



Paper Type: Original Article

Kidney Disease Classification Using Local Statistical Features and Recurrent Neural Networks

Seyed Hosein Razavi*

Department of Electrical Engineering, Ayandegan University, Tonekabon, Iran; hrazavi994@gmail.com.

Citation:

Received: 28 June 2025

Revised: 23 October 2025

Accepted: 03 January 2026

Razavi, S. H. (2026). Kidney disease classification using local statistical features and recurrent neural networks. *Trends in Health Informatics*, 3(1), 36-50.

Abstract

Accurate and rapid diagnosis of kidney cancer from CT images plays a crucial role in timely clinical decision-making and reducing patient risk. In this study, a novel framework based on image preprocessing, edge extraction, local statistical feature analysis, and Deep Learning (DL) using a Recurrent Neural Network (RNN) is proposed for the detection and classification of kidney diseases. The proposed approach eliminates irrelevant image regions and extracts discriminative features through histogram analysis within multiple local windows and by investigating the relationships between image edges and different anatomical regions. These features effectively characterize the structural and textural properties of kidney tissues.

To enhance class separability, adaptive histogram intervals are employed. Specifically, finer histogram bins are assigned to pixels with lower intensity values, while wider intervals are utilized for brighter pixels. This adaptive strategy enables the extraction of more informative statistical descriptors and improves the discrimination capability among different disease categories. Experimental results obtained on the standard Kaggle kidney CT image dataset demonstrate that the proposed method achieves a classification accuracy of 99.76% across four classes, namely normal, cyst, stone, and tumor. The obtained results indicate that the proposed framework outperforms previously reported methods and provides a highly effective solution for automated kidney disease classification.

Keywords: Image processing, Kidney disease, Histogram, Deep learning, Recurrent neural network, Classification.

1 | Introduction

Image processing has emerged as one of the most rapidly evolving scientific fields in recent decades, with its applications expanding continuously across a wide range of industries. Today, there are very few industrial sectors that do not benefit from image processing techniques. Among these, the medical industry has been one of the most significant beneficiaries of advances in image analysis and computer-aided diagnosis systems.

✉ Corresponding Author: hrazavi994@gmail.com

doi: <https://doi.org/10.22105/thi.v3i1.47>



Licensee System Analytics. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0>).

Medical imaging devices represent an important branch of healthcare technology, integrating sophisticated hardware and software components to assist clinicians in disease diagnosis and treatment planning. As technological advancements continue to reduce dependence on manual human intervention, medical imaging systems are becoming increasingly intelligent and efficient [1].

Computed Tomography (CT) imaging has become an indispensable diagnostic tool for the detection and monitoring of abnormalities in various organs, including the heart, liver, kidneys, intestines, and vascular structures. The relatively low operational cost, absence of proven adverse effects for both patients and clinicians, high imaging speed, and the capability to capture moving anatomical structures such as cardiac walls and valves have significantly increased the adoption of CT imaging systems in hospitals and clinical centers [1–3].

Kidney disease and renal failure represent major global public health challenges. The shortage of nephrologists and radiologists further exacerbates the burden associated with Chronic Kidney Disease (CKD). It is estimated that CKD affects more than 10% of the world's population and was ranked as the sixteenth leading cause of death in 2016. Moreover, it is projected to become the fifth leading cause of mortality by 2040. Consequently, early diagnosis of kidney-related disorders is crucial for preventing disease progression and reducing the risk of renal failure [4].

Motivated by this clinical necessity, automated kidney disease diagnosis based on medical image analysis has attracted considerable research attention. Noise reduction, image enhancement, tissue segmentation, and the development of highly accurate disease classification algorithms constitute the primary objectives of current research efforts. Although CT images generally provide superior image quality compared with ultrasound and Magnetic Resonance Imaging (MRI), image quality remains a critical factor affecting the performance of medical image processing techniques, particularly tissue segmentation algorithms. Tissue segmentation is widely recognized as the first and most important step in medical image analysis and computer-aided diagnosis. By accurately identifying tissue location, shape, and size, segmentation methods provide valuable information for detecting and characterizing pathological abnormalities [5].

However, the presence of inherent noise in CT images introduces significant challenges to automated kidney disease diagnosis. Appropriate preprocessing and enhancement techniques can substantially improve image quality, thereby facilitating more reliable disease detection when combined with efficient classification algorithms.

Numerous approaches have been proposed for image enhancement and noise suppression in CT imaging. Representative methods include weighted median filtering [6], the Savitzky–Golay filter [7], and wavelet-based denoising techniques [8]. In wavelet-based approaches, the original image is first decomposed into wavelet coefficients, which are subsequently modified according to a predefined strategy. The enhanced image is then reconstructed through the inverse wavelet transform. Several coefficient modification strategies have been proposed, including Bayesian estimation methods that exploit the statistical properties of wavelet coefficients. Nevertheless, thresholding remains the most widely adopted approach. Threshold-based denoising methods face two major challenges: 1) selecting an appropriate threshold value, which constitutes the most critical design parameter, and 2) addressing the varying statistical distributions of signal and noise across different wavelet scales. Various solutions have been proposed to address these challenges. While some studies apply identical processing strategies across all scales, others employ scale-dependent coefficient modification schemes [9], [10].

Following noise reduction, tissue segmentation constitutes another essential stage in kidney image analysis. Commonly used segmentation techniques include region-growing algorithms [1] and active contour models [2]. Due to the presence of image noise, region-growing methods often suffer from reduced reliability and segmentation accuracy [3]. Similarly, active contour models are not always suitable for kidney segmentation because kidney boundaries are frequently indistinct and exhibit weak contrast with surrounding tissues. Active Shape Models (ASMs) require manually selected landmark points to guide the segmentation process, which

is both time-consuming and operator-dependent. Furthermore, fully automated landmark initialization techniques remain insufficiently mature [3].

A segmentation approach based on shape features and region-based energy functions was introduced in [11] and demonstrated satisfactory performance for MRI tissue segmentation. However, owing to the high level of intrinsic image noise, this method is less effective when applied to kidney CT images. Another group of researchers proposed a shape-based kidney modeling framework combined with a Gabor filter bank for texture feature extraction and segmentation [5]. In this approach, an average kidney shape model is first aligned with the test image. Subsequently, Gabor filters are applied to regions inside and outside the evolving boundary to extract texture features. An energy function is then formulated to maximize the similarity between the texture characteristics inside the boundary and a predefined kidney texture model while minimizing similarity outside the boundary. If the stopping criterion is not satisfied, gradient vectors and energy-transpose matrices are computed to update the boundary location. This iterative optimization process continues until the similarity measure reaches its maximum value. Although effective, the computational complexity of repeatedly applying Gabor filters and extracting texture features results in a relatively high processing time for each image slice [5].

Although semi-automatic segmentation methods have demonstrated promising performance in clinical kidney image analysis, fully automatic kidney segmentation remains challenging because of the considerable variability in kidney morphology and image intensity distributions [12]. Consequently, Deep Learning (DL) techniques have increasingly been adopted for automatic tissue segmentation and medical image analysis tasks [12–14].

In this study, a novel kidney disease classification framework is proposed based on image preprocessing, edge extraction, local statistical feature analysis, and DL using a Recurrent Neural Network (RNN). The proposed approach offers several advantages, including the elimination of complex handcrafted feature extraction procedures, high computational efficiency during the preprocessing stage, and excellent classification performance. The remainder of this paper is organized as follows. Section 2 presents the fundamentals of RNNs, with particular emphasis on Long Short-Term Memory (LSTM) architectures. Section 3 describes the proposed methodology in detail. Experimental results and performance evaluation are presented in Section 4. Finally, conclusions and future research directions are discussed in Section 5.

2 | Recurrent Neural Networks

RNNs differ significantly from conventional Neural Network (NN) architectures in both their structural design and training mechanisms. Unlike feedforward NNs, RNNs incorporate recurrent connections that enable information from previous inputs to be fed back into the network during training. As a result, these networks possess an inherent short-term memory capability, allowing them to retain information from preceding inputs and utilize this information when processing subsequent data. This characteristic makes RNNs particularly suitable for modeling sequential and temporally dependent data.

Among DL architectures, RNNs are especially effective for processing time-series data and image sequences in which each sample may exhibit dependencies on preceding samples. However, standard RNN architectures suffer from several limitations. Due to their relatively simple recurrent structure, constrained hidden-state dynamics, and gradient-based weight update mechanisms, conventional RNNs often fail to effectively learn long-term dependencies when patterns span extended temporal intervals. As the temporal distance between relevant observations increases, the influence of earlier information gradually diminishes during training, ultimately approaching zero. Consequently, the network becomes unable to capture and retain long-range contextual information. This limitation, commonly known as the long-term dependency problem, represents one of the major weaknesses of traditional RNNs.

Various approaches have been proposed to address this issue, including alternative learning strategies and modifications to the recurrent architecture. Among these solutions, improved RNN structures have proven

to be the most effective in simultaneously modeling both long-term dependencies and short-term variations. Owing to their ability to process sequential information efficiently, these architectures have become widely adopted in applications involving image analysis, video processing, speech recognition, and other tasks where the current output depends on both present and historical information.

One of the most successful extensions of conventional RNNs is the LSTM network. LSTM was specifically developed to overcome the long-term dependency limitations of standard recurrent networks by introducing an internal memory mechanism capable of preserving information over extended sequences. In a conventional RNN, the contribution of earlier observations gradually decreases as new inputs are processed, eventually becoming negligible. LSTM networks mitigate this problem through a gated memory structure that selectively stores, updates, and retrieves information. The architecture of an LSTM network is illustrated in Fig. 1.

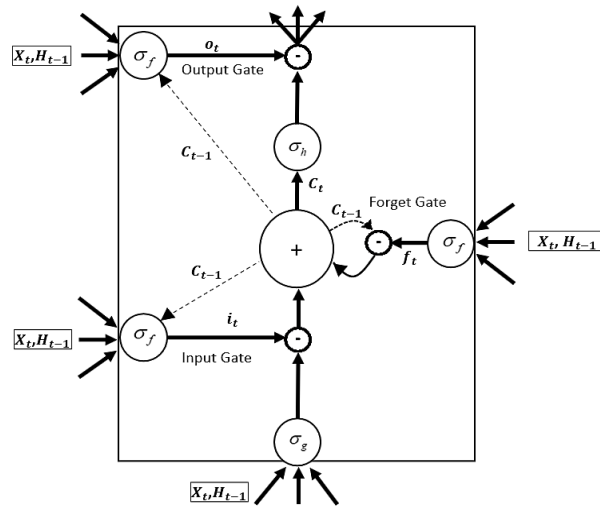


Fig. 1. Architecture of the long short-term memory network.

An LSTM cell consists of several functional components that regulate the flow of information through the network. In Fig. 1, σ denotes the activation function, while σ_g and σ_h represent the input and output activation functions, respectively. These functions are commonly implemented using the hyperbolic tangent (tanh) or logistic sigmoid functions, although the identity function may also be employed in certain configurations.

The dashed connections shown in Fig. 1 correspond to weighted peephole connections. In peephole-based LSTM architectures, the input gate, output gate, and forget gate are influenced not only by the current input and previous hidden states but also by the previous internal cell states. This additional information increases the degrees of freedom of the model and enhances its capability to learn complex temporal relationships.

Among the gating mechanisms, the forget gate plays a particularly important role. It determines which portions of the previous cell state should be retained and which should be discarded during the update process. By selectively preserving relevant historical information, the forget gate enables the network to capture long-term dependencies and maintain meaningful contextual information over extended sequences. Consequently, LSTM networks exhibit superior performance compared with conventional RNNs in applications involving long-range temporal correlations.

3 | Proposed Method

The overall framework of the proposed method is illustrated in Fig. 2.

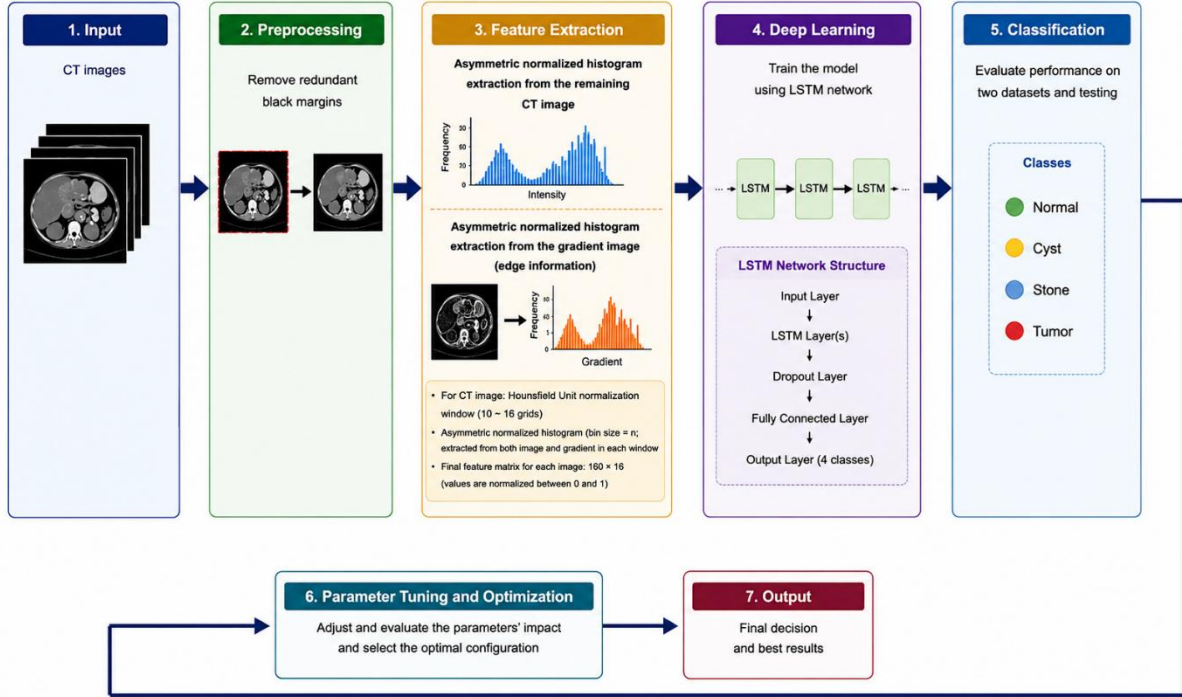


Fig. 2. Block diagram of the proposed kidney disease classification framework.

As shown in Fig. 2, the first step involves removing the black margins surrounding the CT image. This cropping operation eliminates non-informative regions and reduces the image dimensions, thereby decreasing computational complexity and improving feature extraction efficiency.

Following the removal of the surrounding black borders, an asymmetric normalized histogram is computed from the remaining image region. The use of asymmetric histogram intervals is motivated by the non-uniform distribution of pixel intensities within kidney CT images. In particular, darker intensity regions generally exhibit a higher pixel density than brighter regions. Therefore, symmetric histogram partitioning may fail to provide sufficient discrimination between tissue characteristics. To address this issue, asymmetric intensity intervals are employed to achieve a more informative statistical representation of image content.

In addition to the intensity histogram, gradient-based information is extracted from the image. The image gradient highlights edge structures corresponding to anatomical boundaries and pathological abnormalities within the kidney region. Since most renal abnormalities manifest themselves through structural variations and edge-related characteristics, the statistical distribution of edge pixels can provide valuable information regarding the presence and type of disease. Consequently, gradient-based descriptors are incorporated into the feature extraction process to enhance discriminative capability.

To further capture local spatial information, each CT image is partitioned into one hundred non-overlapping regions of equal size, arranged in a 10×10 grid. Subsequently, an asymmetric normalized histogram is extracted from both the original image and its corresponding gradient image for each local region.

The histogram descriptor consists of eight bins. Therefore, for the original image, a feature matrix of size 100×8 is obtained. An additional feature matrix of identical dimensions is extracted from the gradient image. These two feature representations are then concatenated, resulting in a final feature matrix of size 100×16 for each CT image.

All feature values are normalized to the range $[0,1]$, producing a compact and robust representation of both intensity and edge information. The resulting feature matrix serves as the input to the DL classifier.

The classification stage is performed using an LSTM recurrent neural network. The LSTM architecture is selected because of its ability to effectively model sequential dependencies within the extracted feature vectors

while simultaneously preserving both local and long-range contextual information. The configuration of the employed LSTM network is summarized in *Table 1*.

Table 1. Architecture and parameter configuration of the proposed LSTM network.

Layer No.	Layer Name	Specifications
1	Sequence input	Sequence input with 1600 dimensions
2	LSTM	LSTM with 200 hidden units
3	Fully connected	4 fully connected layers
4	Softmax	softmax
5	Classification output	Cross-entropy

The Softmax layer maps the input vector into a set of normalized values within the range of 0 to 1. This activation function provides a smooth and differentiable approximation of the maximum function, making it particularly suitable for gradient-based optimization algorithms. The outputs of the Softmax function represent a probability distribution over the target classes, such that the sum of all output probabilities is equal to one. Since the Softmax layer does not contain trainable parameters, its primary role is to transform the outputs of the preceding layer into normalized class probabilities within the interval [0,1]. The Softmax function is mathematically defined as follows:

$$y_i = \frac{e^{x_i}}{\sum_{i=1}^N e^{x_i}}, \quad (1)$$

where (x_i) denotes the (i) -th element of the input vector (i.e., the output of the fully connected layer), and (y_i) represents the corresponding Softmax output associated with that input.

The Cross-Entropy loss function is widely employed in classification problems where the network output represents a probabilistic value within the range of 0 to 1. This loss function quantifies the discrepancy between the predicted probability distribution and the ground-truth labels. As the predicted probability deviates from the actual class label, the Cross-Entropy value increases rapidly. Consequently, minimizing this loss function effectively drives the predicted probabilities toward the true target values, leading to improved classification performance.

When training a multi-class classifier whose final layer generates probabilistic outputs through the Softmax activation function, the Categorical Cross-Entropy (CCE) loss function is generally preferred due to its favorable optimization properties and probabilistic interpretation. CCE measures the divergence between the predicted probability distribution and the corresponding ground-truth labels, thereby encouraging the network to assign higher probabilities to the correct classes while suppressing incorrect predictions. As the predicted distribution deviates from the target distribution, the loss value increases rapidly, making it an effective objective function for multi-class classification problems [13], [14].

For a dataset containing (N) training samples and (C) target classes, the CCE loss is defined as:

$$L_{CCE} = \frac{-1}{N} \sum_{i=1}^N \sum_{j=1}^C y_{ij} \times \log \hat{y}_{ij}, \quad (2)$$

where:

N : Number of training samples.

C : Number of classes.

y_{ij} : Ground-truth label.

$\hat{y}_{ij}$: Predicted probability for class (j) of sample (i) .

To optimize the network parameters, the Adaptive Moment Estimation (Adam) optimizer was employed. Adam combines the advantages of momentum-based optimization and adaptive learning-rate methods by maintaining exponentially weighted estimates of both the first and second moments of the gradients [15]. The parameter update equations are given by

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t, \quad (3)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2, \quad (4)$$

$$\hat{m}_t = m_t / (1 - \beta_1^t), \quad \hat{v}_t = v_t / (1 - \beta_2^t), \quad (5)$$

$$\theta_{t+1} = \theta_t - \alpha \times [\hat{m}_t / (\sqrt{\hat{v}_t} + \epsilon)], \quad (6)$$

where g_t denotes the gradient at iteration t , m_t and v_t are the first and second moment estimates, respectively, α is the learning rate, and ϵ is a small constant introduced to ensure numerical stability.

The combination of the Adam optimizer and CCE loss provides stable convergence, efficient parameter updates, and improved generalization performance. Consequently, this optimization strategy enables the proposed LSTM-based framework to achieve highly accurate classification results while maintaining robust training behavior across all kidney disease categories.

4 | Experimental Results

4.1 | Dataset Description

The final dataset consists of 12,446 CT images categorized into four classes: Normal (5,077 images), Cyst (3,709 images), Stone (1,377 images), and Tumor (2,283 images). The dataset has been publicly released through the Kaggle platform, enabling researchers to implement, evaluate, and compare their algorithms using a common benchmark dataset.

Fig. 3 presents representative samples from the dataset. The red annotations indicate the lesion regions identified by radiologists and utilized as diagnostic references for distinguishing among different kidney disease categories.

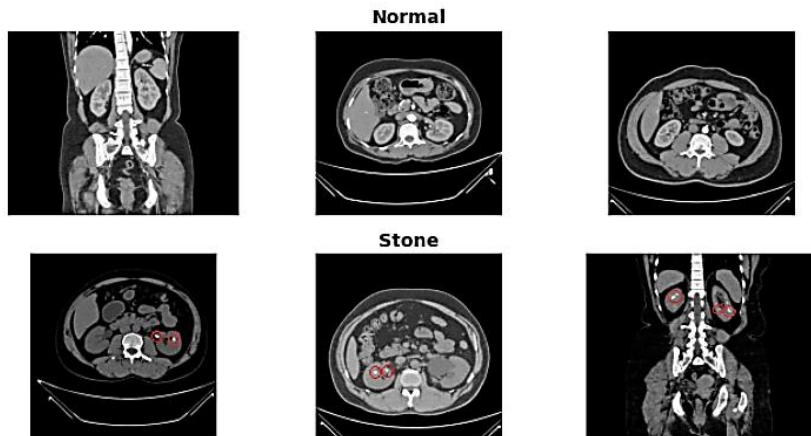


Fig. 3. Representative CT images from the Kaggle kidney disease dataset, including normal, cyst, stone, and tumor cases. Red annotations indicate the lesion regions identified by expert radiologists.

Fig. 4 illustrates the distribution of the mean image intensity values for the entire dataset as well as the corresponding distributions for each disease category. Analysis of these distributions reveals that the overall intensity distribution closely follows the distributions observed within the individual classes (Normal, Stone, Cyst, and Tumor). Furthermore, it can be observed that images belonging to the Stone and Cyst classes generally exhibit lower mean intensity values compared with the other categories.

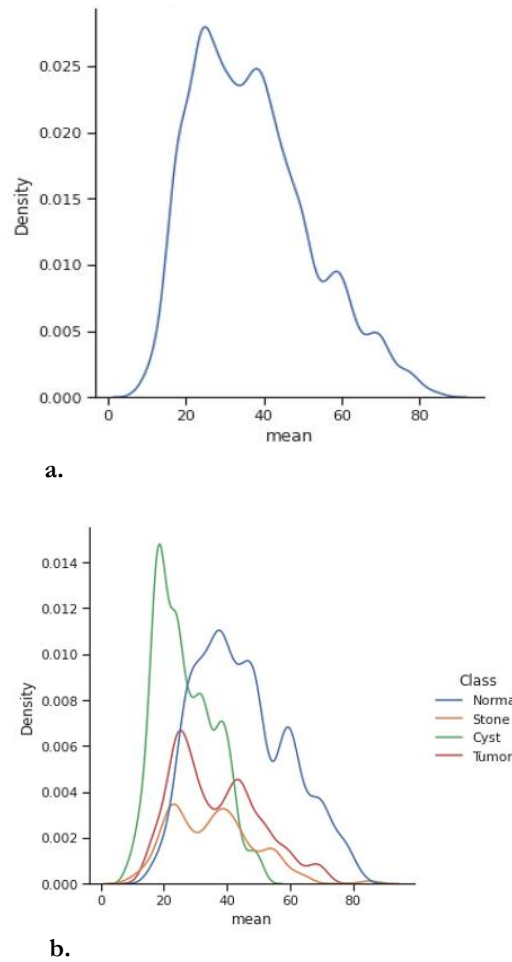


Fig. 4. Distribution of mean intensity values; a. complete dataset, and b. each kidney disease class.

Despite these observable trends, a considerable degree of overlap exists among the intensity distributions of different kidney disease classes. Consequently, reliable discrimination between cyst, tumor, and stone cases cannot be achieved solely through simple statistical analysis of image intensity characteristics. This observation highlights the necessity of employing more sophisticated feature extraction and DL techniques capable of capturing complex structural and textural patterns within CT images.

In the following sections, the outputs obtained from each stage of the proposed framework are presented and analyzed in detail.

4.2 | ROI Extraction (Removal of Redundant Image Margins)

As described in the proposed methodology, the first preprocessing stage involves removing the redundant black margins surrounding the CT images. This operation is performed by applying an appropriate threshold value to identify and eliminate non-informative background regions. The primary objective of this step is to reduce computational complexity while improving processing efficiency.

Fig. 5 presents representative CT images before and after the cropping procedure. As can be observed, the unnecessary black borders have been successfully removed in all cases. As a secondary benefit, the effective image area becomes larger, allowing anatomical details to occupy a greater proportion of the image. Consequently, subsequent feature extraction and statistical analysis can be performed more effectively, since homogeneous zero-valued background pixels no longer influence the extracted descriptors.

The resulting cropped images are subsequently utilized as inputs to the feature extraction and DL stages.

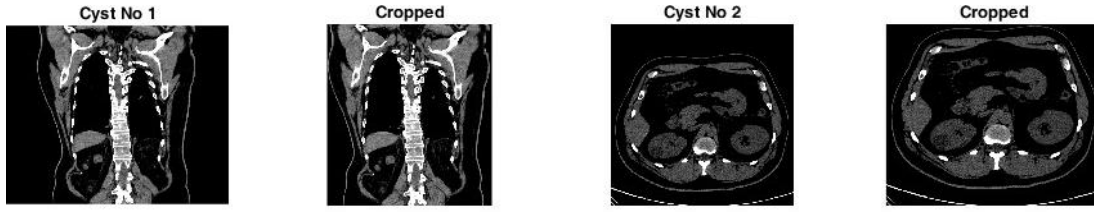


Fig. 5. Representative CT images before and after the removal of redundant image margins.

4.3 | Image Histogram Extraction

In the second stage, each CT image is partitioned into one hundred equal-sized local windows arranged in a 10×10 grid. The windowing strategy employed in the proposed framework is illustrated in Fig. 6.

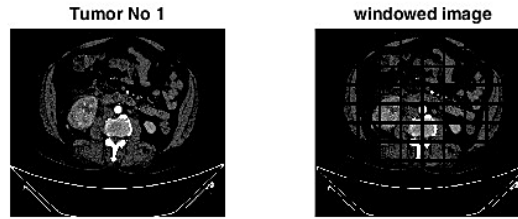


Fig. 6. Window partitioning scheme used for local feature extraction.

Following image partitioning, an asymmetric normalized histogram is computed for each local window. The use of non-uniform histogram intervals is motivated by the uneven distribution of pixel intensities within CT images. Specifically, low-intensity pixels occur considerably more frequently than high-intensity pixels. Therefore, allocating narrower histogram intervals to lower intensity ranges enables the extraction of more detailed statistical information.

To justify the selected histogram configuration, Fig. 7 presents several representative local windows extracted from a sample CT image. As illustrated, the asymmetric histogram design significantly increases the amount of discriminative information captured from certain image regions. In windows where most pixels are concentrated near zero intensity, the proposed histogram behaves similarly to a conventional histogram representation. However, in regions containing richer structural details and brighter tissue characteristics, the proposed adaptive interval selection provides a more informative statistical description of pixel distributions.

Consequently, the proposed asymmetric histogram representation contains substantially more discriminative information than conventional uniform histogram approaches.

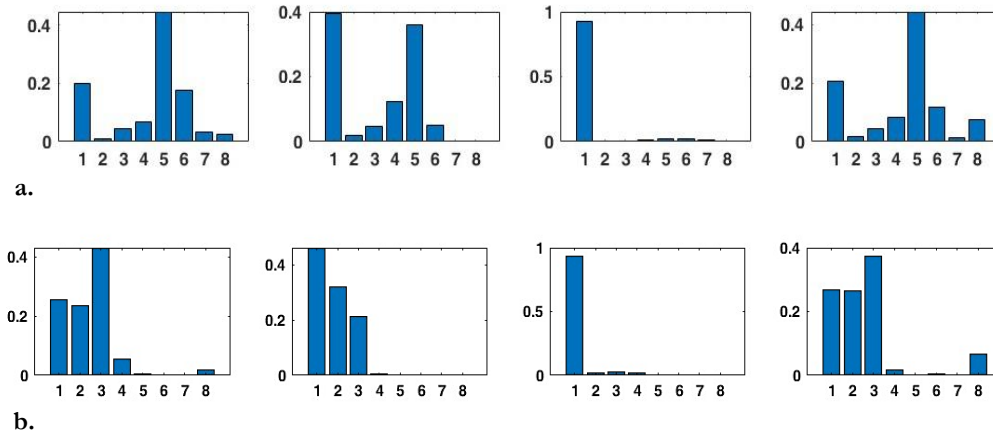


Fig. 7. Histograms extracted from four representative windows; a. proposed asymmetric histogram intervals, and b. conventional uniform histogram intervals.

4.4 | Gradient Histogram Extraction

The next feature extraction stage involves applying a gradient operator to the CT images, followed by histogram extraction from the resulting gradient images using the same local windowing strategy described previously. Gradient images emphasize edge information and tissue boundaries, thereby highlighting structural variations between normal and pathological regions. Since kidney abnormalities often manifest as changes in tissue boundaries and lesion contours, diseased images are expected to contain a larger number of discriminative edge structures than healthy images. Visual inspection of the gradient images reveals that they contain a substantially larger proportion of zero-valued pixels than the original CT images. Furthermore, the intensity distribution of gradient images is generally concentrated near lower intensity values. Consequently, a histogram design that allocates denser intervals near zero and wider intervals at higher intensity levels is more appropriate for capturing subtle edge-related information.

After extensive experimentation, a histogram configuration different from that used for the original CT images was selected to maximize class separability. Therefore, both the original images and their corresponding gradient images are represented using adaptive non-uniform histogram intervals rather than conventional uniform binning strategies. The denser interval allocation near zero enables the extraction of significantly richer statistical information, resulting in improved discrimination among the kidney disease classes.

4.5 | Classification Using LSTM

The input to the DL classifier consists of a 1,600-dimensional feature vector comprising 800 histogram features extracted from the original CT image and 800 features extracted from the corresponding gradient image. All feature values are normalized to the range $[0,1]$.

For model development, 1,000 images from each class were randomly selected for training, while the remaining samples were reserved for testing. The selection of training and testing samples was performed randomly to ensure unbiased evaluation.

Fig. 8 illustrates the convergence behavior of the proposed LSTM-based classification network during training. As can be observed, the network converges smoothly and demonstrates stable learning characteristics throughout the optimization process.

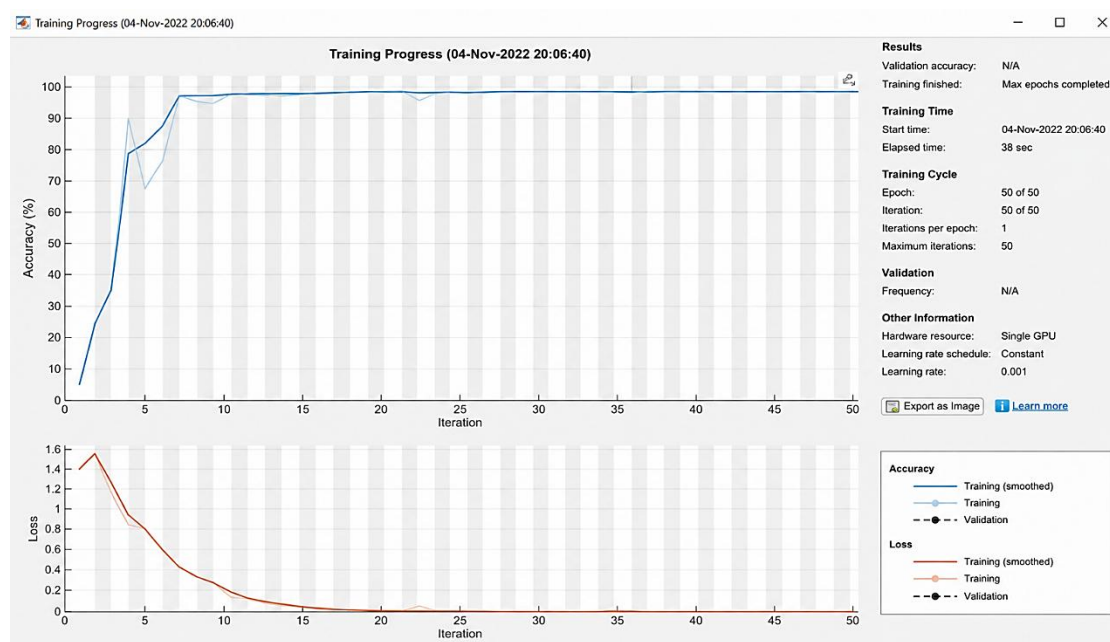


Fig. 8. Convergence behavior of the proposed LSTM network during training.

4.6 | Performance Evaluation

Fig. 9 presents the confusion matrices obtained for both the training and testing datasets.

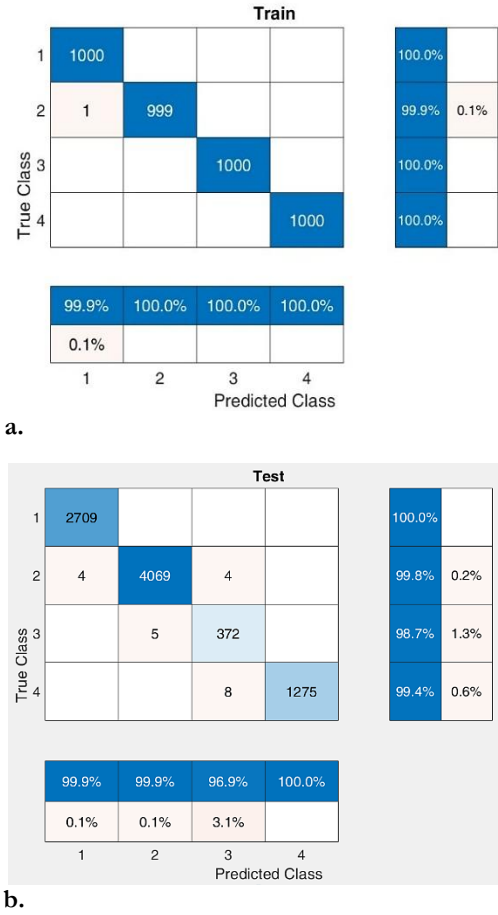


Fig. 9. Confusion matrices; a. training data, and b. testing data.

The classification performance achieved on the training dataset is exceptionally high, with nearly all classes being accurately distinguished. To verify that the model does not suffer from overfitting, the performance on the independent testing dataset was also evaluated. Despite the relatively limited number of training samples and the class imbalance present in the dataset, the proposed method successfully maintains excellent classification performance on unseen test data. One of the primary advantages of confusion matrix analysis is its ability to provide a comprehensive visualization of classification errors in multi-class problems. Examination of the final confusion matrix reveals that the Normal class is classified almost perfectly, while misclassifications mainly occur between neighboring pathological classes. These results indicate that the proposed feature extraction strategy provides highly discriminative representations capable of effectively separating different kidney disease categories. Table 2 compares the proposed method with several recent DL approaches evaluated on the same dataset. According to the reported results, the YOLOv7-based framework proposed by Bayram et al. [16] and Wang et al. [17], which performs classification directly from CT images, achieved an average accuracy of approximately 85% across the four classes. This performance is substantially lower than that obtained by the proposed approach. Furthermore, Islam et al. [18] evaluated six different DL architectures on the same dataset. Their best-performing model, based on the highly sophisticated Swin Transformer architecture, achieved an accuracy of 99.3%, while conventional DL networks attained a maximum accuracy of 98.2%. Despite its significantly lower computational complexity, the proposed framework achieves an accuracy of 99.76%, surpassing both conventional DL models and the Swin Transformer architecture.

Table 2. Performance comparison between the proposed method and recent deep learning approaches evaluated on the Kaggle kidney CT dataset.

Models	Accuracy	Class	Precision	Recall	F1 Score
YOLOv7 [17]	---	Cyst	0.892	0.633	0.74
		Normal	---	---	---
		Stone	0.819	0.855	0.836
		Tumor	---	---	---
		Average	0.882	0.829	0.854
EANet [18]	77.02%	Cyst	0.593	1	0.745
		Normal	0.896	0.848	0.871
		Stone	0.845	0.495	0.624
		Tumor	0.93	0.777	0.847
Swin transformers [18]	99.30%	Cyst	0.996	0.996	0.996
		Normal	0.996	0.981	0.988
		Stone	0.981	0.989	0.985
		Tumor	0.993	1	0.996
CCT [18]	96.54%	Cyst	0.968	0.923	0.945
		Normal	0.989	0.975	0.982
		Stone	0.94	1	0.969
		Tumor	0.964	0.964	0.964
VGG16 [18]	98.20%	Cyst	0.996	0.968	0.982
		Normal	0.985	0.973	0.979
		Stone	0.966	0.988	0.977
		Tumor	0.982	0.996	0.989
Inception v3 [18]	61.60%	Cyst	0.645	0.826	0.724
		Normal	0.584	0.898	0.708
		Stone	0.568	0.462	0.509
		Tumor	0.76	0.295	0.425
Resnet50 [18]	73.80%	Cyst	0.735	0.641	0.685
		Normal	0.77	0.79	0.78
		Stone	0.745	0.692	0.717
		Tumor	0.706	0.827	0.762
Proposed method	99.76%	Cyst	0.999	1.000	0.999
		Normal	0.998	0.999	0.998
		Stone	0.989	0.942	0.965
		Tumor	0.992	1.000	0.996

To further demonstrate the effectiveness of the proposed framework, a broader comparison with state-of-the-art studies published in recent years is presented in *Table 3*. The table summarizes the publication year, employed methodology, and reported classification accuracy for each study. The comparison indicates that traditional transform-based approaches, including Contourlet, Wavelet, and Curvelet methods, generally achieve classification accuracies ranging from approximately 45% to 76%. In contrast, modern DL techniques

have significantly improved performance, reaching accuracy levels close to 99%. The proposed method combines adaptive local histogram analysis and gradient-based edge information within an LSTM-based DL framework and achieves an accuracy of 99.76%. This improvement can be attributed to the adaptive histogram interval design, which provides a richer statistical representation of image characteristics, as well as the ability of RNNs to model complex dependencies among extracted features. Consequently, the proposed framework offers superior classification accuracy while maintaining a relatively simple and computationally efficient architecture.

Table 3. Comparison of the proposed method with recent kidney disease classification approaches reported in the literature.

Ref.	Year	Classification Method	Dataset	Image Processing Technique	Accuracy (%)
[5]	2017	Deep learning	Self-constructed	Contourlet Transform	76
[7]	2020	Neural network	Self-constructed	Wavelet Transform	45
[19]	2019	Boundary distance regression	Self-constructed	Curvelet Transform	70
[16]	2022	Deep learning	Similar to the present study	-	85
[18]	2022	Deep learning	Similar to the present study	-	99.3
Proposed	-	Deep learning (proposed method)	Similar to the present study	Histogram + Original Image Features	99.76

5 | Results and Conclusion

In this paper, a novel method based on image preprocessing, edge detection, and local statistical features of the image, combined with DL (RNN), is proposed for the detection and classification of kidney cancer. Given the critical importance of early detection of various cancers using medical images and timely treatment of patients, a fast and reasonably accurate classification of kidney conditions based on CT images can be highly beneficial. Considering the existing limitations, such as the shortage of specialist physicians in many regions and the high cost of clinical examinations, the proposed method, which eliminates redundant parts of the image, employs statistical histogram analysis in multiple adaptive windows, and simultaneously considers the effect of edges and the relationships between different regions, can detect kidney cancer with high accuracy and determine its type.

Due to the large number of pixels with low intensity in CT images, the histogram bin intervals were selected to be narrower near zero (low brightness) for both the original image and its gradient, and wider for higher brightness values. This adaptive binning strategy increases the discriminative power of the histogram and provides a more representative statistical description of the image, leading to improved classification accuracy for different types of kidney abnormalities.

The proposed method was implemented on images from the standard and publicly available Kaggle dataset, comprising four classes: cyst, normal, stone, and tumor. The classification accuracy achieved across these classes is 99.76%, which demonstrates a highly competitive performance compared to previous studies. Key advantages of the proposed approach include the elimination of complex feature extraction stages and its excellent computational efficiency.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data Availability

All data are included in the text.

Conflicts of Interest

The authors declare no conflict of interest.

References

- [1] Tsai, A., Yezzi, A., Wells, W., Tempny, C., Tucker, D., Fan, A., ... & Willsky, A. (2003). A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2), 137–154. <https://doi.org/10.1109/TMI.2002.808355>
- [2] Chen, G., Dai, Y., Li, R., Zhao, Y., Cui, L., & Yin, X. (2021). SDFNet: Automatic segmentation of kidney ultrasound images using multi-scale low-level structural feature. *Expert Systems with Applications*, 185, 115619. <https://doi.org/10.1016/j.eswa.2021.115619>
- [3] Kalender, W. A. (2011). *Computed tomography: Fundamentals, system technology, image quality, applications*. John Wiley & Sons. <https://www.wiley.com/en-us/shop/general-introductory-biomedical-engineering/computed-tomography-fundamentals-system-technology-image-quality-applications-3rd-edition-p-9783895786440>
- [4] Francis, A., Harhay, M. N., Ong, A. C., Tummalapalli, S. L., Ortiz, A., Fogo, A. B., ... & International Society of Nephrology. (2024). Chronic kidney disease and the global public health agenda: An international consensus. *Nature Reviews Nephrology*, 20(7), 473–485. <https://doi.org/10.1038/s41581-024-00820-6>
- [5] Hao, X., Bruce, C., Pislariu, C., & Greenleaf, J. F. (2000). A novel region growing method for segmenting ultrasound images. *Ultrasonics Symposium* (Vol. 2, pp. 1717–1720). IEEE. <https://doi.org/10.1109/ULTSYM.2000.921653>
- [6] Hwang, H., & Haddad, R. A. (1995). Adaptive median filters: New algorithms and results. *IEEE Transactions on Image Processing*, 4(4), 499–502. <https://doi.org/10.1109/83.370679>
- [7] Ismael, M. A., Alwazzan, M. J., & Mohammed, R. A. (2025). Reducing noise and improving image contrast with a hybrid approach to enhance the quality of medical imaging. *Journal of Medical Physics*, 50(4), 751–759. https://doi.org/10.4103/jmp.jmp_251_25
- [8] Taassori, M. (2024). Enhanced wavelet-based medical image denoising with Bayesian-optimized bilateral filtering. *Sensors*, 24(21), 6849. <https://doi.org/10.3390/s24216849>
- [9] Sulthana, S. L. S., & Sucharitha, M. (2026). Two-phase speckle noise removal in us images: Speckle reducing improved anisotropic diffusion and optimal bayes threshold. *International Journal of Image and Graphics*, 26(01), 2550071. <https://doi.org/10.1142/S0219467825500718>
- [10] Khor, H. G., Ning, G., Zhang, X., & Liao, H. (2022). Ultrasound speckle reduction using wavelet-based generative adversarial network. *IEEE Journal of Biomedical and Health Informatics*, 26(7), 3080–3091. <https://doi.org/10.1109/JBHI.2022.3144628>
- [11] Chalana, V., Linker, D. T., Haynor, D. R., & Kim, Y. (1996). A multiple active contour model for cardiac boundary detection on echocardiographic sequences. *IEEE Transactions on Medical Imaging*, 15(3), 290–298. <https://doi.org/10.1109/42.500138>
- [12] Yin, S., Peng, Q., Li, H., Zhang, Z., You, X., Fischer, K., ... & Fan, Y. (2020). Automatic kidney segmentation in ultrasound images using subsequent boundary distance regression and pixelwise classification networks. *Medical Image Analysis*, 60, 101602. <https://doi.org/10.1016/j.media.2019.101602>
- [13] Goodfellow, I., Bengio, Y., Courville, A., & Bengio, Y. (2016). *Deep learning*. MIT Press. https://books.google.nl/books/about/Deep_Learning.html?id=Np9SDQAAQBAJ&redir_esc=y
- [14] Bishop, C. M., & Nasrabadi, N. M. (2006). *Pattern recognition and machine learning* (Vol. 4). Springer. <https://www.microsoft.com/en-us/research/wp-content/uploads/2006/01/Bishop-Pattern-Recognition-and-Machine-Learning-2006.pdf>
- [15] Kingma, D. P., & Ba, J. (2014). *Adam: A method for stochastic optimization*. <https://arxiv.org/abs/1412.6980>

- [16] Bayram, A. F., Gurkan, C., Budak, A., & Karataş, H. (2022). A detection and prediction model based on deep learning assisted by explainable artificial intelligence for kidney diseases. *Avrupa Bilim ve Teknoloji Dergisi*, (40), 67–74. <https://dergipark.org.tr/en/download/article-file/2635837>
- [17] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 7464–7475). IEEE. <https://doi.org/10.1109/CVPR52729.2023.00721>
- [18] Islam, M. N., Hasan, M., Hossain, M. K., Alam, M. G. R., Uddin, M. Z., & Soylu, A. (2022). Vision transformer and explainable transfer learning models for auto detection of kidney cyst, stone and tumor from CT-radiography. *Scientific Reports*, 12(1), 11440. <https://doi.org/10.1038/s41598-022-15634-4>
- [19] Yin, S., Zhang, Z., Li, H., Peng, Q., You, X., Furth, S. L., ... & Fan, Y. (2019). Fully-automatic segmentation of kidneys in clinical ultrasound images using a boundary distance regression network. *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)* (pp. 1741–1744). IEEE. <https://doi.org/10.1109/ISBI.2019.8759170>