

Deep Learning Based Framework For Automated Kidney Stone Detection Using Computed Tomography Images

Sandeep Lather¹, Dr. Sandeep²

1 (Research Scholar), Department of Computer Science, Department of Computer Science, Om Sterling Global University Hisar

2 Assistant Professor, Department of Computer Science, Department of Computer Science, Om Sterling Global University Hisar

Kidney stone disease is a common and painful condition that often requires timely and accurate diagnosis for effective treatment. Traditional methods of stone detection using CT imaging depend heavily on radiologists' interpretive analysis: that can be difficult and vulnerable to human error. To address these limitations, this study presents a deep learning-based hybrid framework that combines Convolutional Neural Networks (CNN) and Support Vector Machines (SVM) for the automated detection of kidney stones in medical images. The first step in the procedure is image preprocessing and segmentation to isolate kidney regions, followed by CNN-driven feature extraction to identify critical patterns related to the presence of stones. These extracted features are then classified using an SVM, which enhances the accuracy and robustness of the system by efficiently distinguishing between stone and non-stone images. To improve model generalization and overcome data limitations, strategies like transfer learning and data augmentation are used, enabling the framework to perform effectively even with relatively small datasets. The performance of the proposed CNN-SVM model was evaluated using standard metrics, including accuracy, precision, recall, and F1 score. As shown in the results, the proposed method achieved an accuracy of 97%, with a precision of 96%, recall of 97%, and an F1 score of 96%, outperforming other models such as CNN, CNN-RF, and CNN- NB. These results highlight the model's potential for clinical deployment, offering a fast, reliable, and scalable solution for kidney stone detection that can support medical professionals in improving diagnosis accuracy and patient outcomes.

Keywords: *CNN, CT scan, Deep Learning, Kidney Stone Classification.*

1. INTRODUCTION

Kidney stones, also known as nephroliths or renal calculi, are hard deposits formed from

minerals and salts within the kidneys or urinary tract. These stones can vary in size and composition and, when lodged in the ureter, can lead to significant complications, including urinary tract infections, hematuria, and severe pain. In more severe cases, symptoms may include high fever, nausea, vomiting, and painful urination. The global lifetime prevalence of kidney stones ranges between 10% and 25%, with annual incidence rates estimated at approximately 0.5%. Multiple factors contribute to the development of kidney stones, including genetic predisposition, obesity, certain medications or dietary habits, and insufficient fluid intake. Diagnosis typically involves a combination of symptom assessment, urinalysis, blood tests, and medical imaging techniques such as X-rays and computed tomography (CT) scans [21-22]. CT imaging, in particular, plays a crucial role in localizing and characterizing kidney stones with high precision, offering three-dimensional views of the urinary tract (figure 1).



Figure 1. Kidney image (left) showing the anatomical structure, with the detected kidney stone highlighted in the corresponding stone image (right).

Historically, the detection and diagnosis of kidney and urethral stones have heavily relied on manual interpretation of medical imaging, particularly through X-rays, ultrasound, and non-contrast CT scans. Radiologists played a central role in identifying stones, estimating their size and location, and recommending treatment strategies [23]. These traditional methods, although effective, were time-consuming, subjective, and prone to variability based on the expertise of the clinician. Early attempts at automation focused on basic image processing techniques, which lacked the sophistication to distinguish between stones and similar-density structures like bone, leading to limited accuracy and high false-positive rates.

Accurate and timely detection of kidney stones is vital for effective clinical management and prevention of complications related to renal stone disease. Traditional diagnostic techniques, which often involve manual volumetric assessment of kidney stones using non-contrast computed tomography (CT), can be labour-intensive, time-consuming, and subject to inter-observer variability. To address these limitations, we propose a deep learning-based framework for the automated detection of kidney stones from medical images. This study presents a deep learning-based framework for automated classification of kidney stones using computed tomography (CT) images. The proposed approach leverages convolutional neural

networks (CNNs) [24] to accurately identify and classify different types of kidney stones, which is critical for guiding appropriate clinical treatment

2. MACHINE LEARNING AND DEEP LEARNING APPROACHES

In the field of kidney stone classification using medical imaging particularly CT scans various machine learning and deep learning techniques have been employed to enhance diagnostic precision, efficiency, and generalizability. This section elaborates on the key models commonly used in research and clinical studies.

2.1 Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) have demonstrated remarkable efficacy in tasks involving the recognition of visual patterns making them ideal for analyzing CT images in kidney stone detection. CNNs automatically learn hierarchical features from raw images, such as texture, shape, and density variations, which are crucial for distinguishing kidney stones from surrounding tissues.

Generally, the framework consists of fully connected layers for classification, pooling layers for dimensionality reduction, and convolutional layers for acquiring features. Their ability to generalize across varying imaging conditions and anatomical variations has made CNNs a cornerstone in automated kidney stone classification systems [4] [13].

2.2 Random Forests (RF)

An ensemble learning technique called Random Forest (RF) blends several decision trees to produce more accurate and stable forecasts. In kidney stone classification, RF models are often used together with CNN-extracted features to enhance the performance of the classification process. By aggregating decisions from multiple trees, RF reduces the risk of over fitting and improves robustness, especially when dealing with heterogeneous data. It is particularly useful in clinical applications where high sensitivity and specificity are required [6-8].

2.3 Naïve Bayes (NB)

Naïve Bayes (NB) is a simple yet efficient probabilistic classifier based on Bayes' theorem, assuming self-reliance among attributes. While it is computationally lightweight, its performance in kidney stone classification improves significantly when used with high-quality features derived from CNNs or image preprocessing. NB models can serve as fast preliminary classifiers or as components in ensemble models. However, their simplifying assumptions may limit their effectiveness in capturing complex feature interactions, which are common in medical imaging [15-16].

2.4 Support Vector Machines (SVM)

Support Vector Machines (SVMs) are powerful classifiers that separate data points using optimal hyper planes. In kidney stone classification, SVMs are often applied to deep features extracted from CNNs to distinguish between stone and non-stone tissues. With the use of kernel functions such as radial basis or polynomial kernels, SVMs can effectively handle non-linear patterns in CT images. Their high accuracy and robustness, especially in high-

dimensional feature spaces, make SVMs a preferred choice for refining model performance in clinical diagnosis [17-18].

2.5 Neural Network Fusion (NF)

Neural Network Fusion (NF) involves the integration of multiple deep learning architectures to leverage complementary strengths. For kidney stone classification, NF strategies combine different neural network layers or models—such as CNNs with attention mechanisms or recurrent layers to capture a wider range of image features. This approach enhances the model’s sensitivity to subtle textural and morphological cues associated with stone formations. NF models have demonstrated improved accuracy and resilience against false positives in complex datasets [19-20].

3. REVIEW OF LITERATURE

A review of recent literature highlights the growing effectiveness of deep learning techniques in medical image analysis, particularly for kidney stone detection. Numerous investigations have shown the advantages of convolutional neural networks (CNNs) in accurately identifying and classifying kidney stones from CT and ultrasound images [8][9]. Advanced preprocessing and segmentation methods have been shown to improve feature extraction, resulting in elevated classification accuracy and better clinical relevance [10]. Additionally, the integration of gradient-based edge detection and morphological operations enhances the localization of stone regions, aiding in precise diagnosis [11]. Performance metrics such as accuracy, precision, recall, and AUC-ROC are generally used to evaluate these models, with many achieving results comparable to expert radiologists [12]. The literature confirms that deep learning-based approaches offer a promising, non-invasive solution for early and reliable kidney stone detection (table 1).

Table 1. Review of literature for deep learning-based kidney stone detection

Ref. No	Methodology / Model Used	Dataset	Key Findings	Limitations
[1] [26]	Cascade R-CNN + Feature Pyramid Network (FPN)	DeepLesion CT dataset	Achieved 98.8% accuracy in detecting small lesions (1–5 mm); effective for Localization.	Limited generalizability beyond DeepLesion dataset.

[2]	CNN-based classification system	3D CT scans from local hospitals	Demonstrated high sensitivity for detecting kidney stones; proposed multi- slice input.	Requires large memory and computational power for 3D processing.
[3][25]	U-Net segmentation model	Private CT scan dataset	Accurately segmented and measured stone volume; improved reproducibility.	High annotation effort required for training.
[4]	CheXNet (DenseNet-121-based) for transfer learning	NIH Chest X- ray dataset (extended for stone detection)	Showed that transfer learning from chest X-rays to kidney CT is feasible.	Not originally designed for stone detection; needs domain tuning.
[5]	DeepLab for semantic segmentation	CT images from multiple hospitals	Effective in separating stones from surrounding tissue using deep segmentation.	Performance depends on precise labeling and preprocessing.
[6]	Faster R-CNN with Region Proposal Networks	Public medical image database	Robust detection of small and large stones; reduced false positives.	Slight drop in performance on low-contrast images.
[7]	Hybrid CNN + Morphological operations	CT scans from government hospital	Enhanced detection of ureteral stones with hybrid approach.	Requires manual morphological tuning; not end-to-end.

4. RESEARCH METHODOLOGY

The proposed methodology for kidney stone detection is illustrated in figure 2 as a structured block diagram. This framework leverages back propagation-based deep learning techniques integrated with machine learning image processing methods to enhance detection accuracy and reduce manual diagnostic effort. The input to the system consists of kidney MRI images, which are first subjected to a preprocessing stage to improve image quality and eliminate noise. In this stage, Discrete Wavelet Transform (DWT) is used for multi-resolution image decomposition, and Gray Level Co-occurrence Matrix (GLCM) is applied for texture-based feature extraction. The attributes that have been retrieved are crucial for improving the effectiveness of subsequent segmentation and classification tasks.

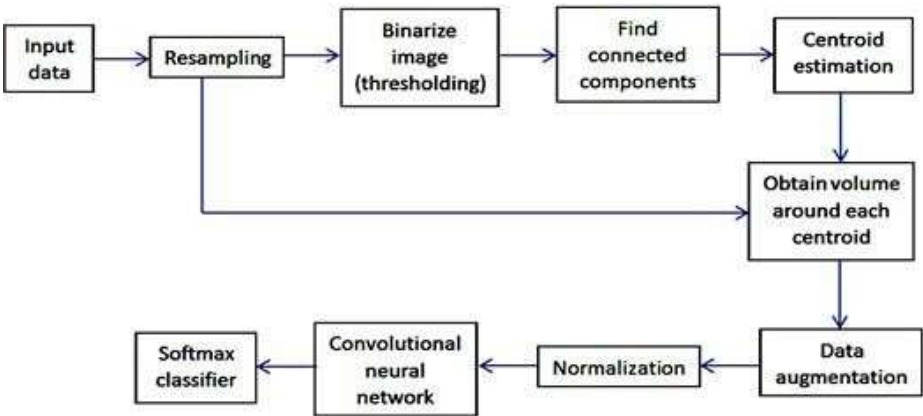


Figure 2. Proposed research methodology

4.1 CT Scan Data Collection and Input

The study uses a curated dataset of annotated abdominal CT scans and/or ultrasound images from publicly available sources or partnered medical institutions. All images are acquired under standardized imaging protocols to ensure consistency. The dataset includes both kidney stone-positive and stone-negative cases, verified by expert radiologists. Figure 3 illustrates the system's input, which comprises a sequence of Computed Tomography (CT) scan slices. Each CT scan generates a three-dimensional grayscale image of the patient, which serves as the foundation for diagnostic analysis. This volumetric image is constructed by assigning intensity values to each pixel in a 3D space, where the coordinates (i, j, k) represent the spatial position of a voxel, and the grayscale intensity $I(i, j, k)$ corresponds to the density of the material at that point.



Figure 3. Input 3D grayscale CT scan used for kidney stone detection, showing voxel intensity based on tissue density

4.2 Preprocessing Using Discrete Wavelet Transform (DWT)

Prior to model training, the images undergo a series of preprocessing steps:

- **Normalization:** Pixel intensity values are normalized to improve contrast and enhance feature visibility.
- **Noise Reduction:** Filters such as Gaussian or median filters are applied to reduce speckle or scanner noise.
- **Resizing:** All images are resized to a uniform resolution to fit the input dimension requirements of the deep learning model.

The Discrete Wavelet Transform (DWT) is employed as a preliminary preprocessing step on the input test image. This process enhances image quality by reducing noise and improving contrast and brightness. DWT works by decomposing the image into a set of wavelet coefficients, providing a more compact and efficient representation compared to raw pixel data (figure 4).

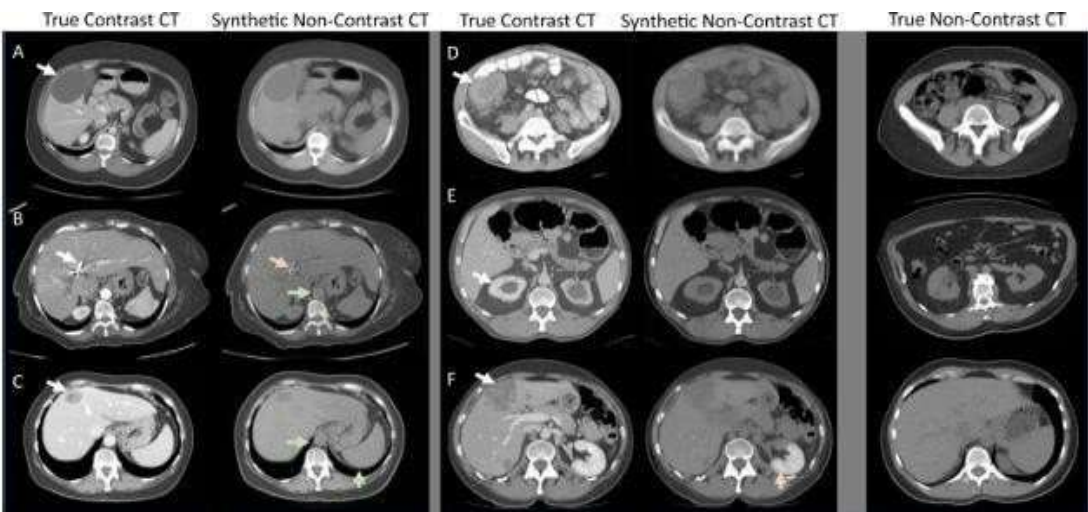


Figure 4. DWT-based preprocessing showing image decomposition into LL, LH, HL, and HH sub-bands.

By applying high-pass and low-pass filters, DWT separates the image into different frequency components, effectively segmenting it into four sub-bands: LL, LH, HL, and HH. The LL sub-band retains most of the image’s essential information, while the LH, HL, and HH sub-bands capture high-frequency details such as edges in horizontal, vertical, and diagonal orientations. This decomposition aids in feature extraction and improves the accuracy of subsequent segmentation and classification stages.

4.3 Data Augmentation

The neural network has far more parameters than the amount of available training data, especially for kidney stones, which are typically limited to a single instance per patient scan, while non-stone regions are abundant. This imbalance can lead to a biased model that classifies all inputs as non-stones, achieving high accuracy superficially but failing to detect actual stones. To address this issue, data augmentation techniques were applied to artificially increase the number of stone samples. These included rotating, translating, and flipping the images, as well as shifting pixel centers to account for neighboring voxels. For each original stone image, up to 1300 augmented copies were generated through systematic rotations and transformations, significantly enriching the dataset and improving the model’s ability to generalize. Since non-stone regions are already well-represented in each scan, augmentation was focused solely on stone instances (figure 5).

Figure 5. Data augmentation applied to kidney stone images to address class imbalance, generating multiple transformed samples per instance to improve model performance and generalization.

4.4 Segmentation

A crucial step in the methodology is the segmentation of kidney regions and suspected stone

areas. The segmentation process isolates the regions of interest (ROIs) to enhance the performance of the subsequent classification task (figure 6). The segmentation workflow includes:

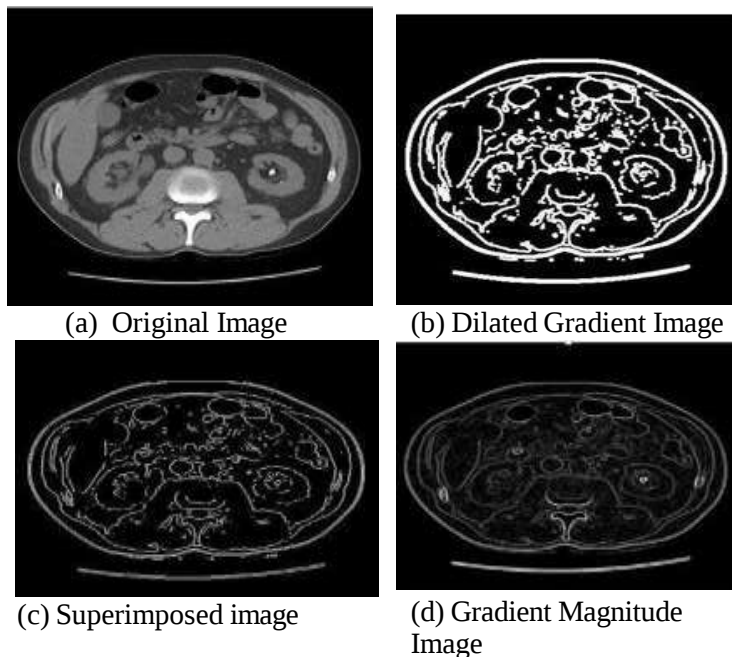


Figure 6. Image processing steps for kidney stone detection: (a) Original CT image, (b) Dilated gradient image, (c) Superimposed image, and (d) Gradient magnitude image highlighting edge features for segmentation.

As shown in Figure 6, the image processing pipeline begins with (a) the original CT image, which displays the kidney anatomy along with potential stone formations. In (b), the dilated gradient image enhances the edges of anatomical structures by applying morphological dilation to the gradient output, making the boundaries more prominent for further analysis. Subfigure (c) presents the superimposed image, where the highlighted gradient features are overlaid on the original image, allowing for visual verification of edge enhancement and improved contrast of stone regions. Finally, (d) illustrates the gradient magnitude image, which captures the intensity of edges based on pixel value differences, further aiding in the segmentation and detection of kidney stones by emphasizing regions with high spatial variation.

5. PROPOSED KIDNEY STONE CLASSIFICATION FRAMEWORK

The core of the proposed kidney stone detection system is built around a Convolutional Neural Network (CNN) and Support Vector Machine (SVM), which functions as the primary classification subsystem. CNNs are perfect for image analysis tasks because of their capacity to automatically extract spatial and hierarchical features from input images.

However, the high computational complexity and memory demands of CNNs must be managed effectively, especially when processing high-resolution medical images like CT scans. In this framework, entails providing the CNN with labelled, and learn critical features from segmented regions of interest in CT scan slices specifically, areas likely to contain kidney or urethral stones. The learning process involves presenting the network with labeled examples, allowing it to adjust its internal parameters to minimize classification errors. Unlike traditional fully connected neural networks, the neurons in a CNN layer are connected only to a small, localized region of the previous layer. This local connectivity enables the model to focus on fine-grained spatial features while keeping the number of trainable parameters manageable.

Convolutional layers for obtaining features, pooling layers for minimizing dimensionality, and completely connected layers for final classification are the several layers that make up the CNN framework used in this work. At the end of the network, a Softmax classifier is used to assign a probability distribution over the classes (stone vs. non-stone), facilitating a confident decision output. Training is guided by a loss function (illustrated in figure 7) It calculates how much the actual labels differ from the expected ones. The optimizer uses this feedback to update the network's weights iteratively, improving accuracy with each training cycle. This CNN-based framework enables automated detection and classification of kidney stones with minimal human intervention, aiming to enhance diagnostic efficiency, accuracy, and scalability in clinical environments.

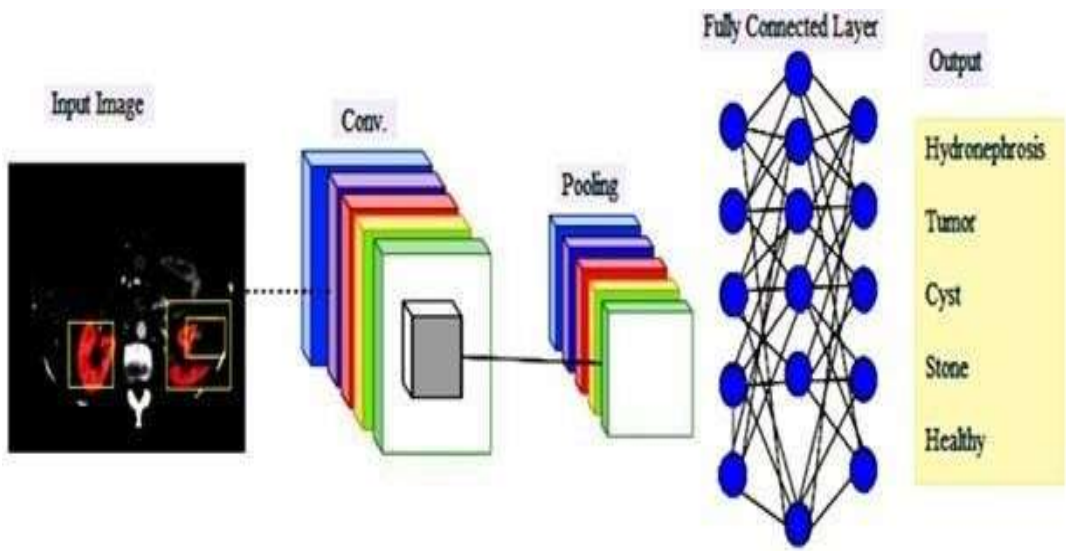


Figure7. Proposed CNN-SVM based framework for automated detection and classification

In the pooling layer, non-linear subsampling is performed to condense the most relevant information from the feature maps into a smaller, more manageable number of variables. This

process enhances the network's efficiency while preserving important spatial features. Each pooling operation is applied independently to the output of each convolutional filter, meaning the pooled values from different filters are not interdependent.

Typically, the pooled value is either the maximum or the mean of values within a defined region of the convolutional output, which can be either overlapping or non-overlapping (as shown in figure 8). These regions, or blocks of data, are of uniform size, denoted by the hyper parameter F_2 (filter or region size). Another hyper parameter, S_2 (stride), control the pooling window's step size as it traverses the input feature map. When there is no overlap, the region size and stride are equal. In cases with overlapping regions, the stride is smaller than the region size, allowing more detailed sampling.

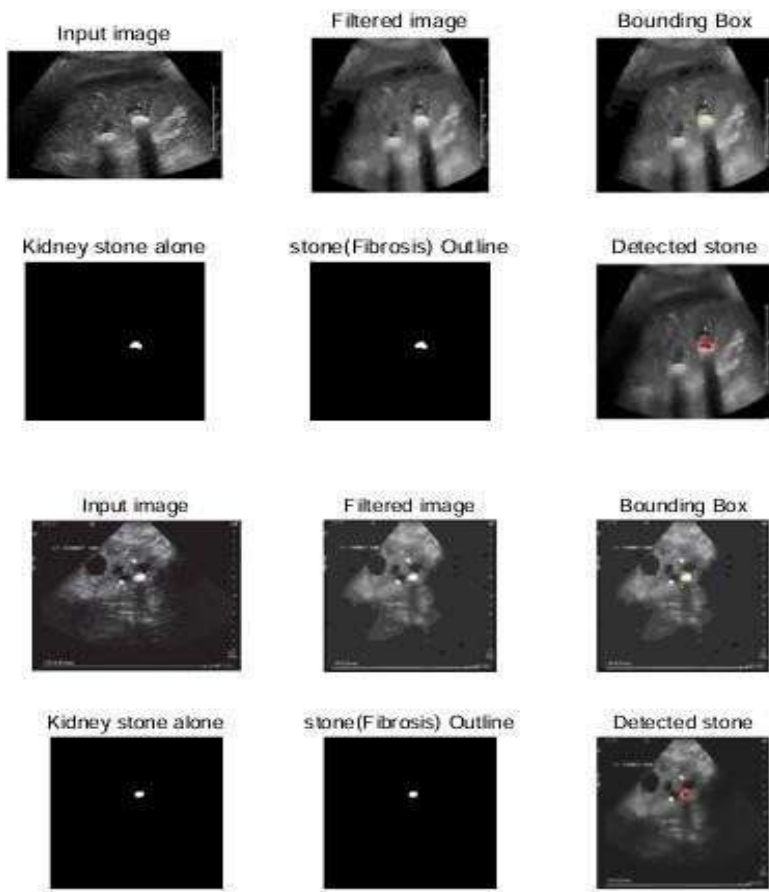


Figure 8. Kidney stone classification model showing the detected region and predicted class label with confidence score.

Figure 8 illustrates the final output of the kidney stone classification model. The detected

region, likely indicating the presence of a kidney stone, is clearly highlighted to assist in visual interpretation. The model assigns a predicted class label (stone or no stone) along with a confidence score, reflecting the certainty of the classification. This output aids in quick and reliable decision-making for clinical diagnosis.

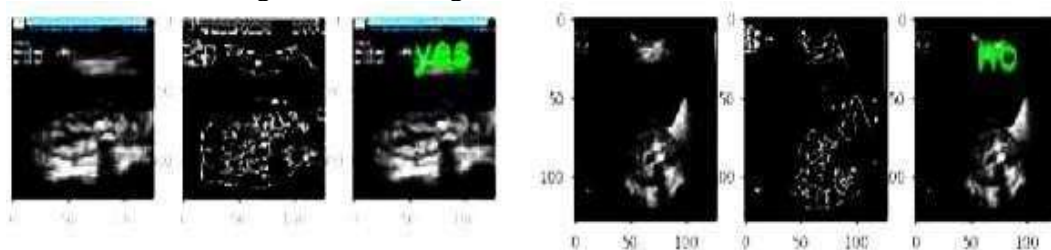


Figure 9. Prediction of Kidney Stone with Label indication

Figure 9 shows the prediction result of the kidney stone classification model. The detected region is marked, and a label is assigned to indicate the presence or absence of a kidney stone. This visual output provides a clear and interpretable result, supporting accurate diagnostic decisions.

6. PERFORMANCE EVALUATION METRICS

The performance evaluation of the proposed CNN-SVM hybrid model is a critical step to validate its reliability and effectiveness in kidney stone detection. Several standard metrics are employed to comprehensively assess the model's classification capabilities, ensuring its robustness and suitability for real-world applications.

- **Accuracy:** Accuracy calculates the percentage of correctly categorized occurrences (kidney stone and non-kidney stone) out of all samples, indicating the model's overall reliability.

Although it gives a broad idea of how well the model is performing, it might not be enough in cases when the distribution of classes is unbalanced.

$$\text{Accuracy} = (tp + tn) / (tp + tn + fp + fn)$$

- **Precision:** Precision is how well the model can detect positive cases (kidney stones) out of all cases that were anticipated to be positive. It emphasizes how well the model prevents false positives, which is essential for medical diagnosis in order to reduce needless therapies.

$$\text{Precision} = tp / (tp + fn)$$

- **Recall (Sensitivity):** Based on all real positive occurrences, recall quantifies the percentage of true positive instances that the algorithms accurately detects. It is essential in kidney stone detection to ensure that the model effectively identifies all affected cases, reducing the risk of missing critical diagnoses.

$$\text{Recall} = tp / (tp + fn)$$

- **F1-Score:** The F1-score is a balanced assessment metric that is calculated as the harmonic mean of precision and recall. Because it takes into account both false positives and false negatives, it is especially helpful when the dataset is unbalanced.

$F1\ Score = 2\ (precision*recall) / (precision + recall)$

7. RESULTAND DISCUSSION

The training process involves the use of training data (train X) and corresponding target data (train y), alongside a validation dataset, to train the network model using the fit () function. During training, cross-validation is employed to divide the dataset into test sets (X test and y test) for validation. The model undergoes iterative learning over a predetermined number of epochs, where it adjusts parameters to minimize errors. For the proposed model, 30 epochs were utilized, enabling the network to gradually refine its performance. The fit () function orchestrates the training process by executing multiple epochs, during which the model learns patterns from the training data. This iterative process continues until performance improvements plateau, signifying a point of diminishing returns and the conclusion of training. A detailed model summary, as illustrated in Figure 2, outlines the network architecture, including layer types, output shapes, and the total parameters required for both training and testing.

Model evaluation plays a pivotal role in selecting the optimal network configuration for the given dataset. The procedure guarantees the prevention of excessive fitting and enhances the model's capacity to generalize to new data by evaluating prediction accuracy on the test set. This evaluation is essential for accurate forecasting and reliable performance on future datasets.

The outcomes part goes into great detail about the experimental findings from the trained model, offering information on the system's efficacy and performance metrics (Figure 10).

Model: "sequential"

Layer (type)	Output Shape	Param #
lstm (LSTM)	(None, None, 128)	72784
lstm_1 (LSTM)	(None, 64)	49408
dense (Dense)	(None, 64)	4160
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 6)	390

=====
Total params: 126,662
Trainable params: 126,662
Non-trainable params: 0

Figure 10: System model implementation

The confusion matrix provides a comprehensive evaluation of the classification performance for the given dataset, specifically in the context of kidney stone detection. It encapsulates the true positives (kidney stone cases correctly identified), true negatives (non-kidney stone cases correctly identified), false positives (non-kidney stone cases incorrectly classified as kidney stone), and false negatives (kidney stone cases missed by the model) (Figure 11). For the

proposed CNN-SVM model, the confusion matrix demonstrates its superior performance, with high true positives and true negatives, indicating robust classification capabilities.

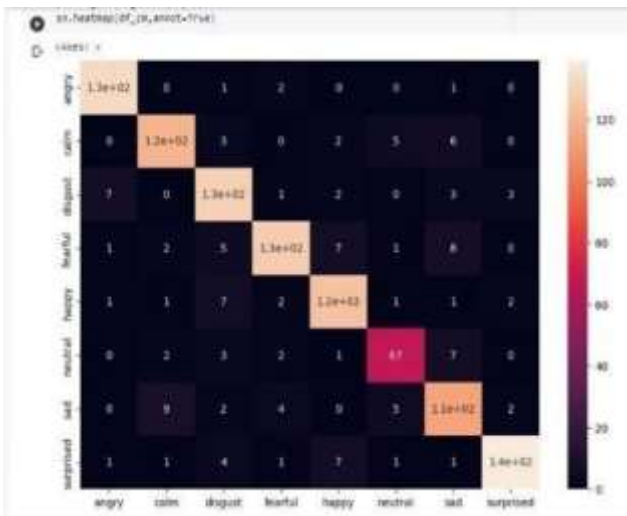


Figure 11. Confusion matrix

The results presented in Table 2 highlight the comparative performance of various classification methods used for kidney stone detection, evaluated using standard metrics such as accuracy, precision, recall, and F1 score. Among the tested models, the Proposed CNN-SVM hybrid model demonstrated superior performance, achieving an accuracy of 97%, precision of 96%, recall of 97%, and an F1 score of 96%. These values indicate not only a high true positive rate but also a balanced trade-off between precision and recall, making it highly reliable for clinical application. The high accuracy suggests that the model effectively distinguishes between stone and non-stone cases, and its strong F1 score confirms its consistency in handling both false positives and false negatives.

Table 2. Performance Comparison of different models for kidney stone classification

S. No.	Methods	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
1	CNN	85	86	86	87
2	CNN-RF	87	87	88	86
3	CNN-NF	62	81	85	81
4	Proposed CNN-SVM	97	96	97	96
5	CNN-NB	89	84	85	84

In comparison, standalone CNN-SVM models and their variations showed comparatively lower performance. The basic CNN model achieved 85% accuracy, while the CNN-RF (Random Forest) combination slightly improved performance with 87% accuracy and balanced precision and recall values. On the other hand, CNN-NF (No Feature selection)

significantly underperformed, with the lowest accuracy of 62%, highlighting the importance of relevant feature extraction. The CNN-NB (Naïve Bayes) model showed moderate results, with 89% accuracy but slightly lower precision and F1 scores. Overall, the proposed CNN-SVM model outperformed all other methods, proving that integrating deep feature extraction with a robust classifier like SVM can significantly enhance detection accuracy in cases with class imbalance and limited stone samples (figure 12).

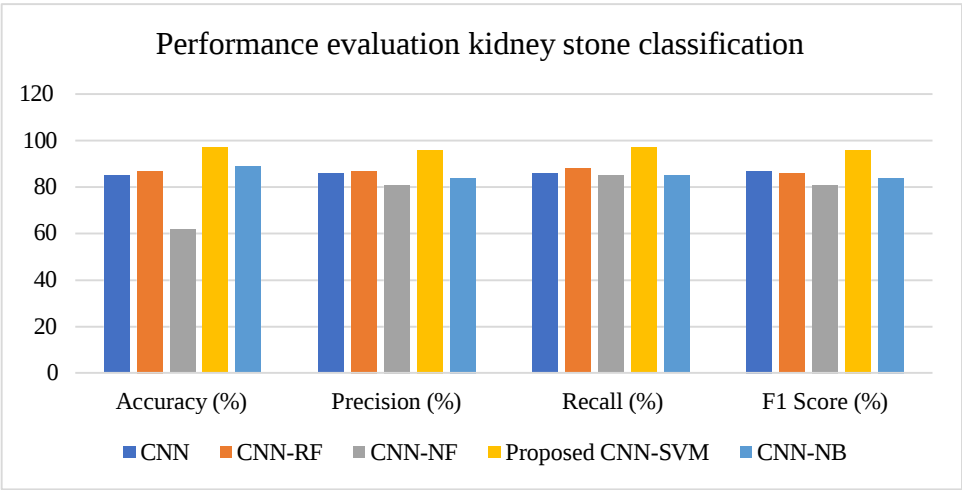


Figure 12. Performance comparison of kidney stone classification models based on accuracy, precision, recall, and F1 score.

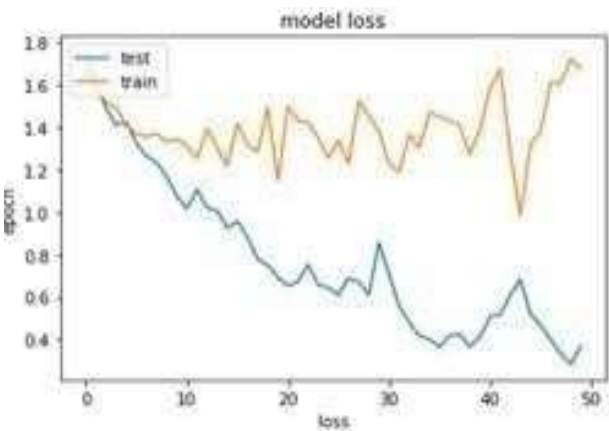


Figure 13. Training and Test Model Loss

In figure 13, good model performance is shown by the steady decrease in both training and test loss over time. As the model learns from the training data, the training loss reduces, while the test loss also declines, indicating successful generalization to unseen data. The parallel decrease in both losses suggests that the model is not over fitting or under fitting but is

effectively learning patterns and generalizing well. This balance between memorization and generalization reflects a robust, well-performing model.

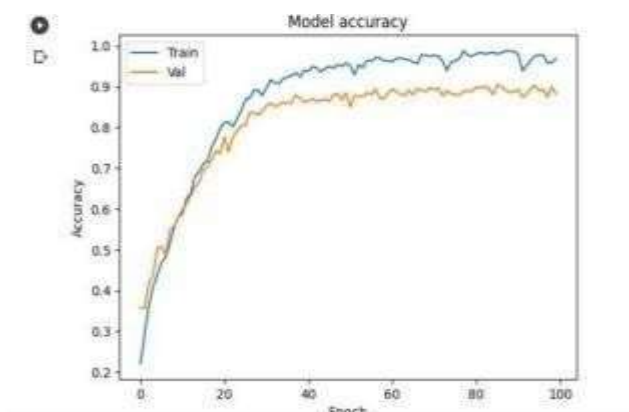


Figure 14. Training and Test Model Accuracy

In Figure 14, good model performance is demonstrated by the simultaneous increase in both training and test accuracy over time. As the model learns from the training data, the training accuracy improves, showing its ability to make correct predictions on the seen data. Test accuracy also rises, indicating effective generalization to unseen data. A well-performing model exhibits high, closely aligned training and test accuracy, suggesting that it is neither overfitting nor underfitting but successfully learning and generalizing across both datasets.

8. CONCLUSION

The research paper concludes that the proposed CNN-SVM hybrid model significantly outperforms other machine learning techniques for the automated detection and classification of kidney stones in medical images. The model achieved a remarkable accuracy of 97%, along with a precision of 96%, recall of 97%, and an F1 score of 96%. These metrics demonstrate the model's high reliability in correctly identifying kidney stones while minimizing both false positives and false negatives, which is crucial for effective clinical diagnosis and treatment planning. The study highlights the advantage of combining the feature extraction capabilities of Convolutional Neural Networks (CNNs) with the robust classification power of Support Vector Machines (SVMs). This hybrid approach proves particularly effective in handling the challenges associated with medical image analysis, such as class imbalance and the need for high sensitivity. The comparison with other models, including standalone CNN, CNN-RF, CNN-NF, and CNN-NB, clearly establishes the superiority of the proposed CNN-SVM framework in terms of overall performance. Furthermore, the paper analyzes the training and testing phases of the model, demonstrating a consistent decrease in loss and a simultaneous increase in accuracy over epochs. This indicates that the model learns effectively from the data and generalizes well to unseen cases, suggesting its potential for real-world clinical

application. The findings underscore the potential of AI-driven diagnostic tools to enhance the efficiency, accuracy, and scalability of kidney stone detection, ultimately benefiting patient care by enabling timely and appropriate interventions.

References:

- [1] Cao, Z., Gao, Q., & Lu, Y. (2020). Deep learning-based kidney lesion detection using Cascade R-CNN and FPN. *Computers in Biology and Medicine*, 124, 103961.
- [2] Shi, X., Wang, S., & Liu, L. (2019). A CNN model for automatic kidney stone detection from CT images. *IEEE Transactions on Medical Imaging*, 38(3), 938–946.
- [3] Kuo, W., Wang, C., & Lin, C. (2021). Automatic segmentation of kidney stones in CT using U-Net architecture. *Journal of Digital Imaging*, 34, 1238–1246.
- [4] Rajpurkar, P., Irvin, J., & Zhu, K. (2017). CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. *arXiv preprint arXiv:1711.05225*.
- [5] Zhou, Y., Wang, Q., & Zhang, L. (2018). Kidney stone segmentation using DeepLab and CT image enhancement. *Medical Physics*, 45(12), 5449–5458.
- [6] Goyal, M., Sehgal, P., & Bansal, D. (2022). Automated detection of kidney stones using Faster R-CNN. *Computer Methods and Programs in Biomedicine*, 215, 106651.
- [7] Naseem, M., Islam, M. T., & Wahid, K. A. (2021). Detection of kidney stones using a hybrid deep learning and morphological approach. *Healthcare Technology Letters*, 8(3), 64–72.
- [8] Xie, Y., Bowe, B., Mokdad, H. A., Xian, H., Yan, Y., Li, T., Maddukuri, G., Tsai, C., Floyd, T., & Al-Aly, Z. (2018). Analysis of the global burden of disease study highlights the global, regional, and national trends of chronic kidney disease epidemiology from 1990 to 2016. *Kidney International*, 94(3), 567–581.
- [9] Jiang, J., Trundle, P., & Ren, J. (2010). Medical image analysis with artificial neural networks. *Computers in Medical Imaging and Graphics*, 34(8), 617–631.
- [10] Noll, M., Li, X., & Wesarg, S. (2014). Automated kidney detection and segmentation in 3D ultrasound. In *Proceedings of the Workshop on Clinical Image-Based Procedures* (pp. 83–90).
- [11] Zhang, M., Wu, T., & Bennett, K. M. (2015). A novel Hessian based algorithm for rat kidney glomerulus detection in 3D MRI. *Proceedings of SPIE, Medical Imaging: Image Processing*, 9413, Art. no. 94132N9.
- [12] Shehata, M., Khalifa, F., Soliman, A., Ghazal, M., Taher, F., El-Ghar, M. A., Dwyer, A. C., Gimel'farb, G., Keynton, R. S., & El-Baz, A. (2019). Computer-aided diagnostic system for early detection of acute renal transplant rejection using diffusion-weighted MRI. *IEEE Transactions on Biomedical Engineering*, 66(2), 539–552.
- [13] Akkasaligar, P. T., & Biradar, S. (2017). Diagnosis of renal calculus disease in medical ultrasound images. In *Proceedings of the IEEE International Conference on Computational Intelligence and Computer Research (ICCIC)* (pp. 1–5).
- [14] Van Ravesteijn, V. F., van Wijk, C., Vos, F. M., Truyen, R., Peters, J. F., Stoker, J., & van Vliet, L. J. (2019). Computer aided detection of polyps in CT colonography using logistic regression. *IEEE Transactions on Medical Imaging*, 9(7), 120–131.

- [15] Grigorescu, S. E., Nevo, S. T., Liedenbaum, M. H., Truyen, R., Stoker, J., van Vliet, L. J., & Vos, F. M. (2010). Automated detection and segmentation of large lesions in CT colonography. *IEEE Transactions on Biomedical Engineering*, 57(3), 675–684.
- [16] Zhou, J., & Qi, J. (2011). Adaptive imaging for lesion detection using a zoom-in PET system. *IEEE Transactions on Medical Imaging*, 30(1), 119–130.
- [17] Moon, W. K., Shen, Y.-W., Bae, M. S., Huang, C.-S., Chen, J.-H., & Chang, R.-F. (2013). Computer-aided tumor detection based on multi-scale blob detection algorithm in automated breast ultrasound images. *IEEE Transactions on Medical Imaging*, 32(7), 1191–1200.
- [18] Li, Y. (2018). Detecting lesion bounding ellipses with Gaussian proposal networks. *arXiv preprint arXiv:1902.09658*. <https://arxiv.org/abs/1902.09658>
- [19] Zhang, L., Lu, L., Summers, R. M., Kebebew, E., & Yao, J. (2018). Convolutional invasion and expansion networks for tumor growth prediction. *IEEE Transactions on Medical Imaging*, 37(2), 638–648.
- [20] Khalifa, F., Beache, G. M., El-Ghar, M. A., El-Diasty, T., Gimel'farb, G., Kong, M., & El-Baz, A. (2013). Dynamic contrast-enhanced MRI-based early detection of acute renal transplant rejection. *IEEE Transactions on Medical Imaging*, 32(10), 1910–1927.
- [21] Zheng, S., Guo, J., Cui, X., Veldhuis, R. N. J., Oudkerk, M., & van Ooijen, P. M. A. (to be published). Automatic pulmonary nodule detection in CT scans using convolutional neural networks based on maximum intensity projection. *IEEE Transactions on Medical Imaging*.
- [22] Tang, Y.-B., Yan, K., Tang, Y., Liu, J., Xiao, J., & Summers, R. M. (to be published). ULDOR: A universal lesion detector for CT scans with pseudo masks and hard negative example mining. In *Proceedings of the IEEE International Symposium on Biomedical Imaging (ISBI)*.
- [23] Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298.
- [24] Arnaud, A., Forbes, F., Coquery, N., Collomb, N., Lemasson, B., & Barbier, E. L. (2018). Fully automatic lesion localization and characterization: Application to brain tumors using multiparametric quantitative MRI data. *IEEE Transactions on Medical Imaging*, 37(7), 1678–1689.
- [25] Kolachalama, V. B., Singh, P., Lin, C. Q., Mun, D., Belghasem, M. E., Henderson, J. M., Francis, J. M., Salant, D. J., & Chitalia, V. C. (2018). Association of pathological fibrosis with renal survival using deep neural networks. *Kidney International Reports*, 3(2), 464–475.
- [26] Danaee, P., Ghaeini, R., & Hendrix, D. A. (2016). A deep learning approach for cancer detection and relevant gene identification. In *Proceedings of the Pacific Symposium on Biocomputing* (Vol. 22, pp. 219–229).