Automated Kidney Stone Detection Using Ultrasound Imaging and Watershed Algorithm

Ch. Satyanarayana Reddy[1], K.Pavani[2], B. Prasanth[3]

#1 Assistant Professor in the Department of MCA, SRK Institute of Technology, Vijayawada.

#2 Assistant Professor & Head of Department of MCA, SRK Institute of Technology, Vijayawada.

#3 Student in the Department of MCA, SRK Institute of Technology, Vijayawada

Abstract: Kidney stone detection using ultrasound imaging is widely preferred due to its non-invasive and cost-effective nature; however, the presence of speckle noise, low contrast, and अस्पష్టమైన boundaries makes accurate identification challenging. This paper proposes an automated kidney stone detection system based on advanced image enhancement and segmentation techniques. Initially, ultrasound images are preprocessed through resizing and normalization to ensure consistency. Multiple image filtering techniques, including Gaussian Blur, Median Filter, Gabor Filter, Laplacian Filter, Bilateral Filter, and Adaptive Histogram Equalization (CLAHE), are applied to enhance image quality. Among these, Bilateral Filtering and CLAHE demonstrate superior performance in preserving edges while improving contrast and visibility of kidney stones.

Further, the enhanced images are processed using the Watershed segmentation algorithm to accurately identify and extract the Region of Interest (ROI) corresponding to kidney stones. The proposed system is evaluated on ultrasound images, and the results indicate significant improvement in visualization and detection accuracy compared to traditional methods. This approach reduces manual effort, minimizes human error, and assists medical professionals in faster and more reliable diagnosis. The system can be further extended by integrating deep learning techniques for real-time and highly precise kidney stone detection.

Index terms - *Kidney Stone Detection, Ultrasound Imaging, Image Processing, Image Enhancement, Bilateral Filter, Adaptive Histogram Equalization (CLAHE), Watershed Algorithm, Image Segmentation, Region of Interest (ROI), Noise Reduction, Medical Imaging*

1. INTRODUCTION

Kidney stones are a common urological disorder caused by the accumulation of minerals and salts in the kidneys. Early detection of kidney stones is very important to prevent severe pain, infection, and complications. Among various medical imaging techniques, ultrasound imaging is widely used because it is non-invasive, cost-effective, and safe for patients. However, ultrasound images often suffer from limitations such as speckle noise, low contrast, and blurred edges, which make accurate detection of kidney stones difficult.

To overcome these challenges, image processing techniques play a crucial role in improving the quality of ultrasound images. By applying different filtering and enhancement methods, it becomes easier to highlight the stone regions and improve visibility. In this work, multiple image enhancement algorithms such as Gaussian Filter, Median Filter, Gabor Filter, Laplacian Filter, Bilateral Filter, and Adaptive Histogram Equalization (CLAHE) are applied to enhance image clarity. Among these, Bilateral Filter and CLAHE show better performance in preserving edges and improving contrast.

After enhancement, image segmentation is performed using the Watershed algorithm to accurately detect the Region of Interest (ROI) where kidney stones are present. This automated approach reduces manual effort and increases diagnostic accuracy. The proposed system provides a reliable and efficient method for kidney stone detection in ultrasound images and can assist medical professionals in faster decision-making.

2. LITERATURE SURVEY

2.1 Performance Evaluation of Deep CNN-Based Crack Detection and Localization Techniques for Concrete Structures:

In order to identify cracks in concrete constructions, a customized convolutional neural network is proposed in this article. Based on training data quantity, data heterogeneity, network complexity, and epoch count, the suggested approach is contrasted with four current deep learning techniques. Eight datasets of varying sizes, generated from two public datasets, are used to assess and compare the performance of the proposed convolutional neural network (CNN) model with pretrained networks, namely the VGG-16, VGG-19, ResNet-50, and Inception V3 models. For each model, the evaluation considered computational time, crack localization results, and classification measures, e.g., accuracy, precision, recall, and F1-score. The number of training data and sample variability have a major impact on model performance, according to experimental studies. On a small amount of diverse training data, all models showed encouraging performance; however, overfitting resulted from increasing training data quantity and decreasing diversity. On a little quantity of data, the suggested customized CNN and VGG-16 models performed better than the other approaches in terms of classification, localization, and computing time. The findings show that these two models exhibit improved crack detection and localization for concrete structures.

2.2 Swin-Unet: Unet-like Pure Transformer for Medical Image Segmentation:

Convolutional neural networks (CNNs) have made significant progress in medical image processing in recent years. In particular, skip-connections and U-shaped architecture-based deep neural networks have been extensively used in a range of medical imaging applications. Despite CNN's outstanding performance, the localization of the convolution process prevents it from learning global and long-range semantic information interaction. In this research, we present Swin-Unet, a pure Transformer for medical picture segmentation that is similar to Unet. The Transformer-based U-shaped Encoder-Decoder architecture with skip-connections for local-global semantic feature learning receives the tokenized image patches. In particular, we extract context features using a hierarchical Swin Transformer with shifted windows as the encoder. Additionally, a symmetric Swin Transformer-based decoder with a patch expanding layer is created to carry out the upsampling process in order to recover the feature maps' spatial resolution. Experiments on multi-organ and cardiac segmentation tasks show that the pure Transformer-based U-shaped Encoder-Decoder network performs better than those with full-convolution or the combination of transformer and convolution under the direct down-sampling and upsampling of the inputs and outputs by 4x. This https URL will make the scripts and trained models accessible to the general audience.

2.3 TransUNet: Transformers Make Strong Encoders for Medical Image Segmentation

For the development of healthcare systems, medical image segmentation is a necessary precondition, particularly for illness diagnosis and treatment planning. The u-shaped architecture, commonly referred to as U-Net, has become the de facto standard and achieved remarkable performance on a variety of medical picture segmentation tasks. However, U-Net typically exhibits limits in clearly representing long-range reliance because of the inherent locality of convolution processes. Transformers, which are alternative architectures with built-in global self-attention mechanisms for sequence-to-sequence prediction, may have limited localization capabilities because of inadequate low-level details. In this study, we offer TransUNet as a powerful alternative for medical picture segmentation, which benefits from both Transformers and U-Net. In order to extract global contexts, the Transformer first encodes tokenized picture patches using a convolution neural network (CNN) feature map. To achieve accurate localization, however, the decoder upsamples the encoded features and combines them with the high-resolution CNN feature maps.

In order to improve finer details by recovering localized spatial information, we contend that Transformers can function as powerful encoders for medical picture segmentation tasks when combined with U-Net. TransUNet outperforms rival techniques in a variety of medical applications, such as cardiac segmentation and multi-organ segmentation. Models and code may be found at this https URL.

2.4 3D TransUNet: Advancing Medical Image Segmentation through Vision Transformers:

In order to diagnose diseases and arrange treatments, medical image segmentation is essential to the advancement of healthcare systems. For a number of medical picture segmentation tasks, the u-shaped architecture, sometimes referred to as U-Net, has shown great performance. However, U-Net's capacity to accurately describe long-range relationships is intrinsically limited by its convolution-based processes. Transformers, which are well-known for their global self-attention processes, have been used by researchers as alternate designs to overcome these constraints. Our prior TransUNet, which uses Transformers' self-attention to supplement U-Net's localized knowledge with the global context, is one well-known network. In this study, we thoroughly explore the possibilities of Transformers in both the encoder and decoder design, building on the state-of-the-art nnU-Net architecture to expand the 2D TransUNet architecture to a 3D network. We present two essential elements: 1) A Transformer encoder that extracts global contexts by tokenizing picture patches from a convolution neural network (CNN) feature map, and 2) A Transformer decoder that uses cross-attention between candidate proposals and U-Net features to adaptively improve candidate areas. Our research shows that diverse architectural styles are beneficial for different medicinal functions. In multi-organ segmentation, where the relationships between organs are critical, the Transformer encoder performs exceptionally well. However, when it comes to tiny and difficult segmented targets, like tumor segmentation, the Transformer decoder works better. The substantial potential of incorporating a Transformer-based encoder and decoder into the u-shaped medical picture segmentation architecture is demonstrated by extensive studies. In a number of medical applications, TransUNet performs better than rivals.

2.5 Deep Residual Learning for Image Recognition:

Training deeper neural networks is more challenging. In order to facilitate the training of networks that are far deeper than those previously employed, we provide a residual learning approach. Rather than learning unreferenced functions, we explicitly reformulate the layers as learning residual functions with reference to the layer inputs. We present thorough empirical data demonstrating that these residual networks may improve accuracy with far greater depth and are simpler to tune. We assess residual nets up to 152 layers deep on the ImageNet dataset, which is 8× deeper than VGG nets [40] but still less complicated. On the ImageNet test set, an ensemble of these residual nets achieves an error of 3.57%. In the ILSVRC 2015 classification task, this outcome took first place. Additionally, we analyze CIFAR-10 with 100 and 1000 layers. For many visual identification tasks, the depth of representations is crucial. We achieve a 28% relative improvement on the COCO object identification dataset only because of our incredibly deep representations. Our contributions to the ILSVRC & COCO 2015 competitions1, where we also took first place in ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation tasks, are based on deep residual nets.

3. METHODOLOGY

i) Proposed Work:

The proposed work focuses on developing an automated system for kidney stone detection using ultrasound images by combining image enhancement and segmentation techniques. Initially, the input ultrasound image is preprocessed through resizing and normalization to ensure uniformity and reduce noise. Then, multiple image enhancement filters such as Gaussian, Median, Gabor, Laplacian, Bilateral, and Adaptive Histogram Equalization (CLAHE) are applied to improve image clarity and highlight stone regions.

Among all enhancement techniques, Bilateral Filter and CLAHE are identified as the most effective in preserving edges and improving contrast. Finally, the enhanced image is processed using the Watershed algorithm to segment the image and accurately detect the Region of Interest (ROI) corresponding to kidney stones. This approach minimizes manual intervention, improves detection accuracy, and provides a reliable solution for medical diagnosis.

ii) System Architecture:

The system architecture is designed as a sequential pipeline that processes ultrasound images through multiple stages to accurately detect kidney stones. Initially, the user uploads an ultrasound image into the system, which acts as the input. The image then undergoes preprocessing, where it is resized and normalized to ensure consistency and reduce noise for further processing.

In the next stage, various image enhancement techniques such as Gaussian, Median, Gabor, Laplacian, Bilateral, and Adaptive Histogram Equalization are applied to improve image clarity and highlight stone regions. Among these, Bilateral Filter and CLAHE provide the best visual results. Finally, the enhanced image is passed to the Watershed segmentation algorithm, which identifies and marks the Region of Interest (ROI) corresponding to kidney stones. The output is a segmented image with clearly detected stone regions, enabling accurate and efficient diagnosis.

iii) MODULES:

1. Upload Ultrasound Image

This module allows the user to select and upload kidney ultrasound images into the system. It serves as the starting point for the entire detection process.

2. Preprocess Image

In this module, the uploaded image is resized and normalized to maintain uniformity. Noise reduction is also performed to prepare the image for enhancement.

3. Image Enhancement Module

This module applies multiple filtering techniques such as Gaussian, Median, Gabor, Laplacian, Bilateral, and Adaptive Histogram Equalization (CLAHE). It improves image contrast and clarity, making kidney stones more visible.

4. Filter Evaluation Module

Different enhanced images are compared to identify the best-performing filters. Bilateral Filter and CLAHE are selected based on superior edge preservation and contrast improvement.

5. Watershed Segmentation Module

The Watershed algorithm is applied to the enhanced image to segment it into regions. It accurately detects and marks the Region of Interest (ROI) where kidney stones are present.

6. Output Visualization Module

This module displays the final output with highlighted kidney stone regions. It helps users clearly identify detected stones and analyze results.

iv) ALGORITHMS:

Adaptive Histogram Equalization (CLAHE)
CLAHE enhances the contrast of images by applying histogram equalization to small regions instead of the entire image. This improves local contrast and makes kidney stones more visible, especially in low-contrast ultrasound images.

Watershed Algorithm

The Watershed algorithm is a powerful image segmentation technique that treats the image as a topographic surface. It divides the image into regions based on intensity gradients and accurately identifies the Region of Interest (ROI), enabling precise detection of kidney stones.

4. EXPERIMENTAL RESULTS

The proposed kidney stone detection system was evaluated using multiple ultrasound images to analyze the effectiveness of different image enhancement techniques. Initially, raw ultrasound images with noise and low contrast were processed using various filters such as Gaussian, Median, Gabor, Laplacian, Bilateral, and Adaptive Histogram Equalization (CLAHE). The results show that traditional filters like Gaussian and Median reduce noise but slightly blur important details, while Gabor and Laplacian filters enhance texture and edges but are less effective in improving overall clarity.

Among all techniques, Bilateral Filter and CLAHE produced the best results by significantly enhancing contrast and preserving edge information, making kidney stones more visible. After enhancement, the Watershed segmentation algorithm was applied to identify the Region of Interest (ROI). The algorithm successfully segmented the images and accurately highlighted the kidney stone regions.

The experimental analysis demonstrates that the combination of Bilateral filtering, CLAHE, and Watershed segmentation improves detection accuracy and visualization compared to traditional methods. The system also reduces manual effort and provides faster and more reliable results, making it suitable for practical medical applications.

Accuracy: The ability of a test to differentiate between healthy and sick instances is a measure of its accuracy. Find the proportion of analysed cases with true positives and true negatives to get a sense of the test's accuracy. Based on the calculations:

Accuracy = TP + TN /(TP + TN + FP + FN)

$$Accuracy = \frac{(TN + TP)}{T}$$

Test Accuracy: 0.9895

Precision: The accuracy rate of a classification or number of positive cases is known as precision. Accuracy is determined by applying the following formula:

Precision = True positives/ (True positives + False positives) = TP/(TP + FP)

$$\Pr e\,cision = \frac{TP}{(TP + FP)}$$

Recall: The recall of a model is a measure of its capacity to identify all occurrences of a relevant machine learning class. A model's ability to detect class instances is shown by the ratio of correctly predicted positive observations to the total number of positives.

$$Recall = \frac{TP}{(FN + TP)}$$

mAP: One ranking quality statistic is Mean Average Precision (MAP). It takes into account the quantity of pertinent suggestions and where they are on the list. The arithmetic mean of the Average Precision (AP) at K for each user or query is used to compute MAP at K.

$$\boldsymbol{mAP = \frac{1}{n}\sum_{k=1}^{k=n} AP_k}$$

$$\boldsymbol{AP_k = \text{ the AP of class } k}$$

$$\boldsymbol{n = \text{ the number of classes}}$$

F1-Score: A high F1 score indicates that a machine learning model is accurate. Improving model accuracy by integrating recall and precision. How often a model gets a dataset prediction right is measured by the accuracy statistic..

$$F1 = 2 \cdot \frac{(Recall \cdot \Pr e\, cision)}{(Recall + \Pr e\, cision)}$$

Fig.8. upload image

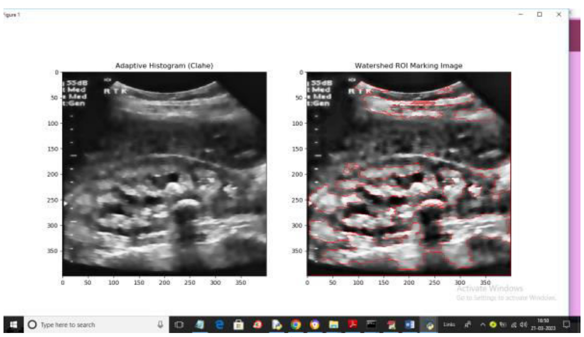

Fig.9. output

5. CONCLUSION

The proposed kidney stone detection system effectively utilizes advanced image enhancement and segmentation techniques to improve the accuracy of ultrasound-based diagnosis. By applying multiple filtering methods, the system enhances image quality and highlights stone regions more clearly. Among all techniques, Bilateral Filter and Adaptive Histogram Equalization (CLAHE) provide the best performance in terms of contrast enhancement and edge preservation.

The integration of the Watershed algorithm enables precise segmentation and accurate identification of the Region of Interest (ROI), ensuring reliable detection of kidney stones. Overall, the system reduces manual effort, minimizes human error, and supports faster medical diagnosis. Hence, the proposed approach offers an efficient and practical solution for kidney stone detection using ultrasound images.

6. FUTURE SCOPE

The proposed system can be further enhanced by integrating advanced deep learning techniques such as Convolutional Neural Networks (CNNs) to improve detection accuracy and automate feature extraction. Real-time kidney stone detection can be achieved by connecting the system directly with ultrasound devices, enabling instant diagnosis during scanning.

Additionally, the system can be extended to detect other kidney-related abnormalities such as tumors or cysts, making it more versatile in medical applications. Deployment as a web or mobile-based healthcare application can improve accessibility for remote and rural areas. Further improvements can also include the use of larger datasets and hybrid models to increase robustness and reliability of the system.

REFERENCES

[1] RajaRAandRanjaniJJ2013Segmentbaseddetectionandquantificationofkidneystonesanditssymmetricanalysisusingtexture properties basedonlogicaloperatorswithultrasoundscann

ingInternationalJournalofComputerApplications18–15
[2] ManikandanSandRajamaniV2008A mathematicalapproachforfeatureselection&imageretrievalofultrasoundkidneyimagedatabasesEuropeanJournalofScientificResearch24 163–71
[3] HafizahWMandSupriyantoE2012AutomaticgenerationofregionofinterestforkidneyultrasoundimagesusingtextureanalysisInternational Journal ofBiologyandBiomedicalEngineering326–34
[4] SunYN,LeeJS,ChangJCandYaoWJ1994Three-dimensionalreconstructionofkidneyfromultrasonicimagesIn:IEEEWorkshoponBiomedicalImageAnalysis43–9
[5] HaoX,GaoSandGaoX1999Anovelmultiscalenonlinearthresholdingmethodforultrasonicspecklesuppressing IEEETrans.Med.Imaging18787–94

Author profiles

Mr. Ch. Satyanarayana Reddy Completed his MCA. He also a web developer and python developer, currently working has an Assistant Professor in the department of MCA at SRK Institute of Technology, Enikepadu, NTR District. His area of interest includes Artificial Intelligence and Machine Learning.

Ms. K. Pavani is Working as an Assistant and Head of Department of MCA, in SRK Institute of technology in Vijayawada. She completed her MCA and M.Tech in Computer Science. She has 10 years of Teaching experience in SRK Institute of technology, Enikepadu, Vijayawada, NTR District. Her areas of interest includes AI and ML, etc

Mr. B.Prasanth is an MCA Student in the Department of Computer Applications at SRK Institute of Technology, Enikepadu, Vijayawada, NTR District. He has Completed Degree in B.Sc. (Computer Science) from Andhra Loyola College, Vijayawada. His area of interest are DBMS and Machine Learning with Python