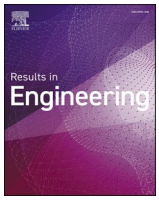
Results in Engineering 26 (2025) 105221

Contents lists available at ScienceDirect

Results in Engineering

journal homepage: www.sciencedirect.com/journal/results-in-engineering



Optimised hybrid deep learning classification model for kidney stone diagnosis

Y Jini Jacob [a], Bethanney Janney J [a,*], Hemalatha RJ [b], Preethi S [c]

[a] Department of Biomedical Engineering, School of Bio and Chemical Engineering, Sathyabama Institute of Science and Technology, Chennai, Tamil Nadu, India
[b] Department of Biomedical Engineering, Vels Institute of Science, Technology and Advanced Studies, Chennai 600117, Tamil Nadu, India
[c] ISE Cambridge Institute of Technology, Bengaluru 560036, India

ARTICLE INFO

Keywords:
Kidney stone detection
Alexnet
GRU model
Elephant herd optimizer
Hyperparameter
performance analysis

ABSTRACT

The kidney plays a vital role in maintaining homeostasis within the human body. In recent years, the prevalence of nephrolithiasis (kidney stone formation) characterized by the accumulation of crystalline solids within the renal system has emerged as a significant health concern. Early detection is critical for effective treatment and prevention of complications. Diagnostic imaging techniques such as computed tomography (CT), ultrasonography, and Doppler imaging are routinely employed for this purpose. To enhance the precision and reliability of early diagnosis, Deep Learning (DL) models are increasingly being integrated into the diagnostic workflow, offering superior accuracy through advanced image analysis and pattern recognition capabilities. The proposed work combines two deep learning models, AlexNet and Gated Recurrent Unit (GRU) for feature extraction and classification. These models are integrated to deliver optimal training parameter performance. An optimized AlexNet-GRU model is introduced in this work for detection of kidney stone, feature extraction, and classification. The Elephant Herding Optimizer (EHO) is utilized to fine-tune the hyperparameters of the AlexNet-GRU model. by performing this EHO fine tuning, the performance metrics of the proposed work have provided a high optimal result. Finally, the proposed evaluation metrics like precision, recall, accuracy, and F1 score are evaluated and compared with the traditional models to prove their efficient performances. The proposed model achieved a precision of 98.67 %, a recall of 97.68 %, an accuracy of 98.82 %, and an F1 score of 97.54 %.

1. Introduction

The kidney is one of the most important organs in the human body. Kidney stones are one of the most widespread and serious issues that affect the kidney. This stone formation occurs in both male and female genders and is also formed in all the category people. The kidney stone is like a solid material piece that is formed due to the minerals present in urine [1]. It is formed by genetic and environmental factors. All individuals who do not care about their health, such as those who are overweight, consume a wide variety of meals, do not drink enough water, and take excessive amounts of medication, are affected by these stones.

To diagnose this kidney stone, blood tests, urine tests, and scans are used. When the stones are recognized in an early stage, the over pain and surgeries are not required [2]. Therefore, earlier detection is the most important factor in it. Image processing is an effective technique that is used for detection in earlier stages. The imaging techniques are processed by examining the internal organs using medical image scanning like Doppler scans, CT scans, and Ultrasound scans. Sometimes the results may provide inaccurate predictions and insufficient methods [3].

Nowadays, for an accurate prediction and to provide an effective result, medical imaging is performed with Machine Learning (ML) algorithms and Deep Learning (DL) techniques [4]. There are several models of ML and DL are used for the feature extraction and classification of image processing techniques [5]. Some of the popular ML models used for medical imaging are Naive Bayes, Decision Tree, K-nearest Neighborhood (KNN), Support vector machine (SVM), Artificial neural network (ANN), and Convolutional Neural Networks [6]. The DL has multiple CNN architectures that provide effective learning and training datasets. Some of the popular DLs are AlexNet, GoogleNet, ShuffleNet, InceptionNet, VGGNet, ResNet, Long Short-Term Memory (LSTM), Recurrent Neural Networks (RNN), GRU (Gated Recurrent Unit) model, and so on [7].

Combining AlexNet and GRU provides a balanced approach to

* Corresponding author.
E-mail addresses: bethanneyj@gmail.com (B.J. J), hemalatharj.se@velsuniv.ac.in (H. RJ), hod.ise@cambridge.edu.in (P. S).

https://doi.org/10.1016/j.rineng.2025.105221
Received 31 January 2025; Received in revised form 16 April 2025; Accepted 5 May 2025
Available online 10 May 2025

kidney stone detection. AlexNet handles complex feature extraction from images, while GRU models temporal or sequential relationships, allowing the model to account for changes over time or context from a series of images. This hybrid model is well-suited for tasks where both image and sequential data need to be considered for accurate diagnosis, offering a more comprehensive solution than traditional models focusing on either one alone.

For kidney stone prediction in this work, the DL model is employed. The hybrid of two different DL models, AlexNet and the GRU model, is provided in the proposed work. To obtain a highly optimal solution in classification, the hybrid AlexNet and GRU model is optimized by using the EHO model. The hyperparameters of the DL models employed in this work are finetuned using the EHO model. Finally, the new model's classification results are assessed and compared to those of the current models to demonstrate its efficacy.

Previous studies have also reported on AlexNet hybrid for kidney disease diagnosis [8] by utilizing a comprehensive dataset of 12,446 CT whole abdomen and urogram images, this study developed an advanced AI-driven diagnostic system specifically tailored for kidney disease classification by integrating AlexNet's robust feature extraction capabilities with ConvNeXt's advanced attention mechanisms, the paper achieved an exceptional classification accuracy of 99.85 %. The model of this study demonstrated outstanding performance across various metrics, with an average precision of 99.89 %, recall of 99.95 %, and specificity of 99.83 %. These results highlight the efficacy of the hybrid architecture and optimization strategy in accurately diagnosing kidney diseases. In contrast, this work introduces a novel hybrid approach by integrating AlexNet with a GRU that enhances pattern recognition for improved kidney stone classification. Additionally, the Elephant Herding Optimizer (EHO) is employed to fine-tune hyperparameters, ensuring optimal performance while reducing computational complexity. Furthermore, comparative analysis against traditional and state-of-the-art models was done, demonstrating superior performance with a precision of 98.67 %, recall of 97.68 %, accuracy of 98.82 %, and an F1-score of 97.54 %.

The remaining contributions are outlined as follows. Section 2 is about the related works of this paper. The preliminary part is described in Section 3. Next, the materials and methods are discussed the proposed model and its workflow in Section 4 and Section 5 carried a result and discussion of proposed and existing models. Section 6 discussed a conclusion followed by its references.

2. Related works

This section carried literature based on kidney stone detection using DL models and image processing. Each work is presented by various author and their different ideas in kidney stone detection.

Karaman, et al. (2022) presented an Aggregate Channel Features (ACF) method. This method is based on the ML method that is used for the kidney's automatic detection. The k-fold cross-correlation and confusion matrix methods increased the detection performance [9]. Next, an automated kidney stone detection using a VGG model which is developed by Mohan, et al. (2022). This method uses an publicly available Computed Tomography (CT) image dataset for stone detection [10]. Another work is discussed by Rajput, et al. (2022) which concentrated on overcoming kidney abnormalities such as stone formation, congenital anomalies, cysts, cancerous cells, and urine blockage. This work is used for automatic detection without human intervention [11].

Then, utilizing coronal computed tomography (CT) images and a cross-Residual Network (XResNet-50) architecture, Yildirim, et al. (2021) presented an automated identification of a kidney stone. The proposed automatic system demonstrated a precision of 96.82 % using CT scans to find kidney stones of any size [12]. Lim EJ, et al. (2022) presented the best potential lies in identifying the stones. Here, clinical, molecular, and imaging technologies are used for stone removal. Finally, the segmented images detected the stone's size and location in the affected area [13].

Suresh et al. (2021) presented a method for pre-processing, segmentation, and Morphological Analysis. This work calculated an output parameter and provided a better result than the existing models [14]. In some cases, the contrast, second angular moment, entropy, and correlation are developed by an author Dave et al. (2022). This work used a KNN classification for the training dataset to provide better accuracy. The confusion matrix can be obtained with higher accuracy [15]. The work developed by Myint, et al. (2020) presented an automatic 3D-visualized kidney stone detection. It has several steps, namely i) intensity-based thresholding is applied to remove a hypodense and isodense region, ii) size-based thresholding is presented to remove a bone abdomen, and iii) the false positive is reduced by a geometric feature-based thresholding [16].

Liu et al. (2022) built a ResNet model to categorize Kidneys, Ureters, and Bladder (KUB) images by assessing whether kidney calculi is present or not. Because of several parameters, the suggested model exhibits outstanding effectiveness in classification and can be applied to the quick identification of kidney calculi from standard film X-ray images [17].

Sharen et al. [18] proposed a stone classification model using Sein transformer models. Swin Transformer uses shftwe windows to learn the features deeply. Chaki et al. [19] proposed an ensemble model for kidney stone detection. It combines different learning models with parameter tuning for accurate detection. Pande et al. [20] introduced the YOLOv8 model-based kidney stone detection model. Compared to other models, the YOLOv8 model requires more number of computational units and memory.

3. Preliminaries

In this work, the hybrid DL model of AlexNet and the GRU model is used for classification. Therefore, the basic ideas of both the models and their architecture are presented in this section.

3.1. Alexnet model

AlexNet is a DL model that is based on the CNN method [21]. This method is most significantly used for objection detection, feature extraction, image classification, and so on. In the year 2012, the AlexNet model scored second place in the Image Net LSVRC-2012 competition with a 15.3 % error rate vs. a 26.2 % error rate. This model was an advanced model of the LeNet network. The AlexNet techniques consist of convolution (Conv), ReLU activations, max pooling, and dropout layers which are shown in Fig. 1.

Fig. 1 shows five Conv layers and three full connection layers. After the three-conv layer, the maximum pooling layer is operated. The Rectified Linear Unit (ReLU) is used as the activation function in the AlexNet architecture instead of the sigmoid and tanh activation functions. The ReLU activation function has controlled the gradient disappearance and gradient explosion. Also, this activation function is very simple for training and learning a deeper network which is expressed in the below equation.

$$ReLU(x) = \max(0, x) \tag{1}$$

The dropout layer is used to minimize an overfitting degree in a training process. The neurons in the layers are stopped with a convincing probability. This process may reduce the local node's dependencies and improve generalization ability. It has a huge Conv kernel that increases the number of parameters and makes a local feature in the feature extraction. Simultaneously, the full connection layer is large and the features extracted from Conv influenced the performances and also improved the Conv parameters proportion.

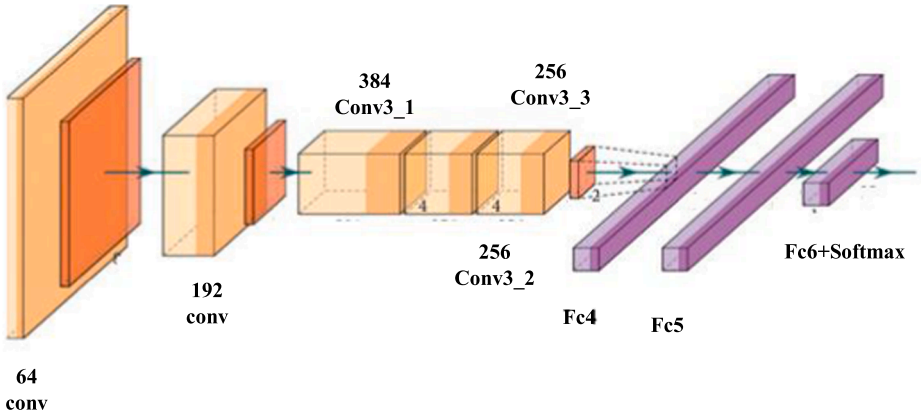


Fig. 1. AlexNet Architecture.

3.2. GRU model

The GRU is an upgraded version of LSTM that was developed by Cho, et al. (2014) [22]. The LSTM model has an input gate, forget gate, and update gate in its architecture. Similarly, the GRU model is simpler than the LSTM model which has only a update gate and reset gate. To recognize the usefulness of data, these gates are used. In the GRU model, relevant data is retained while irrelevant data is discarded [23].

From Fig. 2, the GRU architecture is shown in it. The backward and forward data is processed with an update gate (z), and the reset gate (r) provides the previous data knowledge in it. The reset gate is used as a present memory gate to preserve and retain the essential data from the previous state. The nonlinearity is provided as an input using an input modulation gate with zero mean characteristics. Thus, the GRU model gates are expressed in the following.

$$r_t = \sigma(X_t.\ W_{xr} + H_{t-1}.\ W_{hr} + b_r) \tag{2}$$

$$z_t = \sigma(X_t.\ W_{xz} + H_{t-1}.\ W_{hz} + b_z) \tag{3}$$

where W_{xr} indicates the weight parameter of the reset gate and W_{xz} indicates the weight parameters of update gates. P-value~0.01 to 0.05 and CI value is 95 %. The b_r and b_z denote a biased reset and update gate.

EfficientNet and Vision Transformers are great for tasks where spatial information is critical, like image classification or segmentation of static medical images (e.g., CT scans, X-rays). EfficientNet offers efficiency, while ViTs offers superior performance when there's a need to capture long-range dependencies across the image. U-Net is specialized for segmentation, where pixel-level accuracy is needed, and it excels in the medical imaging field. GRUs, on the other hand, would be particularly useful when the medical imaging task involves temporal sequences (e.g., MRI scans over time or 3D imaging slices) and needs to capture sequential patterns rather than just spatial features.

4. Materials and methods

The materials and methods used in this proposed model are presented in this section. The overall proposed workflow is shown in Fig. 3 which comprises a dataset, pre-processing, proposed methodology, and performance analysis. The proposed methodology presented an optimized AlexNet-GRU model to provide accurate kidney stone detection and efficient classification performances.

4.1. Dataset collection

From the open-source website https://github.com/yildirimozal/Kidney_stone_detection, the kidney stone-affected CT datasets are downloaded. The dataset consists of 1799 images, which include 858 kidney stone images and 941 normal images. The input CT datasets are divided for both the testing and training processes. In this work, 80 % of the datasets are used for training, and 20 % of the datasets are used for testing.

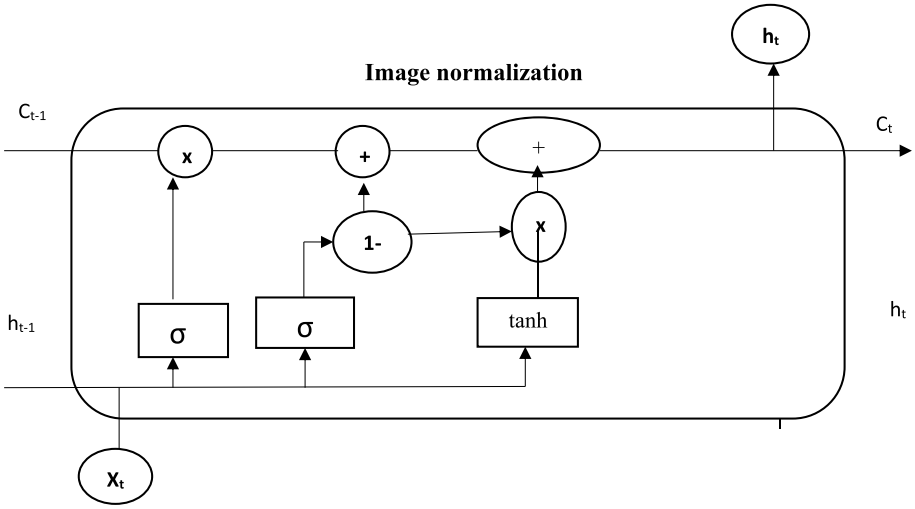


Fig. 2. Architecture of GRU Model.

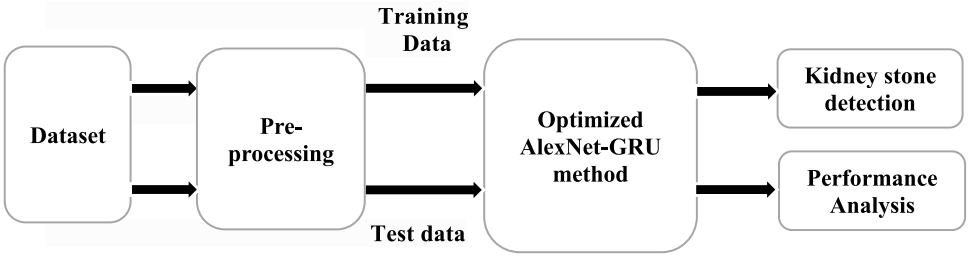


Fig. 3. Overall workflow.

4.2. Pre-processing

By artificially boosting the quantity of training images, data augmentation techniques are essential for improving the effectiveness of DL models. Several augmentation techniques can be used to enhance model generalization and decrease overfitting in kidney stone identification using CT scans. Variations in image orientation and positioning can be produced with the aid of geometric transformations including rotation, flipping, scaling, translation, and shearing. Gamma correction and brightness and contrast adjustments are examples of intensity transformations that mimic various imaging circumstances. By simulating scanner artifacts, noise injection techniques like Gaussian noise and salt-and-pepper noise improve resilience. Affine transformations and elastic deformations create minute distortions to mimic changes in anatomical features. When combined, these augmentation techniques lead to increased performance in kidney stone detection tasks, higher generalization, and increased model accuracy.

5. Proposed methodology

The processed dataset is applied as and input to the proposed block which contains three algorithms namely the AlexNet model, GRU model, and EHO Model as shown in Fig. 4. Firstly, the hybrid AlexNet-GRU model is discussed in it.

5.1. Hybrid Alexnet-GRU model

Initially, the image input is reshaped into image height as 60 pixels and width as 60 pixels with RGB, and the 3 channels i.e., (60, 60, 3). This image is sent to a first Conv layer for feature extraction. The Conv-1 has a feature map output shape of 128 in number. It has a Conv stride of (3 × 3) and a kernel size of 1. The ReLU function is used to reduce and decrease the nonlinearity dimension issue. Next, the pooling layer minimized the output feature map as (58, 58) size for 128 feature maps to speed up the computations. The training parameter of (58, 58, 128) was transferred to the dropout layer. It has a 0.9 dropout in the Conv layer to overcome an overfitting issue. Thus, the training parameter is reduced after the Conv and max-pooling process. Later the Conv and max-pooling are composed into an ID array and transfer the data into the fully connected layer. After every Conv process implementation, the dropout has provided 1024 feature maps. These 1024 feature maps are processed by a GRU model to solve an issue of gradient vanishing. After gradient vanishing, other fully connected layers are performed and SoftMax operations are implemented.

5.2. Optimised AlexNet-GRU model

The optimized AlexNet-GRU Model is proposed for feature extraction and classification. The EHO model is implemented to fine-tune the hyperparameter of the hybrid AlexNet-GRU model to provide an optimal solution. Thus, the EHO model is discussed in the following.

5.3. Elephant herding optimization (EHO) model

The EHO model is a metaheuristics model that is based on the herding behavior of elephants [24]. This model can solve an issue of global optimization based on three rules. 1) The elephant population is estimated to consist of several clans, each containing a number of elephants. 2) Each generation saw a number of male elephants leaving their family and distancing themselves from the main elephant group 3) each clan of elephants lived under the matriarch leadership. The EHO model is evaluated with two processes namely the Clan separating operator and updating operator respectively.

The Clan updating behavior is expressed with an elephant j is given in the following.

$$X_{new,\ ci,j} = X_{ci,j} + \alpha \times \left(X_{best,ci} - X_{ci,j}\right) \times r \tag{4}$$

Where $X_{new,\ ci,j}$ denotes a newly updated position, $X_{ci,j}$ indicates an old position, and $X_{best,ci}$ denotes the best position.

The fittest value is updated by Eq. (1) as $X_{ci,j} = X_{best,ci}$. The fittest

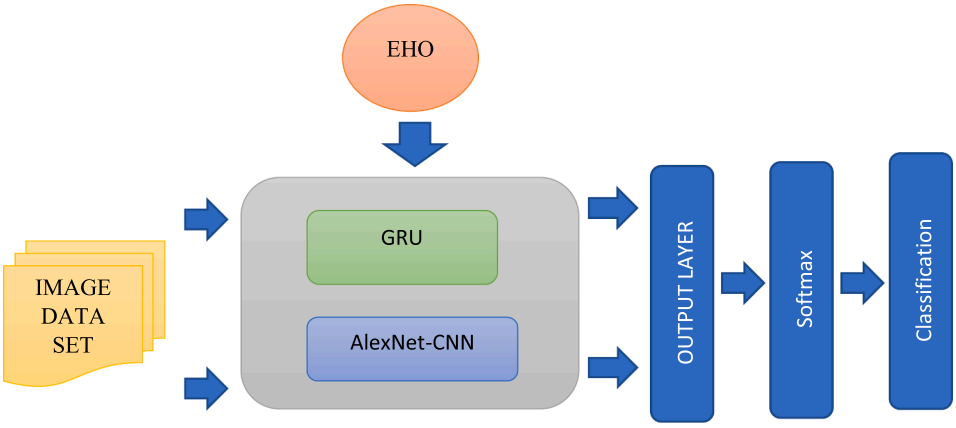


Fig. 4. Proposed Block Diagram.

value is updated as

$$X_{new,\ ci,j} = \beta \times x_{cen,ci} \tag{5}$$

Where beta belongs to [0,1], a factor that influences the xcenter, ci on xnew, ci, j.

$$x_{cen,ci,d} = \frac{1}{n_{ci}} \times \sum_{j=1}^{n_{ci}} x_{ci,j,d} \tag{6}$$

where $1 \leq d \leq D$ represents the D-th dimension, and D - total number of dimensions, while n_{ci} denotes the elephants number in clan ci.

5.4. Separating operator

Consider an elephant entity with the lowest fitness and provide a separating operation at every generation. Thus, the lowest elephant entity in clan ci ($x_{wor,ci}$) at [0, 1] is expressed in Eq. (7)

$$x_{wor,ci} = x_{min} + (x_{max} - x_{min} + 1) \times rand \tag{7}$$

where x_{max} indicates an upper bound position of the elephant andx_{min} represents a lower bound position of the elephant. xworst, ci - lowest elephant entity in clan ci.

Based on the above numerical calculation, the pseudocode of the EHO Model is given in algorithm 1.

Elephant Herding Optimization is a powerful, flexible tool in the context of CT-based kidney stone detection. It can be applied at multiple stages: Pre-processing, Segmentation, Feature Selection, and Model Tuning. It adds value, especially when dealing with high-dimensional data and the need for robust, interpretable, and efficient models.

6. Results and discussion

The results and discussions of the proposed work are presented to prove the effectiveness of its performance. The CT images are used to process a classification and feature extraction based on the proposed strategy. These metrics are compared with the traditional DL models such as AlexNet, ResNet, GoogleNet, DenseNet, MobileNet, CNN-LSTM, and SVM models, respectively. Fig. 5 shows the segmented kidney stone images using the proposed hybrid model. The highlighted regions correspond to stone formations, validating the deep learning model's accuracy. This approach aids radiologists in improving diagnostic efficiency, reducing manual errors, and facilitating early detection of nephrolithiasis. Further evaluation using clinical data is necessary to assess real-world applicability and reliability.

Fig. 6 shows the layers of the proposed model used for training. The model processes input images of size 224 × 224 × 3224 × 224 × 3 through multiple convolutional layers. Each convolutional block enhances feature representation, with increasing filter depths (from 64 to 512) capturing intricate patterns.

The mathematical representation of performance metrics is expressed below.

$$precision = \frac{T_{Positive}}{T_{Positive} + F_{Positive}} \tag{8}$$

$$Recall = \frac{T_{Positive}}{T_{Positive} + F_{Negative}} \tag{9}$$

$$Accuracy = \frac{T_{Positive} + T_{Negative}}{T_{Positive} + T_{Negative} + F_{Positive} + F_{Negative}} \tag{10}$$

Where $T_{Positive}$ indicates true positive, $F_{Positive}$denotes a False positive, $F_{Negative}$ indicates a False Negative and $T_{Negative}$ denotes a True Negative.

$$\text{F1 measure} = \frac{2Precision \times Recall}{Precision + Recall} \tag{11}$$

The comparative analysis of various deep learning and machine learning models shown in Table 1. With an accuracy of 98.82 %, precision of 98.67 %, recall of 97.68 %, and an F1 score of 97.54, the proposed model outperforms well-established architectures such as AlexNet, ResNet50, and DenseNet. While models like AlexNet and DenseNet exhibit competitive accuracy levels above 95 %, their recall and F1 scores remain slightly lower. Traditional machine learning models, such as SVM, show the lowest performance across all metrics, reinforcing the advantages of deep learning-based approaches in handling complex medical imaging tasks. The results highlight the model's robustness in minimizing false positives and false negatives, making it a reliable tool for early diagnosis and detection.

The training performance of the proposed deep learning model is illustrated through accuracy and loss curves over 50 epochs as shown in Fig. 7. The accuracy plot demonstrates a steady increase, with both training and test accuracy surpassing 98 %, indicating high model reliability. The loss graph shows a significant decline in the initial epochs, stabilizing at a minimal value, suggesting effective learning and convergence. Notably, the test loss remains slightly lower than the training loss, implying strong generalization and minimal overfitting.

Fig. 8 visualizes the performance metrics of different models, including precision, recall, accuracy, and F1-score. The proposed model outperforms all other architectures, achieving the highest values across all metrics, indicating superior classification performance. AlexNet, ResNet50, and MobileNet demonstrate competitive accuracy but fall

Algorithm 1
pseudocode of EHO model.

```
Initialize population, Maximum generation MaxG
Set generation t = 1
While t<MaxG do
Based on fitness of all Elephants
Update clan operator
For ci=1to nclan do
  For j = 1 to nci do
    Update Eq. (5)
    If x_ci,j = x_best,ci then
    Update Eq. (5)
    End if
  End for
End for
Perform Separating operator
For ci=1 to nclam do
  Replace worst using Eq. (7)
End for
Calculate population by new position
t = t + 1
end while
```

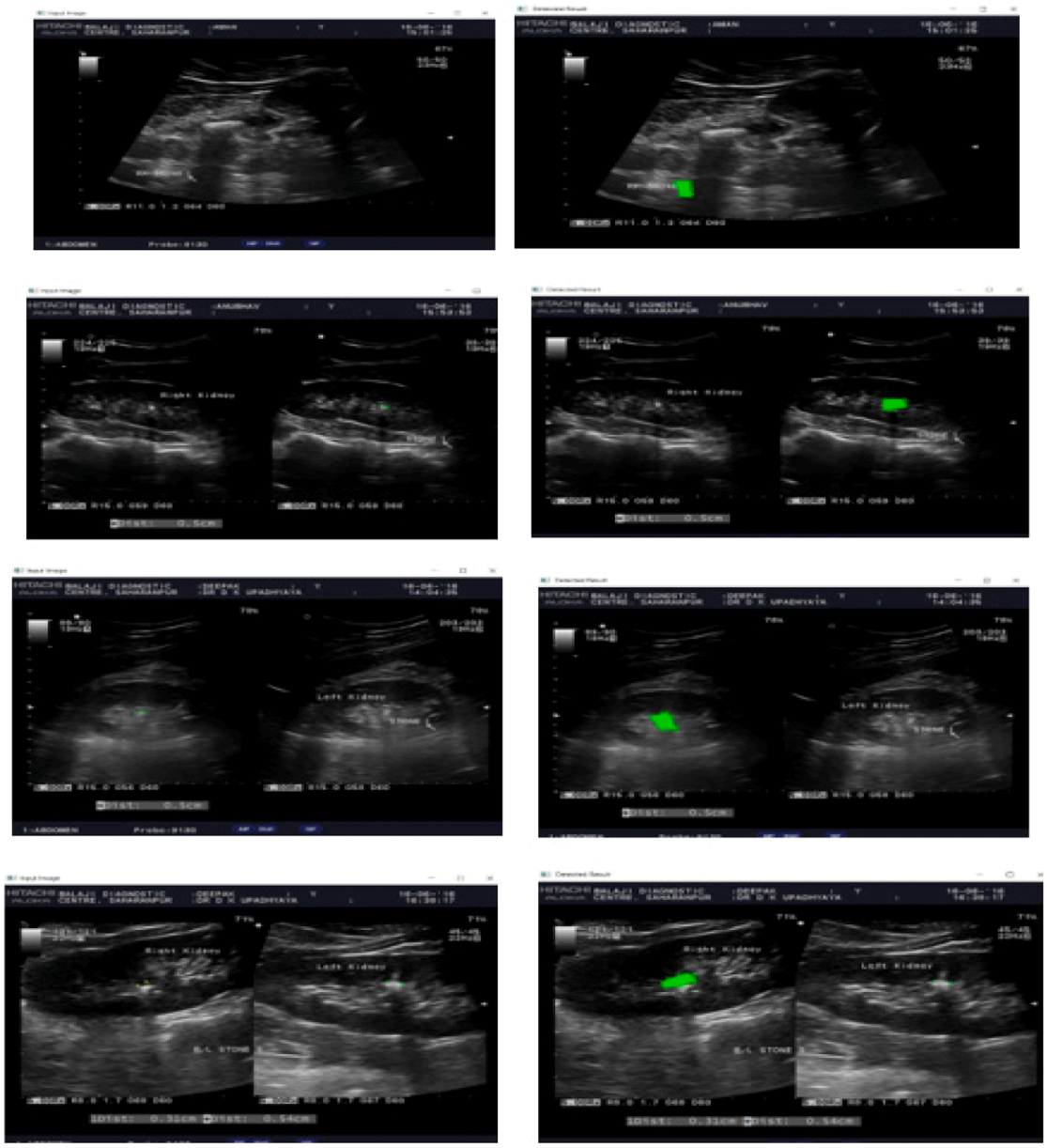


Fig. 5. Input and corresponding processed images.

slightly behind in recall and F1-score. CNN-LSTM, LSTM, and SVM exhibit relatively lower performance, with GoogleNet having the lowest recall.

Fig. 9 illustrates the precision values of various models. The proposed model achieves the highest precision, indicating its ability to minimize false positives effectively. MobileNet and AlexNet also demonstrate strong precision, while ResNet50, LSTM, and CNN-LSTM perform moderately well. In contrast, GoogleNet, DenseNet, and SVM exhibit the lowest precision, suggesting a higher rate of misclassification. The results emphasize the effectiveness of the proposed model in making accurate predictions with minimal errors.

Fig. 10 represents the recall values of different models. The proposed model achieves the highest recall, indicating its strong ability to correctly identify positive cases with minimal false negatives. AlexNet and ResNet50 also show high recall scores, reflecting their effectiveness in capturing true positives. In contrast, models like GoogleNet, DenseNet, MobileNet, CNN-LSTM, and LSTM exhibit comparatively lower recall values, suggesting they may miss more relevant instances. SVM performs moderately well but does not surpass the proposed model. These results highlight the superior recall performance of the proposed model, making it more reliable for applications requiring high sensitivity.

Fig. 11 illustrates the accuracy of different models in classification tasks. The proposed model achieves the highest accuracy, surpassing all other models, demonstrating its superior performance in correctly identifying both positive and negative cases. ResNet50 and MobileNet also exhibit high accuracy, indicating their effectiveness in classification. AlexNet and DenseNet show slightly lower accuracy, but they still perform well. In contrast, CNN-LSTM and SVM display the lowest accuracy among the models, suggesting potential limitations in their classification capability. These results highlight the effectiveness of the proposed model in achieving highly accurate predictions.

Fig. 12 illustrates the F1 measure of different models, which balances precision and recall to evaluate overall performance. The proposed model achieves the highest F1 score, indicating its superior balance between correctly identifying positive cases while minimizing false positives and false negatives. MobileNet and AlexNet also perform well, with competitive F1 scores, followed by ResNet50 and CNN-LSTM,

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 224, 224, 3)]	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808

Fig. 6. Training of layers in the proposed model.

Table 1
performance analysis of proposed and existing models.

Model	Precision	Recall	Accuracy	F1 score
Proposed	98.67	97.68	98.82	97.54
AlexNet	94.23	94.35	95.12	95.29
ResNet50	92.40	92.27	97.59	92.39
GoogleNet	89.43	90.37	94.28	91.68
DenseNet	89.87	90.14	95.43	95.65
MobileNet	95.30	89.09	96.52	93.90
CNN-LSTM	91.85	89.26	89.89	92.87
LSTM	92.87	90.23	95.30	94.48
SVM	89.63	93.81	89.68	90.54

which exhibit slightly lower performance. GoogleNet, DenseNet, and LSTM maintain moderate F1 scores, while SVM records the lowest, suggesting potential inefficiencies in handling classification tasks. These results reinforce the effectiveness of the proposed model in achieving optimal classification performance.

Fig. 13 shows the plot represents the convergence behaviour of an optimization algorithm over multiple iterations. The x-axis denotes the number of iterations, while the y-axis represents the best fitness value obtained so far. Initially, there is a steep decline in the fitness value within the first few iterations, indicating rapid progress in optimization. After this phase, the fitness value stabilizes, suggesting that the algorithm has reached a near-optimal or optimal solution. The near-horizontal trend after the initial drop signifies convergence, meaning further iterations do not significantly improve the solution.

7. Conclusion

In recent times, kidney stones have become the most common disease that occurred for all the category people. The pain is unbearable for anyone and that can be solved in an earlier stage, In this paper, the optimized hybrid DL models are presented for an accurate earlier

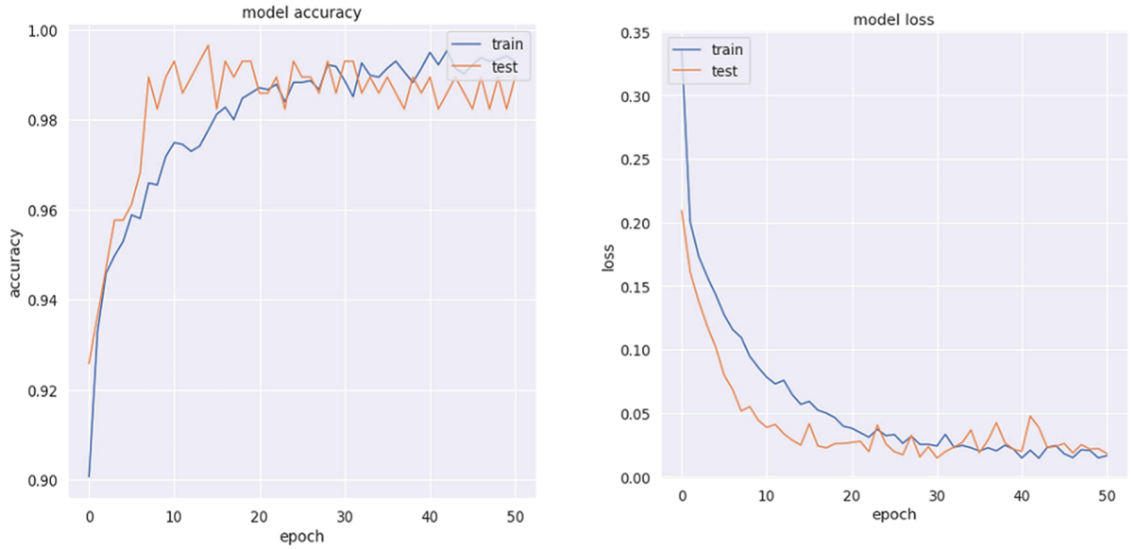


Fig. 7. Accuracy and validation curve of the proposed model.

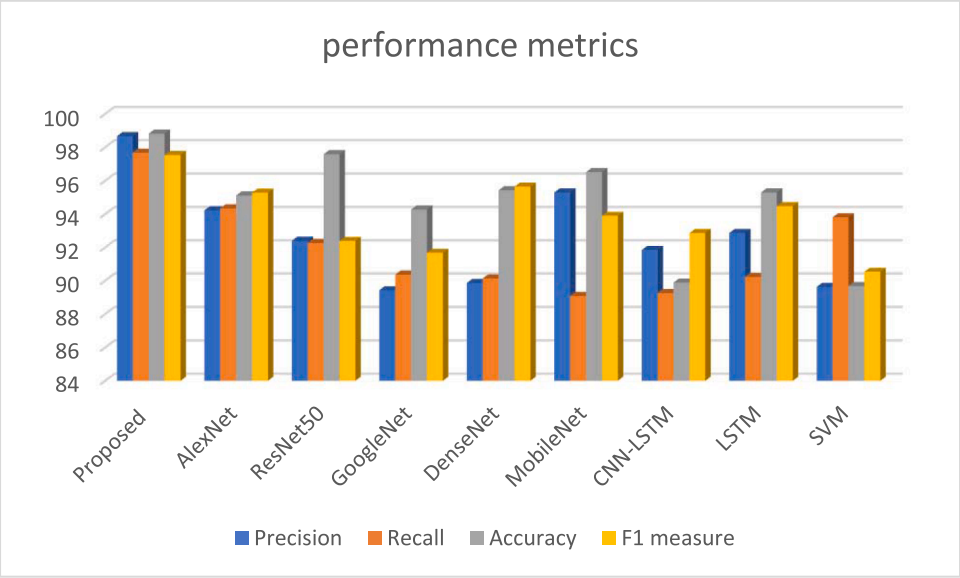


Fig. 8. overall performance results.

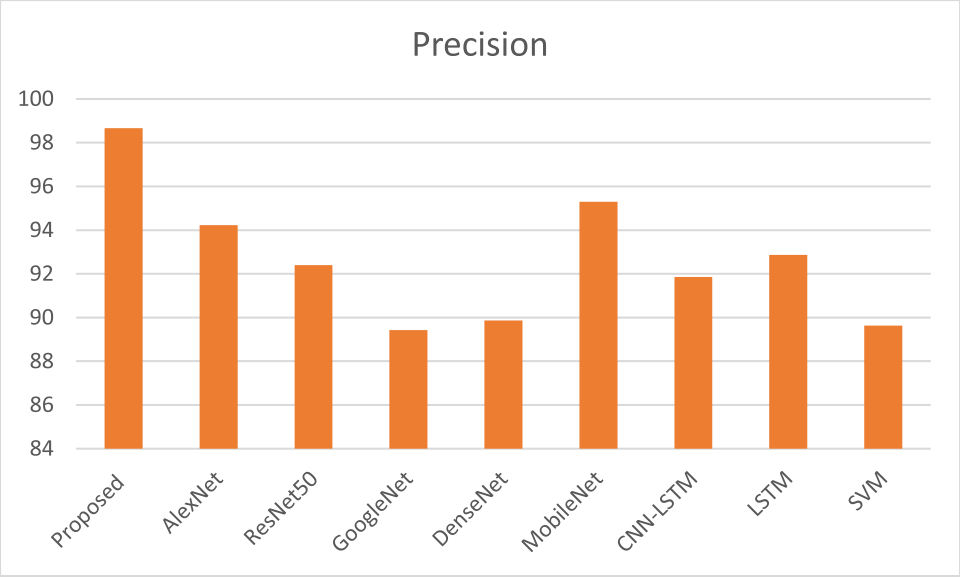


Fig. 9. precision results.

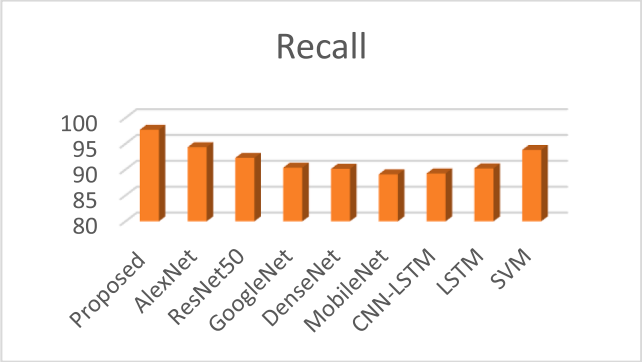


Fig. 10. Recall results.

detection of kidney stones. The CT scan images are used as the dataset. These datasets are used as 80 % of training datasets and 20 % of testing datasets for this work. The proposed work performed a hybrid of AlexNet and GRU model which attained a higher training rate than the traditional methods. Furthermore, the EHO model is used for the hybrid model's hyper parameter tuning to obtain an optimal solution. The experimental results showed that the proposed optimized DL models are suited well for feature extraction and classification. The performance metrics such as Accuracy, precision, F1 measure and recall for the proposed and traditional models are evaluated and compared. The result cleared that the proposed optimized hybrid model is much better than all the traditional methods for kidney stone detection.

CRediT authorship contribution statement

Y Jini Jacob: Formal analysis. **Bethanney Janney J:** Conceptualization. **Hemalatha RJ:** Methodology. **Preethi S:** Validation.

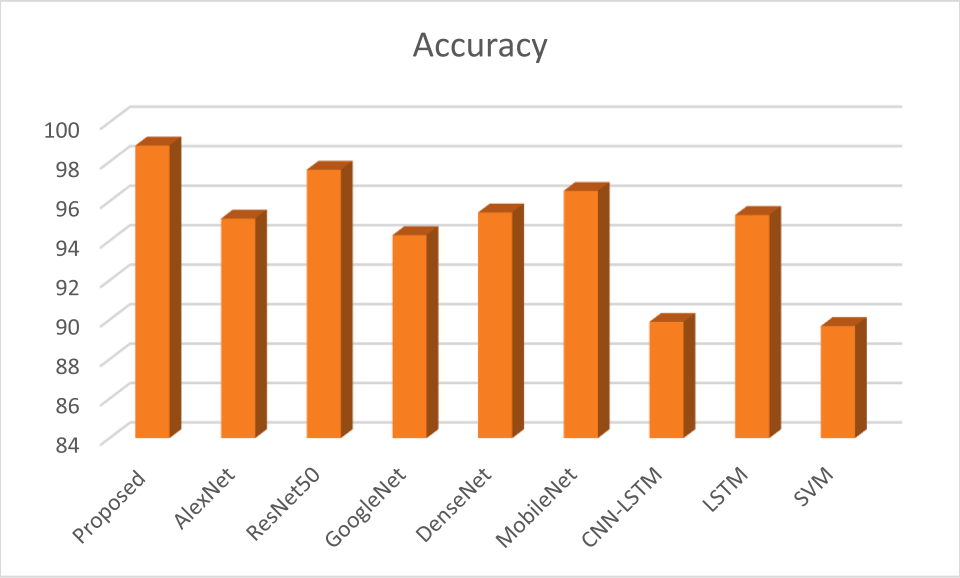


Fig. 11. Accuracy results.

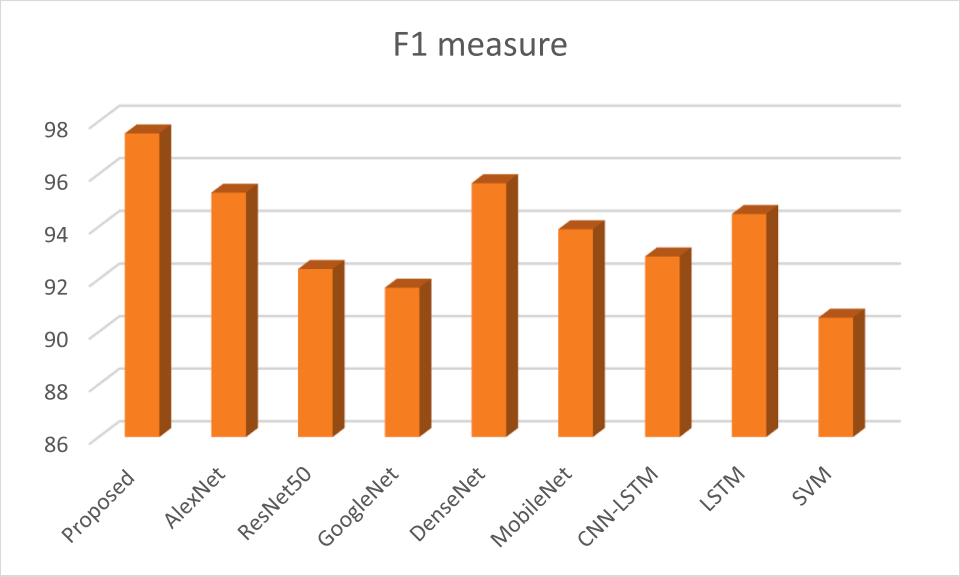


Fig. 12. F1 measure results.

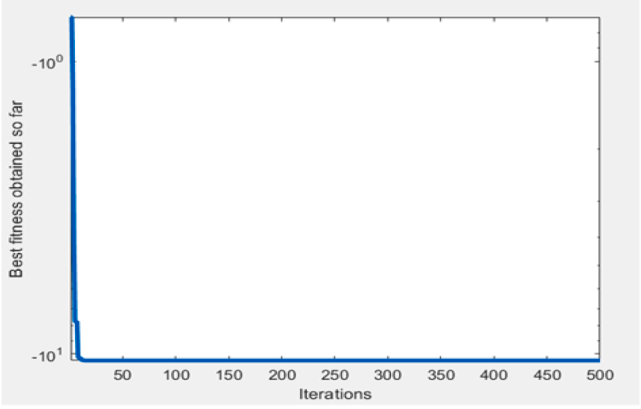


Fig. 13. Best fitness evaluation curve.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] A. Caglayan, M.O. Horsanali, K. Kocadurdu, E. Ismailoglu, S. Guneyli, Deep learning model-assisted detection of kidney stones on computed tomography, Int. Braz. J. Urol. 48 (5) (2022) 830–839. Sep-Oct.

[2] Y. Cui, Z. Sun, S. Ma, W. Liu, X. Wang, X. Zhang, X. Wang, Automatic detection and scoring of kidney stones on noncontrast CT images using S.T.O.N.E. Nephrolithometry: combined deep learning and thresholding methods, Mol. Imaging Biol. 23 (3) (2021) 436–445. Jun.

[3] D.R. Sarvamangala, R.V. Kulkarni, Convolutional neural networks in medical image understanding: a survey, Evol. Intel. 15 (2022) 1–22, https://doi.org/10.1007/s12065-020-00540-3.
[4] S. Shinde, U. Kulkarni, D. Mane, A. Sapkal, Deep learning-based medical image analysis using transfer learning, in: R. Patgiri, A. Biswas, P. Roy (Eds.), Health Informatics: A Computational Perspective in Healthcare. Studies in Computational Intelligence, Springer, Singapore, 2021, https://doi.org/10.1007/978-981-15-9735-0_2 vol 932.
[5] O. Oztuna Taner, A.B. Çolak, Dairy factory milk product processing and sustainable of the shelf-life extension with artificial intelligence: a model study, Front. Sustain. Food Syst. 8 (2024) 1344370, https://doi.org/10.3389/fsufs.2024.1344370.
[6] J. Janney, Bethanney, S. Roslin, Classification of melanoma from dermoscopic data using machine learning techniques, Multimed. Tools. Appl. 79 (2020) 3713–3728, https://doi.org/10.1007/s11042-018-6927-z.
[7] Oznur Oztuna Taner, Hatice Mercan, Andac Batur Colak, Jovana Radulovic, Tolga Taner, AhmetSelim Dalkilic, Application of artificial intelligence techniques for heat exchanger predictions in food industry, Adv. Mater.-Based Therm. Enhanced Phase Change Mater. (2024) 269–325.
[8] S. Elbedwehy, E. Hassan, A. Saber, R. Elmonier, Integrating neural networks with advanced optimization techniques for accurate kidney disease diagnosis, Sci. Rep. 14 (1) (2024) 21740, https://doi.org/10.1038/s41598-024-71410-6. Sep 18.
[9] M. Karaman, S. Çınar, Automatic detection of kidneys on abdominal CT images using aggregate channel features, in: 2022 International Conference on Innovations in Intelligent Systems and Applications (INISTA), 2022, pp. 1–6, https://doi.org/10.1109/INISTA55318.2022.9894149.
[10] M. B, N. Mohan, S.K. S, S.K. P, Automated detection of kidney stone using deep learning models, in: 2022 2nd International Conference on Intelligent Technologies (CONIT), 2022, pp. 1–5, https://doi.org/10.1109/CONIT55038.2022.9847894.
[11] S. Rajput, A. Singh, R. Gupta, Automated kidney stone detection using image processing techniques, in: 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2021, pp. 1–5, https://doi.org/10.1109/ICRITO51393.2021.9596175.
[12] K. Yildirim, P.G. Bozdag, M. Talo, O. Yildirim, M. Karabatak, U.R. Acharya, Deep learning model for automated kidney stone detection using coronal CT images, Comput. Biol. Med. 135 (2021) 104569, https://doi.org/10.1016/j.compbiomed.2021.104569.
[13] E.J. Lim, D. Castellani, W.Z. So, K.Y. Fong, J.Q. Li, H.Y. Tiong, N. Gadzhiev, C. T. Heng, J.Y.C. Teoh, N. Naik, et al., Radiomics in Urolithiasis: systematic review of current applications, limitations, and future directions, J. Clin. Med. 11 (17) (2022) 5151, https://doi.org/10.3390/jcm11175151.
[14] M.B. Suresh, A. M R, Kidney stone detection using digital image processing techniques, in: 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), 2021, pp. 556–561, https://doi.org/10.1109/ICIRCA51532.2021.9544610.
[15] H. Dave, V. Patel, J.N. Mehta, S. Degadwala, D. Vyas, Regional kidney stone detection and classification in ultrasound images, in: 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA), 2021, pp. 1108–1112, https://doi.org/10.1109/ICIRCA51532.2021.9545031.
[16] L.Y. Myint, S.S. Maung, K.T. Zar, Removal of unwanted object in 3D CT kidney stone images and 3D visualization, in: 2020 24th International Computer Science and Engineering Conference (ICSEC), 2020, pp. 1–5, https://doi.org/10.1109/ICSEC51790.2020.9375155.
[17] Y.Y. Liu, Z.H. Huang, K.W. Huang, Deep learning model for computer-aided diagnosis of urolithiasis detection from kidney–ureter–bladder images, Bioengineering 9 (12) (2022) 811, https://doi.org/10.3390/bioengineering9120811.
[18] H. Sharen, M. Narendra, L.J. Anbarasi, MSKd_Net: multi-head attention-based swin transformer for Kidney diseases classification, IEEE Access 12 (2024) 181975–181986, https://doi.org/10.1109/ACCESS.2024.3510634.
[19] J. Chaki, A. Uçar, An efficient and robust approach using inductive transfer-based ensemble deep neural networks for kidney stone detection, IEEE Access 12 (2024) 32894–32910, https://doi.org/10.1109/ACCESS.2024.3370672.
[20] S.D. Pande, R. Agarwal, Multi-class kidney abnormalities detecting novel system through computed tomography, IEEE Access 12 (2024) 21147–21155, https://doi.org/10.1109/ACCESS.2024.3351181.
[21] J. Bethanney Janney, N.R. Krishnamoorthy, S. Divakaran, T. Sudhakar, S. Krishnakumar, V. Akshaya, Diagnosis of skin malignancy using deep learning approaches, in: 2021 International Conference on Advancements in Electrical, Electronics, Communication, Computing and Automation (ICAECA), 2021, pp. 1–4, https://doi.org/10.1109/ICAECA52838.2021.9675722.
[22] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, Yoshua Bengio, Empirical evaluation of gated recurrent neural networks on sequence modeling, neural and evolutionary computing (cs.NE), 2014, doi: 10.48550/arXiv.1412.3555.
[23] M. Pavithra, K. Saruladha, K. Sathyabama, GRU based deep learning model for prognosis prediction of disease progression, in: 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2019, pp. 840–844, https://doi.org/10.1109/ICCMC.2019.8819830.
[24] Monalisa Nayak, Soumya Das, Urmila Bhanja, ManasRanjan Senapati, Elephant herding optimization technique based neural network for cancer prediction, Inform. Med. Unlocked. 21 (2020) 100445, https://doi.org/10.1016/j.imu.2020.100445.