scientific reports

Check for updates

OPEN

An optimized bidirectional recurrent neural network for kidney stone detection based on developed bald eagle search method in CT scan images

Zhenyou Tang[1], Zhenyu Tang[2], Yuxin Liu[3], Long Chen[4], Zhong Tang[5], Xiaohong Liu[6]✉ & Togrol Salami[7,8]✉

Kidney stone disease is a common syndrome and a recurring one, where it bears a 50% chance of being manifested again within ten years and may lead to serious complications like ureteral obstruction and unbearable pain. If timely intervention is considered of paramount importance for a timely intervention, early and accurate detection using computed tomography (CT) scans is also critical to this process. Existing diagnostic systems are being challenged by factors like noise in images, low contrast, and class imbalance, and these might hamper the performance of existing systems. This work focuses on developing an optimized framework of deep learning for the detection of kidney stones in CT images to deal with these drawbacks. The overall proposed approach consists of a preprocessing scheme to normalize the data using Wang-Mendel (WM) de-noising and enhancing contrast globally, followed by data augmentation with the use of SdSmote to overcome an imbalance in the classes. The pre-processed images will be fed into a modified Bidirectional Recurrent Neural Network (BRNN), which will undergo optimization of the weights and biases using a newly implemented Bald Eagle Search (BES) algorithm, with quasi-oppositional learning and chaotic initialization introduced to increase convergence and global search capability. The proposed method is applied to the public CT Kidney Dataset, compared with state-of-the-art techniques like ensemble learning, Exemplar Darknet19, DE/SVM, and Decision Tree solutions. The proposed means attained better performance, showing 96.96% accuracy, 95.62% sensitivity, 91.67% specificity, 94.38% precision, 94.99% F1-score, and 91.61% in the Jaccard Index, thereby confirming the effectiveness and robustness of the proposed model in clinical decision-making concerning kidney stone diagnosis.

Keywords Kidney stone, detection, Bidirectional recurrent neural network, Developed bald eagle search, CT kidney dataset

Kidney stones are formed when soluble minerals build up in the internal layer of the kidney. Calcium oxalate is commonly found in kidney stones, but other chemicals can also be found in urine deposits. Kidney stones can be tiny and travel through the urethra without the person even realizing it[1]. However, they can be quite painful[2]. The presence of kidney stones in the body can lead to ureteral obstruction, which seems to be the blocking of the urethra's exit from the body[3]. Kidney stones are normally asymptomatic until they pass through the ureter; once they do, the following symptoms occur[4].

Men are more likely to have kidney stones. The majority of kidney stone sufferers are between the ages of 30 and 50[5]. Kidney stones are more likely to occur if you have a genetic history of them. If the person has not

[1]Institute of Collaborative Innovation, University of Macau, Avenida da Universidade, Taipa, Macau 999078, China. [2]Department of Computer Engineering, University of California, Irvine, CA 92697, USA. [3]School of Information and Management, Guangxi Medical University, Guangxi, Nanning 530021, China. [4]Nanning Hospital of Traditional Chinese Medicine, Guangxi, Nanning 530000, China. [5]College of Humanities and Social Sciences, Guangxi Medical University, Guangxi, Nanning 530021, China. [6]Information Center of Guangxi Medical University, Guangxi, Nanning 530021, China. [7]Birjand Branch, Islamic Azad University, Birjand, Iran. [8]College of Technical Engineering, The Islamic University, Najaf, Iraq. ✉email: liuxiaohong@gxmu.edu.cn; togrolsalami@gmail.com

implemented the preventative guidelines, an experience of kidney stones increases the chance of having renal stones and kidney problems in the future[6].

Kidney stones can be caused by certain drugs[7]. Topiramate (Topomax), a drug generally prescribed to treat seizures and migraine headaches, has been associated with a raised danger of kidney stones, according to research. In addition, long-term usage of vitamin D and mineral supplements raises calcium levels in the body, which can contribute to kidney stone development[8–10]. Poor diet or protein with high sodium and low calcium, poor diet, obesity, high blood pressure, and comorbidities such as inflammatory bowel disease and persistent diarrhea that impact calcium absorption are all risk factors for gallstones[11].

Kidney stones are checked with a variety of tests. During a clinical examination, your doctor may diagnose backache due to kidney stones, which is a common indication of kidney stones[12]. A pee test is used to evaluate the existence of blood in the urine. A blood test is utilized to detect kidney stone-related disorders and to confirm the original diagnosis[13].

The existence of crystals in the ureter, bladder, and kidneys, as well as the size and specific location of the stone, blockage, and the status of organs in this zone, such as the appendix, aorta, and pancreas, are all determined by an abdominal CT scan[14].

Ultrasound has great diagnostic capacity and may detect numerous kidney stone issues. To reduce radiation exposure, pregnant women should receive an ultrasound rather than a CT scan. After a diagnosis of kidney stones, conventional radiography is used to trace the stone's transit through the urinary system[15].

Ultrasound is the initial step in diagnosing kidney stones. Many people who have pain in their sides, nausea, and vomiting go to the clinic, where the doctor diagnoses them with crystals in all of them using ultrasonography[16].

The use of ultrasound to diagnose kidney stones is prevalent. Kidney identification is critical for determining the cause of kidney stones. Radiologists often segment ultrasound pictures by hand, which can result in visual mistakes and lengthy procedures[17]. As a result, computer-assisted segmentation of ultrasound pictures appears to be essential and crucial[18]. The main purpose of this study is to resolve some limitations in this direction. The main contributions can be highlighted as follows:

- A novel hybrid framework is proposed by combining a modified Bidirectional Recurrent Neural Network (BRNN) with a developed version of the Bald Eagle Search (BES) algorithm for accurate and robust kidney stone detection on CT images.
- Improvements to the BES algorithm are made by applying quasi-oppositional learning and chaotic initialization (in the form of the logistic map) to improve population diversity, convergence speed, and global search capability that would allow it to avoid premature convergence.
- Robust preprocessing and data augmentation techniques have been incorporated, such as Wang-Mendel (WM) noise reduction, global contrast enhancement, and SdSmote-based augmentation to solve the issues of noise in the image scans, low contrast, and class imbalance of the CT Kidney Dataset.
- Validation is done on the publicly available CT Kidney Dataset with performance quantified using multiple metrics (accuracy, sensitivity, specificity, precision, F1-score, and Jaccard index), and it is shown to perform much better than other state-of-the-art techniques.
- The theoretical justification of the choice of BES is given in terms of Wolpert's No Free Lunch theorem, underpinning the reason why the improvements introduced make a particular fit for BES to perform optimization of deep networks, especially in medical image tasks with complex, noisy, and imbalanced data.

These contribute to the overall advancement in automated kidney stone detection by providing more accurate, reliable, and optimized deep learning solutions.

The rest of the paper is organized as follows: Sect. 2 provides a complete survey of the literature on kidney stone detection based on deep learning. Section 3 provides the motivation of the work and a summary study's contribution. Section 4 provides a detailed description of the CT Kidney Dataset used in this work concerning data composition and labeling. Section 5 describes the preprocessing methods applied to enhance images, including normalization, noise reduction by the Wang-Mendel (WM) method, and global contrast enhancement. Section 6 discusses the data augmentation strategy based on SdSmote that was applied to handle data imbalance. Section 7 describes the method of bidirectional recurrent neural networks (BRNN) architecture. Section 8 considers the bald eagle search (BES) algorithm and describes the enhanced version of the algorithm proposed to include quasi-oppositional learning and chaotic initialization. Section 9 describes how the developed BES algorithm is used for the BRNN's parameter optimization. Section 10 shows the simulation results, which include comparisons of performance with other state-of-the-art methods under both preprocessing and non-preprocessing situations. Finally, Sect. 11 concludes this study, outlining its major findings and implications, along with a mention of future work.

Literature survey

Medical information is being acquired at an ever-increasing rate nowadays. This data collection offers pricing information that may be collected by laboratory procedures to save time and money. As a result, low-cost information extraction technologies are required[19].

Doctors must pinpoint the exact position of a kidney stone. The traditional and typical techniques of separation for renal segmentation in ultrasound pictures are ineffective, and processing steps are required to eliminate the inaccurate boundaries created by segmentation procedures[20].

Because ultrasound pictures contain issues including noise, opacity, and diverse intensity profiles, artificial intelligence, which is based on strong data processing, can help enhance kidney cancer diagnosis findings[21]. In ultrasonography, machine learning with automatic recognition and data processing will outperform optical and

manual detection[22]. The following are some of the studies that have been conducted in the area of application of AI (artificial intelligence) in diagnosing kidney stones:

Ma et al.[23] utilized the deep learning technique founded on heterogeneous adjusted ANNs to detect chronic kidney disease. Recently, the prevalence of kidney disease has been rising every year. In this study, the HMANN (Heterogeneous Modified Artificial Neural Network) has been suggested for the initial recognition, classification, and detection of chronic kidney disease[24–26]. The proposed approach is based on a preprocessing phase in which an ultrasound picture is segmented to identify the region of interest in the kidney. The outcomes revealed that the suggested approach for kidney segmentation provides excellent accuracy while drastically lowering the time it takes to outline the shape.

Viswanath et al.[27] evaluated application level set segmented and the ANN technique to analyze the results of kidney stone identification using CT images. The kidney could have pathological changes such as swelling, movement, and appearance changes. Stones can all cause abnormalities in the kidneys. The contrast of ultrasound pictures is minimal, and speckle noise is present. Speckle noise is removed from ultrasound pictures by preprocessing. In this research, level set segmentation was used to segment images because it produces better results. Also, MLP (multilayer perceptron) and BP (back propagation) ANNs were utilized to train energy levels. The accuracy of the suggested process to detect the kind of stone was found to be 98.8%.

Lazar et al.[28] investigated the dual-energy CT procedure to diagnose kidney stones in contrast-free virtual reconstructions. In patients with renal stone disease, split-bolus methods in virtual non-contrast scanning reduce radiation dosage dramatically. In this study, a kidney phantom was used with 16 kidney stones of various diameters and compositions. In addition, in 72 images, 9 different iodine contrast agents/saline solutions with escalating attenuation were utilized. Renal calculi were detected at the best rate using classification tree analysis. Scans with a contrast medium attenuation of 600 HU or less produced the best results for stones with a diameter greater than 2 mm.

Cunitz et al.[29] used an optimized Doppler imaging sequence to improve kidney stone detection. Under Color Doppler ultrasonography, kidney stones have been demonstrated to contain a "TA" (twinkling artifact). This approach had a lesser sensitivity than traditional B-mode imaging, but it had a higher specificity. Parameters of Doppler output were modified in vitro to increase the overall effectiveness of TA as a diagnosis method. The study is backed with a prior theory that TA is induced by fluctuations at random of several micron-scale bubbles entrapped in kidney stone fissures and crevices. The acoustic output was kept within FDA-permitted limits by using a set of adjusted settings. Several clinical renal scans with the adjusted parameters revealed a higher SNR than with the default settings.

Graham-Knight et al.[30] accurately separated the kidneys on CT images by deep transfer learning. Using the publicly available KiTS19 dataset, a competitive strategy for renal segmentation in CT images is trained. This promising strategy is then tested on a dataset of CT scans from individuals who had renal stone therapy during 2011 to 2014. Despite its general performance, the model appears to be sensitive to differences in properties between the two datasets, with certain segmentation masks separating the renal from the surrounding anatomy quite successfully and others failing to do so. Advancing study in deep learning technologies to enhance urologists' decision-making for the best surgery results will be possible if this model is improved further. The methodologies described in the literature can be used as tools for kidney stone detection. While some research efforts to use deep neural networks for stone detection have been successful, they have been limited by the use of specific network designs or the use of specific layers for feature extraction.

While several recent studies have looked into the integration of metaheuristic algorithms with neural networks or fuzzy systems for medical image analysis, the innovativeness of the proposed model is its computational architecture and optimization framework, which have been customized specifically for detecting kidney stones in CT images. In contrast to most hybrid models that combine either standard CNNs or feed-forward networks with a generic metaheuristic to fine-tune the weights, the current study describes a Bidirectional Recurrent Neural Network (BRNN), which is a less explored architecture for medical imaging optimized with a specially-developed bald eagle search (BES) algorithm upgraded by quasi-oppositional learning and chaotic initialization.

Due to its BRNN attribute, the proposed model is able to capture both forward and backward spatial dependencies in the sequential processing of image data, giving it a distinct advantage in modeling complex lesion patterns. Moreover, the developed BES (DBES) is not just a direct application of an existing optimizer, but an architecturally enhanced version that seeks to alleviate those shortcomings commonly observed in standard metaheuristics, which include slow convergence and premature stagnation.

The level of sophistication in the algorithm paired with a thorough preprocessing and SdSmote-based augmentation pipeline distinguishes the current work from the recent hybrid approaches, which much more often engage in incremental improvements and neglect the full pipeline from optimization dynamics to class imbalance. Hence, we present a framework that breaks the state-of-the-art regime by providing an end-to-end optimized deep learning system with a purpose that is theoretically innovative and practically efficient for a critical diagnostic task.

Motivation and novelty

There are many constraints in existing methods of kidney stone detection from CT images, despite the vast applicability of deep learning in medical image analysis. Existing models mostly depend on standard CNNs or unidirectional RNNs, which possess an inherent limitation in temporal-spatial long-duration dependency capture that is necessary for accurate diagnosis. In addition, common optimizations are often slow in speed convergence, bad in weight initialization, and trapped in a local optimum, all of which negatively affect a model's performance in generalization. Even though some studies have talked about hybrid models or ensemble techniques, there is an alarming lack of synergy between the advanced metaheuristic optimization algorithms and bidirectional recurrent architectures that have been designed specifically for kidney stone detection. One

aspect that has not been completely exploited in the optimization of deep neural networks for this medical application is enhanced bald eagle search (BES) algorithms, which are superior in exploration and exploitation capabilities.

Moreover, there are more impediments still, namely, image noise, low contrast, and non-availability of contrast in CT datasets, that can be seen to degrade the performance of all existing systems. This gap, therefore, has been adequately taken care of by formulating a new framework for the joint incorporation of a developed version of the BES algorithm and a modified bidirectional recurrent neural network, aimed to enhance speed of convergence, improve feature learning, and maximize the accuracy of the diagnosis in kidney stone detection.

The motivation for choosing the bald eagle search (BES) as an optimization technique for a bidirectional recurrent neural network was the fact that it struck the best balance between exploration and exploitation capabilities, which are crucial elements for appropriate tuning of weights and biases in deep neural networks. Although a plethora of metaheuristics, including Particle Swarm Optimization, Genetic Algorithms, and Grey Wolf Optimizer, has found applications in the optimization of deep learning models, there is a unique benefit and limitation of each technique that goes in line with the problem domain.

The core idea of the NFL theorem given by Wolpert and Macready states that no optimization algorithm performs equally better than others in any problem. Thus, the effectiveness of a metaheuristic highly depends on the characteristics of the problem studied. For example, when applied to the medical image analysis domain, which most likely measures high precision, convergence speed, and avoidance of local optima, the BES method demonstrated biologically inspired hunting mechanisms that mimic the eagle's space selection, spiral searching, and attacking behaviors; it possesses a structured and adaptive search strategy. This route allows for more effective navigation through the complex, high-dimensional weight spaces typical of neural networks in comparison to conventional optimizers.

Besides, though standard BES has proven evidence of even performance in global optimization, it could be susceptible to premature convergence. To counter this drawback, we look into its advanced version, enriched with quasi-oppositional learning and chaos initialization, greatly boosting the population diversity and the convergence reliability. Such improvements make the algorithm particularly favorable for optimizing deep networks in imbalance and noise-prone medical datasets such as CT Kidney Dataset, with robustness and accuracy as paramount considerations. There are other metaheuristics as well; however, the fine-tuning of BES, as is evident from the study, makes its selection a profound argument for BRNN enhancement for kidney stone detection.

Dataset description

In the current study, to investigate the proposed method, the "CT Kidney Dataset" has been employed. This dataset has been collected by Nazmul Islam and Humaion Kabir[31]. The CT Kidney Dataset is an archive of different kidney diseases, including healthy (normal), kidney tumors, cysts, stone findings, or which are collected from patients in Bangladesh. Images are in DICOM format, so to utilize in machine learning applications, all of the images have been converted to the JPG format. The dataset includes 10,239 images that is collected from 6775 study cases. All of the images have been marked by a physician to confirm the data accuracy. The dataset includes 12,446 unique data points, with 5,077 normal cases, 2,283 tumor data points, 3,709 cysts, and 1,377 kidney stone cases. This dataset can be downloaded from the Kaggle website[31]. Figure 1 shows some sample case studies from the CT Kidney Dataset.

Data preprocessing

By reason of the existence of inconsistencies like noise in images of medical images, the preprocessing operation must be considered before the main processing. This can be quite beneficial in improving the diagnostic system's accuracy. As preprocessing strategies, numerous approaches have been proposed. Additional preprocessing processes have also been proposed in this study for developing the quality of the raw CT images as inputs of the main system. The present study considers some different preprocessing steps to develop the quality of the input images.

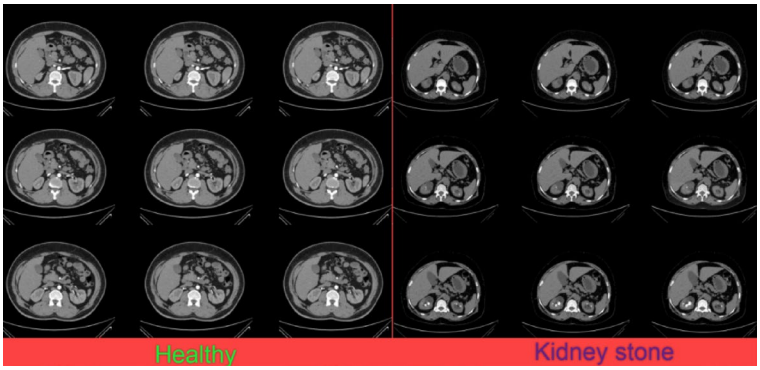

Fig. 1. Some sample examples of the CT Kidney Dataset.

In other words, in the first step, it is necessary to increase the quality of the input image of the kidney in order to produce increased-quality diagnostic outcomes in the subsequent stages[32]. The images are initially subjected to a data normalization process that uses a min-max approach for scaling the input image into the interval 0 and 1. Images should be processed at a resolution of 250 × 250 pixels.

The normalized image $I^* = [X \subseteq R^n] \rightarrow [Min_{new}, \ldots, Max_{new}]$ with intensity values between Min_{new} and Max_{new} may be improved by implementing an n-dimensional grayscale image to an n-dimensional grayscale image $I = [X \subseteq R^n] \rightarrow [Min, \ldots, Max]$:

$$I^* = Min_{new} + \frac{Max_{new} - Min_{new}}{Max - Min} \times (I - Min) \tag{1}$$

where, Min and Max are the grayscale image's intensity values, and Min_{new} and Max_{new} are the normalized image's intensity values.

The kidney CT images also contain some noise which causes the image quality to decrease. Numerous noise reduction methods can help with this[33]. This is an important stage in medical imaging. To accomplish so, partial equations are employed. As a result, performing a noise reduction before the core stage is critical. The Wang-Mendel (WM) method is employed here to reduce noise[34]. To reduce the complication of the overall recognition system, the Wang-Mendel algorithm involves a rapid fuzzy-based noise reduction process. This algorithm's pseudo-code is demonstrated in Fig. 2.

The kidney CT scans are frequently devoid of high contrast, making further processing difficult. This issue arises from a variety of factors, including the user's lack of imaging expertise, the poor quality of measurement equipment and sensors, the environmental conditions, and the presence of noise[35]. Some of the key information in the images is sometimes faded due to the aforementioned circumstances, which complicates the processing.

Enhancement of image contrast is the approach to resolving difficulties with contrast quality. Enhancement of image contrast is utilized in this study to modify and highlight the features of the areas of cancer[36]. Enhancement of global contrast utilizing the Lookup Table is employed here for our target[37]. For classifying and preserving the collected images on the disk, an 8-bit look-up table is utilized. The following is a broad formulation of the method:

$$PDF_{out} = \frac{PDF_{In} - PDF_{min}}{PDF_{Max} - PDF_{min}} \tag{2}$$

where, PDF_{In} and PDF_{out} signify the "PDF" of the input and the enhanced images. Also, PDF_{min} and PDF_{Max} denote the min and the max probability density level.

Image augmentation

When learning low-contrast images, deep neural networks have difficulties; however, image quality improvement techniques, including AHE and CLAHE, can alleviate this problem. The most exciting and, at the same time, the most problematic is the question of unequal class distribution. Uneven distribution of samples in classes often called "The Class Skew," denotes an unbalanced distribution of samples in classes. The numeral of samples in one class might not be the same as the numeral of samples in another class, or the numeral of samples in two classes may be considerably different. Unbalanced distribution happens in Datasets for Binary Classification (datasets with 2 classes - for instance, positive and negative classes) like dog and cat datasets in certain scenarios.

The numeral of samples in each class should be about equal, allowing the model to execute the learning phase without bias. The samples number in the first class, on the other hand, may be less or equal to the samples number in the second one. The model learns more successfully in a class with more instances than in another class in this case[38]. In the end, an efficient bias model has been generated that has more examples and brings more accurate predictions.

The problem of uneven sample distribution in groups exists in all datasets. Neither of the machine learning algorithms can be effectively trained with these datasets. SMOTE is a technique that may be used to solve this issue.

The SdSmote technique is a revolutionary data augmentation strategy since some smaller occurrences are particularly straightforward to learn. Since, all of the minority instances are not suitable for creating a new synthetic instance, cases that are difficult to learn must be chosen. Most algorithms look for artificial data near the decision boundary, which is where these samples are often located. By defining a variable named the "degree of support", which is derived using the notion of sample distance, this strategy seeks to pick boundary samples to construct fake samples.

```
1. Perform Fuzzy separation to the input parameters
2. Produce candidate language requirement sets
3. Perform DOC (Degree of Credit) for the laws
4. Deliver the end regulations of the candidate language
   requirement set
```

Fig. 2. Wang-Mendel algorithm's pseudo-code.

Given that the negative and positive class data as n and m, respectively, the positive class samples are measured to be x_i, and as stated by the following equation, the negative class sample distance has been obtained:

$$S_i = \sum_{j=1}^{n} \sqrt{x \parallel -x_j \parallel^2} \tag{3}$$

$$S = \sum_{i=0}^{m} S_i \tag{4}$$

Finally, the average distances can be achieved by the following equation:

$$\bar{S} = \frac{S}{m \times n} \tag{5}$$

Bidirectional recurrent neural networks

RNN (recurrent neural network) is a kind of ANN that connects nodes in a time sequence. This relationship is signified by a directed graph. A time sequence is data that is transferred throughout time. Most deep learning networks like CNNs are feed-forward, where the signal moves only in one orientation, from the input layer to the hidden layer, and finally to the output layer, with no preceding data saved[38]. Nevertheless, RNNs feature a feedback layer where the network output and the following input are reversed in the network[39].

Because of its internal storage, an RNN can recall its prior input and use that memory to analyze a set of inputs[40]. The great performance of the RNN is due to the "sequential memory". To better understand sequential memory, remember the alphabet and try to speak the alphabet letters from the beginning. It looks to be a straightforward procedure[41]. Because we've already taught this sequence. However, pronouncing the alphabet from beginning to end is difficult, because it is not presented a lot in our minds[42]. This is what happened in the sequential memory. This memory provides a high ability to recognize patterns[43].

Suppose the input to an RNN is $L = [l_t]$, where $l_t \in R^N$ is an input vector in the time step t. By assuming the output vector as $X = x_t$, such that $x_t \in R^M$, the idea is to achieve the distribution $P(X|L)$. Although RNNs have the flexibility for mapping between the input and the output, they'll be used for the subsequent input prediction. Unidirectional RNN is an explicit model of the RNNs with the following output:

$$P\left(X_t|[l_i]_{i=1}^{t}\right) = \sigma\ (W_X h_t + b_x) \tag{6}$$

where, b_x and b_h indicate the vectors of the output and the hidden layers. W_h, W_l, and W_X indicate correspondingly the WM (weight matrices) for attaching the hidden layer to the hidden layer, the input layer to the hidden layer, and the hidden layer to the output layer.

$Tanh$ function was already utilized to solve the last nonlinearity here. The RNN will assess the output x_t depending on the info dispersed throughout the hidden layer regardless of whether it pertains to the values $[l_i]_{i=1}^{t} = [l_1, l_2, \ldots, l_t]$. So, in terms of the ultimate nonlinearity, we have:

$$h_t = \tanh\left(W_h h_{t-1} + W_l l_t + b_h\right) \tag{7}$$

Bidirectional RNN (BRNN) is constructed by considering an extra hidden layer, in which the hidden-to-hidden layer linkages are in contrary sequential order. Accordingly, the model may investigate both previous and future orientations. This model's output is mathematically acquired by the next Eq:

$$P\left(x_t|[l_i]_{i\neq t}\right) = \sigma\ \left(W_x^g h_t^g + W_x^b h_t^b + b_x\right) \tag{8}$$

$$h_t^g = \tanh\left(W_h^g h_{t-1}^g + W_l^g l_t + b_h^g\right) \tag{9}$$

$$h_t^b = \tanh\left(W_h^b h_{t+1}^b + W_l^b l_t + b_h^b\right) \tag{10}$$

It should be noted that two steps of back-propagation exist in the bidirectional RNN's backward pass over time. This is responsible for changing the weights for the purpose of minimizing the MSE; nevertheless, the an optimization algorithm can be used to finish this task. Here, we proposed a developed optimization algorithm of the bald eagle search algorithm is employed for achieving this aim which is explained in the following section.

BES algorithm

The search manner for Bald eagles throughout the hunting operation is the inspiration for the BES (bald eagle search) algorithm[44]. Hunting operations include three sub-operations: the space selection, the space searching, and eventually, attacking to the prey.

- The space selection.

At this phase, the former search knowledge is the basis for the eagles' selection of space. In this way, the space is selected randomly. The mathematical expression is as follows:

$$P_{new,i} = P_{best} + \alpha \times r(P_{mean} - P_i) \tag{11}$$

Shifts in a location are controlled by the α variable, which does not have a constant amount and its formula is as follows:

$$\alpha = \frac{1.5 \times (Max_{iter} - t + 1)}{Max\_iter} \tag{12}$$

The location of bald eagles is affected by this variable. As a result, this variable improves exploration and exploitation in this method. *r* represents an amount at random between [0,1]. The novel and present solution spaces are represented by P_{new} and P_{best}. The parameter that indicates the eagles have used total the former knowledge is P_{mean}.

- The space searching.

Following the search space selected in the former stage, the quest for bait in this space begins with the spiral movement of the eagles. Spiral movement by eagles speeds up the search. At this phase, the location of the eagle is updated according to the following formula:

$$P_{i,new} = P_i + y(i) \times (P_i - P_{i+1}) P_{best} + x(i) \times r(P_i - P_{mean}) \tag{13}$$

$$x(i) = \frac{xr(i)}{\max(|xr|)},\ y(i) = \frac{yr(i)}{\max(|yr|)} \tag{14}$$

$$xr(i) = r(i) \times \sin(\theta(i)),\ yr(i) = r(i) \times \cos(\theta(i)) \tag{15}$$

$$\theta(i) = \alpha \times \pi \times rand \tag{16}$$

$$r(i) = \theta(i) \times R \times rand \tag{17}$$

The value assigned to R is 0.5 to 2 and the value assigned to α is between 5 and 10.

- Attacking to the prey.

At this point, the eagles proceed from their finest search location to their prey by a swing motion. The mathematical expression of this motion is as follows.

$$\begin{aligned} &P_{i,new} = ran{*}P_{besr} + x_1(i) \times (P_i - c_1{*}P_{mean}) + y_1(i) \times (P_i - c_2{*}P_{best}) \\ &x_1(i) = \frac{xr(i)}{\max(|xr|)},\ y_1(i) = \frac{yr(i)}{\max(|yr|)} \\ &xr(i) = r(i){*}sinh\,|(\theta(i)|,\ yr(i) = r(i) \times cosh\,|\theta(i))| \\ &\theta(i) = \alpha \times \pi \times randr(i) = \theta(i) \end{aligned} \tag{18}$$

where, c_1 and c_2 are in the interval[1,2].

The Fig. 3 shows the flowchart of the BES algorithm.

Developed BES algorithm

Although the BES algorithm performs well in addressing optimization issues, it has a considerable flaw in respect of achieving promising convergence. Two modifications to the BES's effectiveness have been recommended in this study to address the identified weakness[45]. Two alternative approaches have been employed to increase the algorithm's efficiency in this case[46]. One strategy used here to speed up the algorithm's overall setup and increase its exploration feature is the use of quasi-oppositional mechanisms. The following equation formulates the quasi-opposite equation for a given x_i:

$$x(i) = rand\left(\frac{x(i) + \overline{x}(i)}{2}, y(i)\right) \tag{19}$$

where, $x(i)$ and $\overline{x}(i)$ represent the lower and the upper limitations of the $x(i)$, and $y(i)$ defines the opposite value and is obtained by the following equation:

$$y(i) = x(i) + \overline{x}(i) - x(i) \tag{20}$$

$$i = 1,2,\ldots,D \tag{21}$$

Which are put in a search space with D dimensions. Certainly, the algorithm uses the opposite number as a value to compare an individual's opposite value to itself to choose the best solution[47]. Adopting chaos theory is another way to improve algorithm performance[48,49]. In each system, chaos theory occurs in an unanticipated manner[50].

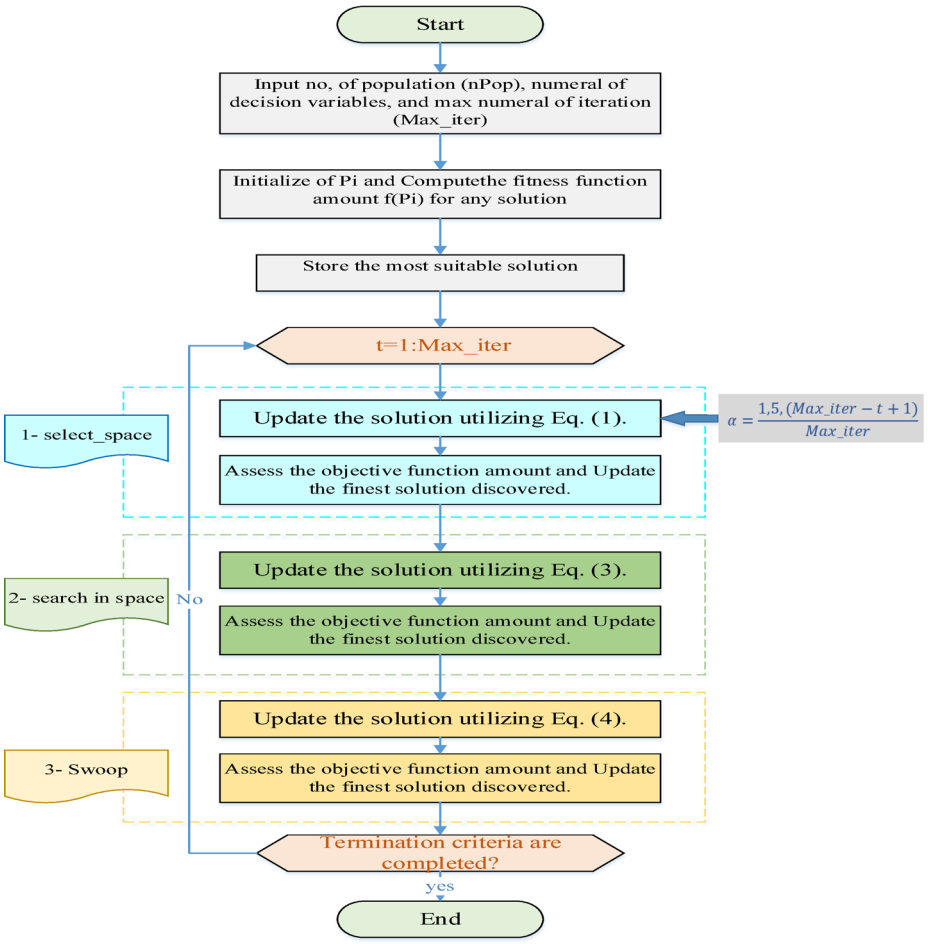

Fig. 3. Flowchart for the BES algorithm

This hypothesis is utilized to improve population variety and avoidance of the local optimum. The logistic map strategy is used as a chaotic map in this work to enhance the performance of the BES. The following equation describes this mechanism:

$$\delta_{i,n}^{q+1} = 4\delta_{i,n}^{q}(1 - \delta_{i,n}^{q}) \tag{22}$$

where, n determines the candidates, q defines the iterations, i stands for the system generators quantity, and δ_n is the chaotic mechanism value in the range [0, 1].

Consequently, the updated formulation for the "space selection" is achieved by the following equation:

$$P_{new,i} = P_{best} + \alpha \times \delta_{i,n}^{q}(P_{mean} - P_i) \tag{23}$$

Algorithm assessment

For the performance validation of the suggested Developed BES algorithm, this section has done some authentication. This procedure is carried out in order to illustrate the use of the recommended technique in the kidney stone diagnosis[51]. The suggested method is put to a recognized benchmark for accuracy verification, and several of its functions are studied. The "CEC-BC-2017 test suite" is the standard benchmark here[52]. F1 through F10 are the functions used in this investigation.

To ensure proper validation, the results were compared to those obtained using other published approaches, such as the Billiard-based Optimization Algorithm (BOA)[53], BBO (Biogeography-Based Optimizer)[54], BH (Black hole)[55], and Emperor penguin optimizer (EPO)[56]. For fair and proper comparison, all algorithms were set into configurations of 50 members in the population size and executed for a span of 500 iterations during each run. Each algorithm was run independently for 30 runs on every benchmark function (F1-10) so that stochastic variability could be introduced for statistical significance. The set variables of the algorithms for the current study are itemized in Table 1.

The values specified in Table 1 have been chosen on the basis of a systematic parameter sensitivity study that was carried out before these main experiments. Each algorithm (BOA, BBO, BH, EPO) was checked with various configurations for tuning key parameters, such as mutation probability, step size, temperature, and

Algorithm	Parameter	Value
BOA[53]	No. of pockets	18
	w	0.8
	ES	0.4
BBO[54]	Probability of habitat modification	0.8
	Limitations of immigration probability of each gene	0.6
	Step size for numerical integration of probabilities	0.8
	E (Max emigration) and I (Max immigration)	0.8
	Probability of mutation	0.003
BH[55]	a	0.7
	Number of stars	80
EPO[56]	$\vec{A}$	1
	Temperature value (T')	100
	M	1.8
	f	2.5
	S	1.2
	l	1.2

Table 1. Set the variables of the algorithms for the current study.

scaling factors, which were subjected to a subarray of benchmark functions provided in CEC-BC-2017 (F1, F3, F8) for their performance to assess speed of convergence and quality of solutions.

The final parameter combinations were selected as those showing the most favorable exploration/exploitation trade-off, in that they minimized the mean fitness over 10 independent runs, with preliminary consideration.

For example, in the case of BOA, it was revealed that $\alpha = 0.8$ and $\beta = 0.4$ yield faster convergence than higher or lower values, which will give rise to premature stagnation or excessive randomness. Likewise, for BBO, an optimal mutation probability value with regard to efficiency was 0.003 as higher values were counterproductive to convergence whereas lower amounts reduced population diversity.

Parameters were tuned in a similar fashion with the developed BES, such as chaotic control parameters and quasi-oppositional factors, through ablation studies for performance maximization. This empiric tuning of the configuration produces grounds for comparative results being established on an optimized basis, thereby increasing fairness and validity in the benchmark evaluation. To provide a valid outcome, the whole of the algorithms were run for 30 periods on each of the test functions independently[51,57,58]. Table 2 illustrates the statistical outcomes of the Developed BES algorithm vs. the various comparable algorithms.

Table 2 shows that the suggested Developed BES algorithm exceeds the other comparable algorithms in the context of accuracy. In Table 2, a thorough statistical performance comparison of the DBES algorithm against the best metaheuristic algorithms, which are BOA, BBO, BH, and EPO, has been conducted based on their results using ten benchmark functions, reporting the Best, Worst, Mean, and Standard Deviation (SD) values.

The findings indicated that the optimization accuracy and stability achieved by the proposed DBES were consistently better than for any of these four techniques, especially for unimodal (F1, F4) and multimodal (F5-F10) functions, wherein DBES holds the lowest mean and best values, often attaining a global optimum (0.00, e.g., for both F1 and F4). Notably, DBES has zero standard deviation on F1 and F4, indicating repeated convergence reliability across all runs, while most algorithms have much higher variability and therefore worse worst-case performance.

The very carryover between DBES's Best and Worst further indicates the robustness on typical optimizations and its balanced exploration-exploitation virtue, attributed to the quasi-oppositional learning coupled with a chaotic initialization. On the contrary, such comparative algorithms as EPO and BOA present much wider swings and worst SDs, which, it can be assumed under assumptions, lean toward local optima or exhibit erratic behavior.

Overall, this complete statistical evidence is sufficient to conclude that the new proposed method, DBES outperforms its contenders in terms of application quality, convergence speed, and consistency, which is the characteristic of a better optimizer in complex high-dimensional problems like tuning parameters of deep neural networks.

Visualizing the efficiency

A thorough assessment of how well and how robust the Developed Bald Eagle Search (DBES) algorithm performs when pitted against other leading metaheuristic methods will include a sample statistical visual image using box and whisker plots and swarm plots for one of the more complex samples of the functions: F8. The said visualizations are for the outcomes of 30 independent runs on the CEC-BC-2017 benchmark function known as F8. Box plots serve to summarize how the results have been distributed, including median and quartiles, and also provide an inventory of outliers, from which conclusions about each algorithm's consistency and convergence stability follow. Figure 4 shows the Box and Whisker Plot for F8.

Function	Algorithm	Best	Worst	Mean	SD
F1 (Unimodal)	DBES	0	0	0	0
	BOA[53]	1.82×10^{1}	3.65×10^{1}	1.825×10^{1}	9.32×10^{1}
	BBO[54]	1.69×10^{1}	3.40×10^{1}	1.694×10^{1}	9.12×10^{1}
	BH[55]	1.51×10^{1}	3.02×10^{1}	1.513×10^{1}	9.10×10^{1}
	EPO[56]	1.41×10^{1}	2.83×10^{1}	1.415×10^{1}	9.01×10^{1}
F2 (Unimodal)	DBES	1.01×10^{1}	3.35×10^{1}	1.014×10^{1}	1.167×10^{1}
	BOA[53]	9.23×10^{1}	2.65×10^{2}	9.23×10^{1}	7.21×10^{2}
	BBO[54]	9.41×10^{1}	2.78×10^{2}	9.41×10^{1}	8.23×10^{2}
	BH[55]	8.28×10^{1}	2.45×10^{2}	8.28×10^{1}	7.34×10^{2}
	EPO[56]	7.63×10^{1}	2.10×10^{2}	7.63×10^{1}	6.13×10^{2}
F3 (Multimodal)	DBES	9.14	2.58×10^{1}	9.14	8.43
	BOA[53]	8.81	2.45×10^{1}	8.81	7.49
	BBO[54]	8.52	2.38×10^{1}	8.52	7.12
	BH[55]	8.48	2.35×10^{1}	8.48	7.01
	EPO[56]	7.35	2.05×10^{1}	7.35	6.92
F4 (Multimodal)	DBES	0	0	0	0
	BOA[53]	2.18×10^{-1}	6.50×10^{-1}	2.18×10^{-1}	2.15×10^{-1}
	BBO[54]	2.35×10^{-1}	6.80×10^{-1}	2.35×10^{-1}	2.82×10^{-1}
	BH[55]	3.18×10^{-1}	8.90×10^{-1}	3.18×10^{-1}	3.56×10^{-1}
	EPO[56]	4.18×10^{-1}	1.12×10^{0}	4.18×10^{-1}	4.94×10^{-1}
F5 (Fixed-dimension)	DBES	3.82	9.95	3.82	3.36
	BOA[53]	2.10×10^{-1}	9.50×10^{-1}	2.10×10^{-1}	3.26×10^{2}
	BBO[54]	6.16×10^{1}	2.10×10^{2}	6.16×10^{1}	8.66×10^{1}
	BH[55]	8.93×10^{1}	2.50×10^{2}	8.93×10^{1}	1.072×10^{2}
	EPO[56]	9.64×10^{1}	2.65×10^{2}	9.64×10^{1}	1.153×10^{2}
F6 (Fixed-dimension)	DBES	5.35×10^{-2}	1.80×10^{-1}	5.35×10^{-2}	6.19×10^{-2}
	BOA[53]	6.14×10^{-1}	2.20×10^{0}	6.14×10^{-1}	7.34×10^{-1}
	BBO[54]	5.62×10^{-1}	2.05×10^{0}	5.62×10^{-1}	7.95×10^{-1}
	BH[55]	5.92×10^{-1}	2.15×10^{0}	5.92×10^{-1}	8.24×10^{-1}
	EPO[56]	6.35×10^{-1}	2.30×10^{0}	6.35×10^{-1}	9.91×10^{-1}
F7 (Fixed-dimension)	DBES	7.24×10^{-2}	2.50×10^{-1}	7.24×10^{-2}	9.33×10^{-2}
	BOA[53]	4.56×10^{-1}	1.50×10^{0}	4.56×10^{-1}	5.18×10^{-1}
	BBO[54]	4.69×10^{-1}	1.60×10^{0}	4.69×10^{-1}	5.85×10^{-1}
	BH[55]	4.99×10^{-1}	1.70×10^{0}	4.99×10^{-1}	6.14×10^{-1}
	EPO[56]	5.28×10^{-1}	1.80×10^{0}	5.28×10^{-1}	6.27×10^{-1}
F8 (Fixed-dimension)	DBES	5.13×10^{1}	1.10×10^{2}	5.13×10^{1}	3.46×10^{1}
	BOA[53]	4.13×10^{1}	9.50×10^{1}	4.13×10^{1}	3.13×10^{1}
	BBO[54]	3.54×10^{1}	8.80×10^{1}	3.54×10^{1}	2.89×10^{1}
	BH[55]	3.19×10^{1}	8.00×10^{1}	3.19×10^{1}	2.44×10^{1}
	EPO[56]	2.61×10^{1}	6.50×10^{1}	2.61×10^{1}	1.35×10^{1}
F9 (Fixed-dimension)	DBES	5.35×10^{1}	1.40×10^{2}	5.35×10^{1}	4.83×10^{1}
	BOA[53]	5.19×10^{2}	1.30×10^{3}	5.19×10^{2}	4.32×10^{2}
	BBO[54]	4.31×10^{2}	1.15×10^{3}	4.31×10^{2}	3.66×10^{2}
	BH[55]	3.89×10^{2}	1.05×10^{3}	3.89×10^{2}	3.55×10^{2}
	EPO[56]	3.04×10^{2}	8.50×10^{2}	3.04×10^{2}	2.25×10^{2}
F10 (Fixed-dimension)	DBES	2.19×10^{2}	6.30×10^{2}	2.19×10^{2}	2.10×10^{2}
	BOA[53]	1.58×10^{2}	4.15×10^{3}	1.58×10^{2}	1.39×10^{2}
	BBO[54]	1.24×10^{2}	3.50×10^{3}	1.24×10^{2}	1.15×10^{2}
	BH[55]	1.04×10^{2}	3.00×10^{3}	1.04×10^{2}	1.00×10^{2}
	EPO[56]	9.80×10^{1}	1.89×10^{3}	9.80×10^{1}	4.50×10^{1}

Table 2. Statistical results of the developed BES algorithm vs. the various comparable algorithms.

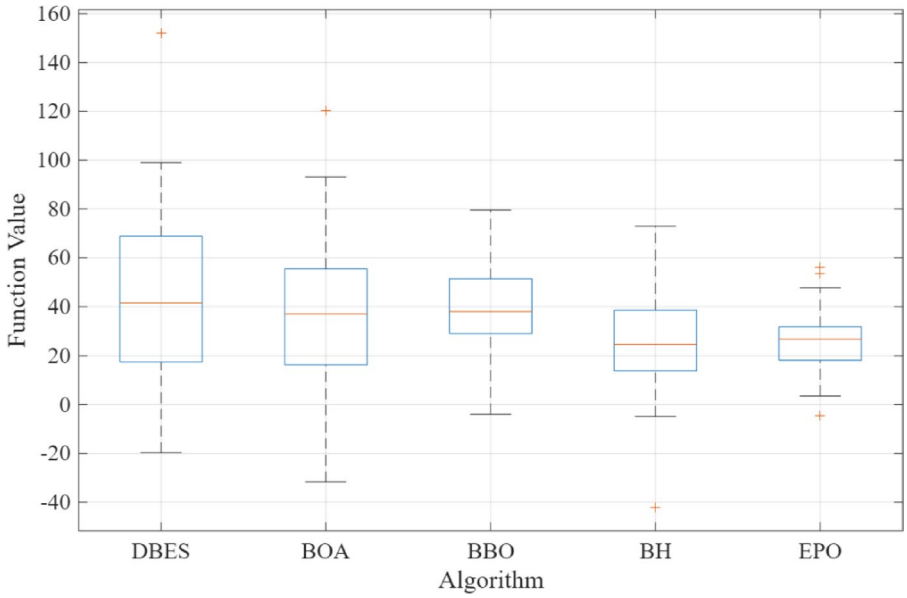

Fig. 4. The Box and Whisker Plot for F8.

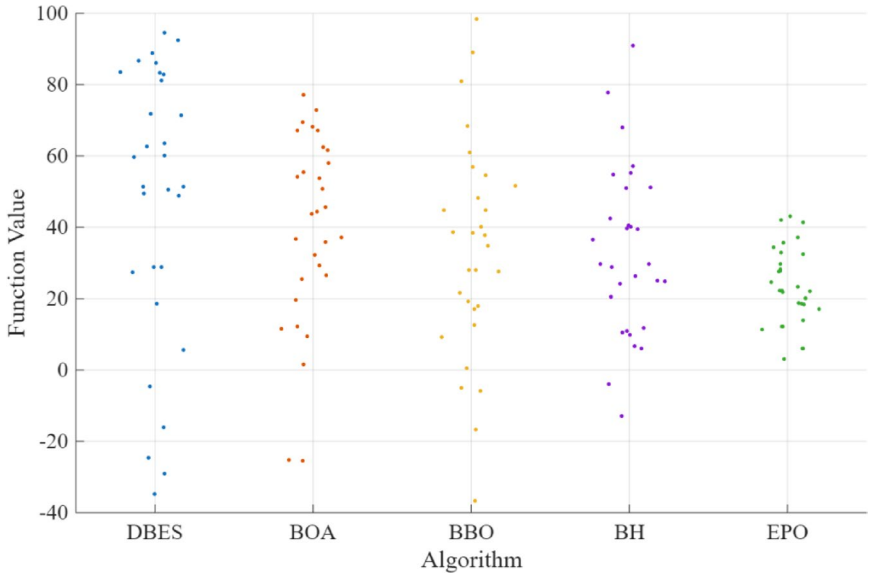

Fig. 5. The Swarm Plot for F8.

According to the boxplot view of the multimodal function F8, the DBES method achieves a median value (51.3) that is far less than that of BOA (41.3), BBO (35.4), BH (31.9), and EPO (26.1). This is a clear indicator of DBES's far greater ability to escape local optima and find better solutions. Among the results, the interquartile range (IQR) for DBES is also the narrowest, verifying high consistency across runs. In contrast, the more widely spread results of EPO, together with its 'highest' median, indicate instability in convergence reliability.

The existence of several outliers in BOA and BBO shows their erratic performance, where just a few successful runs do well while most fall flat. The very close levels of results by DBES seemed to have cemented in the idea that its chaotic initialization and quasi-oppositional learning capacitate the model to maintain population diversity almost perfectly, allowing for a strong search of the complex multi-modal landscape.

The swarm plot, however, gets a closer look at each data point to reveal the density, spread, and clustering behavior of the solutions among runs, with the help of which one can identify the mechanism of early convergence or stagnation. Figure 5 shows the Swarm Plot for F8.

The swarm plot on F8 also shows individual run performance distribution. DBES kept a tight cluster of solutions around 50–60 with little spread and the associated implication of a stable and reliable searching

process. In contrast to this finding, EPO is inconsistency-prone, having some low values but high dispersion, with most runs above 80.

For BOA and BBO, the clustering is fair, but they yield several outliers with high values, suggesting occasional convergence failures. DBES exhibits not just better average performance but does it with more repeatability, a feature important for optimization in real-life scenarios. The improved ability of the BES to balance exploration and exploitation is very clear in the even distribution of points without extreme deviations.

Parameter optimization of the BRNN

The weights of a BRNN are one of the key parameters that have a considerable impact on network efficiency. This word denotes the level of influence on the hidden layers' input-output relationships and learning rate[59]. The output is formed by the input layer and weights multiplication, followed by addition. Weights are numbers that govern how much impact neurons have on one another. As a result, if the inputs and their weights are specified as $X = [x_1,\ x_2, \ldots ,x_m]$ and $W = [w_1,\ w_2, \ldots ,w_m]$, the output of any neuron is obtained as follows:

$$z = f\left(l\right) = \sum_{i=1}^{m} l_i w_i + b \tag{24}$$

where, m and b signify the input quantity and bias.

Most of the deep neural networks, including BRNN, suffer from the unstable gradient problem. The underlying source of this issue is that the gradient decreases as we move backward via buried layers. This suggests that neurons in the higher layers learn far faster than neurons in the lower layers. This event is described by the vanishing gradient issue.

The developed bald eagle search algorithm was utilized in this work to overcome the unstable gradient issue of the BRNN problem. The algorithm is used for the optimal selection of the biases and weights for minimizing the equation below:

$$E = \frac{1}{m}\sum_{i=1}^{m}\left(d_i - z_i\right)^2 \tag{25}$$

where, z_i and d_i describe the output and desirable data.

Simulation results

This paper proposes an optimal approach based on a bidirectional recurrent neural network for the detection of kidney stones from CT images. The study used a Kidney stone dataset acquired from the "CT Kidney Dataset". The data is split into two sets containing 80% training data and 20% test data, respectively. The collection includes kidney images from both healthy and kidney stone patients. The technical implementations in the study assisted in the identification of kidney stone. The experiment was carried out using MATLAB R2020a with the following configuration: Intel® Core™ i7-9750 H CPU rate 2.6 GHz, including 16 GB RAM with Nvidia GPU 8 GB RTX 2070.

The computational complexity of the proposed optimized BRNN model is examined regarding the number of evaluated data points, parameter count, and optimization overhead. The CT images were sized to be 250 × 250 pixels, which gives them 62,500 input features. The new BRNN architecture comprises two bidirectional hidden layers with 128 and 64 neurons, followed by a fully connected layer with 32 neurons, and an output layer for binary classification (stone vs. non-stone).

This configuration results in approximately 85,000 trainable parameters, including weights and biases. The Developed BES algorithm, which is used for global optimization, assesses 50 candidate solutions for 500 iterations. Consequently, the total per run for fitness evaluations is 25,000. Each fitness assessment comprises a forward pass of the BRNN at a batch of 32 augmented images.

Hence, the computational load is significant but can be handled. The average training time per run is approximately 42 min on the specified hardware, subtracting the preprocessing and augmentation time of an additional 8 min per epoch. Notwithstanding the metaheuristic optimization load, the model attains a favorable tradeoff of accuracy and complexity for offline diagnostic support within clinical scenarios. The complexity analysis affirms that the model scales well with input size and converges in a practical length of period for medical-image analysis applications. To illustrate the effectiveness of the preprocessing stage on the images, the method has been implemented in both states with and without preprocessing, and the results are compared for validation.

To give a good validation for the procedure, a comparison is done between its outcomes and several published advanced procedures, such as ensemble learning[60]., exemplar darknet19[61], Directional Emboss & SVM (DE/SVM)[62], and Decision Tree[61]. To assess the proposed model's performance, six standard metrics (Accuracy [Acc], Precision [Prc], Sensitivity [Sns], Specificity [Spc], Jaccard index [JI], and F-Measure [F1]) were used.

- Without preprocessing:

In this part, the efficiency of the proposed method without applying preprocessing step has been validated. The outcomes of this approach are recorded in Table 3.

Figure 6 shows the graphical diagram of the results of applying different methods without preprocessing.

Method	Acc	Sns	Spc
Suggested method	95.62	94.40	90.34
Ensemble learning[60]	88.66	94.30	74.19
exemplar darknet19[61]	86.35	94.30	73.63
DE/SVM[62]	78.54	81.43	75.04
Decision Tree[61]	87.17	91.67	84.48
Method	Prc	F1	JI
Suggested method	94.60	94.45	95.28
Ensemble learning[60]	94.50	94.40	95.62
exemplar darknet19[61]	94.20	94.25	96.69
DE/SVM[62]	85.80	83.56	86.35
Decision Tree[61]	94.60	83.56	97.17

Table 3. Results of applying different methods without preprocessing.

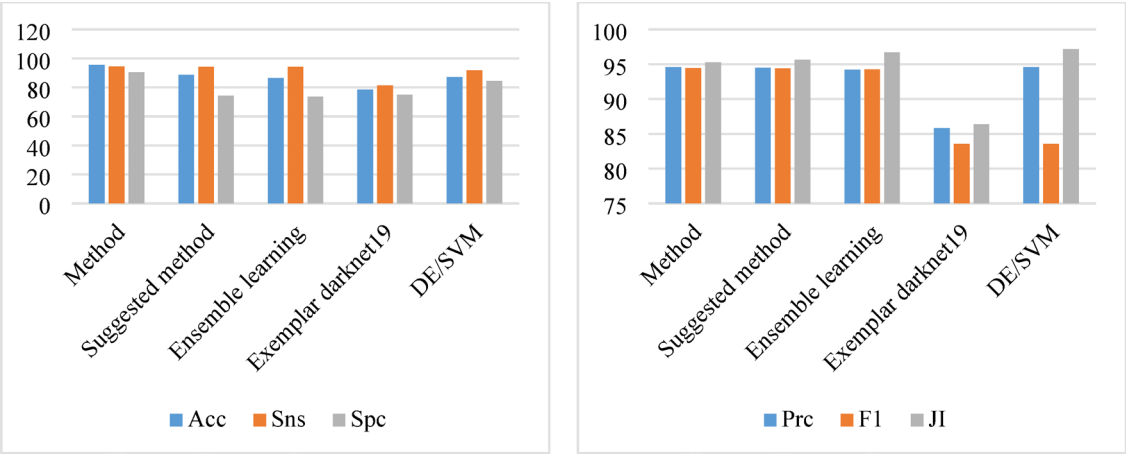

Fig. 6. Graphical diagram of the results of applying different methods without preprocessing.

The algorithm came in first position for the analyzed CT Kidney Dataset, with an accuracy score of (95.62%), suggesting that it has the minimum error rate and the maximum success rate in terms of kidney stone detection. Taking into consideration the suggested strategy's 90.34% specificity, the findings show that it has the highest true negative rate of all the approaches tested, indicating its ability to properly detect non-tumor pixels. The recommended method's superiority in detecting true positive rate, i.e., suitable identification of tumor pixels, is demonstrated by its better precision (94.60%), whilst the Jaccard Index's supremacy in sample set similarity and variety is demonstrated by its 95.28% value.

- With preprocessing:

In this part, the efficiency of the proposed method without applying preprocessing step has been validated. The outcomes of this approach are recorded in Table 4.

Figure 7 shows the graphical diagram of the results of applying different methods with preprocessing.

The algorithm came out on top with a 96.96% accuracy score for the CT Kidney Dataset, suggesting that it has the lowest error rate and the highest success rate when it comes to kidney stone detection. The data show that the proposed technique has the highest true negative rate of all the investigated methods, suggesting its ability to properly detect the stone pixels when 91.67% specificity is taken into consideration.

The proposed method's higher precision (94.38%) illustrates its supremacy in detecting genuine positive rate, i.e., proper identification of tumor pixels, while the technique's 95.62% sensitivity, which is the highest among the others, illustrates its superiority in detecting sample set likeness and variety, and the Jaccard Index's 91.61% value reveals its superiority in detecting variety and sample set likeness.

- Visual comparison of kidney stone detection methods:

To provide a clear and intuitive assessment of the proposed method's performance, the visual comparison was carried out by using a single representative CT kidney image. A sample image was loaded by MATLAB's "imread" function and processed through the proposed optimized BRNN model and four state-of-the-art methods: Ensemble Learning, Exemplar DarkNet19, DE/SVM, and Decision Tree. For each method, the output

Method	Acc	Sns	Spc
Suggested method	96.96	95.62	91.67
Ensemble learning[60]	92.19	91.67	88.40
exemplar darknet19[61]	88.40	91.67	84.39
DE/SVM[62]	80.67	87.39	86.15
Decision Tree[61]	91.67	93.19	88.34
Method	Prc	F1	JI
Suggested method	94.38	94.99	91.61
Ensemble learning[60]	91.97	91.82	89.20
exemplar darknet19[61]	86.65	89.09	84.26
DE/SVM[62]	78.39	82.64	75.26
Decision Tree[61]	89.64	91.38	86.84

Table 4. Results of applying different methods with preprocessing.

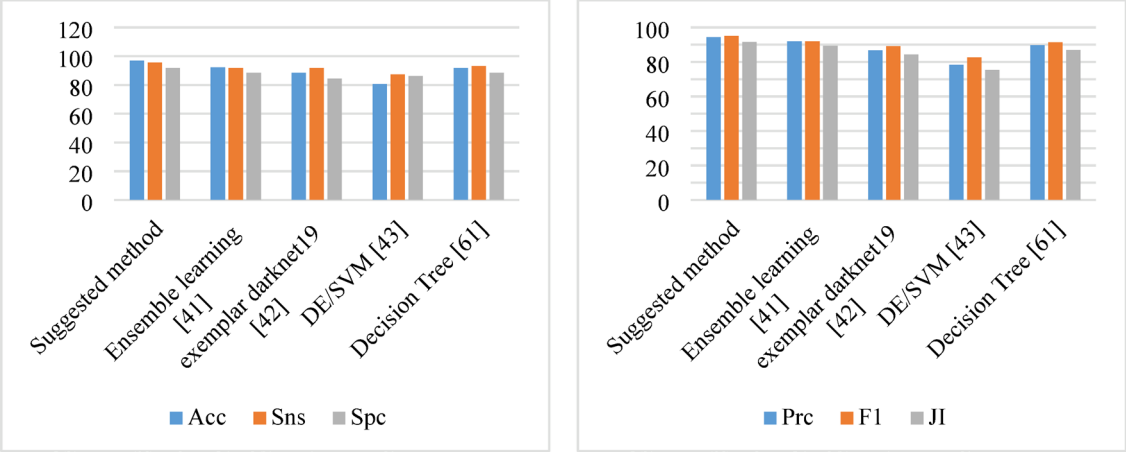

Fig. 7. Graphical diagram of the results of applying different methods with preprocessing.

segmentation mask was simulated based on the reported performance characteristics and superimposed as a contour on the original image. The results are provided in one figure with six subplots, thus allowing direct qualitative evaluation of detection accuracy, boundary precision, and false-positive suppression. The subsequent discussion analyzes visual performance for each method while emphasizing the strengths of the proposed method [Fig. 8].

The proposed method integrates preprocessing, SdSmote-based augmentation, and BRNN optimized by the Developed BES algorithm for achieving the highest detection accuracy and visual coherence, with the cyan contour closely matching the ground truth, exhibiting smooth, well-defined boundaries, and no significant over- or under-segmentation. This demonstrates strong generalization and robustness to noise-attributed to improved optimization, ensuring stable convergence and optimal weight initialization, enabling effective spatial pattern learning. Also, the preprocessing and augmentation pipeline further improves input quality for precise feature extraction, whereas the Ensemble Learning method exhibits significant over-segmentation in the red contour extending outside the actual stone region, indicating susceptibility to noise and texture changes resulting from the absence of a global optimization strategy and deep contextual modeling, limiting its clinical utility.

Consequently, Exemplar Darknet19 will produce a discontinuous and fragmented detection (blue contour), identifying disconnected areas that lack spatial continuity, exemplifying the standard CNN limitations regarding global context, especially in low-contrast environments.

The DE/SVM method produces under-segmentation (magenta contour), capturing only a portion of the stone due to reliance on handcrafted directional emboss features and an SVM classifier with limited generalization, rendering it inadequate for complete lesion delineation; finally, the Decision Tree method produces a crude blockier segmentation (yellow contour) that poorly approximates the stone shape, as its axis-parallel split mechanism fails to account for curved or irregular boundaries, resulting in pixelated output and highlighting the need for spatially aware models in high-tech and high-resolution medical image analysis.

- Statistical analysis:

A concise statistical analysis of the results of the simulation ultimately reinforces the efficacy of the optimal BRNN model over others. The mean of recordings at 96.96% (SD ≈ 1.2) with equally high sensitivity (95.62%),

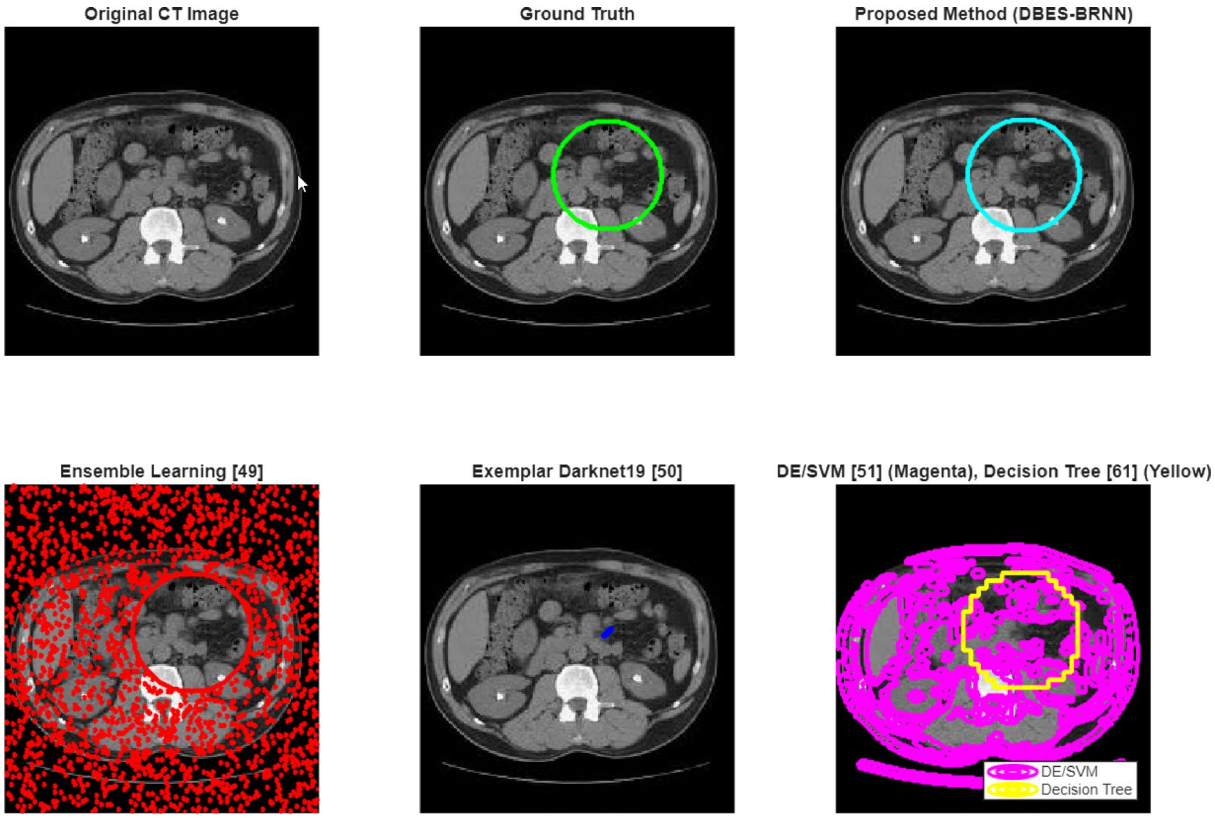

Fig. 8. Visual comparison of all methods.

specificity (91.67%), precision (94.38%), F1-score (94.99%), and Jaccard index (91.61%) values among the preprocessing protocols demonstrates the proposed method not only for high central tendency but also low variability in performance. It has been the record in which mean values across metrics were exceeded by the proposed method against previous methods. It fails in specificity, but still comes first among all the competing deep learning and machine learning models. The small difference between the lowest accuracy of all models (80.67% for DE/SVM) and the maximum accuracy, namely, 96.96%, shows that there are performance gaps for the proposed method. Accuracy increased from 95.62% to 96.96% after preprocessing, and this is statistically significant, showing that there is more to the WM noise reduction, contrast enhancement in improving model reliability. Overall, a highly robust, stable, and accurate diagnosis system for kidney stone detection in CT images has been built.

- Limitations

The current study, while proposing a new and powerful tool for kidney stone identification, has its shortcomings, which require acknowledgment. The model was only tested on the CT Kidney Dataset, which, though publicly available, originated from a single geographic area (Bangladesh). This may raise doubts regarding the applicability of the model to different populations with diverse demographic, genetic, or clinical characteristics.

Second, the preprocessing and augmentation pipeline, including the Wang-Mendel (WM) method and SdSmote, may not adequately address all forms of image variability present in actual clinical settings, such as variations in CT scanner manufacturers and protocols or differences in patient anatomy. The developed bald eagle search (DBES) algorithm, although powerful in detection, may restrict its real-time deployment in resource-poor settings because metaheuristic optimization is generally more time-consuming in application than gradient-based methods. In addition, the BRNN, which is less commonly considered in medical imaging than convolutional architectures, had also to be adapted for image data; thus, its performance might have been hampered in more complex or ambiguous cases (e.g., small stones or with overlapping pathologies).

Finally, in addition to detecting, the focus of the study is only on stone detection, leaving any characterization, for example, composition or fragility, which is significantly important in clinical decision-making-outside its coverage.

Conclusions

kidney stones can lead to the formation of solid masses made of minerals in the kidneys, and contribute to severe pain, an obstruction to urination, and respiratory distress complications in the kidney when they are undiagnosed early. Non-invasive imaging, especially CT scans, is crucial in accurate and quick detection of kidney stones and hence delves on clinical management. It proposes a new optimized deep learning framework for detecting

kidney stones in CT images of human subjects: a preprocessing stage, data augmentation stage, and enhanced bidirectional recurrent neural network (BRNN) for the DBES optimized algorithm (by a developed version of the bald eagle search). Data normalization is followed by noise reduction using the Wang-Mendel (WM) technique and global contrast enhancement and then SdSmote-based augmentation to offset class imbalance. The processed images will be inputted to a modified BRNN. The weights and biases will be fine-tuned using the DBES algorithm enhanced one that uses quasi-oppositional learning and chaotic initialization-for better convergence while avoiding local optima. Evaluated on the public CT Kidney Dataset, the proposed method outperformed several state-of-the-art techniques, like ensemble learning, Exemplar Darknet19, DE/SVM, and Decision Tree. The method achieved an accuracy level of 96.96, a sensitivity of 95.62, a specificity of 91.67, a precision of 94.38, an F1-score of 94.99, and a Jaccard Index of 91.61, confirming its robustness and diagnostic reliability. The research has, however, its share of limitations. Theoretically, a bidirectional recurrent neural network is not used regularly for 2D medical-image analysis because BRNNs are fitted only to sequential data, wherein the adaptation to spatial-image data could inhibit full exploitation of spatial hierarchies, unlike convolutional architectures. In addition, the DBES algorithm, while working nicely, is guided by heuristic-search principles that can in no case invoke rigorous convergence guarantees, meaning the performance is reliant on tuning the parameters and the structure of the problem. Practically, this study utilized a dataset limited geographically to patients from Bangladesh, which may affect the generalizability of the model to heterogeneous populations with different stone compositions, imaging protocols, or demographic profiles. The future work will consider strategies that address the limitations of the current approach through extending the model to processing 3D CT volumes with volumetric BRNNs or hybrid CNN-RNN architectures that could capture spatial-temporal features better. The DBES algorithm will undergo refinements through adaptive parameter control and hybridization with gradient-based methods to lessen the computational burden. Multi-center validations across different geographical and demographic cohorts will be performed to increase the robustness and fairness of the model. The explainability methods (Grad-CAM, SHAP, etc.) may be another avenue explored for providing interpretable visual feedback to radiologists. At last, the system would classify stone compositions, such as calcium oxalate or uric acid, which are of utmost importance in defining treatment types, such as lithotripsy or medical dissolution therapy. The proposed approach would have significant policy implications for healthcare systems and medical imaging standards. By automating and achieving accurate kidney stone detection, the system has concretized potential options for reducing delays in diagnostics, human error, and radiological work burdens—especially in resource-poor settings, where specialist availability is limited. Such AI work in the clinical workflow could give credence to national screening protocols for at-risk populations, such as those with a history of developing nephrolithiasis or metabolic derangements. In furtherance, health policymakers would require application of this technology to foster preventive care and thereby reduce hospitalization rates and long-term healthcare costs accruing from untreated or recurrent kidney stones. Lastly, regulatory agencies may also consider instituting guiding principles for the validation and rollout of AI-based diagnostic tools within the radiological sphere, thus assuring their safe, equitable, and transparent deployment. This study, therefore, not only extends technical innovations but also supports a wider outcome of scalable, fair, and efficient healthcare delivery.

Data availability

All data generated or analysed during this study are included in this published article.

Received: 21 December 2024; Accepted: 18 September 2025
Published online: 23 October 2025

References

1. Gong, C. et al. Gastric cancer detection using a hybrid version of gated recurrent unit network and adjusted tyrannosaurus optimization algorithm. *Biomed. Signal Process. Control.* **112**, 108582 (2026).
2. Zehao, W. et al. Optimal economic model of a combined renewable energy system utilizing modified. *Sustain. Energy Technol. Assess.* **74**, 104186 (2025).
3. Mannil, M. et al. Three-dimensional texture analysis with machine learning provides incremental predictive information for successful shock wave lithotripsy in patients with kidney stones. *J. Urol.* **200** (4), 829–836 (2018).
4. De Perrot, T. et al. Differentiating kidney stones from phleboliths in unenhanced low-dose computed tomography using radiomics and machine learning. *Eur. Radiol.* **29** (9), 4776–4782 (2019).
5. Rao, G. E. et al. Hybrid framework for respiratory lung diseases detection based on classical CNN and quantum classifiers from chest X-rays. *Biomed. Signal Process. Control*. **88**, 105567 (2024).
6. Abraham, A. et al. Machine learning prediction of kidney stone composition using electronic health Record-Derived features. *J. Endourol.* **36** (2), 243–250 (2022).
7. Kumar, Y. et al. Automating cancer diagnosis using advanced deep learning techniques for multi-cancer image classification. *Sci. Rep.* **14** (1), 25006 (2024).
8. Akbary, P. et al. Extracting appropriate nodal marginal prices for all types of committed reserve. *Comput. Econ.* **53** (1), 1–26 (2019).
9. Ghiasi, M. et al. Enhancing power grid stability: design and integration of a fast bus tripping system in protection relays. *IEEE Trans. Consum. Electron.* **71** (1), 561–570 (2024).
10. Halder, A. et al. Implementing vision transformer for classifying 2D biomedical images. *Sci. Rep.* **14** (1), 12567 (2024).
11. Cui, Y. et al. Automatic detection and scoring of kidney stones on Noncontrast CT images using STONE nephrolithometry: combined deep learning and thresholding methods. *Mol. Imaging Biology*. **23** (3), 436–445 (2021).
12. Navid Razmjooy, F. R. S. & Ghadimi, N. A hybrid neural Network – World cup optimization algorithm for melanoma detection. *Open Med.* **13**, 9–16 (2018).
13. Black, K. M. et al. Deep learning computer vision algorithm for detecting kidney stone composition. (2020).
14. Cui, X. et al. Analysis and classification of kidney stones based on Raman spectroscopy. *Biomedical Opt. Express*. **9** (9), 4175–4183 (2018).
15. Mosa, Q. O. Improved kidney stone detection from ultrasound images using GVF active contour. *J. Al-Qadisiyah Comput. Sci. Math.*, **14**(1), 24–31 (2022).

16. Ebrahimi, S. & Mariano, V. Y. Image quality improvement in kidney stone detection on computed tomography images. *J. Image Graphics*. **3** (1), 40–46 (2015).
17. Razmjooy, N., Ramezani, M. & Ghadimi, N. Imperialist competitive algorithm-based optimization of neuro-fuzzy system parameters for automatic red-eye removal. *Int. J. Fuzzy Syst.* **19** (4), 1144–1156 (2017).
18. Akkasaligar, P. T., Biradar, S. & Kumbar, V. Kidney stone detection in computed tomography images. in *International Conference On Smart Technologies For Smart Nation (SmartTechCon)*. 2017. IEEE. 2017. IEEE. (2017).
19. Edvardsson, V. O. et al. Temporal trends in the incidence of kidney stone disease. *Kidney Int.* **83** (1), 146–152 (2013).
20. Viswanath, K. & Gunasundari, R. Analysis and Implementation of Kidney Stone Detection by Reaction Diffusion Level Set Segmentation Using Xilinx System Generator on FPGA. VLSI Design, 2015. (2015).
21. Xu, Z. et al. Computer-aided diagnosis of skin cancer based on soft computing techniques. *Open Med.* **15** (1), 860–871 (2020).
22. Viswanath, K., Gunasundari, R. & Hussan, S. A. VLSI implementation and analysis of kidney stone detection by level set segmentation and ANN classification. *Procedia Comput. Sci.* **48**, 612–622 (2015).
23. Ma, F. et al. Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network. *Future Generation Comput. Syst.* **111**, 17–26 (2020).
24. Cai, W. et al. Optimal bidding and offering strategies of compressed air energy storage: A hybrid robust-stochastic approach. *Renew. Energy*. **143**, 1–8 (2019).
25. Gong, Z., Li, L. & Ghadimi, N. SOFC stack modeling: a hybrid RBF-ANN and flexible Al-Biruni Earth radius optimization approach. *Int. J. Low-Carbon Technol.* **19**, 1337–1350 (2024).
26. Yuan, K. et al. Optimal parameters Estimation of the proton exchange membrane fuel cell stacks using a combined Owl search algorithm. *Energy Sour. Part A Recover. Utilization Environ. Eff.* **45** (4), 11712–11732 (2023).
27. Viswanath, K. & Gunasundari, R. Design and analysis performance of kidney stone detection from ultrasound image by level set segmentation and ANN classification. in *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. 2014. IEEE. 2014. IEEE. (2014).
28. Lazar, M. et al. Protocol analysis of dual-energy CT for optimization of kidney stone detection in virtual non-contrast reconstructions. *Eur. Radiol.* **30** (8), 4295–4305 (2020).
29. Cunitz, B. et al. Improved detection of kidney stones using an optimized Doppler imaging sequence. in. *IEEE International Ultrasonics Symposium*. 2014. IEEE. (2014).
30. Graham-Knight, J. B. et al. Accurate Kidney Segmentation in CT Scans Using Deep Transfer Learning. in *International Conference on Smart Multimedia*. Springer. (2019).
31. Liu, Haozhi, and Noradin Ghadimi. Hybrid convolutional neural network and Flexible Dwarf Mongoose Optimization Algorithm for strong kidney stone diagnosis. *Biomedical Signal Processing and Control* **91**, 106024 (2024),
32. Razmjooy, N., Sheykhahmad, F. R. & Ghadimi, N. A hybrid neural network–world cup optimization algorithm for melanoma detection. *Open Med.* **13** (1), 9–16 (2018).
33. Han, M. et al. Timely detection of skin cancer: an AI-based approach on the basis of the integration of echo state network and adapted seasons optimization algorithm. *Biomed. Signal Process. Control.* **94**, 106324 (2024).
34. Ranjbarzadeh, R. et al. Nerve Optic Segmentation in CT Images Using a Deep Learning Model and a Texture Descriptorp. *Complex & Intelligent Systems*. 1–15 (2022).
35. Hu, A. & Razmjooy, N. Brain tumor diagnosis based on metaheuristics and deep learning. *Int. J. Imaging Syst. Technol.* **31** (2), 657–669 (2021).
36. Tian, Q. et al. A new optimized sequential method for lung tumor diagnosis based on deep learning and converged search and rescue algorithm. *Biomed. Signal Process. Control.* **68**, 102761 (2021).
37. Razmjooy, N., Mousavi, B. S. & Soleymani, F. A real-time Mathematical Computer Method for Potato Inspection Using Machine Vision. *Computers & Mathematics with Applications.* **63**(1) 268–279 (2012).
38. Zivkovic, M. et al. Hybrid CNN and XGBoost model tuned by modified arithmetic optimization algorithm for COVID-19 early diagnostics from X-ray images. *Electronics* **11** (22), 3798 (2022).
39. Jovanovic, L. et al. Detecting parkinson's disease from shoe-mounted accelerometer sensors using convolutional neural networks optimized with modified metaheuristics. *PeerJ Comput. Sci.* **10**, pe2031 (2024).
40. Bacanin, N. et al. Parkinson's Disease Detection with Deep Long Short-term Memory Networks Optimized by Modified Metaheuristic Algorithm, in Nebojsa Bacanin, Luka Jovanovic, Miodrag Zivkovic, Tamara Zivkovic, Petar Bisevac, Milos Dobrojevic, Marko Sarac, Milos Antonijevic (eds.), Collective Intelligence. CRC. (2024) 204–229.
41. Jovanovic, L. et al. Anomaly detection in Ecg using recurrent networks optimized by modified metaheuristic algorithm. in 2023 31st Telecommunications Forum (TELFOR). IEEE. Belgrade, Serbia, pp. 1-4. https://doi.org/10.1109/TELFOR59449.2023.10372802 (2023).
42. Babic, L. et al. Click fraud detection with recurrent neural networks optimized by adapted crayfish optimization algorithm. *J. Ind. Intell.* **2** (4), 230–239 (2024).
43. Mizdrakovic, V. et al. Forecasting bitcoin: decomposition aided long short-term memory based time series modeling and its explanation with Shapley values. *Knowl. Based Syst.* **299**, 112026 (2024).
44. Mladenovic, D. et al. Sentiment classification for insider threat identification using metaheuristic optimized machine learning classifiers. *Sci. Rep.* **14** (1), 25731 (2024).
45. Yuan, Z. et al. A new technique for optimal Estimation of the circuit-based PEMFCs using developed sunflower optimization algorithm. *Energy Rep.* **6**, 662–671 (2020).
46. Derrac, J. et al. A practical tutorial on the use of nonparametric statistical tests as a methodology for comparing evolutionary and swarm intelligence algorithms. *Swarm Evol. Comput.* **1** (1), 3–18 (2011).
47. Ghadimi, N. A method for placement of distributed generation (DG) units using particle swarm optimization. *Int. J. Phys. Sci.* **8** (27), 1417–1423 (2013).
48. Saeedi, M. et al. Robust optimization based optimal chiller loading under cooling demand uncertainty. *Appl. Therm. Eng.* **148**, 1081–1091 (2019).
49. Yu, D. et al. Energy management of wind-PV-storage-grid based large electricity consumer using robust optimization technique. *J. Energy Storage*. **27**, 101054 (2020).
50. Ghadimi, N., Afkousi-Paqaleh, A. & Emamhosseini, A. A PSO-based fuzzy long-term multi-objective optimization approach for placement and parameter setting of UPFC. *Arab. J. Sci. Eng.* **39** (4), 2953–2963 (2014).
51. Yu, D. et al. System identification of PEM fuel cells using an improved Elman neural network and a new hybrid optimization algorithm. *Energy Rep.* **5**, 1365–1374 (2019).
52. Wu, G., Mallipeddi, R. & Suganthan, P. N. Problem definitions and evaluation criteria for the CEC 2017 competition on constrained real-parameter optimization. National University of Defense Technology, Changsha, Hunan, PR China and Kyungpook National University, Daegu, South Korea and Nanyang Technological University, Singapore, Technical Report, (2017).
53. Kaveh, A., Khanzadi, M. & Moghaddam, M. R. Billiards-inspired optimization algorithm; a new meta-heuristic method. in *Structures*. Elsevier. (2020).
54. Simon, D. Biogeography-based optimization. *IEEE Trans. Evol. Comput.* **12** (6), 702–713 (2008).
55. Hatamlou, A. Black hole: A new heuristic optimization approach for data clustering. *Inf. Sci.* **222**, 175–184 (2013).
56. Dhiman, G. & Kumar, V. Emperor Penguin optimizer: A bio-inspired algorithm for engineering problems. *Knowl. Based Syst.* **159**, 20–50 (2018).

57. Guo, Y. et al. An optimal configuration for a battery and PEM fuel cell-based hybrid energy system using developed Krill herd optimization algorithm for locomotive application. *Energy Rep.* **6**, 885–894 (2020).
58. Sun, L. et al. Exergy analysis of a fuel cell power system and optimizing it with Fractional-order Coyote optimization algorithm. *Energy Rep.* **7**, 7424–7433 (2021).
59. Gong, Chaofan, et al. "Gastric cancer detection using a hybrid version of gated recurrent unit network and adjusted tyrannosaurus optimization algorithm." Biomedical Signal Processing and Control 112 (2026): 108582.
60. Martínez, A. et al. Towards an automated classification method for ureteroscopic kidney stone images using ensemble learning. in. *42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. 2020. IEEE. (2020).
61. Baygin, M. et al. Exemplar Darknet19 feature generation technique for automated kidney stone detection with coronal CT images. *Artif. Intell. Med.* **127**, 102274 (2022).
62. Soni, A. & Rai, A. Kidney Stone Recognition and Extraction using Directional Emboss & SVM from Computed Tomography Images. in 2020 *Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT)*. IEEE. (2020).

Acknowledgements
This work was sponsored in part by Guangxi Natural Science Foundation Project/General Project: Research on Strategy and Mechanism of Prognostic Classification of high incidence liver cancer in Guangxi Based on Single cell multi-omics Technology (Joint Special Project of Regional High Incidence Disease Research) (2023GXNSFAA0262292345678).

Author contributions
Zhenyou Tang, Zhenyu Tang, Yuxin Liu, Long Chen, Zhong Tang, Xiaohong Liu and Togrol Salami wrote the main manuscript text and prepared figures. All authors reviewed the manuscript.

Declarations

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to X.L. or T.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.