuffleNetV2 model with an accuracy of 76.32%. ResNet18 attain an accuracy of 88.29%, whereas AlexNet attain an accuracy of 81.49%.

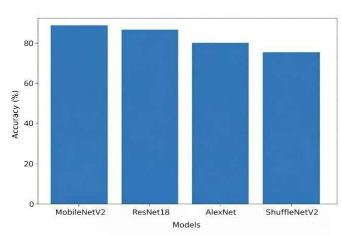


Fig.2. Model Accuracy Comparison

The most successful model would be the one that is shown in Fig.3. Based on the matrix, it is also evident that most of the predictions have been established to be correct as the values in the diagonal are high. The model is demonstrating the capability of distinguishing the four classes as it is capable of distinguishing between the Cyst normal Stone and Tumor classes that are alike in certain aspects. The model is doing well as there are minimal mistakes against the number of right predictions. MobileNetV2 demonstrates its effectiveness by the assessment metrics that provide more than 0.95 values of precision, recall, and F1-score of their macro averages. Lightweight of ShuffleNetV2 was the reason why it had a low performance in comparison with other models, whereas AlexNet and ResNet18 received average performance results. Moreover, the time and computation efficiency of the models were considered in its evaluation process as well.
Though the training time of the MobileNetV2 was a little more than the training time of ShuffleNetV2 model, its provided accuracy was significantly greater. ResNet18 model also offered good accuracy at the expense of increased computational complexity. The AlexNet model was an intermediate accuracy and low training time model. The outcome of testing prove that MobileNetV2 is the best model to be used in this particular application as it provides the trade-off between the efficiency and the accuracy. The model is proved to be applicable in the real-time detection of kidney disease through the establishment of a web-based system that integrates the proposed model.

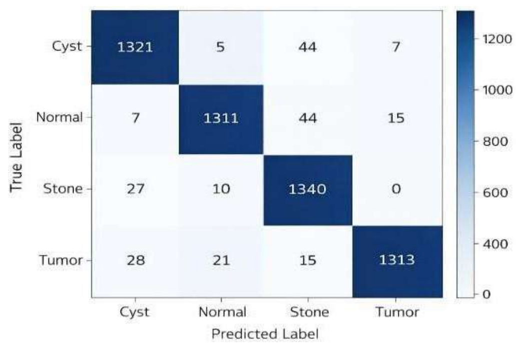


Fig.3. Confusion Matrix MobileNetV2

## VI. CONCLUSION

This study suggested using deep learning to improve kidney disease diagnosis using medical imaging. Different CNN models were explored and compared in details, among which are MobileNet V2, ResNet18, AlexNet, and ShuffleNet V2. The model MobileNetV2 received the highest accuracy 95.95% and equivalent values of recall and F1-score and thus supporting the conclusions. A confusion matrix analysis was conducted to ensure that the model is capable of classifying the four categories (Cyst, Normal, Stone and Tumor). Contrarily AlexNet and ResNet18 performed averagely while ShuffleNetaV2's weaker low weight architecture led to its underperformance. The MobileNetV2 model was implemented with the help of a web application on Flask that will allow making live prediction in terms of image uploaded by the user. This incorporation explains how proposed system can be helpful in detection early stage kidney disease diagnosing system. At all the level presented this system is well reliable effective way of detecting kidney diseases. Larger datasets with complex structures and additional clinical features can be used in further research to improve the system's accuracy.

## REFERENCES


[1] M. N. H. Chowdhury et al., "Deep learning for early detection of chronic kidney disease stages in diabetes patients: A TabNet approach," Artif. Intell. Med., vol. 166, p. 103153, Aug. 2025, doi: 10.1016/j.artmed.2025.103153.

[2] S. M. Awad Yousif, H. T. Halawani, G. Amoudi, F. M. Osman Birkea, A. M. R. Almunajam, and A. A. Elhag, "Early detection of chronic kidney disease using eurygasters optimization algorithm with ensemble deep learning approach," Alexandria Engineering Journal, vol. 100, pp. 220– 231, Aug. 2024, doi: 10.1016/j.aej.2024.05.011.

[3] R. Sawhney, A. Malik, S. Sharma, and V. Narayan, "A comparative assessment of artificial intelligence models used for early prediction and evaluation of chronic kidney disease," Decision Analytics Journal, vol. 6, p. 100169, Mar. 2023, doi: 10.1016/j.dajour.2023.100169.

[4] Y. Lin et al., "An interpretable XGBoost model for risk prediction of progression from sepsis-associated acute kidney injury to chronic kidney disease," Inform. Med. Unlocked, vol. 58, p. 101685, 2025, doi: 10.1016/j.imu.2025.101685.

[5] N. M. To et al., “Refine XGBoost with SHAP explainability for non-invasive early detection of diabetic kidney disease: Estimated cardiac output as a potential indicator,” Comput. Methods Programs Biomed., vol. 273, p. 109122, Jan. 2026, doi: 10.1016/j.cmpb.2025.109122.

[6] I. Rehan and M. Aamir, “AI-enhanced non-invasive diagnosis of chronic kidney disease using LIBS of fingernail biomarkers,” Spectrochim. Acta A Mol. Biomol. Spectrosc., vol. 355, p. 127687, Jul. 2026, doi: 10.1016/j.saa.2026.127687.

[7] I. I. Ayogu, C. F. Daniel, B. A. Ayogu, J. N. Odii, C. L. Okpalla, and E. C. Nwokorie, “Investigation of ensembles of deep learning models for improved chronic kidney diseases detection in CT scan images,” Franklin Open, vol. 11, p. 100298, Jun. 2025, doi: 10.1016/j.fraope.2025.100298.

[8] K. Ramu et al., “Hybrid CNN-SVM model for enhanced early detection of Chronic kidney disease,” Biomed. Signal Process. Control, vol. 100, p. 107084, Feb. 2025, doi: 10.1016/j.bspc.2024.107084.

[9] S. Asif, X. Zheng, and Y. Zhu, “An optimized fusion of deep learning models for kidney stone detection from CT images,” Journal of King Saud University - Computer and Information Sciences, vol. 36, no. 7, p. 102130, Sep. 2024, doi: 10.1016/j.jksuci.2024.102130.

[10] A. Alzughaibi, A. A. Alanazi, M. Alshahrani, I. Hilali Jaghdam, and A. A. Hassaballa, “Synergizing vision transformer with ensemble of deep learning model for accurate kidney stone detection using CT imaging,” Alexandria Engineering Journal, vol. 127, pp. 357–373,Aug. 2025, doi: 10.1016/j.aej.2025.05.025.

[11] F. V.M., A. P.J.A., U. Fiore, and G. G.R., “RLTNT: An explainable residual learning-based transformer model for kidney disease classification,” Image Vis. Comput., vol. 163, p. 105781, Nov. 2025, doi: 10.1016/j.imavis.2025.105781.

[12] D. N. Sudha and N. Senthamarai, “Attention U-Net model and Vision transformer-based for segmenting and classifying kidney disease,” Biomed. Signal Process. Control, vol. 120, p. 109914, Jul. 2026, doi: 10.1016/j.bspc.2026.109914.

[13] P. A. Abdalla, M. Y. Shakor, A. K. Ameen, B. S. Mahmood, and N. R. Hama, “Kidney stone detection via axial CT imaging: A dataset for AI and deep learning applications,” Data Brief, vol. 59, p. 111446, Apr. 2025, doi: 10.1016/j.dib.2025.111446.

[14] J. Liu, B. Sun, J. Ye, X. Zhao, and X. Chu, “Hybrid deep- learning prediction model based on kernel multi-granularity fuzzy rough sets and its application in the diagnosis and treatment of chronic kidney disease,” Eng. Appl. Artif. Intell., vol. 147, p. 110297, May 2025, doi: 10.1016/j.engappai.2025.110297.