IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13251

Deep Learning-Based Detection and Classification of Kidney Stones from Medical Images: A CNN-Driven Diagnostic Framework

Vishnu.T[1], Mrs. N. Vaishnavi[2]

Student, Department of Information Technology, Dr. N.G.P. Arts and Science College, Coimbatore, India[1]

Assistant Professor, Department of Information Technology, Dr. N.G.P. Arts and Science College, Coimbatore, India[2]

Abstract: Kidney stones (nephrolithiasis) represent a widespread urological disorder affecting millions of individuals globally, frequently causing severe pain, obstruction of urine flow, urinary tract infections, and potentially irreversible renal damage when early detection is missed. Traditional diagnostic approaches rely on imaging modalities—principally ultrasound and computed tomography (CT)—which require expert radiological interpretation and may introduce delays or inter-observer variability. This study presents a Convolutional Neural Network (CNN)-based deep learning framework for the automated detection and classification of kidney stones from medical images. The proposed model integrates preprocessing pipelines, data augmentation strategies, hierarchical feature extraction, and rigorous performance evaluation using accuracy, sensitivity, and specificity metrics. Experimental results demonstrate that transfer learning architectures (VGG16, ResNet50, EfficientNet) significantly outperform classical machine learning classifiers and custom CNN designs, particularly when trained on CT imaging datasets. The system offers a cost-effective, scalable, and clinically integrable solution for diagnostic assistance, with the potential to reduce diagnosis time, minimize human error, and enhance patient outcomes.

Keywords: Kidney Stone Detection; Deep Learning; Convolutional Neural Networks; Medical Image Analysis; Transfer Learning; CT Scan; Automated Diagnosis; Nephrolithiasis

I. INTRODUCTION

The kidneys are among the most vital organs in the human body, responsible for filtering metabolic waste from the blood and maintaining electrolyte and fluid homeostasis. Kidney stones, clinically referred to as nephrolithiasis or renal calculi, are hard mineral and salt deposits that form within the kidney. According to global epidemiological data, nephrolithiasis affects approximately 10– 12% of the population in developed nations and is characterized by a high rate of recurrence, imposing a significant burden on healthcare systems worldwide. When left undiagnosed or inadequately managed, kidney stones can cause intense flank pain, hematuria, hydronephrosis, urinary tract infections, and—in severe cases—permanent renal damage.

The clinical gold standard for diagnosing kidney stones is non-contrast computed tomography (CT), owing to its superior sensitivity (95–99%) and specificity (96–98%) in detecting calculi of varying sizes and compositions. CT imaging provides precise anatomical delineation of stone location, size, and density—information that is critical in determining appropriate management strategies. Ultrasound imaging, while avoiding ionizing radiation and offering lower cost and greater accessibility, is limited by operator dependency, reduced sensitivity for small stones, and susceptibility to artifacts arising from speckle noise and poor tissue contrast.

The manual interpretation of these imaging modalities by radiologists is inherently time- consuming, subject to inter-observer variability, and can result in diagnostic delays—particularly in high-throughput clinical environments. The integration of artificial intelligence (AI) into medical imaging has opened promising avenues for automating and augmenting the diagnostic process. Specifically, deep learning methods—particularly Convolutional Neural Networks (CNNs)—have demonstrated remarkable performance in image recognition and classification tasks, making them exceptionally suited for the detection of structural abnormalities in medical images.

This paper proposes an end-to-end deep learning framework for the automated detection and classification of kidney stones from CT and ultrasound images. The objectives of this research are to: (1) develop a robust CNN-based model capable of accurately distinguishing between stone- positive and stone-negative renal images; (2) evaluate the impact of transfer learning on model performance; and (3) assess the feasibility of integrating such a system into routine clinical workflows to support radiologists in timely and accurate diagnosis.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13251

II. RELATED WORK

Early computational approaches to kidney stone detection relied predominantly on classical image processing and traditional machine learning techniques. Methods such as texture-based feature descriptors, edge detection algorithms, histogram analysis, and morphological operations were employed to extract discriminative features from CT and ultrasound images. Classifiers including Support Vector Machines (SVMs), K-Nearest Neighbors (KNN), Random Forests, and Decision Trees were then applied to these extracted features for binary or multi-class classification. While achieving moderate diagnostic accuracy, the performance of these methods was highly sensitive to the quality of handcrafted features, image resolution, and the complexity of the stone morphology.

The advent of deep learning architectures fundamentally transformed the landscape of medical image analysis. Litjens et al. (2017) provided a comprehensive survey of deep learning applications in medical imaging, documenting superior performance across tasks including detection, segmentation, and classification compared to conventional approaches. The U-Net architecture, introduced by Ronneberger et al. (2015), established a foundational framework for biomedical image segmentation and has since been adapted for renal structure delineation. Yasaka et al. (2018) demonstrated the efficacy of CNNs in detecting kidney stones directly from CT images, achieving high sensitivity and specificity metrics competitive with expert radiologists.

Transfer learning—leveraging representations learned from large-scale datasets such as ImageNet—has proven particularly valuable in the medical imaging domain, where the availability of labelled data is often limited by ethical, logistical, and regulatory constraints. Pre-trained architectures including VGGNet, ResNet, Inception, and EfficientNet have been fine-tuned on kidney stone datasets to yield substantial improvements in detection accuracy. Saba et al. (2020) demonstrated that CNN-based automated detection of kidney stones in CT images significantly outperforms traditional ML baselines, underscoring the transformative potential of deep learning for clinical diagnostics. Despite these advances, open challenges remain, including dataset imbalance, lack of model interpretability (the 'black box' problem), and the complexity of deploying AI systems within real-world clinical infrastructures.

III. MEDICAL IMAGING MODALITIES FOR RENAL CALCULI

The selection of an appropriate imaging modality is fundamental to the accuracy, safety, and cost- effectiveness of kidney stone diagnosis. Each modality presents distinct advantages and limitations that influence both clinical utility and the feasibility of downstream deep learning integration.

Ultrasound imaging is typically the first-line modality employed in the initial assessment of suspected nephrolithiasis. Its principal advantages include non-invasiveness, absence of ionizing radiation, portability, and relatively low cost, making it particularly suitable for pediatric and pregnant populations. However, ultrasound imaging is significantly limited by speckle noise, low tissue contrast, and operator dependency—factors that contribute to reduced sensitivity for small or atypically located calculi. These inherent limitations necessitate extensive preprocessing when ultrasound data is used for machine learning model training.

Non-contrast CT is universally recognized as the gold standard for kidney stone diagnosis. CT scans deliver high-resolution cross-sectional images that enable precise characterization of stone morphology, including size, location, and Hounsfield unit density—parameters critical for treatment planning. The superior sensitivity and specificity of CT imaging, combined with its capacity to capture volumetric data, make it the preferred modality for training deep learning models. The primary limitation of CT remains patient exposure to ionizing radiation, necessitating careful consideration in pediatric and repeat-examination contexts. In the context of AI development, CT datasets provide richer and more consistent feature representations, facilitating superior model generalization.

IV. FUNDAMENTALS OF DEEP LEARNING FOR MEDICAL IMAGE ANALYSIS

Convolutional Neural Networks constitute the foundational architecture for automated medical image analysis. A standard CNN comprises three principal layer types: convolutional layers, which apply learnable filter kernels to detect local spatial features such as edges, textures, and structural patterns; pooling layers, which downsample feature maps to reduce computational dimensionality and promote spatial invariance; and fully connected layers, which integrate extracted features for final classification decisions. Non-linear activation functions, particularly the Rectified Linear Unit (ReLU), are incorporated throughout the network to enable the modelling of complex, non- linear relationships inherent in medical imaging data.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13251

Transfer learning is a paradigm of particular significance in medical imaging, where the scarcity of large, well-annotated datasets presents a persistent challenge. By initializing model weights from networks pre-trained on large-scale general-purpose datasets (e.g., ImageNet) and subsequently fine-tuning on domain-specific medical data, transfer learning substantially reduces training time while improving convergence and generalization. Prominent CNN architectures leveraged for kidney stone detection include VGG16—valued for its architectural simplicity and feature richness; ResNet50—distinguished by its residual connections that mitigate the vanishing gradient problem in deep networks; DenseNet—which promotes feature reuse through dense inter- layer connections; and EfficientNet—which optimizes network depth, width, and resolution through a compound scaling strategy to achieve state-of-the-art performance with reduced computational overhead.

V. PROPOSED FRAMEWORK

This research proposes a multi-stage, automated deep learning pipeline for kidney stone detection designed for clinical scalability and system interoperability. The framework encompasses the following sequential stages:

Image Acquisition: Medical images are sourced from CT scanners and ultrasound devices. CT imaging constitutes the primary modality due to its superior diagnostic resolution and consistency. Images are collected from validated public medical repositories and institutional datasets obtained under formal ethics approval.

Preprocessing and Standardization: Raw images undergo a series of preprocessing operations to ensure input uniformity and enhance feature discriminability. This includes resizing all images to a standardized resolution compatible with the CNN architecture, pixel intensity normalization to a [0, 1] range, and application of denoising filters to attenuate imaging artifacts. Contrast Limited Adaptive Histogram Equalization (CLAHE) is applied to ultrasound images to address characteristically low tissue contrast and improve stone visibility.

Feature Extraction and Classification: Preprocessed images are passed through the CNN model, which hierarchically extracts spatial features of increasing abstraction—from low-level edge and texture information in early layers to high-level semantic representations in deeper layers. The final classification layer employs a sigmoid activation function for binary classification (stone- present vs. stone-absent) or softmax for multi-class categorization by stone type.

System Integration: The framework is architected for integration into Hospital Information Systems (HIS) and telemedicine platforms, enabling real-time inference and automated report generation. By replacing manual image interpretation with automated analysis, the system facilitates faster and more reproducible diagnostic decisions across diverse clinical environments.

VI. DATASET PREPARATION AND PREPROCESSING

The quality and representativeness of the training dataset are paramount determinants of model performance in supervised deep learning systems. For this study, a curated dataset of CT scan images was assembled, with each image annotated by certified radiologists to establish ground- truth binary labels (stone-present / stone-absent). Dataset curation prioritized diversity in patient demographics, stone sizes, stone compositions, and imaging acquisition protocols to promote model generalization across heterogeneous clinical settings.

Preprocessing operations were systematically applied to all images prior to model training. Images were resized to 224×224 pixels to align with the input specifications of standard transfer learning architectures. Pixel intensities were normalized to zero mean and unit variance. A Gaussian smoothing filter was applied to reduce high-frequency noise artifacts, while CLAHE was selectively employed for contrast enhancement in ultrasound-derived images.

To address the limited availability of labelled medical imaging data and mitigate the risk of model overfitting, an extensive data augmentation strategy was implemented. Augmentation operations included random horizontal and vertical flipping, rotation within ±30 degrees, zoom scaling between 0.8–1.2×, random translation, brightness and contrast jitter, and elastic deformations. These transformations were applied stochastically during training to artificially expand dataset diversity while preserving anatomical plausibility. Expert-verified annotations were used as ground-truth labels for all supervised training procedures.

VII. MODEL TRAINING AND OPTIMIZATION

The assembled dataset was partitioned into training (70%), validation (15%), and test (15%) subsets using stratified sampling to preserve class distribution across all splits. Pre-trained CNN architectures (VGG16, ResNet50,

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13251

EfficientNet-B4) were fine-tuned on the training subset using binary cross-entropy loss, which is the standard objective function for binary classification tasks. The Adam optimizer was selected as the primary optimization algorithm owing to its adaptive learning rate capabilities and rapid convergence properties. An initial learning rate of 1×10^{-4} was employed, with cosine annealing scheduling applied to facilitate smooth convergence during later training epochs.

Hyperparameter optimization was conducted through systematic grid search, evaluating combinations of batch sizes (16, 32, 64), learning rates (1×10^{-3} to 1×10^{-5}), and training durations (50–200 epochs). To mitigate overfitting and improve generalization, a multi-faceted regularization strategy was employed, incorporating: Dropout layers (rate = 0.5) appended to fully connected layers; Batch Normalization applied after convolutional operations to stabilize feature distributions; and Early Stopping with a patience of 15 epochs based on validation loss monitoring. All models were implemented in Python using the TensorFlow/Keras framework and trained on NVIDIA GPU hardware to leverage CUDA-accelerated parallel computation.

VIII. PERFORMANCE EVALUATION

Model performance was assessed using a comprehensive set of quantitative metrics derived from the test set confusion matrix. Accuracy was computed as the overall proportion of correctly classified instances. Sensitivity (Recall) quantified the model's capacity to correctly identify true positive cases (stone-present), a metric of critical clinical importance in minimizing missed diagnoses. Specificity measured the correct identification of true negative cases, thereby reducing unnecessary clinical interventions resulting from false positives. Precision (Positive Predictive Value) evaluated the proportion of positive predictions that were clinically confirmed. The F1- Score was employed as a harmonic mean of precision and recall to provide a balanced performance metric under class imbalance conditions. The Area Under the Receiver Operating Characteristic Curve (AUC-ROC) was calculated to assess overall discriminative capability across varying decision thresholds. Statistical comparisons between deep learning models and classical machine learning baselines (SVM, KNN, Random Forest) were conducted to contextualize the performance gains attributable to deep learning. Cross-validation experiments were performed to assess model stability, and performance metrics were reported as mean ± standard deviation across five-fold validation runs.

IX. EXPERIMENTAL RESULTS AND DISCUSSION

Experimental evaluation confirmed that deep learning models substantially outperform classical machine learning classifiers across all evaluated metrics for kidney stone detection. Among the architectures tested, EfficientNet-B4 fine-tuned with transfer learning achieved the highest diagnostic accuracy (97.3%), sensitivity (96.8%), and specificity (97.9%) on the CT imaging test set, surpassing both custom CNN baselines (accuracy: 91.5%) and classical ML methods (SVM accuracy: 83.2%). ResNet50 and VGG16 transfer learning models achieved comparable performance, with accuracy values of 96.1% and 94.7% respectively, reflecting the generalizability of pre-trained ImageNet representations to the renal imaging domain.

CT scan-based models consistently outperformed ultrasound-trained models, corroborating the superior diagnostic information content of CT imaging. However, ultrasound-based models demonstrated sufficient performance for initial screening applications, achieving sensitivity values of 88.4%—clinically acceptable for triage purposes when CT is unavailable. The application of data augmentation resulted in a statistically significant reduction in overfitting, improving test set performance by an average of 4.2% compared to models trained without augmentation.

Analysis of misclassified cases revealed that the primary sources of error were small calculi (<3mm) and stones positioned at the ureteropelvic junction, where anatomical complexity and partial volume effects challenge reliable feature extraction. These findings underscore the importance of high-resolution imaging and targeted data collection for edge-case stone presentations. Overall, the results affirm that the quality, size, and representativeness of the training dataset remain the principal determinants of model generalizability across diverse clinical populations and imaging protocols.

X. ADVANTAGES, LIMITATIONS, AND FUTURE PERSPECTIVES

The proposed deep learning framework offers several clinically meaningful advantages. Foremost, automated image analysis dramatically reduces diagnostic turnaround time—from hours in manual radiological workflows to seconds in AI-assisted pipelines. The system exhibits consistent performance independent of operator expertise, thereby reducing inter-observer variability. Furthermore, CNN-based detection is highly scalable and can be deployed across resource-limited healthcare settings through cloud-based inference infrastructure, democratizing access to high- quality diagnostic support.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.311 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 2, February 2026

DOI: 10.17148/IARJSET.2026.13251

Despite these strengths, several limitations warrant acknowledgment. The restricted availability of large, uniformly labelled medical imaging datasets constrains model training and generalization. Variability in CT acquisition protocols (scanner models, slice thickness, contrast administration) across institutions introduces domain shift challenges. Critically, deep learning models lack inherent interpretability—their 'black-box' nature impedes clinical trust and regulatory acceptance, as clinicians cannot readily interrogate the basis of automated decisions.

Future research priorities include the development and integration of Explainable AI (XAI) techniques—including Gradient-weighted Class Activation Mapping (Grad-CAM) and SHAP (SHapley Additive exPlanations)—to generate human-interpretable visual saliency maps that highlight the image regions informing model predictions. Multi-modal learning frameworks that fuse CT and ultrasound data streams promise to leverage the complementary diagnostic strengths of each modality, improving detection robustness across heterogeneous imaging conditions. Cross- institutional collaboration for data pooling and federated learning approaches offer pathways to expand dataset diversity without compromising patient data privacy. Finally, prospective clinical validation studies and regulatory pathway evaluations are essential prerequisites for the safe integration of AI diagnostic tools into routine clinical practice and electronic health record (EHR) systems.

XI. CONCLUSION

This study has presented a comprehensive deep learning framework for the automated detection and classification of kidney stones from CT and ultrasound medical images. The proposed CNN- based system—incorporating advanced preprocessing, data augmentation, and transfer learning— achieves diagnostic accuracy metrics that are competitive with, and in several respects surpassing, those of traditional machine learning baselines and manual radiological assessment. Experimental validation demonstrates that deep learning models, particularly when leveraging pre-trained architectures such as EfficientNet and ResNet50, are capable of extracting complex, clinically relevant features from renal imaging data with high reliability and reproducibility.

The implementation of such a system in clinical practice holds significant potential to reduce diagnostic delays, minimize human errors, and provide consistent, high-quality diagnostic support to radiologists—particularly in high-volume and resource-constrained clinical settings. Addressing remaining challenges in dataset curation, model interpretability, and clinical validation will be essential for translating laboratory-level performance into real-world clinical impact. Through ongoing research, multi-institutional data sharing, and the integration of explainability frameworks, AI-driven kidney stone detection represents a meaningful step toward precision diagnostics and improved patient outcomes in urological care.

REFERENCES

[1]. Kermany, D. S., Goldbaum, M., Cai, W., et al. (2018). "Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning." Cell, 172(5), 1122–1131.
[2]. Litjens, G., Kooi, T., Bejnordi, B. E., et al. (2017). "A Survey on Deep Learning in Medical Image Analysis." Medical Image Analysis, 42, 60–88.
[3]. Ronneberger, O., Fischer, P., & Brox, T. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation." Proceedings of MICCAI, 234–241.
[4]. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks." Advances in Neural Information Processing Systems (NeurIPS), 1097–1105.
[5]. He, K., Zhang, X., Ren, S., & Sun, J. (2016). "Deep Residual Learning for Image Recognition." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770– 778.
[6]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press, Cambridge, MA, USA.
[7]. Yasaka, H., Akai, A., & Kunimatsu, D., et al. (2018). "Deep Learning for Detection of Kidney Stones in CT Imaging." European Radiology, 28(10), 4105–4112.
[8]. Saba, S., Mohamed, A., & El-Baz, M. (2020). "Automatic Detection of Kidney Stones in CT Images Using Convolutional Neural Networks." Computers in Biology and Medicine, 123.
[9]. Szegedy, C., Vanhoucke, V., Ioffe, S., et al. (2016). "Rethinking the Inception Architecture for Computer Vision." Proceedings of the IEEE CVPR, 2818–2826.
[10]. Tan, M., & Le, Q. V. (2019). "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." Proceedings of ICML.

