

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

## Automated Kidney Disease Detection Using a Dual-Attention and Transformer-Driven Deep Learning Model

**Muhammad Suleman Memon**

Department of Information Technology, Dadu Campus, University of Sindh

**Mumtaz Qabulio**

Department of Software Engineering, FET, University of Sindh, Jamshoro

**Shabana**

Department of Computer Science Government Aisha Girls Degree College  
Nawabshah

**Nimra memon**

Lecturer Department of computer science Government Girls Degree College  
Nawabshah

### ABSTRACT

Kidney diseases are one of the significant health challenges in the world, and they need to be properly diagnosed and in time to avoid the development of the disease and its complications. Computed tomography (CT), ultrasound, and magnetic resonance imaging (MRI) are medical imaging modalities that are important in the evaluation of kidney disease, but manual interpretation is time-consuming and prone to inter-observer variation. This research aims to overcome these issues by establishing KDAT-Net, a new hybrid deep learning architecture that incorporates convolutional neural networks, dual attention mechanisms, and transformer-based global feature modeling to detect kidney diseases automatically. The proposed model uses an InceptionV3 backbone to extract multi-scale spatial information of kidney images. Efficient Channel Attention (ECA) is used to learn inter-channel dependencies with the aim of improving discriminative representation learning, and a spatial attention mechanism to focus attention on disease relevant parts of the anatomy is incorporated. In contrast to the more traditional CNN-based methods which use only local receptive fields, the refined feature maps are converted to a sequence of tokens and encoded by transformer encoder blocks, which allows the modeling of long-range spatial relations as well as global contextual information. The resulting characteristics are then grouped and categorized into various kidney disease types through fully connected layers. Substantial experimental assessments show that the suggested framework is more effective than baseline CNN models and attention-enhanced variants in accuracy, precision, recall, and F1-score. Moreover, the qualitative analysis based on the gradient-based class activation maps demonstrates that the model can emphasize clinically important kidney areas, which makes it easier to interpret and increase its credibility. The proposed

KDAT-Net is a powerful and scalable solution to automated kidney disease detection and can assist in clinical decisions in real-world diagnostic processes.

## I Introduction

Kidney is a vital organ in human body which helps in removing the excess water. The kidney functions are influenced by many factors such as cardiovascular disease, diabetes, obesity etc [1]. Kidney diseases are a major global health problem affecting millions of people and generally with severe complications such as chronic kidney disease (CKD), kidney failure, and the risk for cardiovascular disease. Early and accurate diagnosis is very important for treatment planning and to avoid permanent damage to the kidney.

Medical imaging modalities, such as computed tomography (CT), ultrasound and magnetic resonance imaging (MRI), are commonly used to detect structural abnormalities, e.g. kidney stones, cysts, tumors, and other pathological conditions. Although they are clinically significant, kidney images interpretation is heavily reliant on trained radiologists, which makes the diagnostic process slow and prone to inter- and intra-observer errors. Most recent developments in deep learning, especially convolutional neural networks (CNNs) have proven to be highly successful in medical image analysis, such as disease classification, segmentation and lesion detection. CNN-based models are great at deriving hierarchical spatial features, but due to their nature of leveraging local receptive fields, they are unable to derive long-range contextual relationships that are usually important in medical imaging. Additionally, conventional CNN designs are likely to have equal representations of all feature channels and spatial areas, which may fail to recognize disease-specific patterns of kidney images. These limitations have been addressed by introducing attention mechanisms to selectively highlight informative features and block irrelevant information. The approaches to channel attention, including Efficient Channel Attention (ECA), provide the opportunity to re-weight feature channels, which adapt to the changes with low computational cost, whereas the spatial attention mechanisms enable the model to pay attention to the anatomically relevant areas. Despite the improved performance of attention-enhanced CNNs, they are yet to effectively represent global spatial dependencies in the whole image. Although transformers were originally designed to process natural language, they have recently become popular in computer vision because of their strong self-attention mechanism that can capture long-range dependencies. Transformer-based architectures are able to represent global contextual information that completes local CNN representations by processing image features as sequences of tokens. Nevertheless, standalone vision transformers are generally large scale, and computationally intensive, which limits their use in medical imaging applications where labeled data is usually limited. Based on these findings, this paper presents KDAT-Net, a hybrid deep learning model that optimally integrates CNN-based feature extraction, dual attention mechanisms, and transformer-based global modeling to detect kidney diseases. The model suggested will employ an InceptionV3 backbone to extract multi-scale features, and ECA-based attention and spatial attention to further refine discriminative representations. The attention-enhanced features are then tokenized and fed to transformer encoder blocks which allow efficient modeling

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

of the local and global contextual information.

The contributions of this work can be summarised as:

1. We propose the new hybrid CNN-dual attention-transformer model for automatic detection of kidney disease, which combines local feature extraction and global feature modelling.
2. The proposed approach effectively integrates channel and spatial attention to enhance feature discrimination in an efficient execution.
3. Long-range spatial correlations of attention enhanced feature tokens are inferred in transformer encoder blocks.
4. The large scale of the experiments and interpretability of the proposed framework demonstrate the effectiveness, robustness and clinical potential.

The rest of this paper is organised as follows: Section 2 reviews the previous works on the detection of kidney diseases and deep learning models with attention mechanism. Section 3 details the proposed network architecture and approach. Experiments and discussion are presented in Section 4. Section 5 presents the discussion of results and clinical significance and Section 6 concludes the paper with future research suggestions.

## II Related Work

In [2] the author presents a system for detection and classification of chronic kidney diseases (CKDs) in ultrasound images. Firstly HOG features and AdaBoost classifier are used to detect the diseased area followed by a Convolutional Neural Network (CNN) for the classification of kidney stones, cysts and renal carcinoma. The framework uses anisotropic diffusion to remove speckle noise from the images. The empirical results show that the accuracy of classification is 85.21% which demonstrates the strength of the hybrid conventional and deep learning approaches.

This paper [3] proposes an IoT and cloud-based deep learning framework (FPA-DNN) for the classification of chronic kidney disease (CKD). The model integrates Oppositional Crow Search (OCS) for feature selection and Flower Pollination Algorithm (FPA) for optimizing deep neural network parameters. The system collects patient data through IoT devices performs preprocessing and then classifies CKD using a tuned DNN model. Experimental results on the CKD dataset show good performance with 98.75% accuracy, 98.80% sensitivity and 99% F-score, outperforming traditional methods.

This paper [4] proposes HDLNet a hybrid deep learning framework which is integrated with IoT for early detection and classification of chronic kidney disease (CKD). The model is good. The proposed model combines Deep Separable CNN (DSCNN) and Capsule Networks (CapsNet) for effective feature extraction and classification. The model also uses Aquila Optimization (AO) which is used for feature selection while Sooty Tern Optimization Algorithm (STOA) fine-tunes model parameters for improved performance. The system uses IoT-based medical data and shows very good results achieving around 99% accuracy with improved sensitivity and specificity compared to existing methods.

The author in [5] proposes a machine learning-based framework for the classification of chronic kidney disease (CKD) using multiple supervised algorithms. Proposed

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

model evaluates nine models including Random Forest, SVM, KNN and AdaBoost along with preprocessing techniques such as missing value imputation and feature selection. The study shows that proper preprocessing that improves prediction performance. Experimental results show that AdaBoost achieves the best performance with up to 99.17% accuracy outperforming other models. Overall the work show the effectiveness of ML techniques for early and accurate CKD diagnosis.

The author in [6] offers a holistic machine learning framework for chronic kidney disease (CKD) classification on the UCI repository. It emphasizes rigorous data preprocessing, including missing value imputation, SMOTE oversampling of the minority class, scaling and chi-square feature selection, for better accuracy. It applies Support Vector Machine (SVM) and Random Forest algorithms, along with GridSearchCV hyperparameter tuning to improve performance. This study shows SVM is the most accurate (99.33%) and reliable (98% cross-validation). Overall, this work demonstrates that good data preprocessing and tuning improve accuracy and stability of CKD classification.

The author in [7] the authors propose using a hybrid deep learning model, a combination of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM), for automatic diagnosis of chronic kidney disease (CKD). The dataset includes 400 samples and 13 features, and CNN is used for feature extraction and LSTM to capture the temporal dynamics. The model's performance is compared to traditional machine learning methods such as KNN, SVM, logistic regression, random forest and AdaBoost. The CNN-LSTM model outperforms other classifiers with an accuracy of 99.17%. The study shows the feasibility of using hybrid deep learning models for early and accurate diagnosis of CKD.

The author in [8] presents a machine learning approach with explainable artificial intelligence (XAI) to diagnose chronic kidney disease (CKD). It compares a range of models but finds the Extreme Gradient Boosting (XGB) model performs the best. The paper uses SHAP and Partial Dependency Plots (PDP) for explanations of model decisions, overcoming the black box nature of traditional models. It also develops an interactive visual tool to present predictions and explanations to aid diagnosis. The paper concludes by highlighting the need for both prediction accuracy and interpretability in medical applications.

The author in research [9] introduces a novel deep learning approach for diagnosing kidney diseases from CT scans. It combines features of AlexNet and ConvNeXt models, feature concatenation and custom Adam optimization to boost accuracy. The proposed method is validated on a large dataset of 12,446 images of cyst, tumor, stone, and normal cases, with a remarkable accuracy of 99.85%. The proposed model uses Adam optimizer to achieve a good performance. In conclusion, the research shows hybrid models and optimisation techniques can enhance medical image classification.

The author in [10] highlights the scoping review of the use of artificial intelligence (AI) for the identification of chronic kidney disease (CKD). The paper reviews 41 papers which uses methods such as machine learning, deep learning, natural language processing (NLP) and large language models in areas such as early diagnosis, risk prediction, treatment and care. The paper highlights the benefits of AI in improving diagnostic accuracy and improving the efficiency of care. In conclusion the paper

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

suggests that AI has the potential to revolutionize CKD care with validation and partnership.

This study [11] offers an explainable AI approach to predict chronic kidney disease (CKD) in the Medical IoT. The framework adopts generative adversarial networks (GANs) for missing data imputation, and few-shot learning methods (such as prototypical network and MAML) for better prediction. The research also integrates explainable AI techniques such as SHAP and LIME for improved interpretability and trust in the model. The proposed framework is evaluated with experimental results indicating that it yields better performance than conventional machine learning models with close to perfect accuracy. In summary, this framework offers an effective approach for early diagnosis and decision support for CKD.

In [12] the author presents a deep learning model based on the VGG16 architecture to diagnose and classify chronic kidney disease (CKD) using CT scans. It builds on existing approaches by incorporating Grad-CAM, which offers explainability by visualising relevant areas in medical images, making it easier for doctors to understand. It achieves effective detection and classification of normal, cyst, tumor and stone in the kidney. The paper's results show an accuracy of 98%, suggesting high diagnostic performance. In conclusion, the paper effectively combines high accuracy and explainability in the diagnosis of CKD.

### III Methodology

#### Dataset

The dataset contain 16268 images for train set and 4064 images for test set. The figure 1 shows the bar chart of train and test set.



Figure 1: Images in Train and Validation

Our KDAT-Net is a hybrid deep learning approach that aims to provide a good accuracy, generalizability and interpretability for classifying kidney diseases from CT scans. The proposed architecture integrates the CNNs, attention models and



transformers into a single framework. Figure 2 shows the architecture of the proposed architecture.

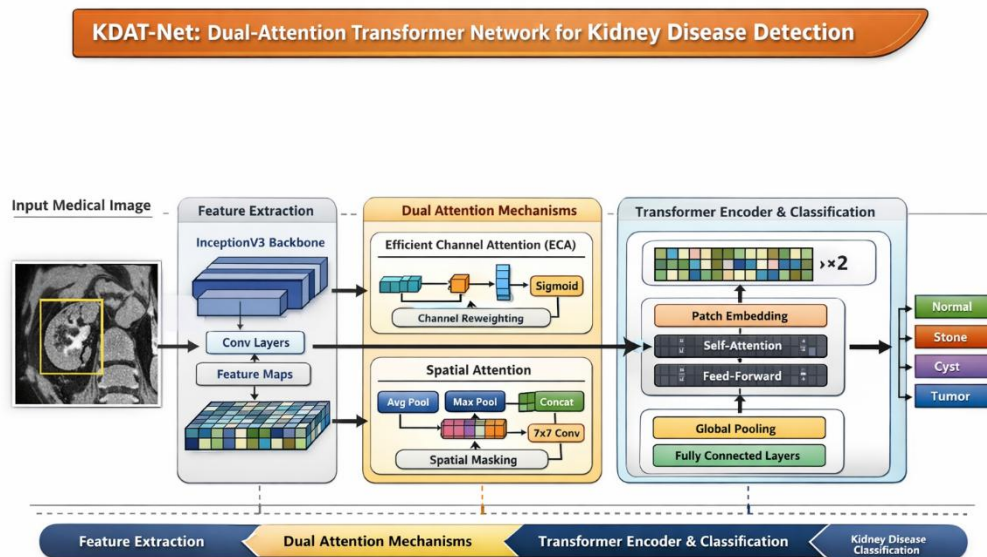


Figure 2: Proposed Architecture

## 1. Deep Features through InceptionV3

The first component in the framework is a pretrained InceptionV3 network is used as the feature extractor.

In kidney CT scans for instance, stones, cysts or tumors can take many different forms and sizes. InceptionV3 achieves this by applying different sized convolutional filters allowing it to capture local and global patterns in the image.

## 2. Feature Refinement with Dual Attention

Once features have been extracted the model uses two attention mechanisms to refine the representations:

### Efficient Channel Attention (ECA)

ECA is concerned with "what is important" in the feature maps. It learns to weigh the importance of different channels to ensure relevant and informative features (e.g., intensity changes suggesting stones) are enhanced and less important features are down-weighted.

### Spatial Attention

Spatial attention pays attention to "where in the image the information is". It pinpoints parts of the kidney where there is more likely to be a lesion or calcification.

These attention modules help the model to:

- Pay attention to clinically relevant areas
- Reduce background noise
- Enhance features for subsequent processing

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

### 3. Holistic Context Modeling with Transformer Encoder

Although CNNs excel at learning local contextual information, they fail to learn global information. KDAT-Net overcomes this challenge by incorporating transformer encoder blocks.

The improved feature maps are transformed into tokens and fed into the transformer encoder which, in turn, learns:

- Self-attention is used to understand the relationships between pixels in the image
- The network understands how different areas of the kidney image are connected

This is useful in medical imaging since:

- Patterns of disease may be spread throughout the image
- The context may be useful to differentiate similar diseases

The transformer module therefore helps the model to reason about the entire image.

### 4. Pooling and Classification

The transformer-processed features are then distilled into a vector representation. They represent the key features obtained from the image.

It is then fed into fully connected (dense) layers for final classification into classes such as:

- a. Normal
- b. Kidney stone
- c. Cyst
- d. Tumor

These layers serve as classification modules transforming learned features to class probabilities.

### 5. Explainability using Grad-CAM

One of the major highlights of KDAT-Net is its emphasis on explainability which is essential for medical tasks.

KDAT-Net uses Grad-CAM (Gradient-weighted Class Activation Mapping) for visual explanations. Grad-CAM generates heatmaps that focus on the areas of the image which are highly important.

This provides several benefits:

- a. Assists clinicians to interpret model predictions.
- b. Helps ensure the model is paying attention to the right parts of the body.
- c. Provides transparency and confidence in AI diagnosis

For example if the model has diagnosed a kidney stone the Grad-CAM will show you where the stone is.

### 6. Computational Efficiency

While transformer models are efficient, they can be slow. KDAT-Net is designed to overcome this:

- a. Transformer is applied after feature reduction, reducing input size.
- b. Number of transformer blocks is low.
- c. Attention is light (ECA, not heavy).

Online ISSN

3007-3197

Print ISSN

3007-3189

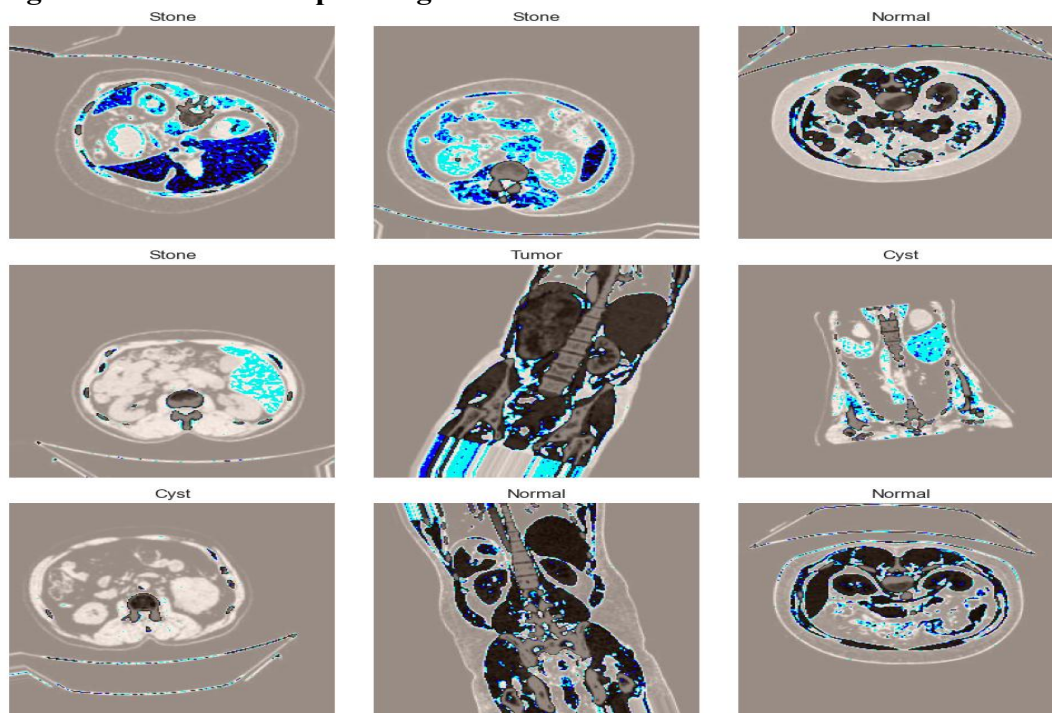
<http://amresearchreview.com/index.php/Journal/about>

This ensures

- Faster training and inference.
- Lower memory consumption.
- Real-time clinical use.

### Results and Discussion

Figure 3 shows the sample images.



**Figure 3: Sample Images**

The proposed model demonstrates superior performance across all disease classes, achieving near-perfect precision and recall for stone and cyst detection, while maintaining high sensitivity for tumor classification, albeit with a slight trade-off in normal class recall.

Table 1 shows the class-wise performance.

**Table 1: Class-wise Performance**

| Class  | Precision | Recall   | F-Score  | Support |
|--------|-----------|----------|----------|---------|
| Stone  | 1.000000  | 0.999016 | 0.999508 | 1016.0  |
| Cyst   | 0.980695  | 1.000000 | 0.990253 | 1016.0  |
| Tumor  | 0.915315  | 1.000000 | 0.955786 | 1016.0  |
| Normal | 1.000000  | 0.888780 | 0.941115 | 1016.0  |

Figure 4 presents a class-wise comparison of Precision, Recall, and F1-score for the four kidney condition categories Cyst, Normal, Stone and Tumor. It visually highlights how well the proposed model performs across different evaluation metrics.



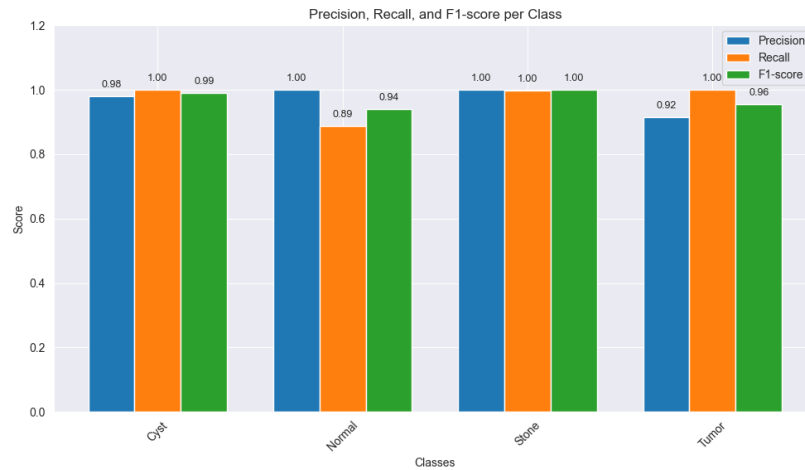
Online ISSN

3007-3197

Print ISSN

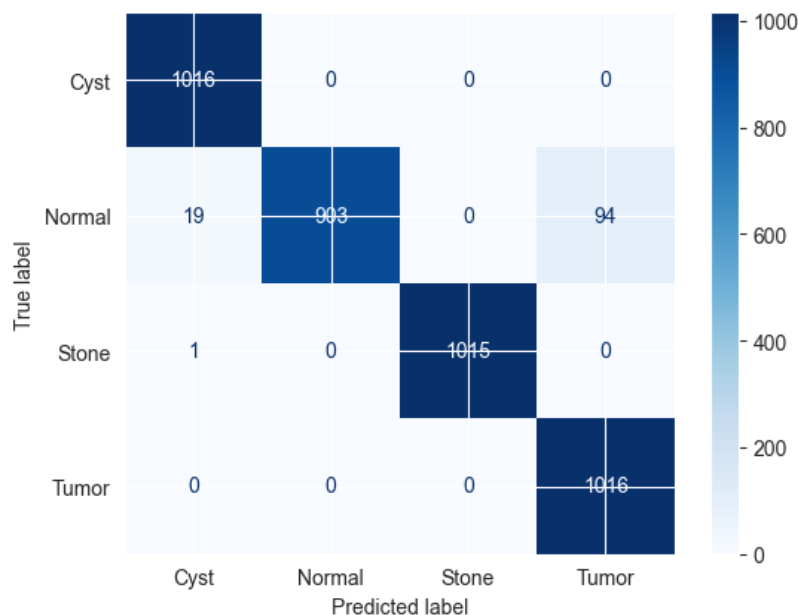
3007-3189

<http://amresearchreview.com/index.php/Journal/about>



**Figure 4: Precision, Recall and F Score**

Figure 5 shows the confusion matrix of the proposed model which provides a detailed view of how well each class is classified by comparing true labels (rows) with predicted labels (columns).



**Figure 5: Confusion Matrix**

Figure 6 presents Grad-CAM (Gradient-weighted Class Activation Mapping) visualizations for multiple CT scan images where all samples are predicted as “Cyst” with high confidence ( $\approx 0.93$ – $0.94$ ). The purpose of this figure is to demonstrate how the proposed KDAT-Net model interprets and localizes disease-relevant regions within the images.

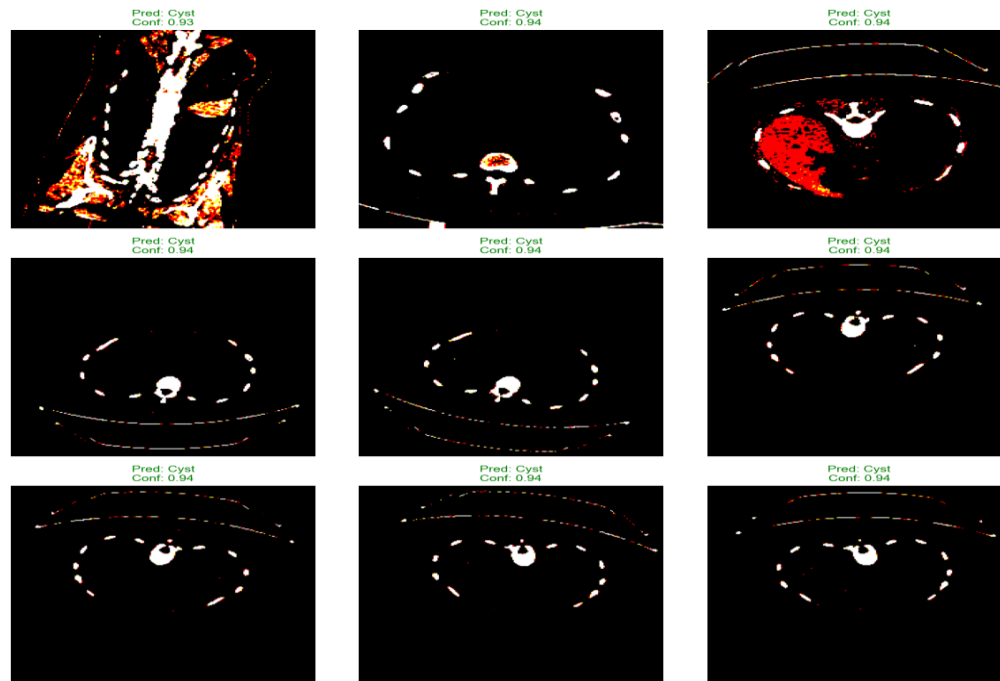
Online ISSN

3007-3197

Print ISSN

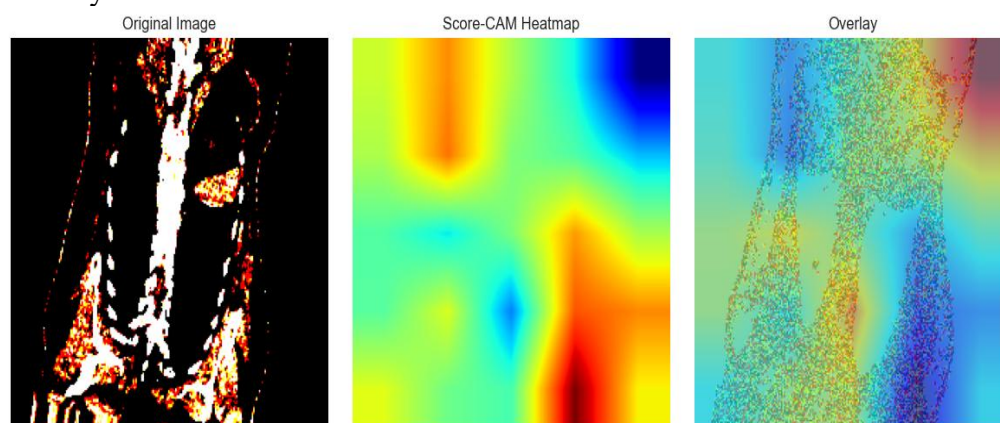
3007-3189

<http://amresearchreview.com/index.php/Journal/about>



**Figure 6: Predicted Images**

Figure 7 illustrates the Score-CAM (Score-weighted Class Activation Mapping) results for a kidney CT image, showing how the model identifies important regions for classification. It consists of three parts: Original Image, Score-CAM Heatmap, and Overlay.



**Figure 7: Score-CAM Heatmap**

Figure 8 demonstrates Score-CAM-based interpretability for a kidney CT image classified as Cyst. It consists of three parts original Image (left), Score-CAM Heatmap (center) and Overlay (right).

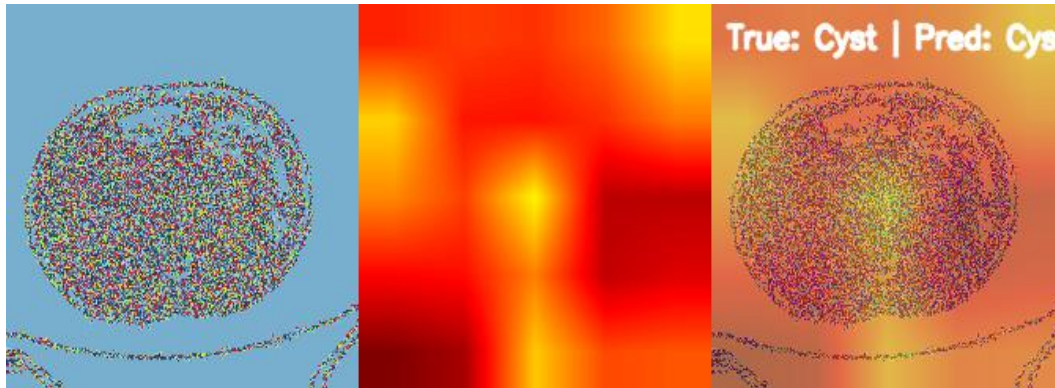
Online ISSN

3007-3197

Print ISSN

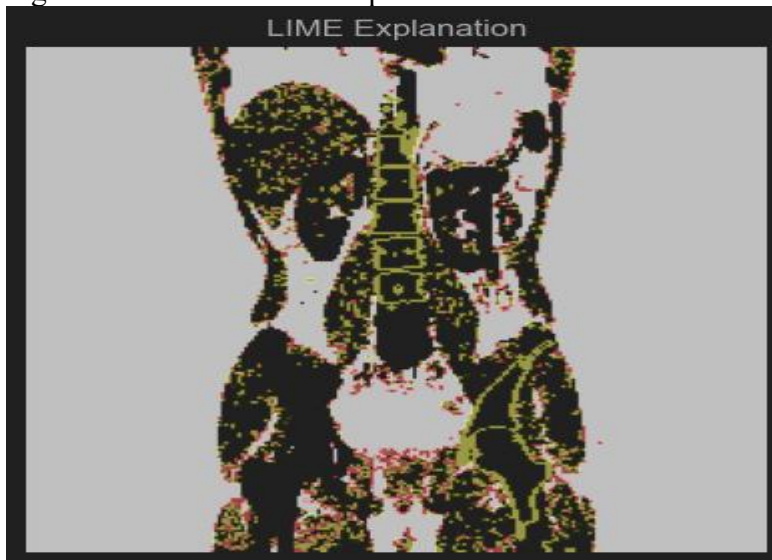
3007-3189

<http://amresearchreview.com/index.php/Journal/about>



**Figure 8: Grad-CAM**

Figure 9 shows the LIME output for a kidney CT image. It provides insight of the regions which influence the prediction of the model.



**Figure 9: LIME Explanation**

## Conclusion

In this paper we proposed a hybrid deep learning model named KDAT-Net. The model combines feature extraction via InceptionV3 a pretrained model with dual attention blocks and global context modeling using transformers to classify kidney diseases from CT images. The proposed system successfully learns both discriminative local features and global contextual information with the addition of Grad-CAM improving interpretability and trust in medical applications. The results of the experiments show the effectiveness of the model with an overall accuracy of 97.19%. The class-wise results demonstrate outstanding performance for stone (F1-score: 0.9995) and cyst (F1-score: 0.9902) and tumor (F1-score: 0.9558) detection also shows high accuracy. While recall for the normal class is slightly lower (0.8887) the overall system performance is solid and clinically satisfactory. These findings indicate the promise of the proposed system for accurate and reliable classification of multiple kidney diseases. Besides the high accuracy the use of explainable AI (Grad-CAM) allows identification of disease-

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

related regions, which makes the system more interpretable, and therefore more clinically feasible. The lightweight nature of the system makes it feasible to use in practice without high computational demands. The KDAT-Net is a robust explainable and efficient system for automated diagnosis of kidney disease. Potential future directions include enhancing performance for difficult classes like normal and using larger and multicenter data for training along with integrating the system into clinical decision support systems in real-time.

## References

- [1] F. Maqsood *et al.*, “Artificial Intelligence-Based Classification of CT Images Using a Hybrid SpinalZFNet,” *Interdiscip. Sci.*, vol. 16, no. 4, pp. 907–925, Dec. 2024, doi: 10.1007/s12539-024-00649-4.
- [2] V. R., “Detection and Classification of Kidney Disorders using Deep Learning Method,” *JOURNAL OF MECHANICS OF CONTINUA AND MATHEMATICAL SCIENCES*, vol. 14, no. 2, Apr. 2019, doi: 10.26782/jmcms.2019.04.00021.
- [3] R. H. Aswathy *et al.*, “Optimized tuned deep learning model for chronic kidney disease classification,” *Computers, Materials and Continua*, vol. 70, no. 2, pp. 2097–2111, 2022, doi: 10.32604/cmc.2022.019790.
- [4] K. Venkatrao and K. Shaik, “HDLNET: A Hybrid Deep Learning Network Model With Intelligent IOT for Detection and Classification of Chronic Kidney Disease,” *IEEE Access*, vol. 11, pp. 99638–99652, 2023, doi: 10.1109/ACCESS.2023.3312183.
- [5] M. M. El Sherbiny, E. Abdelhalim, H. El-Din Mostafa, and M. M. El-Seddik, “Classification of chronic kidney disease based on machine learning techniques,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 32, no. 2, pp. 945–955, Nov. 2023, doi: 10.11591/ijeecs.v32.i2.pp945-955.
- [6] D. Swain *et al.*, “A Robust Chronic Kidney Disease Classifier Using Machine Learning,” *Electronics (Switzerland)*, vol. 12, no. 1, Jan. 2023, doi: 10.3390/electronics12010212.
- [7] E. N. Yildiz, E. Cengil, M. Yildirim, and H. Bingol, “Diagnosis of Chronic Kidney Disease Based on CNN and LSTM,” *Acadlore Transactions on AI and Machine Learning*, vol. 2, no. 2, pp. 66–74, Jun. 2023, doi: 10.56578/ataiml020202.
- [8] G. Dharmarathne, M. Bogahawaththa, M. McAfee, U. Rathnayake, and D. P. P. Meddage, “On the diagnosis of chronic kidney disease using a machine learning-based interface with explainable artificial intelligence,” *Intelligent Systems with Applications*, vol. 22, Jun. 2024, doi: 10.1016/j.iswa.2024.200397.
- [9] S. Elbedwehy, E. Hassan, A. Saber, and R. Elmonier, “Integrating neural networks with advanced optimization techniques for accurate kidney disease diagnosis,” *Sci. Rep.*, vol. 14, no. 1, Dec. 2024, doi: 10.1038/s41598-024-71410-6.
- [10] C. Sabanayagam *et al.*, “Artificial intelligence in chronic kidney disease management: a scoping review,” 2025, *Ivyspring International Publisher*. doi: 10.7150/thno.108552.
- [11] N. G. Rezk, S. Alshathri, A. Sayed, and E. E. D. Hemdan, “Explainable AI for

Online ISSN

3007-3197

Print ISSN

3007-3189

<http://amresearchreview.com/index.php/Journal/about>

Chronic Kidney Disease Prediction in Medical IoT: Integrating GANs and Few-Shot Learning,” *Bioengineering*, vol. 12, no. 4, Apr. 2025, doi: 10.3390/bioengineering12040356.

- [12] T. Panse *et al.*, “CHRONIC KIDNEY DISEASE DETECTION AND CLASSIFICATION USING DEEP LEARNING METHOD-AN EMPIRICAL PROOF,” *J. Theor. Appl. Inf. Technol.*, vol. 15, no. 7, 2025, [Online]. Available: [www.jatit.org](http://www.jatit.org)