Check for
updates

Multi-classification of kidney abnormalities in sonography using the LOA-MFO and long-term recurrent convolutional network

F. Josephine Nijofi Mactina[1] · S. Neduncheliyan[1]

Abstract

CKD is a medical condition that affects people all around the world and is a major health concern. It increases the risk of developing cardiovascular and cerebrovascular diseases, which can lead to serious illness and even death. Ultrasound imaging is typically the initial and most widely used diagnostic technique for individuals at risk for CKD. The existing methods are restricted by features with high dimensions, computational hurdles, and extended processing times. To address these issues, this article proposes the development of an enhanced deep-learning model with an optimum selection of features for the accurate diagnosis of CKD. The proposed technique begins with pre-processing, involving image filtering and contrast enhancement. Then, presented an improved Otsu's algorithm for segmenting kidney masses, followed by a stage of feature extraction. A hybrid Lion Optimization Algorithm (LOA) and Moth flame Optimization (MFO) is a novel contributions to improve the convergence rate of the MFO algorithm. The hybrid FS algorithms select the optimal subset of features for disease classification. Finally, a novel Long-Term Recurrent Convolutional Network (LRCN) is introduced for detecting kidney impairment. The models are developed and validated utilizing a database of ultrasonography (US) images including four classes of Chronic kidney images: stone, cyst, tumor, and normal. The efficiency of the framework is assessed based on its of accuracy, precision, recall, F1-score, and specificity of 98.7%, 96.6%, 96.4%, 97.9%, and 96.2% respectively. In addition, testing results show that the framework obtains the best overall performance when compared to existing methods for the classification of ultrasound images.

Keywords Deep learning · Chronic kidney disease · Feature selection · Classification · Ultrasound images · Long-term recurrent convolutional network · Feature extraction · Lion optimization algorithm · Moth flame optimization

✉ F. Josephine Nijofi Mactina
  josephinenijofimactinaf78@gmail.com

[1] Bharath institute of higher education and Research, Chennai, Tamil Nadu 600126, India

1 Introduction

1.1 Background of CKD

Chronic kidney disease is a significant public health issue with serious complications such as renal damage and premature mortality. The CKD condition leads to a steady decline in renal function over time. Commonly diagnosed disorders such as tumors, stones and cyst formation significantly impair kidney function [11, 18]. Untreated kidney stones cause the loss of kidney function and the advancement of CKD. Abnormal kidney stones result from crystal accumulation. Multiple research suggests that persons with reoccurring kidney anomalies will be at an increased risk for CKD. Even though cysts are not the major cause of CKD, they may be a risk. Patients with a history of kidney stones have an elevated chance of developing CKD and requiring dialysis therapy [10, 19]. In Polycystic kidney disease (PKD), fluidfilled cysts develop in the kidney region, leading to multiple cysts formation. The expansion of these cysts can lead to compression of the nephrons and subsequently impair the functioning of the kidneys. It is the most prevalent cause of renal failure. Cysts, stones, and tumors that lead to renal function loss can cause ESRD, CKD etc. To avoid such catastrophic effects, early detection of renal abnormalities is vital [29, 34]. A computer-aided diagnostic (CAD) technology is thus suggested to identify many classes of kidney problems using ultrasound pictures [23, 31]. Complicated kidney illnesses may be averted by recognizing renal abnormalities in their earliest stages.

US imaging is a widely used diagnostic approach for detecting kidney disorders. Ultrasound diagnosis categorizes kidney abnormalities into four types based on their echogenicity: normal, cystic, calculi, or neoplastic. Irregularly shaped, fluid-filled structures characterize kidney cysts [30, 36, 40]. In ultrasonic imaging, they look dark or hypoechoic because the majority of the waves pass through them without being reflected back to the transducer. Most ultrasound waves bounce back to the transducer, making kidney stones hyperechoic and brighter. The irregularly shaped tumors are formed of solid tissues and seem hyperechoic [13, 14, 28].

Furthermore, there are not enough nephrologists, and as a result, many patients may never visit one. Depending on their clinical experience, radiologists may have a distinct perception of ultrasound imaging [17, 37]. Radiologists manually evaluate ultrasound images to get accurate diagnoses from many kidney screenings. Due to their poor contrast, speckle noise in US images will reduce the resolution in images. This makes ultrasonography diagnosis difficult and may affect the findings of the diagnostic [2, 35].

The tumor has been classified using a variety of ML approaches, such as SVMs, MLPs, DTs, k-nearest neighbor and Fisher linear discriminant analysis. Furthermore, a straightforward multi-category classification issue may be elegantly approached using artificial neural network (ANN) approaches. With the decision tree classification approach and the association rule classification approach, medical professionals now have a more effective method for distinguishing between benign and malignant images. Future encoding methods will have more reliability and include 2-D error correction, error encoding, ASQED, FPGA, and radiation in CMC [6, 25].

1.2 Motivation

Early diagnosis of CKD is crucial for the effective treatment and management of the disease. Medical imaging, such as ultrasound, has been widely used for CKD diagnosis. However, the high dimensionality and complexity of ultrasound images present significant challenges for accurate diagnosis. Feature selection is a critical step in medical image analysis, as it aims to determine the most significant attributes that can improve the accuracy of disease diagnosis. Hybrid optimization algorithms have emerged as a promising approach for feature selection in medical image analysis. These algorithms combine multiple optimization techniques to improve the accuracy and efficiency of feature selection. LOA and MFO are two such hybrid optimization algorithms that can show promising results in feature selection for medical imaging analysis.

In this approach, the search process begins with the LOA algorithm, which is a population-based optimization algorithm that imitates the hunting behavior of lions in nature. The LOA algorithm is used to search the solution space and find the global optima of the objective function. Once the LOA algorithm has been run for a specified number of iterations, the MFO algorithm is used to further refine the solution obtained by the LOA algorithm. MFO is also a population-based optimization technique that simulates the behavior of moths in nature. MFO uses a series of search mechanisms, including attraction, repulsion, and diffusion, to explore the solution space and locate the global optima of the objective function. The integration of LOA and MFO involves using the best individuals from the LOA population as the initial population for the MFO algorithm. By integrating the strengths of both algorithms, the hybrid approach can overcome some of the limitations and challenges associated with each individual algorithm and provide more accurate and reliable results.

1.3 Research contribution

This study proposes an automated CAD to diagnose CKD to address the aforementioned issues. Initial components of the proposed CAD system include a pre-processing module that reduces speckle noise in ultrasound images. A segmentation approach is then required to extract the renal area from its surrounding tissues. Thirdly, in the feature extraction step, the collection of statistical, texture and geometrical characteristics are calculated for the selected masses. The previous approaches have a few inadequacies in terms of higher-dimensional characteristics, computing complexity, greater execution time, etc. This work proposes a unique classification approach for CKD prediction using a hybrid metaheuristic algorithm and optimum feature selection. The addition of LOA to MFO increases the convergence rate of the Moth flame optimization algorithm. LRCN was recommended for classifying CKD cases for diagnosis. The following are the paper's contributions:

- Perform pre-processing steps to improve the data quality.
- For the segmentation of the diseased region proposed is an improved Otsu's algorithm.
- In the third stage, the segmented kidney images undergo feature extraction, including geometric features, histogram features, and texture features.
- To choose the significant features from the US image dataset, employed a feature selection algorithm. Hence, employed hybrid LOA-MFO to improve the classification of optimal features from that helps in accurate diagnosis of CKD.

- Finally, DL-based LRCN is utilized to classify the four CKD classes, viz., normal, Cyst, Stone and Tumor.
- To demonstrate the effectiveness of this research, the performance of the proposed algorithm has been compared with previous techniques.

The rest of the sections are structured as follows. Section 2 discussed the existing methods for CKD prediction. The overall proposed work is explained in Section 3. Analysis and explanations of the experimental results are included in Section 4. Section 5 contains the presentation of both the conclusions and the future work.

2 Related work

In this portion, a summary of prior studies on the diagnosis of kidney disorders is provided from four perspectives: methodologies, associated datasets, data modality, and experimental results. Table 1 provides a summary of various existing schemes for normal and abnormal kidney images in the US images.

Pal et al. [27] developed a predictive model for CKD by analyzing the CKD dataset from the UCI Machine Learning Repository that has 25 features. They employed three well-known ML classifiers, namely LR, DT, and SVM, to investigate the dataset. The authors further utilized the bagging ensemble method to enhance the performance of the predictive model, which was trained using the clusters of the CKD dataset. Lastly, the non-linear features of the Kidney Disease Collection were grouped and summarized.

Sudharson et al. [39] introduced a CAD method for identifying multi-class kidney anomalies from US images. They employed ResNet-101 which has been pre-trained to extract features and SVM classifier for classification. Speckle noise affects ultrasound image quality and CAD performance. Therefore, speckle noise must be eliminated from the US images. Pre-processing US images with DRLN improved CAD classification ability.

Chaudhuri et al. [7] presented a novel approach to differentiate chronic kidney disease from non-CKD, detecting the critical factors that affect the illness while maintaining prediction accuracy, and creating a predictive model that prevents the possibility of overfitting or underfitting. The study employed the recursive feature elimination (RFE) method to select the most suitable set of features and utilized the enhanced decision tree (EDT) algorithm in combination to accurately predict the disease.

To predict chronic kidney disease (CKD) in patients, Chittora et al. [8] developed a model which was tested for its effectiveness using both all attributes and a selection of specific features. To identify the most important features, the researchers utilized various methods for feature selection such as Wrapper, Filter, and Embedded techniques. To train the model, a variety of machine learning classifiers including RT, KNN, LSVM, LR, C5.0, and ANN, were utilized, and the IBM SPSS tool was employed in the development of the model.

Integration of pre-training models ACWGANGP and MobileNetV2 were presented by Liao et al. [22] to classify CKD stages. MobileNetV2 concurrently learned the new training photos created by the ACWGAN-GP and merged them with the input images. Deep learning may help an automated system model overcome imbalanced and insufficient training data samples. By employing a GAN model for data augmentation, the classification accuracy of CKD stage significantly improved. This was particularly evident in the case of renal US images with extremely poor quality, where the GAN-generated samples effectively

Table 1 Summarized methodologies from existing papers on ultrasound images

References	Task	Methods	Limitations
Pal et al. [27]	Normal and Abnormal	LR, DT, and SVM	This may result in limited generalizability of the developed predictive model to other datasets
Sudharson et al. [39]	Cyst, stone, tumor and Normal	ResNet 101 + SVM	• Computationally expensive. • Lack of transparency of results.
Chaudhuri et al. [7]	Normal and Abnormal	EDT classifier	It was conducted on a relatively small sample size.
Chittora et al. [8]	Normal and Abnormal	LSVM	• A limited number of ML algorithms for training the model
Liao et al. [22]	Four stages	ACWGAN-GP + MobileNetV2	Prone to over-fitting on small training dataset.
Kim et al. [21]	Normal, Mild and Moderate, Severe	GLCM+ANN	Difficult in segmentation.
Ahmad et al. [1]	Five stages	GLCM+LDA	Results are obtained interms of classification accuracy. Other performance metrics can also be taken in account such as Recall, Precision, Specificity etc.
Sudharson et al. [38]	Cyst, stone, tumor and Normal	DNN	The system does not apply any feature technique to remove redundant features
Nithya et al. [26]	Stone and Tumor	ANN	Large Computational time.
Wu et al. [42]	Normal and Abnormal	mf-Net	Accuracy improves only when the testing and training ratio increases.
Hosseinzadeh et al. [15]	Normal and Abnormal	Naïve Bayes, MLP,J48, and SVM methods	Limited explanation of the feature selection process and its impact on the performance of the model
Khamparia et al. [19]	Normal, and Abnormal	SAE	High computational time
Raju et al. [33]	Cyst and Tumor	FB-FCM-WOA	The implementation of feature selection might have improved classification accuracy.

enhanced the dataset. This approach was applicable to a variety of medical imaging applications, particularly under situations of unbalanced datasets.

Kim et al. [21] revealed that applying the GLCM algorithm to US images categorized as normal, mild, moderate, and severe CKD, as well as severe CKD images, resulted in a range of distinct values. Classifying chronic kidney disease using an ANN utilizing GLCM values and kidney size. This study's objective is to categorize CKD into three groups in order to discover if early detection is achievable.

Ahmad et al. [1] employed GLCM to extract local renal imaging features. The feature extraction block obtains a fourteen-point feature vector from each US kidney image using the specified approach. The classifier then utilizes this vector to predict the stage of CKD.

Sudharson et al. [38] developed an automated categorization of B-mode renal ultrasonography pictures using transfer learning and an ensemble of DNNs. Speckle noise influences ultrasound images, and image quality is selected using a perception-based score. Pre-trained DNN models are provided with three distinct datasets for feature extraction, followed by SVM for classification. By aggregating the predictions of many DNNs, the ensemble model demonstrates superior performance in classification over the individual models.

The kidney stone detection technique developed by Nithya et al. [26] utilizes an ANN and a multi-kernel k-means clustering approach. Normal system modules include pre-processing, feature extraction, classification and segmentation. Utilizing the median filter, they mostly removed the noise in the input image. Then, the image's essential GLCM characteristics were retrieved. Subsequently, a NN classifier was employed to categorize the image as normal or abnormal. The irregular images are then segmented using multi-Kernel K-means clustering.

Wu et al. [42] presented an automated architecture for detecting kidney problems based on abnormalities using ultrasonography and CNNs. The model has three parts: Initially chooses relevant US images of kidneys, the second uses a detection model to find the kidney in the image, and finally uses Mf-Net to discriminate between normal and defective kidneys by recognizing aberrant symptoms. To boost performance, the presented model was supplemented with a weighted ensemble technique. Mf-Net additionally includes a multi-feature fusion layer for extracting distinguishing features from several picture perspectives. The three components collaborate to automatically identify kidney-related problems.

Hosseinzadeh et al. [15] introduced a model that utilizes sensor technology within an IoT platform to predict the severity level of CKD. The model was built upon analyzing the correlation between various vital signs and clinical symptoms that may have an impact on the severity level of CKD. The study aimed to tackle the problem of low efficiency and prolonged execution time in CKD prediction processes caused by the presence of an excessive number of features. To address this issue, the researchers proposed a feature set comprising influential parameters. These parameters were obtained from previous studies on CKD prediction, as well as clinical observations made by physicians. The classification methods employed in the study included J48, SVM, MLP, and Naïve Bayes methods.

Khamparia et al. [19] introduced a DL-based approach for classifying CKD using multimedia data. The approach proposed in the study involves using a stacked autoencoder model for extracting features, and a softmax classifier for the final classification of CKD. The study utilizes various multimedia data, such as serum creatinine, age, gender, blood pressure, etc. to train the model.

Raju et al. [33] developed an effective Optimal GLCM attribute-related FCM segmentation method to classify kidney tumors and cysts from ultrasound kidney images. The FCM

makes use of certain properties of the GLCM texture extractor and the Whale optimization method to optimally attach the cluster centroids of the FCM. The suggested method was implemented on the Matlab platform.

Despite the comprehensive literature survey and limitations outlined above, the current state of research on predicting CKD using ultrasound imaging still presents several research gaps and are listed below:

- Most existing studies on CKD diagnosis using ultrasound imaging have used standard image filters and contrast enhancement techniques. However, these techniques may not be optimal for enhancing the diagnostic features of CKD in ultrasound images.
- The studies have focused on segmenting the kidney mass in ultrasound images for CKD diagnosis. However, accurate segmentation of other relevant anatomical structures, such as the renal cortex, medulla, and pelvis, may also provide important diagnostic information.
- Previous studies have employed several techniques to extract features for CKD diagnosis, including texture analysis and shape-based features. However, these techniques may not be suitable for capturing the distinctive diagnostic characteristics of CKD in ultrasound images. Therefore, there is a need for research into developing novel feature extraction techniques that can capture the unique and complex diagnostic features of CKD in ultrasound images.
- The studies have used conventional feature selection techniques, such as PCA and RFE, to select the most relevant features for CKD diagnosis. However, these techniques may not be optimal for capturing the complex and nonlinear relationships between the features and the disease. Therefore, there is a need for research into developing more advanced feature selection techniques that can capture these complex relationships and improve the accuracy of CKD diagnosis.
- The majority of studies have used supervised learning techniques, such as SVMs and ANNs, for CKD diagnosis using ultrasound images. There is a need for research into developing more advanced and efficient classification techniques, such as deep learning models and ensemble classifiers, that can handle these challenges and improve the accuracy of CKD diagnosis.

Our research focuses on a unique feature selection and classification method that has been developed for the CKD classification in order to address the research gaps identified in the literature survey. In addition, the results of the proposed classifications show the maximum accuracy. Our study provides a cost-effective and accurate solution for medical examination of CKD. Our approach requires less than 20 features for model training and classification of the condition, demonstrating its potential as a promising solution.

3 Methodology

The proposed method for classifying US images of the human kidney consists of several phases, including pre-processing, segmentation, extraction of features, selection of features, and classification. Initially, US images are evaluated to enhance image quality, followed by a feature extraction technique to extract the attributes (texture, geometric and histogram) of the images based on certain strategies. Following feature extraction, the method of feature selection is explored to choose the most important characteristics

for the classification stage. In the proposed classification approach, a unique hybrid LOA-MFO-based feature selection strategy is used to exclude irrelevant data.

On the basis of the extracted characteristics, the US Chronic kidney images are classed as normal, cyst, stone, or tumor during the classification phase. The classification problem typically consists of two stages ie., testing and training phases; The classifier is trained utilizing training data features. In contrast, the results of the classification technique during the testing phase indicate whether the images include kidney cancer areas [cyst, stone, tumor] or non-cancer regions. The present investigation employed an LRCNN-optimized feature set. This method is shown in Fig. 1 by its superior efficiency and minimal effort in both training and testing stages for the detection of US Chronic kidney images.
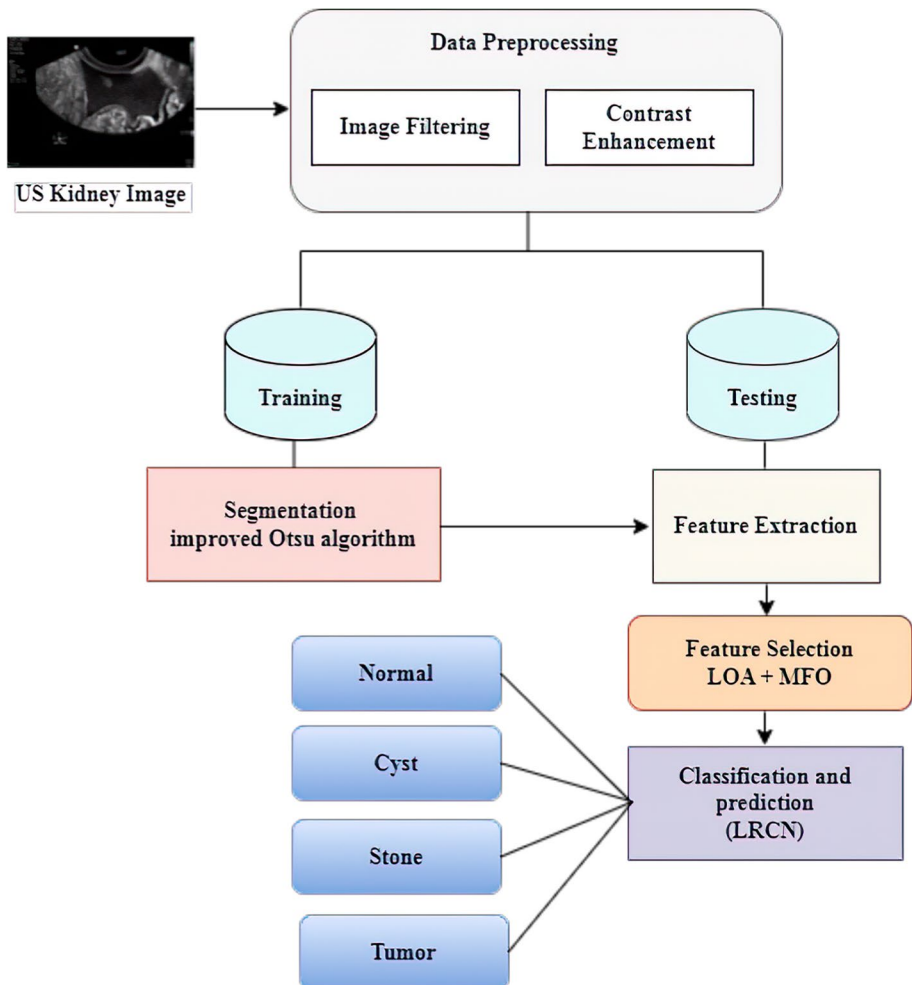

Fig. 1 Architecture of the proposed methodology

3.1 Pre-processing

Ultrasound images must undergo pre-processing because noise is more prevalent in ultrasound images than in other imaging modalities, such as MRI and CT. US images are mostly affected by speckle noise. Therefore, noise reduction is an important process in medical US images.

3.1.1 Image filtering

In this study, Bicubic-Interpolation (BI) is used to enhance raw images by removing noise. Using Bicubic Interpolation, a high-resolution picture is obtained from low-resolution data. Then, pre-processing involves the use of filters such as Gaussian, guided filter, median filter, top & bottom hat, etc., to minimize image noise and smooth the picture. A Guided filter, Gaussian filter, top & bottom hat, and median filter are used to smooth the image. The pre-processing steps for kidneys are as follows:

Step 1.   To form I_{r*t}, BI is performed,

$$
\begin{aligned}
&I(i,j) \leftarrow a_{ij}x^i y^j \\
&I(i,j) - \quad image\ at\ location \\
&a_{ij} - \quad low\ resolution\ image \\
&x^i y^j - high\ resolution\ image
\end{aligned}
\tag{1}
$$

Step 2.   Utilizing top hat – bottom hat, image is enhanced

$$
I_{Tho} \leftarrow I_{guassian} - \left[I_{guassian}\theta\ strel) \oplus strel \right]
\tag{2}
$$

$$
I_{Bhc} \leftarrow \left[(I_{guassian} \oplus strel)\theta strel \right] - I_{guassian}
\tag{3}
$$

$$
I_{Th-Bh} \leftarrow I_{guassian} + I_{Tho} - I_{Bhc}
\tag{4}
$$

$$
I_{enhance} \leftarrow I - I_{guassian} + I_{Th} - Bh
\tag{5}
$$

$$
I \leftarrow I_{enhance}
\tag{6}
$$

$Strel - strcturing\ element$
$I_{guassian} \gg Gaussian\ filter\ (I)perform\ Gaussian\ filter\ on\ I$
$I_{Tho} - image\ after\ performing\ top\ hat\ opening$
$I_{Bhc} - image\ after\ performing\ bottom\ hot\ closing$
$I_{Th-Bh} - image\ after\ performing\ subtract\ (Th - Bh)$

Step 3.   Subtracting weighted Laplacian from blurred image, sharpen can be computed

$$
I_{laplacian} \leftarrow I_{grad\ X} + I_{grad\ Y}
\tag{7}
$$

$$
I_{sharpen} \leftarrow I_{enhance} - C * I_{laplacian}
\tag{8}
$$

$C-$ *is multiplicative coefficient*

$I_{\text{grad } X}-$ *image gradient in X − direction*

$I_{\text{grad } Y}-$ *image gradient in mage gradient in Y − direction*

3.1.2 Contrast enhancement phase

Adaptive histogram equalization is the technique for contrast enhancement used after eliminating database noise

$$contrast(i,j) = \text{rank} * \text{max\_int } ensity(i,j). \because Initially rank = 0$$
$$\text{rank} = 0 + 1 \tag{9}$$

The trailing column carrying the new leading row is subtracted from the main position of the preceding row to produce the histogram in the major role of each line. The difficulty of US images must be increased and limited such that the image's gray level is identified and modified adaptively in the new histogram. The filtered image and the original image are presented in Fig. 2.

3.2 Kidney segmentation image technique utilizing improved Otsu's approach

Identification and segmentation of the kidneys in ultrasound images are important for disease diagnosis. A segmentation technique's precision has a direct effect on the feature extraction and classification process. Consequently, it is difficult to precisely segment the kidney from ultrasound images owing to the influence of several variables.

Here, $j(x,y)$ is the kidney image along with levels of grey 0, 1...$L-1$. Input imagery j can described as $F = \{f_0, f_1......f_{L-1}\}$, where $f_0, f_1......f_{L-1}$ represents frequency of every greyscale input image i.

$$p_j = \frac{n_i}{N}, and\, p_j \geq 0, \sum_{j=0}^{L-1} p_j = 1 \tag{10}$$

Where, $N = \sum_{j=0}^{L-1} p_j$ is the total number of kidney images in pixels.

(a)   (b)
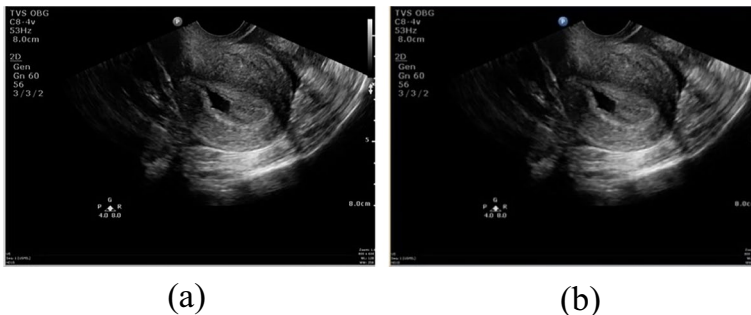
Fig. 2 **a** Input image **b** Pre-processed image

Probability for j greyscale level is calculated as

$$P_1(t) = \sum_{i=0}^{t} p_j, P_2(t) = \sum_{j=t+1}^{L-1} p_j = 1 - P_1(t) \tag{11}$$

This segmentation technique is an extension of Otsu's method [16]. They can be computed as:

$$\mu_1(t) = \frac{\sum_{j=0}^{t} jp_j}{P_1(t)}, \mu_2(t) = \frac{\sum_{j=t+1}^{L} jp_j}{P_2(t)} \tag{12}$$

$$\sigma_1^2(t) = \sum_{j=0}^{t} \left(j = \mu_i(t)\right)^2 \frac{p_j}{p_1(t)}, \quad \sigma_2^2(t) = \sum_{j=t+1}^{L} \left(j - \mu_2(t)\right)^2 \frac{p_j}{p_2(t)} \tag{13}$$

The overall mean of an image's pixels is denoted by μ:

$$\mu = \sum_{j=0}^{L-1} jp_j = P_1(t)\mu_1(t) + P_2(t)\mu_2(t) \tag{14}$$

The within-class variance $\sigma_\omega^2(t)$ *of* C_1 *and* C_2 is defined as follows:

$$\sigma_\omega^2(t) = P_1(t)\sigma_1^2(t) + P_2(t)\sigma_2^2(t) \tag{15}$$

The optimal threshold T is obtained by minimizing the within-class variance $\sigma_\omega^2(t)$ as follows:

$$T = \underset{0<t<L-1}{argmin} \sigma_\omega^2(t) \tag{16}$$

In this model, a threshold approach is described for automating image segmentation of the kidney. In the process of over-segmentation, the inter-class variation between background and foreground intents and the noise level is examined.

Our proposal involves incorporating a weight parameter into the within-class variance of the Otsu technique, resulting in an enhanced approach. Following is a description of the specifics of the enhanced Otsu technique.

The Otsu method's within-class variance, as given by Eq. (15), has two classes. In order to achieve a specified threshold, the initial component must have a lower impact on the variation within the class. To weight the initial item, the parameter, which ranges from 0 to 1, is introduced. Now, the background variance is multiplied by, which effects the within-class variance. By utilizing the value of α, it is possible to rephrase Eq. (15) as shown in Eq. (1):

$$\sigma_\omega^2(t) = \alpha P_1(t)\sigma_1^2(t) + P_2(t)\sigma_2^2(t) \tag{17}$$

Under α = P_2(t) Above Eq. (17) is rewritten as follows:

$$\sigma_\omega^2(t) = P_2(t)P_1(t)\sigma_1^2(t) + P_2(t)\sigma_2^2(t) \tag{18}$$

The function below is used to select the optimal threshold T:

$$T = \underset{0<t<L-1}{argmin}\sigma_\omega^2(t) \tag{19}$$

$$= P_2(t)P_1(t)\sigma_1^2(t) + P_2(t)\sigma_2^2(t) \tag{20}$$

Adjusting the alpha value of n parameters yields n segmentation thresholds. In this method, the selection range of the threshold may be increased, which not only assures the largest variation between the target class and the background class, but also ensures the threshold's uniformity within the class.

3.3 Feature extraction

Feature extraction compresses and separates a vision into a matrix-vector or single values. In image processing, feature extraction computes dimensionality reduction depending on which image may be utilized for classifications. It involves reducing the input data to a subset of representative attributes. Classifiers employ the attributes to assign them to the class they represent. The objective of feature extraction is to decrease the amount of data by predicting positive characteristics. In the present work, geometric, texture and histogram characteristics are extracted from separate ultrasound image bands, while these features are retrieved from various US image bands.

3.3.1 Geometric features

The segmented binary image of each kidney mass is used for computing the geometric characteristics. The geometric characteristics used to calculate distinction are:

Size In our technique, the main axis length of the nodule is used to determine its size. The excised nodule of interest is measured in centimeters for size. Calculating size will assist in the detection of small cancerous tumours.

Solidity Solidity describes a form's convexity or concavity.

$$Solidity = \frac{A}{H} \tag{21}$$

where, convex hull area of the object is denoted as H, area of the object region is denoted as A.

Equivalent diameter The circle's diameter equals the tumor's area. Since it acts on area, it may be used with binary images.

Perimeter The perimeter of a nodule of interest is determined by the number of pixels present along its boundaries. The results will be accurate with the binary image format.

Area The tumor's binary image determines this property. Area calculated in binary essentially counts the tumor's pixels. It improves accuracy.

Convex area This function counts pixels in a convex nodule or tumor. ROI is calculated in its binary version.

$$D = \frac{A}{\left(4\pi R^2\right)} \tag{22}$$

Spherical disproportion

$$R = \sqrt[3]{\frac{3V}{4\pi}} \tag{23}$$

where, area of the object as *A*, radius of the sphere as *R*, volume as *V*.

$$E = \frac{100.n}{V} \tag{24}$$

Spherical density
where, number of voxels as *n*.

Irregularity index It measures border-induced form irregularity.

$$Irregularity\ index = \frac{4\pi.Area}{(Perimeter)^2} \tag{25}$$

3.3.2 Histogram features

The image is rendered in pixels according to the criteria of the histogram. The histogram displays the total number of image pixels at each intensity level. The transformation of the power values of the image's histogram approximates a predefined histogram. The histogram technique evaluates the complete range of grayscale values in a picture. Here, there are 256 shades of gray, ranging from 0 to 255. Variance, mean, skewness, kurtosis, and standard deviation are among their frequent characteristics.

Mean The mean gives the average gray level of each zone and is not a power measure.

$$\mu = \frac{1}{N \times M} \sum_{i=1}^{M} \sum_{j=1}^{N} p(i,j) \tag{26}$$

where, size of the image as $M \times N$, $p(i, j)$ is the pixel's intensity value at the point *(i, j)*.

Kurtosis It measures a real-valued random variable's potential distribution and detects visual abnormalities.

$$Kurtosis = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[\frac{p(i,j) - \mu}{\sigma} \right]^4 \tag{27}$$

where, standard deviation as σ.

Skewness Using the histogram tail, skewness is calculated. The value of the histogram's tail is separated into two groups, negative and positive.

$$Skewness = \frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} \left[\frac{p(i,j) - \mu}{\sigma} \right]^3 \tag{28}$$

Variance Variance is the number of gray level differences. The statistical distributions might be used to discern low-profile texture differences.

Standard deviation SD is the square root of visual contrast variance. The contrast level of an image is evaluated using high and low variance values. The variance of an image is a measure of its contrast level. A higher variance indicates a higher contrast, while a lower variance indicates a lower contrast.

$$\sigma = \sqrt{\frac{1}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} (p(i,j) - \mu)^2} \tag{29}$$

where, mean intensity as μ.

Skewness and Kurtosis are used in statistical analysis to determine the distribution's form.

3.3.3 Texture features

Histogram features are taken from the input image before texture features are extracted. Since the irregularity is widely spread, each class's textural orientation is different, improving classification accuracy. GLCM considers the spatial relationship of pixels while analyzing a surface.

GLCM functions define an image's texture by calculating identical pixel pairs. In general, these characteristics are computed using GLCM probability values, and there are somewhere between 18 and 22 of them; however, only a small subset of these features is used for the classification of US chronic kidney images in the present investigation.

$$G_{P_{ij}} = \frac{F_{ij}}{\sum_{i,j=0}^{L-1} F_{ij}} \tag{30}$$

Energy Highest consistent values or intermittent consistency in gray level distribution represent the most essential surface.

$$Energy = \sum_{i=0}^{N_{p-1}} \sum_{j=0}^{N_{p-1}} P^2(i,j) \tag{31}$$

Entropy It is the amount of picture data necessary for compression. Highest consistent values or intermittent consistency in gray level distribution represent the most essential surface.

$$H = - \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} G(i,j).\left(\log_2 G(i,j) \right) \tag{32}$$

Homogeneity The homogeneity constraint is sometimes referred to as the contrast minute, which analyzes the homogeneity of an image on the assumption that

the predominant values for minor gray-tone variations in pair components exist. Consequently, homogeneity is an assessment that describes prevalent values for images with low contrast.

$$Homegeneity = \sum_{i=0}^{n_{1p-1}} \sum_{j=0}^{n_{1p-1}} \frac{p(i,j)}{1 + |i - j|} \tag{33}$$

Contrast This one calculates spatial recurrence and GLCM moments. It compares nearby pixels' greatest and lowest values.

$$Co = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} (i - j)^2 G(i,j) \tag{34}$$

Correlations Correlation measures the linear correlation between adjacent pixel gray levels. The tracking of the digital image correlation is an optical process that makes use of tracking and approaches image registration for measuring visual changes.

$$R_c = \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2} (i - j)G(i,j) - \mu_i \mu_j}{\sigma_i \sigma_j} \tag{35}$$

$$\sigma_i = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left(i - \mu_i\right)^2 G(i,j) \tag{36a}$$

$$\sigma_i = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \left(i - \mu_j\right)^2 G(i,j) \tag{36b}$$

It is very challenging to classify kidney masses based on a single computed feature set, such as geometry or textural or statistical factors only. Consequently, a set of attributes are necessary for the proper categorization of CKD cases. In contrast, increasing the number of characteristics increases the computing cost. To balance accuracy and computation performance, we developed a nature-inspired optimization technique for CKD classification.

3.4 Kidney image feature selection using hybrid Meta-heuristic algorithm (LOA-MF optimization)

After the feature extraction process, the significant characteristics are identified, which are often performed by relevant search algorithms. Several search strategies have previously been suggested. SVM-RFE, PSO, GA, branch and bound, sequential backward selection and sequential forward selection are popular feature selection techniques. In contrast, increasing the number of characteristics increases the computing cost. We have developed a nature-inspired optimization approach, the LOA [43] & MFO [24] algorithm, to find the best attributes for the accurate classification of kidney disease while maintaining the balance between computational time and accuracy. The proposed feature selection procedure is described in detail below.

To avoid premature convergence and speed up convergence, enhance population variety. As it was previously mentioned, the proposed model contains a hybrid (LOA-MF). Here, the section mainly covers a hybrid (LOA-MF) is briefly described. In the case of Moth-flame, it might progressively promote population diversity in contrast to early convergence. Moreover, it successfully permits the local optimal. Following this section are mathematical models of initiation, hunting, migrating to a safe location, wandering, mating, defense, migration, equilibrium, and convergence.

3.4.1 Initialization

The initial stage involves randomly generating the M population over the solution space. Here, there are fifty inhabitants. The remaining populations were split into two prides of 20 lions each, and 20% of the solutions were selected at random to represent nomad lions. Seventy percent of each pride is considered female, while the remainder is masculine. 30% of nomad lions are considered females, while the remaining 70% are males.

In a kidney disease classification problem, a lion symbolizes a group of traits used to categorize kidney disease samples. Let ($n = 123$) represent the no. of attributes associated with $k = 50$ objects (lion). If the attributes j is chosen for I, then assigned as '1'. otherwise, 'assigned as 0' ($j < = n, i < = k$).

3.4.2 Hunting

Each pride contains a small number of females that hunt and bring back prey for the rest of the pride. A roulette wheel selects female lions. Occasionally, prey may elude the hunter, and a new location of prey can be determined using Eq. (37):

$$prey' = prey + \text{ran}(0, 1)PreI \times (prey - Hunter) \tag{37}$$

Hunter's fitness improvement is denoted as *PreI*, Prey's current location as $prey'$. Hunter groups employ Eqs. (38) and (39) to surround targets.

$$Hunter' = \begin{cases} \text{rand}((2 \times prey - Hunter), prey), (2 \times prey - Hunter) < prey \\ \text{rand}(prey, (2 \times prey - Hunter)), (2 \times prey - Hunter) > prey \end{cases} \tag{38}$$

$$Hunter' = \begin{cases} \text{rand}(Hunter, prey), Hunter < prey \\ \text{rand}(prey, Hunter), Hunter > prey \end{cases} \tag{39}$$

Hunter's new location is denoted as $Hunter'$.

3.4.3 Advancing towards a safe location

Each pride has several females hunting, while the rest are located in protected areas. The most efficient solutions are preserved and used to develop future events. Consequently, the new location of a female lion is calculated using the (40)

$$New\ Position\ Female\ Lion' = Current\ Position\ Female\ Lion + 2D$$
$$\times\ \text{rand}(0,1)\{R_1\} + U(-1,1) \times (tan\theta) \times D \times \{R_2\} \tag{40}$$

3.4.4 Roaming

Roaming helps the algorithm get a better solution via local search. Every male lion of a pride roams the area of pride. To simulate this behavior of resident male prides, 20% of the region is randomly picked and visited by the lion. As a male resident roams, the best-visited solution is updated whenever an improved position is found. Both female and male nomadic lions roams in search space randomly, as shown by the Eq. (41)

$$Lion_{ij} = \begin{cases} Lion_{ij} & if\ \text{rand}_j > pr_i \\ \text{rand}_j & Otherwiswe \end{cases} \tag{41}$$

Where $Lion_{ij}$ is current of ith named lion in j^{th} dimension. Probability (pr_i) is given in Eq. (42)

$$pr_i = 0.0 + \min\left(0.5, \frac{nomad_i - Best_{nomed}}{Best_{nomed}}\right) \tag{42}$$

$$i = 1,2,\ldots no.of\ nomad\ lions$$

3.4.5 Mating

Thirty percent of female lions in each pride mate with random men to have progeny. In nomad lions, females mate with randomly chosen males. Roulette selects 30% of female pride lions. Using Eqs. (43) and (44), mating a female lion with a male produces cubs.

$$offspring_j1 = \beta \times Female\ Lion_j + \sum \frac{(1-\beta)}{\sum_{i=1}^{noR} S_i} \times male\ lion_j^i \times S_i \tag{43}$$

$$offspring_j2 = (1-\beta) \times Female\ Lion_j + \sum \frac{\beta}{\sum_{i=1}^{noR} S_i} \times male\ lion_j^i \times S_i \tag{44}$$

where j is a dimension. Number generated randomly is represented as β. Each gene of one offspring is mutated with a 1% chance.

3.4.6 Defence

Male lions engaged in fighting with other males within their pride as they grow. Men whose pride has been destroyed become nomads. If a nomad male lion engages in male-on-male combat with a pride, the defeated male is expelled from the pride and becomes a wanderer.

3.4.7 Migration

The percentage of migratory females in each pride is 40 %. Selected ladies abandon their pride to become nomads. The old and new nomad female lions are classified by RMSE (fitness), and the fittest lion is chosen to replace the vacancy left by the moving female.

3.4.8 Lions population equilibrium

During the completion of every cycle, the number of living lions are compared to the maximum size of every gender permitted by nomads; Lions with the lowest fitness value are eliminated in the nomadic stage.

3.4.9 Convergence

To improve population diversity, it is necessary to avoid premature convergence; the convergence rate must be increased using the Moth-flame optimization procedure. The moth-flame has the unusual ability to successively increase population diversity. This procedure aids in escaping the local maximum. Furthermore, this approach is effective for the MFO exploitation and exploration process.

Based on the application of equation, the convergence process of the modified Moth-flame optimization technique is included in this (45). Below is the mathematical formulation of this procedure.

$$X_i^{i+1} = X_i^t + usign[\text{rand} - 0.5] \oplus \text{Levy}(\beta) \tag{45}$$

The representation is denoted by the sign [rand −0.5]. The integration of Modified MF into the Lion Optimization method allows for the reduction of local minima and the enhancement of global search capabilities.

Additionally, the Moth-flame process depends mostly on a random walk, in which the step length defines the process step, while the Jump process validates levy distribution approach. It is expressed as,

$$Levy(\beta) \sim \mu = t^{-1-\beta}, \qquad (0 \le \beta \le 2) \tag{46}$$

Levy random numbers are computed in Eq. 47

$$Levy(\beta) \sim \frac{\theta \times \mu}{|v|^{1/\beta}} \tag{47}$$

Where, $\beta = 1.5$, standard Gamma Function as Γ, standard normal distributions as μ and v- . Then, ϕ is expressed as:

$$\phi = \left(\Gamma(1 + \beta) \times \sin\left(\pi \times \frac{\beta}{2} \right) |\Gamma\left(((1 + \beta)/2) \times \beta \times 2^{(\beta-1/2)} \right) \right)^{1/\beta} \tag{48}$$

Random Walk with Levy-Flight is used to improve the proposed algorithm's Global Search Ability. Using this method, the local minimum process may be reduced in order to provide a good outcome, particularly for multi-modal benchmark functions and uni-model functions. Table 2 shows the pseudo code of

Table 2 Pseudo-Code of the LOA-MFO

Input: Max- num of iterations, hunting, moving towards, Roaming, Défense

Output: *Objective function's -Accuracy*

      //Initialization//

           1. *Initialize the random collection of lions* $N_{pop}(N_{pop}$ *is the number of initial population)*

           2. *Initiate nomad and pride lions*

      // Hunting//

Randomly split hunters into three groups

Create a target for predators to hunt

For i = 1: H

According to its relevant group move ith hunter towards prey

If the ith hunter's new location is better using Eqs. (37-39)

Prey escapes from hunter

end

End

  //Moving Toward Safe Place//

For i = 1 to P

    for ith pride determine the size of tournament

For j = 1 to R

     Tournament selection determines pride's region using eqs (40)

     Move the jth female to the specified location

End

end

 // Roaming//

For i=1 to Rm

    Choose % R of region to be visited by an ith male

For j=1 to S

    Towards the direction of jth

  If a males's new spot is superior, it is the best visited location.

     Assigning location (best-visited location is updated)

 End

 End

    Best-visited position is selected by the ith male as its current location using Eqs. (41)

End

For i=1 to NN (No. of nomadic lions as NN)

Randomly shift ith nomad

If the new place of the ith nomad is superior to its best-visited position using Eqs.(42)

Update ith nomad's best-visited position

 End

 End

//Défense//

For i=1 to NN

    BT[1,P] = Generate binary template ([1× P]) and allot Binary (0-1) to every cell

    (prides denoted as P) that has generated randomly

For j=1 to P

  If jth element of BT==1

For z=1 to NR

  If nomad male of i is superior to resident male of z in the pride j

Resident male 'z' in pride 'j' is driven out of the pride, that becomes a nomad

    And nomadic male 'i' becomes resident

 Go next i

 End

 End

 End

End

   // Updating//

   1. *Update the population using moth- flame optimization (MFO) using eqs. (45-48)*

   2. *Returning the best solution;*

Return the best solution from pop;

Fig. 3 Hybridization of LOA and MFO
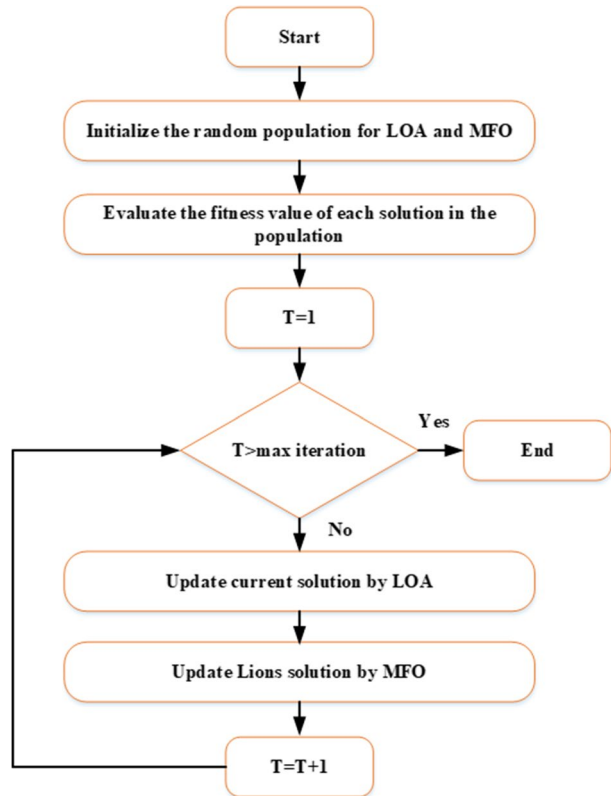
hybridization of ALO and CSO and Fig. 3 displays the flowchart of the hybrid algorithm.

3.5 Classification stage

The selected features are fed into the classification process after the feature selection stage. LRCN is utilized in this study to classify the presented image into four categories: tumor, stone, cyst and normal. A Spatio-temporal model called LRCN is constructed by fusing CNN and LSTM. Similar to behavior recognition, we created an 18-layer LRCN network with 101-time steps to recognize epileptic events. The CNN component of the LRCN network is responsible for extracting spatial features from the input image. CNN layers are typically used for image recognition tasks as they can learn and identify relevant image features by sliding a set of filters (kernels) over the input image. In this study, we used 15 CNN layers, each with varying number of filters and sizes, to extract features from the input image. The output of the CNN layers is a set of feature maps, where each map corresponds to a particular filter and captures the presence of a specific feature in the image. The LSTM component of the LRCN network is responsible for capturing temporal dependencies between the feature maps extracted by the CNN layers. LSTM layers are a type of Recurrent Neural Network (RNN) that are specifically designed to model sequences of data by learning long-term dependencies. In this study, we used 17 LSTM layers that are

connected to the CNN layers. The output of each LSTM layer is a hidden state that summarizes the temporal information encoded in the corresponding sequence of feature maps. The categorization is done in the last layer, a fully linked layer that links the previous time step to an LSTM. For convolution and max-pooling, the stride is set at 1 or 2, respectively. The illustration of proposed model is shown in Fig. 4.

Each time step, the input to the network is a 227×227 picture, which differs from the VGG input size of 224×224. The fusion of 32 kernels, each with a stride of 2 and a size of 3×3, is applied to Layer 1. Next, a convolution layer with 32 kernels of size 3×3 and a stride of 1 is used. Each feature map is then subjected to a max-pooling of size 2 Layer 2). Through these two actions, the output shape's size is lowered from $100 \times 227 \times 227$ to $100 \times 56 \times 56$. In layer 16, after the final max-pooling layer, the feature maps are flattened to 4680 neurons by 0.5 drop-outs. Layer 17 contains their connection to LSTM with 256 units. The final network layer utilizes the last time step of 100 time steps to access a four-classifier softmax layer (normal, cyst, stone and tumor classes). The structure of the LRCN is illustrated in Fig. 5. Equations (49)–(53) illustrate the mathematical formulas.

$$P(x) = f(x) * P(x-1) + i(x) * \tanh\left(W_c * [\boldsymbol{conv}(A(x)), B(x-1)] + b_c\right) \tag{49}$$

$$f(x) = \sigma\left(W_f * [\boldsymbol{conv}(A(x)), B(x-1), P(x-1)] + b_f\right) \tag{50}$$

$$o(x) = \sigma\left(W_o * [\boldsymbol{conv}(A(x)), B(x-1), P(x)] + b_o\right) \tag{51}$$

$$i(x) = \sigma\left(W_i * [\boldsymbol{conv}(A(x)), B(x-1), P(x-1)] + b_i\right) \tag{52}$$

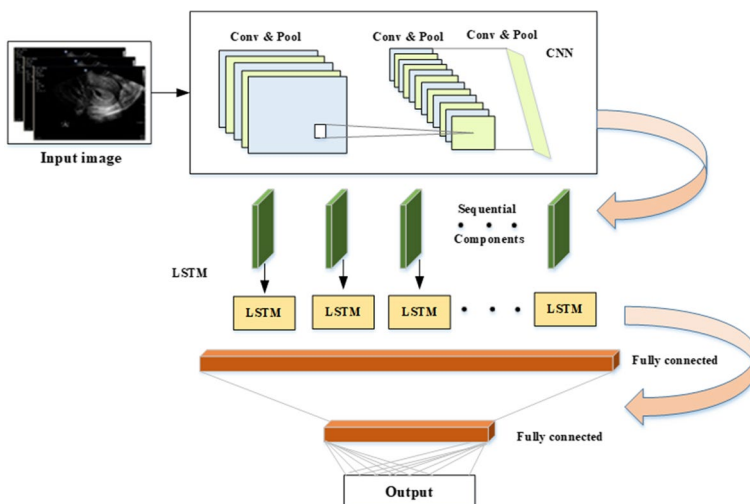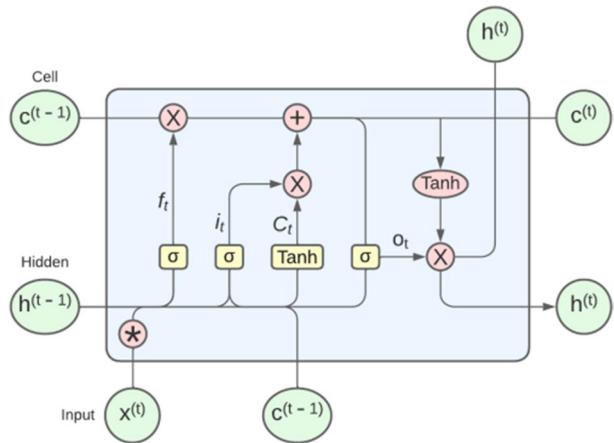$$B(x) = o(x) * \tanh(P(x)) \tag{53}$$

Fig. 4 Illustration of proposed architecture

Fig. 5 Architecture of LRCN

where,

Convolution operator as conv.; Element – wise product as *; Sigmoid function as σ; Vector of bias as b; Matrix of weights as W; Output LRCN cell as H(x); Imput LRCN cell as C(x); Forget gate as f(x); Output gate as o(x); Input gate as i(x).

For each class, a unique threshold value TH is allocated. On the basis of the output value, the class of the provided picture is determined. The images satisfy the condition which is present in the Eq. (54).

$$Output \in \begin{cases} N_1 \, for \, 0 \le TH < 0.25 \\ N_2 \, for \, 0.25 \le TH < 0.5 \\ N_3 \, for \, 0.5 \le TH < 0.75 \\ N_4 \, for \, 0.75 \le TH \end{cases} \tag{54}$$

Here, Normal image is denoted as N_1, Cyst image is denoted as N_2, Stone image is represented as N_3 and tumor image is represented as N_4 .

4 Results and discussion

In this part, the developed system setup, including the configuration and execution of the proposed models and assessment criteria, is first discussed. The outcomes are then explored in depth and debated. In addition, in order to provide a benchmark for evaluating the proposed LRCN methodology, this study employs cutting-edge methodologies. Classifiers are trained utilizing both the whole feature set and a subset of features to show the effect of feature selection. To assess the efficacy of the different models, the 10-fold cross-validation technique is employed. This method included randomly dividing the dataset into 10 folds, testing the classifier on each fold once, and then training the classifier using the remaining (10–1) folds. Each fold represented one iteration, and the mean is used to calculate the overall accuracy after computing the model's

accuracy in each iteration; this approach is the most effective for evaluating the model's performance based on several metrics.

4.1 Experimental setup

This section covers the outcomes of the proposed segmentation and classification of CKD. Various settings have been used in the system's development. The environment configuration for the system design includes a Core i5 Gen 6 CPU, 8 GB of RAM, and a 4 GB GPU. The system design will be implemented using Python software.

4.2 Database description

Ultrasound images of the kidney are obtained from databases [5, 12] and radiologists. The clinical images are gathered from August to December 2019 at the Radiology department of the Chettinad Health City Hospital in Chennai, India. The dataset comprises a total of 4940 images, with 4420 images allocated to the training set and 520 images to the testing set. Both the training and testing datasets consist of four categories (normal, cyst, stone, and tumor), with each category in the training set containing 1110 images and each category in the testing set containing 125 images.

4.3 Training phase

The network is trained utilizing an LRCN network and a learning rate of 0.0001. The parameters of the proposed networks are updated in a minibatch mode with a batch size of 32. Dropout is used before the softmax layer, with a retain probability of 0.2. After 100 epochs, the training is terminated, and the parameters at each epoch are saved. Finally, the model with the best validation set performance is assessed on the test set. The parameters are applied throughout all experiments on the basis of the model presented by Zhang et al. [44]. Figure 6a illustrates the accuracy versus epoch curve. This is clear evidence that no over-fitting has occurred, since the training and accuracy curves are quite comparable. The learning rate begins at 0.0001 and the batch size is 100 epochs. Figure 6b depicts the loss curve, which indicates a minimal amount of lost value.

4.4 Evaluation metrics

The training procedure is done on the training set, while the validation and testing sets are used to assess performance. There are two sorts of experiments: categorization and detection. Five assessment criteria are utilized for categorization tasks: Accuracy, F1-score, precision, specificity, recall, PPV, NPV, False discovery rate (FDR) and False Positive Rate (FPR) were used to assess the model's performance. The predicted quantities of true positive (Tp), false positive (Fp), true negative (Tn), and false negative (Fn) samples are denoted by these terms. This research employed the ROC curve and AUC to
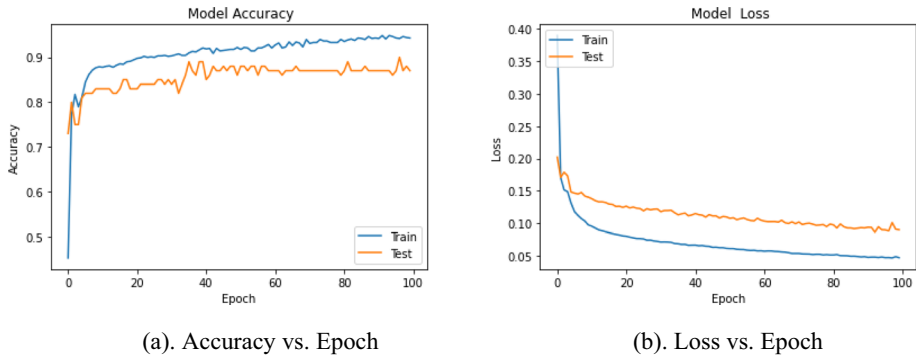
(a). Accuracy vs. Epoch   (b). Loss vs. Epoch

Fig. 6 a Accuracy vs. Epoch b Loss vs. Epoch

measure classifier performance. At varying thresholds, the ROC curve illustrates the relationship between the Tp and Fp rates. It demonstrates the classifier's capacity to differentiate between the four classes. The metrics are mathematically expressed as,

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN} \tag{55}$$

$$\text{Re}\,call = \frac{TP}{TP + FN} \tag{56}$$

$$\Pr ecision = \frac{TP}{TP + FP} \tag{57}$$

$$Specificity = \frac{TN}{FP + TN} \tag{58}$$

$$F1 - Score = 2 \times \frac{\Pr ecision \times \text{Re}\,call}{\Pr ecision + \text{Re}\,call} \tag{59}$$

$$NPV = \frac{TN}{TN + FN} \tag{60}$$

$$FDR = \frac{FP}{TP + FP} \tag{61}$$

$$FPR = \frac{FP}{FP + TN} \tag{62}$$

$$PPV = \frac{TP}{TP + FP} \tag{63}$$

Also, we are leveraging image quality metrics such as SSIM, PSNR, edge preservation index (EPI), MSE, and RMSE to quantitatively assess the efficacy of our image denoising algorithms. The metrics are mathematically expressed as,

$$PSNR = 10\log_{10}\frac{255^2}{MSE} \tag{64}$$

$$SSIM(X, Y) = \left[l(X, Y)^\alpha . c\left(X, Y\right)^\beta . s\left(X, Y\right)^\gamma \right] \tag{65}$$

$$PSNR = 10\log_{10}\frac{255^2}{MSE} \tag{66}$$

$$EPI = \frac{\left(\sum_{m=1}^{M} \sum_{n=1}^{N-1} |X'(m, n+1) - X'(m, n)| \right)}{\left(\sum_{p=1}^{M} \sum_{q=1}^{N-1} |X(m, n+1) - X(m, n)| \right)} \tag{67}$$

$$MSE = \frac{1}{PQ} \sum_{x=1}^{P} \sum_{y=1}^{Q} \left(f'(x, y) - f\left(x, y\right) \right)^2 \tag{68}$$

$$RMSE = \sqrt{MSE} \tag{69}$$

4.5 Result analysis and extensive discussion

The original ultrasound image are processed using image filtering and Histogram equalization. Histogram equalization may enhance the quality of a picture by creating a homogeneous distribution of light and dark. Next, Renal segmentation is a crucial step in several other processing pipelines, such as enhancing the precision of cortical medullary segmentations or lowering computation times by only fitting quantitative maps to voxels inside the kidney. Utilizing the GLCM technique, feature parameters are retrieved after Segmentation. The neural network is developed with 1000 hidden units, and its output is intended to be categorized into four kinds of CKD.

On the ultrasonic images, Fig. 7 presents a complete classification of CKD phases. It is obvious from this diagram that the first row represents the image input and the second row shows the images acquired from the pre-processing step. The third and fourth lines show the segmented and classified images, respectively. As seen in Fig. 7, the first column indicates that the input image is normal, while the categorized image also indicates that the result of the classification is normal. It means that the presented model classifies US images effectively. Likewise, the second, third, and fourth column represent the various phases (Cyst, Stone, and tumor) of CKD. On each of the applied US photos, the results demonstrated that the given method successfully classifies the US into distinct phases.
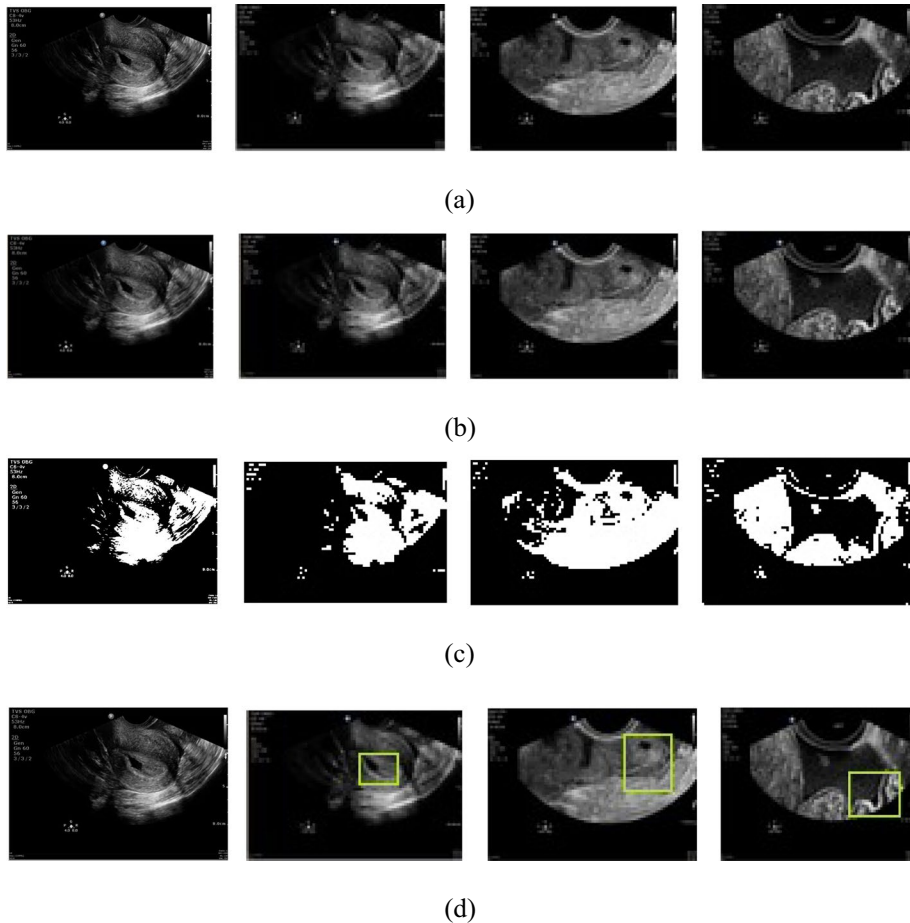
(a)

(b)

(c)

(d)

Fig. 7 Illustration of CKD classification as Normal, Cyst, Stone and Tumor (**a**) Input US kidney images (**b**) Filtered US kidney images (**c**) Segmentation of US kidney images (**d**) Prediction result for CKD stages: Normal, cyst, stone and Tumor

4.6 Analysis of preprocessing results

Bicubic-Interpolation (BI) is a commonly used image filtering technique that aims to improve the quality of an image by reducing its noise and blurring. This technique works by interpolating the pixel values of an image based on its neighboring pixels. To examine the efficacy of our technique, several metrics are utilized, including PSNR, SSIM, EPI, MSE, and RMSE. Several state-of-the-art techniques, such as DeepRLN, LBP-DTD, GTDF, LMMSE, CS-WT, Bilateral, and NLM, can be evaluated to compare the effectiveness of the developed preprocessing techniques with existing methods.

Figure 8 represents the performance of proposed methods regarding PSNR. From the Fig. 8, the proposed technique has the highest PSNR value of 55 dB, indicating that the processed images using this technique have the highest quality among all the techniques compared. LBP-DTD has the lowest PSNR value of 24 dB, indicating that the quality of the processed images using this technique is comparatively lower than

Fig. 8 Performance of proposed and existing methods regarding PSNR
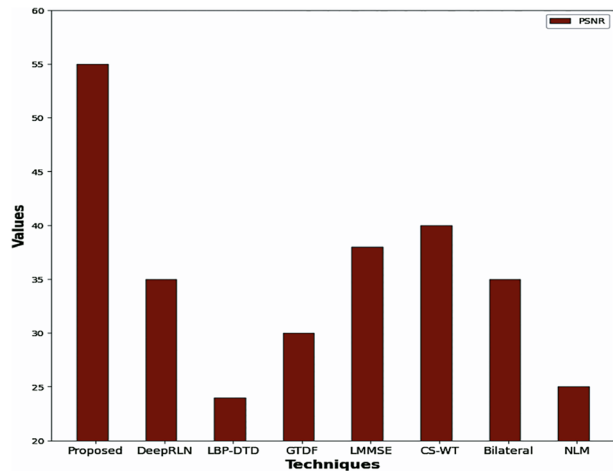
other techniques. The PSNR values for GTDF, LMMSE, CS-WT, Bilateral, and NLM are 30 dB, 38 dB, 40 dB, 35 dB, and 25 dB, respectively. This suggests that these techniques perform better than LBP-DTD but are not as effective as the proposed technique in preserving the information in the original image during the processing stage.

Figure 9 represents the performance of proposed methods regarding SSIM. From Fig. 9, the proposed technique has the highest SSIM value of 0.9751, indicating that the processed images using this technique have the highest level of similarity to the original images. GTDF has the second highest SSIM value of 0.8847, indicating that it is also effective in preserving the structural information of the original image. CS-WT has an SSIM value of 0.8654, indicating that it is better at preserving the original image's structural information than some of the other techniques, but not as effective as the proposed technique. LMMSE and Bilateral have SSIM values of 0.8145 and 0.7958 respectively, indicating that they may result in some loss of structural information during the processing stage. NLM has an SSIM value of 0.8624, indicating that it is relatively effective at preserving the original image's structural information compared to LBP-DTD and DeepRLN, which have the

Fig. 9 Performance of proposed and existing methods regarding SSIM
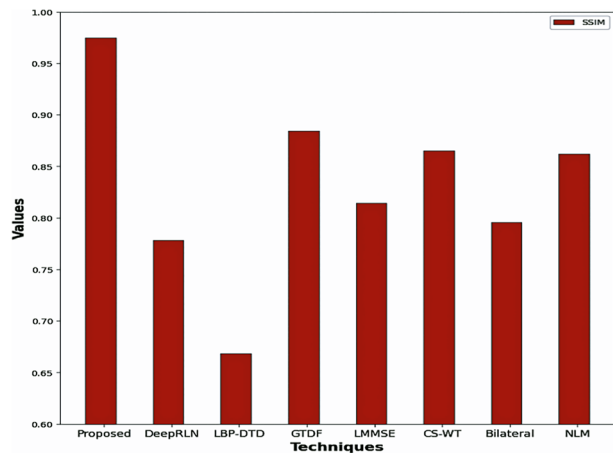
lowest SSIM values of 0.6685 and 0.7784 respectively, indicating that they result in significant loss of structural information during the processing stage.

Figure 10 represents the performance of utilized methods regarding EPI. From the Fig. 10, the proposed technique (0.9852) is significantly higher than all other techniques, indicating that the processed images using this technique have a higher level of edge preservation compared to the original images. Among the compared techniques, Bilateral and NLM have higher EPI values of 0.8865 and 0.8765, respectively, compared to other techniques. The EPI value of CS-WT is the lowest among all techniques at 0.8245, indicating that this technique may result in a significant loss of edge information during the processing stage. Overall, when it comes to edge preservation, the proposed technique displays superior performance in comparison to other techniques.

MSE and RMSE values for each technique are shown in the Figs. 11 and 12. The lower the values for both MSE and RMSE, the better the technique is at preserving the original image quality. According to the results, the proposed technique achieved the lowest values for both MSE and RMSE, which were 0.005 and 0.070710678, respectively. This suggests that the technique is highly successful in maintaining the quality of the original image. The DeepRLN technique has the highest MSE and RMSE values of 0.015 and 0.122474487, respectively, indicating that it results in the greatest loss of image quality during processing. The other techniques have MSE and RMSE values ranging from 0.023 to 0.06 and from 0.151657509 to 0.244948974, respectively, indicating varying levels of effectiveness in preserving the original image quality.

Overall, the proposed technique showed significant improvements over other techniques in terms of PSNR, SSIM, and EPI. It can be concluded from the outcomes that the proposed model successfully preserves the structural information of the original image, while also reducing noise and enhancing contrast to a significant extent. The MSE and RMSE values further support the effectiveness of the proposed technique.

4.6.1 Impact of Despeckling module on state-of-the-art methods for Noisy US images

Deviationsin the quality of training and testing images can lead to a degradation in the performance of a network. This is particularly true for existing methods that are trained on high-quality images and are therefore susceptible to noise when presented with noisy

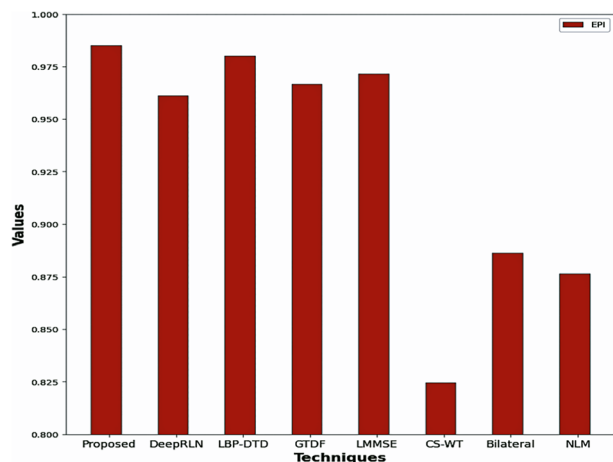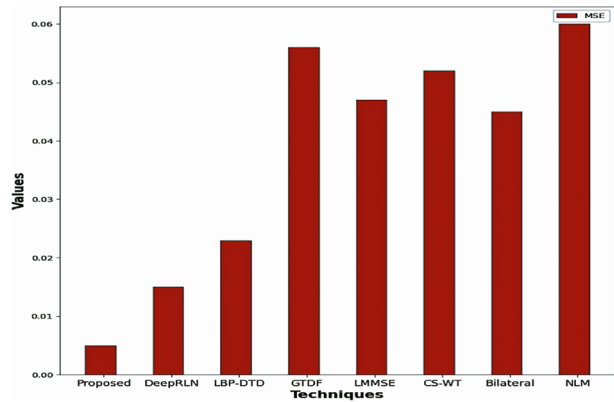**Fig. 10** Performance of proposed and existing methods regarding EPI

Fig. 11 Performance of proposed and existing methods regarding MSE

images. We tackled this issue by integrating a despeckling module into the current state-of-the-art methods and confirmed its efficacy in enhancing the quality of noisy US images. By utilizing the proposed technique to include a despeckling module, the state-of-the-art methods were able to achieve improved classification accuracy across various levels of speckle noise, as shown in Table 3. The significance of the despeckling module in improving the classification performance of existing methods is evident from these findings.

Figure 13 shows the classification accuracy of different image denoising techniques at different levels of speckle noise, denoted by σ^2. The proposed technique has the highest accuracy among all the methods, with an accuracy of 98.7% at $\sigma^2=0.02$ and decreasing to 85.65% at $\sigma^2=0.1$. DeepRLN, LBP-DTD, GTDF, LMMSE, CS-WT, Bilateral, and NLM are other techniques evaluated in the study. At lower noise levels, all techniques achieve relatively high accuracy, but their performance degrades as the noise level increases. NLM has the lowest accuracy among all the techniques, indicating that it is not a suitable technique for denoising speckle noise.

4.7 Confusion matrix result

The generation of the confusion matrix is an essential step in determining classification performance. Figure 14 depicts the confusion matrix produced from the applied test

Fig. 12 Performance of proposed and existing methods regarding RMSE
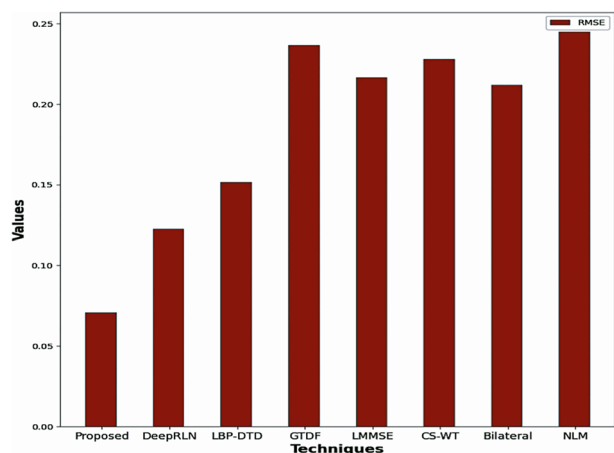
Table 3 Performance Comparison of Proposed System and State-of-the-Art Methods with various Speckle Noise Levels

Techniques	Classification accuracy (%)				
	$\sigma^2 = 0.02$	$\sigma^2 = 0.04$	$\sigma^2 = 0.06$	$\sigma^2 = 0.08$	$\sigma^2 = 0.1$
Proposed	98.7	95.2	91.45	88.12	85.65
DeepRLN [39]	87.31	85.77	77.5	73.65	70.19
LBP-DTD [3]	85.4	81.45	77.54	74.65	69.25
GTDF [9]	85.12	81.65	77.95	75.45	70.15
LMMSE [32]	86.87	84.65	79.65	73.56	98.35
CS-WT [41]	84.56	81.95	76.85	71.12	68.35
Bilateral [45]	85.65	82.45	76.95	71.45	63.42
NLM [20]	78.65	69.77	65.65	61.85	57.31

images. Considered here are the data acquired from various individuals for identifying the four kinds of CKD. It is evident from Fig. 14 that the proposed model distinguishes the diagonally displayed phases of CKD with increased precision. The remaining figures indicate the number of incorrect predictions made for each step.

4.8 Classifier performance without and with feature selection

The results in Table 4 indicate the accuracy considering all and chosen attributes for each train-test split. In addition, the reduced set of features are utilized for training the proposed classifiers in order to illustrate the efficacy of feature selection. The extensive experiments shown in Table 4 indicate that the proposed model achieved its highest level of accuracy by training the classifier with selected attributes, which outperformed training the classifier with all available data. The proposed hybrid feature selection algorithm is thus an excellent approach for predicting CKD. Also, the execution time of classifier on full feature set is 1.25 seconds and on best feature set is 0.02 seconds.

Figure 15 shows a comparison of the performance of three different models, namely the LOA-MF, LOA, and MF, with respect to different features (F1-F19) under various testing conditions. The values in the figure represent the percentage of accuracy achieved by each algorithm for each feature. For instance, the proposed
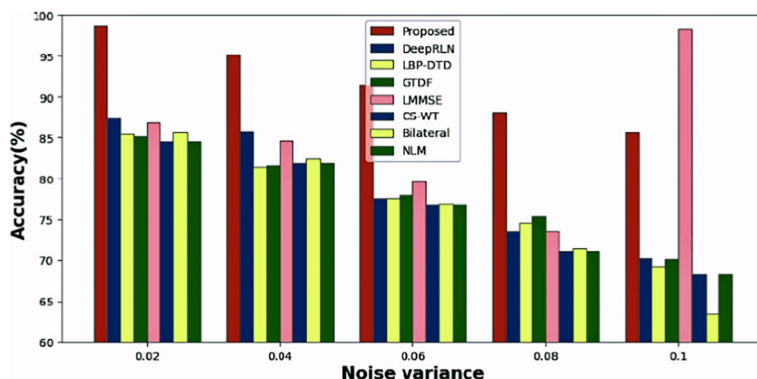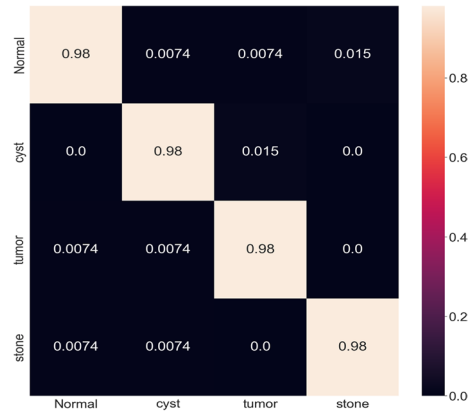

Fig. 13 Performance comparison of different classifiers

Fig. 14 Confusion matrix for each class

algorithm achieved an accuracy of 97% for Feature 1 (F1), while LOA and MF achieved accuracies of 92.35% and 78.14%, respectively. Similarly, for Feature 19 (F19), the proposed algorithm demonstrated a precision of 98.7%, surpassing the accuracy rates of LOA and MF, which were 89.8% and 86.35%, respectively. Overall, the proposed algorithm outperformed the LOA and MF systems for most of the features, with higher accuracy rates.

4.9 Analysis of the proposed classifier model

The performance parameter evaluations are derived using ultrasound images of the kidney during training and assessment. Table 5 provide the outcomes of the proposed methodology. The model has an accuracy of 98.7%, indicating that it is making correct predictions for a large majority of instances. The high precision of 97.6% suggests that the model is accurately predicting positive instances. The recall of 96.4% indicates that the model is correctly identifying positive instances. The F1-Score of 97.9% provides a balanced measure of precision and recall. The specificity of the model is 96.2%, indicating that it is correctly identifying negative instances. The positive predictive value (PPV) of the model is 97.6%, implying that when the model predicts a positive instance, it is correct 97.6% of the time. The negative predictive value (NPV) of the model is 92.5%, indicating that when the model predicts a negative instance, it is correct 92.5% of the time. The false positive rate (FPR) of the model is 0.26%, suggesting that the model is making very few false positive predictions. The false discovery rate (FDR) of the model is

Table 4 Accuracy comparison among Feature Selection Methods and Without Feature Selection

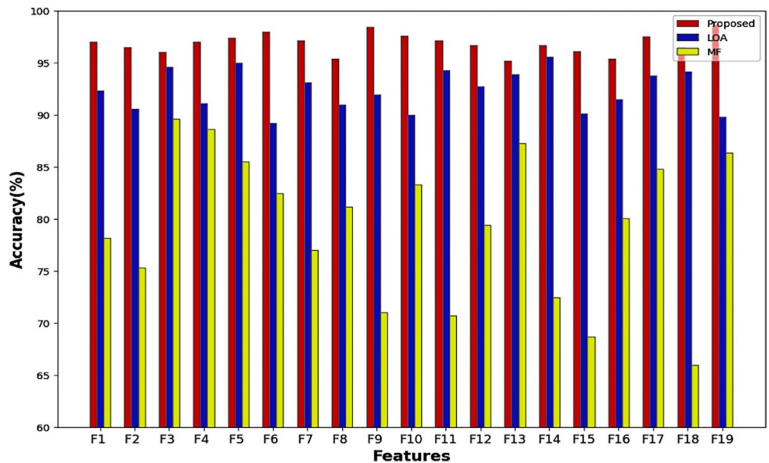| Training - Testing partition	Accuracy with all features (%)	Accuracy with selected features (%)
50–50	92.5	96
65–35	95.2	97
75–25	94	96.5
Tenfold cross validation	96.5	98.7
Execution time (s)	1.25	0.02

Fig. 15 Performance Comparison on Multiple Features under Various Testing Conditions

0.024, indicating that only 2.4% of positive predictions are incorrect. Overall, these performance metrics strongly reveals that the model is performing well and making highly accurate predictions on the dataset.

4.9.1 Comparison results

This section discusses the efficacy of ultrasound image categorization using LRCN. Currently, the features are optimally selected using the LOA-MF Optimization method. The effectiveness of the proposed CKD classification approach is evaluated through a literature review, in comparison to state-of-the-art methodologies.

Table 6 presents a performance comparison between the proposed technique and existing methods in terms of accuracy, F1 Score, recall, precision, and specificity for detecting chronic kidney disease (CKD). Results reveal that the proposed model achieves a high accuracy rate of 98.7%, outperforming existing models [1, 21, 22, 33, 36, 38, 39, 42], which show decreased accuracy rates due to system unreliability and complexity. The high accuracy rate of the proposed model demonstrates superior classification performance. Precision, which is crucial for measuring the effectiveness

Table 5 Analysis of certain performance criteria	Performance Metrics	Percentage
	Accuracy	98.7%
	Precision	97.6%
	Recall	96.4%
	F1-Score	97.9%
	Specificity	96.2%
	PPV	97.6%
	NPV	92.5%
	FPR	0.26
	FDR	0.024

of results, achieved 96.6% accuracy with the proposed technique and improved false detection rate reduction. The recall metric achieved 96.4% with the proposed strategy compared to existing learning algorithms, which achieved 91%, 80%, 96%, 100%, 96%, and 87% for [22, 26, 33, 38, 39, 42], respectively. While the proposed method yielded superior results in most cases, the existing method [21] showed better performance. The F1-Score in Table 6 also supports the proposed method's better ability to classify CKD than existing techniques, achieving a value of 97.9%. By selecting optimal features, the proposed method enhanced the classification of various phases by the F1 measure. The overall specificity performance of the proposed method was 96.2%, indicating high specificity in classification performance. Conversely, existing methods like [21, 26, 33] achieved 93%, 90%, and 96% performances, respectively. Therefore, the proposed method provides improved outcome predictions based on input attributes.

4.9.2　Receiver operating characteristic (ROC)

The ROC curve illustrates the result graphically, but the AUC demonstrates quantitatively the measure. A finite number of points are extracted from the ROC curve image. The receiver operating characteristic (ROC) curves were plotted for various methods, including the proposed method, Sudharson et al., Liao et al., Kim et al., Ahmad et al., Nithya et al., Wu et al., and Raju et al. The area under the ROC curve (AUC) is calculated for each method, with the following results: the proposed method achieved an AUC of 0.99, Sudharson et al. achieved 0.95, Liao et al. achieved 0.87, Kim et al. achieved 0.86, Ahmad et al. achieved 0.81, Nithya et al. achieved 0.79, Wu et al. achieved 0.75, and Raju et al. achieved 0.71. These results indicate that the proposed method outperformed the other methods in terms of classification accuracy. The ROC/AUC curve of our technique is displayed in Fig. 16. It is evident from the preceding tables and figures that the proposed model is effective in detecting CKD. The LOA-MFO-based feature used in the LRCN model contributes to the higher classification rate achieved. Accuracy is 96.5% without feature selection and 98.7% with feature selection for tenfold cross validation. The LRCN classifier achieves a maximum accuracy of 98.7% and can detect the presence of CKD accurately and at a greater speed. The results of preprocessing, segmentation, and classification show that the proposed technique is highly promising in classifying CKD in US images.

Table 6　Comparing the proposed model to existing models in the literature review

Techniques	Accuracy	F1 Score	Recall	Precision	Specificity
Proposed	0.98	0.98	0.96	0.97	0.96
Sudharson et al. [39]	0.98	–	0.91	–	–
Liao et al. [22]	0.81	–	0.8	–	0.93
Kim et al. [21]	0.95	–	–	–	–
Ahmad et al. [1]	0.97	–	–	–	–
Sudharson et al. [38]	0.95	–	0.96	–	–
Nithya et al. [26]	0.93	–	1	–	0.9
Wu et al. [42]	0.95	–	0.96	–	–
Raju et al. [33]	0.95	–	0.87	–	0.96

Fig. 16 Receiver Operating Characteristic Curve
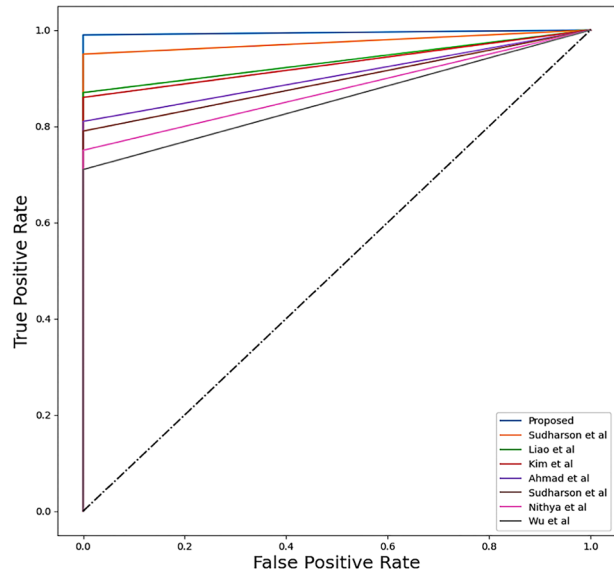
5 Conclusion and future works

This article provides a unique deep neural network architecture for the categorization of chronic kidney disease (CKD) using LRCN. Eventhough LRCN model that combines CNN-LSTM are able to extract features from raw input data, some features may still be irrelevant or redundant. Therefore, a hybrid of Lion Optimization algorithm and Moth flame optimization is proposed that selects the relevant features to effectively optimize the model's performance. Utilizing all available features may result in overfitting issues. The novel hybrid algorithm is based on dimensionality reduction, leading to better model performance. The key contributions of this study are as follows: Firstly, an advanced image processing technique is developed to enhance the diagnostic features of CKD in US images. Secondly, a segmentation technique is proposed that accurately segments all relevant structures in ultrasound images. Thirdly, novel feature extraction techniques are developed to capture the unique diagnostic features of CKD in US images. Additionally, a feature selection approach is introduced to minimize large data dimensionality. Finally, an efficient classification technique is proposed with the potential to speed up computing. On ultrasound images, the effectiveness of the proposed study is measured using Recall, precision specificity, accuracy, and F1-score. The experimental results suggest that the proposed framework is better to existing classifiers in terms of high classification accuracy of 98.7%. In the future, the proposed method may be updated to achieve effective diagnostic outcomes. In addition, they may be utilized to evaluate additional high-dimensional chronic illness datasets, and a method with a 100% classification rate can be designed.

Declarations

References

1. Ahmad R, Mohanty BK (2021) Chronic kidney disease stage identification using texture analysis of ultrasound images. Biomed Signal Process Control 69:102695
2. Ahmed, TI, Bhola, J, Shabaz, M, Singla, J, Rakhra, M, More, S, Samori, IA (2022) Fuzzy logic-based systems for the diagnosis of chronic kidney disease. BioMed Res Int, 2022
3. Alex DM, Christinal AH, Chandy DA, Singh A, Pushkaran M (2020) Speckle noise suppression in 2D ultrasound kidney images using local pattern based topological derivative. Pattern Recogn Lett 131:49–55
4. Alkurdy NH, Aljobouri HK, Wadi ZK (2023) Ultrasound renal stone diagnosis based on convolutional neural network and vgg16 features. Int J Electr Comput Eng 13(3):3440–3448
5. Antony J, Ultrasound-images.com (2015) https://www.ultrasound-images.com/, Accessed: April, 2020
6. Aprilianto D (2020) SVM optimization with correlation feature selection based binary particle swarm optimization for diagnosis of chronic kidney disease. J Soft Comput Explor 1(1):24–31
7. Chaudhuri AK, Sinha D, Banerjee DK, Das A (2021) A novel enhanced decision tree model for detecting chronic kidney disease. Netw Model Anal Health Inf Bioinf 10:1–22
8. Chittora P, Chaurasia S, Chakrabarti P, Kumawat G, Chakrabarti T, Leonowicz Z, … Bolshev V (2021) Prediction of chronic kidney disease-a machine learning perspective. IEEE Access 9:17312–17334
9. Cui W, Li M, Gong G, Lu K, Sun S, Dong F (2020) Guided trilateral filter and its application to ultrasound image despeckling. Biomed Signal Process Control 55:101625
10. Eddy S, Mariani LH, Kretzler M (2020) Integrated multi-omics approaches to improve classification of chronic kidney disease. Nat Rev Nephrol 16(11):657–668
11. Elhoseny M, Shankar K, Uthayakumar J (2019) Intelligent diagnostic prediction and classification system for chronic kidney disease. Sci Rep 9(1):1–14
12. Geertsma T, Ultrasoundcases.info (2011) http://www.ultrasoundcases.info/, Accessed: April, 2020
13. Han X, Zhang S, Chen Z, Adhikari BK, Zhang Y, Zhang J, … Wang Y (2020) Cardiac biomarkers of heart failure in chronic kidney disease. Clin Chim Acta 510:298–310
14. Harimoorthy K, Thangavelu M (2021) Multi-disease prediction model using improved SVM-radial bias technique in healthcare monitoring system. J Ambient Intell Humaniz Comput 12(3):3715–3723
15. Hosseinzadeh M, Koohpayehzadeh J, Bali AO, Asghari P, Souri A, Mazaherinezhad A, … Rawassizadeh R (2021) A diagnostic prediction model for chronic kidney disease in internet of things platform. Multimed Tools Appl 80:16933–16950
16. Huang C, Li X, Wen Y (2021) AN OTSU image segmentation based on fruitfly optimization algorithm. Alexandria Eng J 60(1):183–188
17. Jain D, Singh V (2020) A novel hybrid approach for chronic disease classification. Int J Healthcare Inf Syst Inf (IJHISI) 15(1):1–19
18. Jerlin Rubini L, Perumal E (2020) Efficient classification of chronic kidney disease by using multi-kernel support vector machine and fruit fly optimization algorithm. Int J Imaging Syst Technol 30(3):660–673
19. Khamparia A, Saini G, Pandey B, Tiwari S, Gupta D, Khanna A (2020) KDSAE: chronic kidney disease classification with multimedia data learning using deep stacked autoencoder network. Multimed Tools Appl 79(47):35425–35440
20. Khare S, Kaushik P (2021) Speckle filtering of ultrasonic images using weighted nuclear norm minimization in wavelet domain. Biomed Signal Process Control 70:102997
21. Kim DH, Ye SY (2021) Classification of chronic kidney disease in sonography using the GLCM and artificial neural network. Diagnostics 11(5):864
22. Liao YT, Lee CH, Chen KS, Chen CP, Pai TW (2021) Data augmentation based on generative adversarial networks to improve stage classification of chronic kidney disease. Appl Sci 12(1):352
23. Ma F, Sun T, Liu L, Jing H (2020) Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network. Futur Gener Comput Syst 111:17–26

24. Mirjalili S (2015) Moth-flame optimization algorithm: a novel nature-inspired heuristic paradigm. Knowl-Based Syst 89:228–249
25. Navaneeth B, Suchetha M (2020) A dynamic pooling based convolutional neural network approach to detect chronic kidney disease. Biomed Signal Process Control 62:102068
26. Nithya A, Appathurai A, Venkatadri N, Ramji DR, Palagan CA (2020) Kidney disease detection and segmentation using artificial neural network and multi-kernel k-means clustering for ultrasound images. Measurement 149:106952
27. Pal, S (2022) Chronic kidney disease prediction using machine learning techniques. Biomed Mater Device, 1–7
28. Parthiban R, Usharani S, Saravanan D, Jayakumar D, Palani DU, StalinDavid DD, Raghuraman D (2021) Prognosis of chronic kidney disease (CKD) using hybrid filter wrapper embedded feature selection method. Eur J Mol Clin Med 7(9):2511–2530
29. Pradeepa, P, Jeyakumar, MK (2022) Modelling of IDBN with LSNN based optimal feature selection for the prediction of CKD using real time data. Multimed Tools Appl, 1–36
30. Puttagunta M, Ravi S (2021) Medical image analysis based on deep learning approach. Multimed Tools Appl 80:24365–24398
31. Qin J, Chen L, Liu Y, Liu C, Feng C, Chen B (2019) A machine learning methodology for diagnosing chronic kidney disease. IEEE Access 8:20991–21002
32. Rahimizadeh N, Hasanzadeh RP, Janabi-Sharifi F (2021) An optimized non-local LMMSE approach for speckle noise reduction of medical ultrasound images. Multimed Tools Appl 80:9231–9253
33. Raju P, Rao VM, Rao BP (2019) Optimal GLCM combined FCM segmentation algorithm for detection of kidney cysts and tumor. Multimed Tools Appl 78(13):18419–18441
34. Schena FP, Anelli VW, Abbrescia DI, Di Noia T (2022) Prediction of chronic kidney disease and its progression by artificial intelligence algorithms. J Nephrol 35(8):1953–1971
35. Senan EM, Al-Adhaileh MH, Alsaade FW, Aldhyani TH, Alqarni AA, Alsharif N, … Alzahrani MY (2021) Diagnosis of chronic kidney disease using effective classification algorithms and recur- Shobana, S., Rajaram, A., SteffenRaj, T., & MoonSon, A. L. (2022). Detection Of Kidney Stones Using Machine Learning. Tierarztliche Praxis,42(4).ealthcare Eng 2021
36. Shobana S, Rajaram A, SteffenRaj T, MoonSon AL (2022) Detection Of kidney stones using machine learning. Tierarztliche Praxis, 42(4)
37. Singh V, Asari VK, Rajasekaran R (2022) A deep neural network for early detection and prediction of chronic kidney disease. Diagnostics 12(1):116
38. Sudharson S, Kokil P (2020) An ensemble of deep neural networks for kidney ultrasound image classification. Comput Methods Prog Biomed 197:105709
39. Sudharson S, Kokil P (2021) Computer-aided diagnosis system for the classification of multi-class kidney abnormalities in the noisy ultrasound images. Comput Methods Prog Biomed 205:106071
40. Viswanath K, Anilkumar B, Gunasundari R (2022) Design of deep learning reaction–diffusion level set segmentation approach for health care related to automatic kidney stone detection analysis. Multimed Tools Appl 81(29):41807–41849
41. WangNo, N, Chiewchanwattana, S, Sunat, K (2020) An efficient adaptive thresholding function optimized by a cuckoo search algorithm for a despeckling filter of medical ultrasound images. J Ambient Intell Humaniz Comput, 1–26
42. Wu Y, Yi Z (2020) Automated detection of kidney abnormalities using multi-feature fusion convolutional neural networks. Knowl-Based Syst 200:105873
43. Yazdani M, Jolai F (2016) Lion optimization algorithm (LOA): a nature-inspired metaheuristic algorithm. J Comput Des Eng 3(1):24–36
44. Zhang X, Yu Y, Gao Y, Chen X, Li W (2020) Research on singing voice detection based on a long-term recurrent convolutional network with vocal separation and temporal smoothing. Electronics 9:1458 [CrossRef]
45. Zheng Q, Furth SL, Tasian GE, Fan Y (2019) Computer-aided diagnosis of congenital abnormalities of the kidney and urinary tract in children based on ultrasound imaging data by integrating texture image features and deep transfer learning image features. J Pediatr Urol 15(1):75–e1