2022 2nd International Conference on Intelligent Technologies (CONIT)
Karnataka, India. June 24-26, 2022

Automated Detection of Kidney Stone Using Deep Learning Models

Manoj B, Neethu Mohan, Sachin Kumar S, Soman K.P
Center for Computational Engineering and Networking (CEN),
Amrita School of Engineering, Coimbatore,
Amrita Vishwa Vidyapeetham, India
manojbalu190998@gmail.com, m_neethu@cb.amrita.edu, s_sachinkumar@cb.amrita.edu, kp_soman@amrita.edu

***Abstract*—Kidney stone is the most common disease nowadays. Proper diagnosis of the disease is required to cure and lead a healthy lifestyle. Various methods are posed for kidney stone detection using different imaging techniques. This paper proposes an automated system for detecting kidney stones using deep learning models. The experiments are performed using an open-source Computed Tomography (CT) image dataset. These datasets are made to perform on deep learning models. On analyzing the efficiency of different deep learning models, it is found that VGG series performs the best. The accuracy obtained for kidney stone detection using VGG16 architecture is 99%. The performance of the model has also been evaluated by using a method called stratified K-fold cross validation. Further, the area of kidney stone is detected using Gradient-weighted Class Activation Mapping, which is referred to as Grad-CAM. In short, here we are giving the CT image to VGG model server for classifying and then using Grad-Cam to find the area of interest and finally, the expert checks the output to verify the result.**

***Index Terms*—Kidney stone, CT image, deep learning, stratified K-fold, Grad-Cam**

I. Introduction

Due to lifestyle changes, primarily due to modern food habits, people suffer from new lifestyle and life-threatening diseases. One such lifestyle disease common nowadays is the formation of stones in the kidney. Kidney stones are formed due to the undiluted solution that is not diluted out from the human body. It is mainly generated due to more undiluted crystal substances than liquid substances in the urine. Stones are getting generated because these undiluted crystal substances stick together due to the absence of proper enzymes to dissolve them. The hard crystal kind of solid stones have no shape and have sharp edges to prickle the tissues inside the kidney, which is soft in nature. Thus, the patients may suffer from severe pain because of stones in the kidney. Based on the severity of the stone, patients are admitted to the emergency ward for immediate treatment.

Kidney stones are classified into different types based on their formation. It is important to identify the stone type to get proper treatment. The common symptoms of kidney stone are pain in the kidney area, vomiting, bad urine smell, and blood presence while urinating. Many risk factors are associated with kidney stones and even lead to kidney failures if it is not treated properly. Based on the patient's condition, experts will suggest a scan to identify the area of the stone. The correct and proper treatment is required to cure the disease completely. Detection of kidney stone from scanned images need experts assistance. Even though the specialists classify stones using scanned images, sometimes specialists might make the wrong decision as they deal with several patients daily. Hence, many algorithm models have proposed that automatically detect the stone from the scanned images.

Recently advancements in artificial intelligence are helpful in the medical field to identify malfunctions in our body easily. Deep learning based model has been used for various tasks such as classification, noise removal, detection and identification [1]–[3]. Deep learning models can be used to handle small sets of data and also, it works on more numbers of features. These model works well on unstructured datasets as well.

Motivated by the above facts, this paper proposes an automated methodology for kidney stone detection using deep learning algorithms. The proposed automated methodology assists doctors in the clinical diagnosis of stones. In this work, various deep learning models are explored for the automated detection of kidney stones and followed by a stratified k-fold cross validation procedure. This cross validation is performed to confirm that train data and the test data have a similar feature as compared with the original set data. Further Grad-CAM technique is employed to detect the area of the stone accurately. This technique heats the area where the stone is located, which confirms the location of the stone. If it does show any response, it means that there is no stone found in that area.

II. Related Works

Several works have been proposed for kidney stone detection using various imaging techniques. In [4], a deep learning model using Xresnet50 for kidney stone diagnosis. Xresnet50 consist of 50 layers with the activation function of softmax layer. Even though the model has obtained an accuracy of 96.86%, the false positive rate is higher. In 2021 the paper [5] based on image processing techniques and neural network have been used for kidney stone detection. The Back Propagation algorithm used for the classification has met with an accuracy of 98.8%. However, the problem associated with the model is that the increase in the noise ratio affects the classification task and thus, some cases are not correctly classified. In [6], CNN based model is proposed for kidney stone detection. It consists of a convolution, max pooling, and a fully connected layer. This work used the 128X128 CNN-ML model and 256X256 CNN-ML for the classification of kidney stone images. The accuracies reported are 86.8% with validation and 85.1% without validation. However, the model has failed to identify the stones accurately and also accuracy is low compared with other existing models for both 128X128 and 256X256 CNN-ML models. The article [7] developed a scoring model to identify the area of the stone. Here, they have used NCCT images for the detection task. The hydronephrosis classification algorithm showed up an accuracy of 97%. Even though there are many boons, some images are wrongly classified in the output result. CNN based methodology for ureteral stones detection detection is proposed in [8] with an overall accuracy of 92%.

The article [9] shows a result of 98.8% on the classification of kidney stone using back propagation ANN algorithm and multi-layer perception. To remove noise, a smoothing filter is used and also contrast enhancement is performed to enhance the visibility mode. The model is implemented on FPGA. In [10], hardware devices were used to classify normal kidney and stone images. Here 14.5K images were used for the task. By performing classification 63% of accuracy is reported. The reason for the low accuracy is the high rate of misclassification. In [11], kidney stone classification was done using weka on different machine learning models. The authors have used sampling methods such as under-sampling and oversampling. This paper also used a method called SMOTETomek, a cluster of two methods. The main disadvantage of this work is the class imbalance of the dataset used. Due to this, the method cannot achieve an accurate result and the accuracy reported for the best decision tree model is 85.3%. Another disadvantage associated with this model is that clinical experts are needed to accurately identify the area of the stones as the model fails to identify automatically. Hence it requires a lot of time as it is a manual task. In [12], KNN and SVM based classification is proposed to classify kidney stones. Here, two filters, unsharp masking, morphological operation and segmentation process, are employed to spot the region of interest. Finally, KNN and SVM are used for the classification of each images. However, the model has more misclassifications and failed to identify kidney stones. It has obtained an accuracy of 89% for KNN and 84% for SVM.

III. Proposed Methodology

In this work, the classification of normal and kidney stone images is done automatically using deep learning models and identified the areas of interest. In our proposed work, the Computed Tomography (CT) scan is fed to the model and gives the desired classification output. The block diagram of the proposed methodology is given in Fig 1 and the process flow is described in detail.

First, the CT images of normal and kidney stones are collected, and data augmentation is performed to avoid class imbalance issues. Then, training is done using various deep learning models and identified the model with high accuracy.

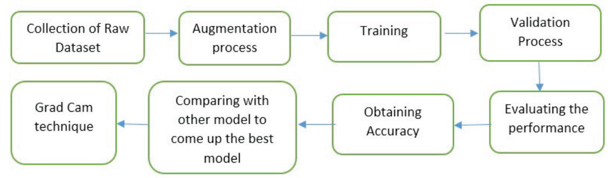

Fig. 1. Block Diagram of Proposed Methodology

K-fold cross validation is performed to evaluate the models. Finally Grad-CAM technique is used to find the locations of the stones inside the kidney. The following sections briefly narrate each of the steps.

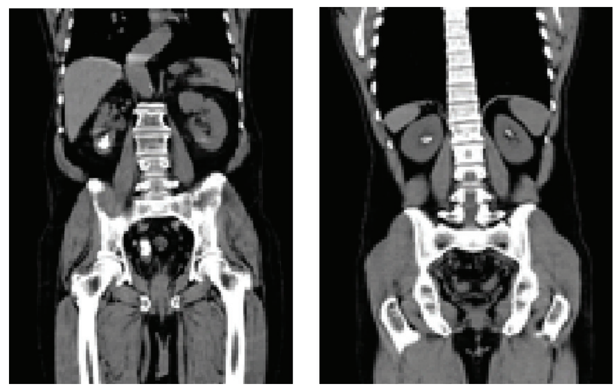

Fig. 2. Examples of CT images with kidney stones

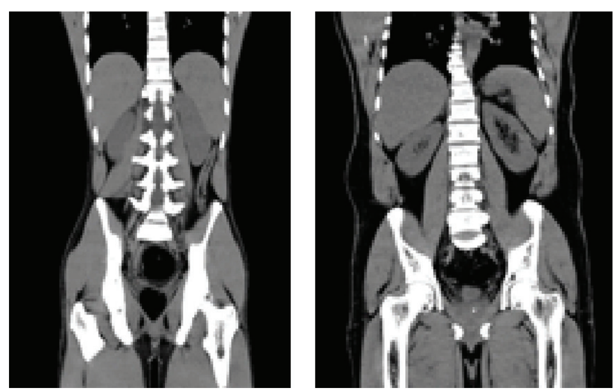

Fig. 3. Examples of normal CT images

A. Dataset Description

Computed tomography (CT) imaging is widely used for abnormality detection associated with the kidney. The open-source dataset used in this study consists of 1799 CT images of normal and kidney stone. Class 0 is the ‘normal’ image and Class 1 is the ‘kidney stone’ image. The dataset is split so that 80% of the images on both the classes are used to train the model and the remaining 20% are used as test data. There are 828 normal images and 625 images with kidney stones in training dataset. There are 181 normal and 165 kidney stone

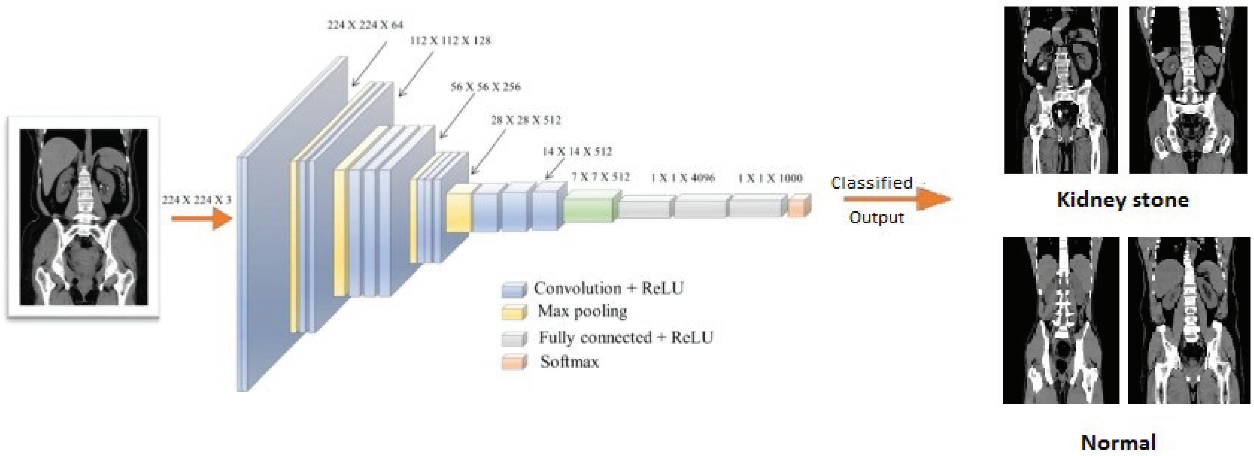

Fig. 4. Architecture details of VGG16

images in the test set. Sample images from the dataset with and without kidney stones are given in Fig 2 and Fig 3 respectively.

B. Data Augmentation

Various augmentation techniques are performed to increase the number of images in the dataset. Augmentation helps to overcome class imbalance problems in general. This work performs augmentation processes such as rotation and zooming on raw CT images. A rotation of 10 degree is performed. After augmentation, the dataset size has been increased to 3265 from 1453.

C. Training the algorithm

After data augmentation, training has been done with 80% of the images using different deep learning algorithms namely VGG, Xresnet, Densenet and Resnet. The model training were carried out using the Fastai library. After the process, evaluations are done on the test image set to verify the performance.

D. Stratified K-fold cross validation

After training, cross validation is performed to estimate the performance of the unseen images in the dataset. Stratified K-fold validation is basically an extension of the cross validation technique. The main difference is that, in stratified K-fold, for each fold, the percentage of each sample of class is preserved and fed to the next fold. 'Stratified' means splitting the dataset into the number of folds.

E. Gradient-weighted Class Activation Mapping (Grad-CAM)

Grad-CAM technique is used to identify the location of the stone area from given set of images. It works in such a way that it produces rough localization map of the area of interest on the image by using gradient information in the final output layer. So the area of the stone will be detected correctly. This method points out the area of the stone in heated and if the Grad-CAM does not show any result in the output image, it means that the image is a person who has a healthy kidney. It can also find the area even the stone is present on both sides (left and right) of the kidney.

IV. Architecture of VGG16

Out of all the models considered in this study, VGG16 has attained the highest accuracy by rectifying the limitations of all the previous works. Figure 4 shows the architecture details of VGG16. The 16 in "VGG16" refers to the total number of layers it has. VGG16 comes under the family of Convolution Neural Network (CNN), which is the best model for vision related applications [1]. It is a simple architecture and is mainly popular due to its easy implementation. VGG16 will not focus on having a large set of hyper-tuning parameters. These are still considered as a great architecture because it particularly shows their interest on the layers which they are using:

- Same padding with 3X3 convolution layer filter which has a stride one.
- Max pooling layer of 2X2 filter with stride two.

This kind of layer follows the entire architecture. The last layer is a softmax layer. Before the softmax layer, two fully connected layers help reduce the loss of data and also help to transfer data to the subsequent activation layer. This VGG16 is basically a vast network that consists of several million parameters.

V. Results and Discussions

The experiments are conducted using CT images of normal and kidney stones. After the augmentation process, 80% of the data has been trained using different deep learning models. The deep learning models used in this study are VGG, Xresnet, densenet and resnet. The performance is evaluated by precision, recall and F1 score for all models and are tabulated in Table I.

TABLE I
PERFORMANCE EVALUATION OF PRECISION, RECALL AND F1 SCORE

Model	Precision	Recall	F1-score
VGG16	0.982	0.994	0.987
VGG19	0.976	0.994	0.984
Xresnet50	0.959	0.994	0.976
Xresnet101	0.976	0.988	0.981
Densenet201	0.965	0.994	0.979
Resnet34	0.95	0.988	0.968
Resnet50	0.965	1	0.982
Resnet101	0.97	0.994	0.979

Out of all the deep learning models, VGG16 has given best results as the value of precision and recall are near to 1 and so as F1- score value. The accuracy obtained for different deep learning models are tabulated in Table II. From Table II, it is confirmed that VGG16 has obtained the highest accuracy of 99%. The confusion matrix for the best VGG16 model for classifying the images is given in Fig. 5. Grad-CAM technique is used on the trained model of VGG16 to identify the area of kidney stones for each image in the data. An illustration of Grad-CAM technique is given in Fig 6. The stone locations are detected correctly and are highlighted using heat maps.

TABLE II
PERFORMANCE COMPARISON WITH DIFFERENT DL MODEL

Model	Train Accuracy	Test Accuracy	Cross validation
VGG16	99%	99%	96.24%
VGG19	98%	98%	91.53%
Xresnet50	97%	97%	62.99%
Xresnet101	98%	98%	62.71%
Densenet201	98%	97%	98.02%
Resnet34	97%	96%	91.24%
Resnet50	98%	98%	91.81%
Resnet101	98%	97%	94.07%

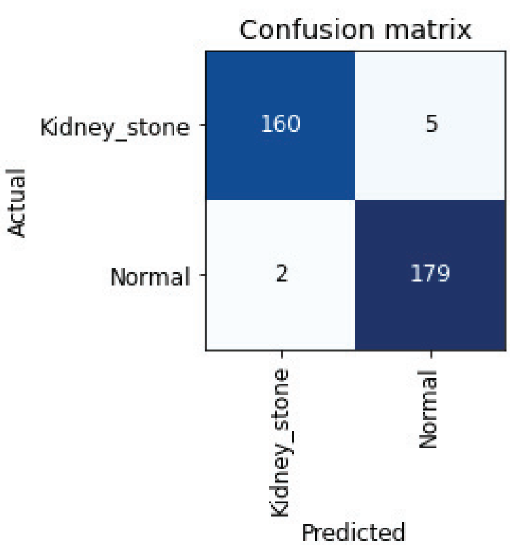

Fig. 5. Confusion Matrix for VGG16 model

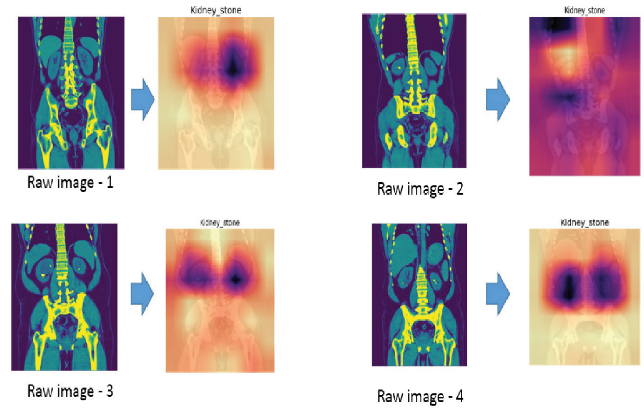

Fig. 6. Illustration of Grad-CAM output

The performance of the proposed methodology is compared with other existing models in the literature for kidney stone detection. The results are tabulated in Table III. The other models used for comparison from literature are Xresnet50 [4], CNN [6], decision tree [11], random forest [10], KNN [12], and SVM [12]. Xresnet50 achieved good accuracy, but there is a high false positivity rate. The models proposed based on CNN, decision tree and random forest have a similar limitation of very low accuracy and thus, the model is not effective for accurate detection. KNN and SVM have obtained an accuracy of 89% and 84%, respectively for detecting and identifying the stone. The main drawback is the high misclassification rate. From the comparison results, it is clear that VGG16 rectified the major limitations of the previous work and came out with good accuracy for kidney stone detection.

TABLE III
PERFORMANCE COMPARISON WITH EXISTING MODELS

Model	Accuracy (%)
Xresnet50	96.82
CNN	86
Decision Tree	85.3
Random Forest	63
KNN	89
SVM	84
VGG16(Proposed work)	99

VI. CONCLUSION

The number of patients diagnosed with kidney stone disease gradually increases every year. Thus there is always a need for highly accurate systems for stone detection and identification. This paper proposes an automated system for the accurate detection of kidney stone using deep learning models. This study successfully classified the computed tomography images into normal and kidney stones and identified the stone area in the kidney using the Grad-CAM technique. It helps the experts work easily, making them less or no work. Experiments with a large set of datasets and new models is devoted as a future scope for this work.

References

[1] R Anand, V Sowmya, EA Gopalakrishnan, KP Soman, et al. Modified vgg deep learning architecture for covid-19 classification using biomedical images. In *IOP Conference Series: Materials Science and Engineering*, volume 1084, page 012001. IOP Publishing, 2021.

[2] Neethu Mohan, KP Soman, and R Vinayakumar. Deep power: Deep learning architectures for power quality disturbances classification. In *2017 international conference on technological advancements in power and energy (TAP Energy)*, pages 1–6. IEEE, 2017.

[3] S Sachin Kumar, M Anand Kumar, and KP Soman. Sentiment analysis of tweets in malayalam using long short-term memory units and convolutional neural nets. In *International Conference on Mining Intelligence and Knowledge Exploration*, pages 320–334. Springer, 2017.

[4] Kadir Yildirim, Pinar Gundogan Bozdag, Muhammed Talo, Ozal Yildirim, Murat Karabatak, and U Rajendra Acharya. Deep learning model for automated kidney stone detection using coronal ct images. *Computers in biology and medicine*, 135:104569, 2021.

[5] A Lalith Vikas. Kidney stone detection using image processing and neural networks. *Annals of RSCB*, 25(6):13112–13119, 2021.

[6] Waliu Olalekan Apena and Orobosa Libert Joseph. Convolutional neural network model layers improvement for segmentation and classification on kidney stone images using keras and tensorflow.

[7] Yingpu Cui, Zhaonan Sun, Shuai Ma, Weipeng Liu, Xiangpeng Wang, Xiaodong Zhang, and Xiaoying Wang. Automatic detection and scoring of kidney stones on noncontrast ct images using stone nephrolithometry: combined deep learning and thresholding methods. *Molecular Imaging and Biology*, 23(3):436–445, 2021.

[8] Johan Jendeberg, Per Thunberg, and Mats Lidén. Differentiation of distal ureteral stones and pelvic phleboliths using a convolutional neural network. *Urolithiasis*, 49(1):41–49, 2021.

[9] Kalannagari Viswanath and Ramalingam Gunasundari. Design and analysis performance of kidney stone detection from ultrasound image by level set segmentation and ann classification. In *2014 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pages 407–414. IEEE, 2014.

[10] Joan Serrat, Felipe Lumbreras, Francisco Blanco, Manuel Valiente, and Montserrat López-Mesas. mystone: A system for automatic kidney stone classification. *Expert Systems with Applications*, 89:41–51, 2017.

[11] Işıl AKSAKALLI, Sibel KAÇDIOĞLU, and Y Sinan HANAY. Kidney x-ray images classification using machine learning and deep learning methods. *Balkan Journal of Electrical and Computer Engineering*, 9(2):144–151, 2021.

[12] Jyoti Verma, Madhwendra Nath, Priyanshu Tripathi, and KK Saini. Analysis and identification of kidney stone using kth nearest neighbour (knn) and support vector machine (svm) classification techniques. *Pattern Recognition and Image Analysis*, 27(3):574–580, 2017.