Technical Journal, University of Engineering and Technology (UET) Taxila, Pakistan *Vol. 30 No. 1-2025*
ISSN:1813-1786 (Print) 2313-7770 (Online)

An Automatic Approach for The Detection and Segmentation of Kidney Stone In Kub CT Images Using Mask R-CNN

F. Alam[1], K. Alam[2], N. Khan[3]

[1,3]*Departmant of Computer Science and IT, University of Malakand, Dir (Lower), Pakistan*
[2]*Department of Urulogy, Miangul Abdul Haq Kidney Hospital, Swat, Pakistan*

[1]fakhre.uom@gmail.com

Abstract- Kidney stones, medically known as nephrolithiasis, are a common urological condition affecting individuals worldwide. Developing advanced techniques based on deep learning models is essential for achieving accurate diagnosis and effective treatment planning of kidney stones. In recent years, deep learning algorithms have been developed for the accurate detection and segmentation of stones in medical images. In this study, we propose an automatic method using the Mask R-CNN (Masked Region-based Convolution Neural Network) framework to accurately identify and segment kidney stones from KUB (Kidneys, Ureters, and Bladder) CT images. The proposed approach leverages both ResNet50 and ResNet101-based Mask R-CNN architectures to achive high accuracy in the detection ans segmentation of stones. A proposed model used 15,000 KUB CT images which contained stone-positive and stone-free conditions. The performance of proposed model was evaluated using Accuracy, Precision, Recall, and F1 Score. Experimental results on the curated dataset shows 96.9% precision, 88.0% recall and 92.0% F1 score. Overall 91% mean average precision (mAP) was achieved outperforming existing approaches. The developed system demonstrates potential to help radiologists enhance their diagnostic performance along with reliability levels.

Keywords- Kidney Stones, CT KUB, Mask R-CNN, ResNet101, ResNet50, Computed Tomography (CT), Deep Learning, Detection and Segmentation.

I. INTRODUCTION

The multiple complex substances which are present in human body are vital and responsible for performing various functions such as blood circulation and digestion. The significant responsibility of regulating these chemical materials and their density in the human body is secreted out by a bean-shaped pair of organs called Kidneys [1]. Kidney stones or renal calculi are tiny chunks of waste mineral deposits, and these lumps of minerals are the greatest dominant factors that affect the function of the urination system. The presence of tiny chunks of deposit in the kidneys is called kidney stones, which is the common cause of kidney malfunction. Sometimes surplus amounts of salt and minerals in the urinary tract stay back and over time build up to hard crystals or lumps known as kidney stones [2]. These crystals cause severe pain in the lower back region and urinary tract infections if untreated [3]. It is estimated that 1 in 11 individuals will experience kidney stones in their lifetime, with a prevalence rate of up to 13% in the United States alone [4].

The detection and segmentation of kidney stone using deep learning techniques has become progressively necessary in the domain of medical imaging. The formation of kidney stones if untreated they can cause kidney complications and severe pain. The standard procedures for kidney stone detection depend heavily on manual medical image assessments which are delayed by human mistakes and lengthy screening times.. Modern algorithms of deep learning sustained by artificial intelligence (AI) provide more suitable and accurate solution. These methods hold CNNs networks to examine and purify medical images, detect the existence and exact location of kidney stones. The significance of such automated detection and segmentation deceits in the capability to accelerate diagnosis, allowing timely interference and treatment arrangement. Furthermore, the use of deep learning in this framework gives to improved diagnostic precision, decreasing the hazard of mistake and refining complete patient consequences. By way of medical capability continues to change, the addition of deep learning procedures in kidney stone detection platforms a talented way for more real and dependable healthcare clarifications.

In recent years, deep learning-based approaches have been successfully applied in computer vision and image analysis tasks [5]. Mask R-CNN is deep

learning algorithm widely used in computer vision applications for the accurate object detection and instance segmentation. It is an enhancement to Faster R-CNN in that it introduces a segmentation branch, which allows it not only to detect objects within an image but also to create pixel-level masks for each object within the image. It is useful for many applications in medical research, medical diagnoses and interventional treatments. Mask R-CNN has been applied in various medical image analysis tasks, including lung nodule detection [6] and liver tumor segmentation [7]. The validation of Mask R-CNN remains a difficult task for clinical data processing. The work of detecting stones within CT KUB scans using automated methods remains an active field of research [8]. The current practice of stone detection along with segmentation either depends on human involvement or traditional image processing protocols. This research employs deep learning techniques to analyze KUB CT images because of recent developments in the field. These advances lead to the recognition of clinical requirements for stone detection methodologies.

Advanced deep learning object detection techniques together with segmentation algorithms permit the enhancement of both accuracy and efficiency levels. Our research develops an automatic system based on Mask R-CNN for detecting and segmenting kidney stones within CT KUB images. The Mask R-CNN method proposed in our study lacked application for kidney stone detection and segmentation in CT KUB images but we identified a solution through this research work. The proposed research advances diagnostic accuracy while making diagnosis more effective by removing manual work procedures and decreasing associated errors. Through its provision of an accurate detection tool for medical staff this investigation enhances patient care because it enables rapid diagnostic responses and customized treatment protocols.

Several conventional machine models are available to identify and segment objects and other region of interest in medical imaging. In the study of Verma et al. [9] the renal calculi were detected using SVM and K-NN techniques on ultrasound images, with 89% accuracy for KNN and 84% accuracy for SVM. However, the ultrasound images suffers from easy disturbance by speckle noises which affect the model performance. In 2020, Soni et al. [10] used Support Vector Machine (SVM) for classification and achieved 98% accuracy. In this work, the preprocessing techniques have been applied for the detection of kidney stones to enhance the visual quality. Nugroho and Thein et al. [11] used a 3-stage technique for segmentation. They achieved 95% sensitivity from 30 patents on Computed Tomography images in different pre-processing procedures. The limitations of this method include the use of three types of thresholding models based on size, intensity, and location while removing unwanted regions and the use of limited number of dataset images. Another method was developed by Hafizah et al. [12] was based on histogram intensity analysis. This technique extracted features successfully by using Ultrasound images. The Ultrasound images were split into groups based on nephritic abnormalities.

Successful outcomes were demonstrated in tasks such as segmentation of the model (R-CNN), a region-based convolutional neural network [13]. Selective search is a popular techniques which can be implemented in R-CNN to detect the region within the image. The popular deep learning algorithm Mask Region-based Convolutional Neural Network (Mask R-CNN) is specially designed for instance segmentation and detection tasks. It was developed by He, Kaiming, et al. [14]. Mask R-CNN, extends Faster R-CNN by adding a branch for predicting an object mask in parallel with the existing branch for bounding box recognition. Mask R-CNN is easy to generalize to other tasks, it is also simple to train and adds only a small overhead to Faster R-CNN. The Mask R-RCNN algorithm is used widely in different types of applications of computer vision, which are autonomous driving vehicles, video and image analysis, and medical imaging [15-17]. The framework of the Mask R-CNN algorithm integrates the ability of instance segmentation and object detection. It enables the algorithm to recognize object recognition and segmentation at the pixel level within an image of each object's specific location [18-19].

Mask R-CNN is famous for its capacity to hold numerous tasks at once, like Faster R-CNN, is excellent at identifying objects in an image by generating bounding boxes that encompass the locations of each object [20-22]. It offers pixel-level segmentation, which represents that each instance's accurate boundaries are defined, pixel by pixel, in addition to object detection [23-24]. Hence, the more desirable model Mask R-CNN for object detection and instance segmentation is our preferred model. In this work we select the Mask R-CNN model for two reasons, the first one is their most accurate object detection compare to faster RCNN and COCO models and the second consideration is their libraries and pre-trained weights which make the process of implementation more suited.

II. PROPOSED APPROACH

We have conduct extensive experiments to evaluate the proposed segmentation method. In the experiments, several CT KUB images were used to evaluate the proposed segmentation method. These images were processed to evaluate the segmentation accuracy and efficiency. An overview of the proposed detection and segmentation method is

shown in Fig. 1. Figure 1 introduces a graphical overview of the research methodology to display the data collection and preprocessing stages as well as model design and training and performance evaluation steps. The workflow for developing and validating the proposed system receives methodical description through each development phase in the system.

A. Data Collection

To develop an automated detection and segmentation methodology employing the Mask R-CNN framework the dataset of kidney stones was acquired from a well-known famous platform Github [25]. The procedure for selecting the dataset entailed a careful evaluation of various factors, including the size of the dataset, the quality of images, and the accessibility of exactly verified ground truth annotations. Our dataset includes two types of KUB CT images, images with stones and images without stones. Medical professionals supervised the process of kidney stones detection which involved identification of stones or absence of stones. Several KUB CT image types served as test subjects for evaluating the proposed method throughout the experiments.

B. Data Preprocessing

Noise reduction techniques are used in the first step of proposed method along with with dataset arrangement through various filtering algorithms in order to improve quality of input images. Pixel values are standardized through normalization which expedites model training. Morovere, histogram equalization improves image contrast for kidney stone visibility. Mask R-CNN algorithm is used with Python, NumPy, and OpenCV, for preprocessing for high efficiently and data readiness.

C. Data Augmentation

To achive more expension of the dataset, we have used multiple augmentation techniques. The nommon techniques are rotation, scaling and translation to the data in order to generate multiple viewing angles and magnification levels. For the replication of tissue distortion, elastic deformation was used and for intensity adjustment, we have mimicked fluctuations in lighting conditions. Due to the proper use of the above augmentation techniques, diverse image variations are vailable in our data set. Therefore, the dataset provide extensive and anotted collection of KUB CT images which will ultimately facilitate effective model learning and stone detection and segmentation.

D. Model Mask R-CNN

In the Mask R-CNN model, the backbone network is the pre-trained networks which are ResNet101, ResNet50, and VGG. The architecture of ResNet101 performs the task of extracting features from images that enter the system. This research experiment uses ResNet101 as well as ResNet50 as backbone networks. The Region Proposal Network (RPN) in Mask R-CNN generate region proposals for objects of interest. These proposals are determined from the feature maps extracted by the backbone network. Region Proposal Network (RPN) adopts multi-task loss described in equation 1 to predict objectness using Lcls while refining box locations with Lreg. The classification loss operates across all anchors whereas the regression loss works only on positive anchors and each loss component gets its normalization through Ncls and $Nreg$ at the same time receiving stabilization through weight factor λ.

$$L\text{rpn}\,(xi, yi) = N\text{cls}1\sum iL\text{cls}(xi, xi*) + \lambda N\text{reg}1\sum ixi * L\text{reg}(yi, yi*) \quad (1)$$

Where, Ncls and Lcls represent normalization terms, and λ is a balancing parameter between classification and regression losses.

The Region of Interest (ROI) Align layer in Mask RCNN plays a pivotal role in guaranteeing precise alignment between features from the backbone network and the region proposals. This alignment is maintained consistently, irrespective of the sizes and shapes of the objects under consideration. The final output from this process is a binary mask that intricately delineates the contours of the kidney stone. The binary cross-entropy loss $Lmask$ mask operates for pixel-level mask predictions within the Mask R-CNN architecture per equation 2. The optimization loss includes the predicted mask fi along with the corresponding ground truth $fi*$ which operates through pixel-based summation p. The loss measure gets normalized by f^2 to maintain scale independence of the output.

$$L\text{mask}(fi, fi*) = -f21\sum p[fi * \log(fi) + (1 - fi*)\log(1 - fi)] \quad (2)$$

To measure the performance of the proposed Mask-RCNN model loss function measure the difference between the ground truth which is kidney and predicted output which are kidney stones.

$$L_{\text{total}} = L_{\text{rpn}} + L_{\text{cls}} + L_{\text{mask}} \quad (3)$$

Here in equation 3, the validation loss is represented by L with classification loss L_{cls} while the bounding box is L_{rpn} and mask loss is $L_{msk.}$

$$Lrpn = L_{cls}^{rpn} + L_{reg}^{rpn} \quad (4)$$

The rpn loss is responsible for creating region proposals (bounding boxes) that possibly hold kidney stones.

$$L\text{mask} = -1/m2\ \sum p[mi * \log(mi) + (1 - mi*)\log(1 - mi)] \quad (5)$$

Here mi is the predicted mask $mi*$ ground truth and $m2$ represents the total pixels number.

E. Training of the Model

The training process begins with loading and preprocessing a dataset of annotated images, split into 9000 training and 3000 validation images. The Mask R-CNN model is constructed with a ResNet-101 backbone, region proposal network, and prediction branches, leveraging pre-trained weights from the COCO dataset for enhanced performance. Transfer learning techniques are applied to reduce training time, with specific RPN commands and parameters adjusted to optimize kidney stone detection and segmentation. As shown in Table 1, for kidney stone detection, the weights were fine-tuned, and the mask head and classifier RPN were trained up to 50 epochs. Learning rate set to 0.001 to refine the weights of kidney stone detection the epochs were set up to 50 epochs while 0.0001, 0.9 are weight decay and momentum respectively.Input images were resized to a maximum dimension of 1280 x 1024 pixels. The model was trained on an NVIDIA GeForce RTX2080 Tix2 graphics processing unit (GPU) with memory of 64 GB.

Throughout the model training phase, our primary tools for implementing and optimizing the model were TensorFlow 2.0 and Keras 2.4. We conducted the training process utilizing Google Colab GPU, enabling accelerated model convergence and reduced training times. This setup effectively optimized the overall training process.

F. Evaluation

The evaluation process involove the systematic assessment of model's performance. Precision, recall, and F1-score were used as evaluation metrices. Qualitative visual error identification and learning curve analysis was assessed in each epoch. Our finding, confirmed the generalizability and robustness of the model when tested on independent validation datasets.

G. Algorithm of the Proposed Research

The procedure for kidney stone detection and segmentation in KUB CT images adopts the following algorithmic approach.

1. Data collection: Collection of KUB CT image dataset.
2. Data Pre-processing: Data cleaning, annotation, normalization, feature scaling, dimensionality reduction and feature engineering
3. Input image: Prepare the input images by resizing or cropping them to the required dimensions.
4. Backbone: Extract features from the input image
5. RPN: Generates candidate bounding boxes
6. ROI: Extract and localize specific regions within an image
7. Mask: Predicting pixel-level masks for each detected stone in an KUB CT image
8. Model Training: Optimizing model's parameters using labeled dataset.
9. Model Evaluation: Assessing the performance and effectiveness of a trained machine learning model on unseen data.
10. Output of the model: Generate bounding boxes, Class labels and instance Segmentation Masks of the kidney stones.
11. End

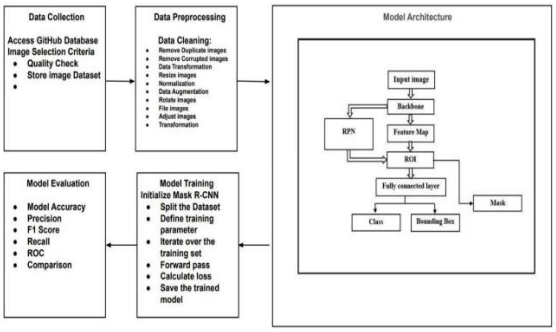

Fig. 1 Graphical overview of the research methodology of the proposed approach.

Table 1 The Hyper-Parameter Training Tuning

S.No.	Backbone	ResNet-101
1	Batch Size	1
2	Epochs	50
3	Learning Rate	0.001
4	Momentum	0.9
5	Optimizer	100
6	Steps Per Epoch	50
7	Validation Steps	0.0001
	Decay	

III. RESULTS AND DISCUSSION

The proposed kidney stone detection and segmentation model yields results through a detailed experimental investigation described in this section. Our evaluation process starts with an overview of the examined dataset before discussing preprocessions steps alongside model architectural decisions. The performance evaluation of the model depends on accuracy alongside precision and recall along with F1-score metrics. The obtained results are contrasted with baseline techniques to demonstrate the success of this proposed technique..

A. Evaluation Metrics

This research uses established metrics including accuracy, precision, recall and F1 score to evaluate the kidney stone detection

and segmentation performance results. These widely recognized indicators serve as robust measures to evaluate the performance of the models.

Accuracy: The segmentation model demonstrated exceptional accuracy, achieving an impressive overall score of 96% in its task of kidney stone detection and segmentation.

Evaluation Criteria: The performance evaluation of kidney stone detection depends on recall, precision together with F-1 score. The ratio between accurate kidney stone recognition in CT KUB images and visual image detections becomes known as P.

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

Mathematically, P is expressed as the ratio of true positives (TP) to the sum of true positives and false positives (FP). Meanwhile, recall, represented by R, measures the ratio of all relevant stone detections to the total number of existing kidney stones.

$$Recall = \frac{TP}{TP + FN} \quad (7)$$

It is calculated as the ratio of TP to the sum of TP and false negatives (FN). F-1 score, denoted by $F1$, is the harmonic mean of precision and recall, offering a balanced assessment of the model's performance.

$$F1\ score = \frac{2 \text{ x } Precision \text{ x Recall}}{Precision + Recall} \quad (8)$$

In this study, kidney stone is identified by matching predicted and actual bounding boxes for kidney stones with an intersection over union value above 0.5 which demonstrates notable overlap between detected and real findings.

$$IoU = \frac{Area\ of\ Overlap}{Area\ of\ Union} \quad (9)$$

The Mean Average Precision (mAP) plays a vital role in evaluating the performance of the Mask R-CNN model in kidney stone detection and segmentation work. Mathematically, mAP is calculated as the mean of the Average Precision (AP) scores across all classes. The Average Precision for each class is computed by first calculating the precision-recall curve and then computing the area under this curve.

The equation for mAP is as follows:

$$\text{mAP} = f1\sum j = 1f\text{AP}(j) \quad (10)$$

Where f is the number of object classes and AP(j) Denotes the Average Precision for class j. A higher mAP score indicates better performance in accurately detecting and segmenting kidney stones across different classes. Thus, mAP serves as a vital metric for assessing the effectiveness of the Mask R-CNN model in kidney stone detection and segmentation tasks. The reserved threshold accommodates the small dimensions of kidney stones which the model needs to handle. The precision- recall and F-1 score together with IoU metrics receive computation using Equations (6), (7), (8), (9) and (10).

These evaluation metrics provide a rigorous framework for quantitatively assessing the effectiveness of kidney stone detection algorithms and guiding further research efforts in this domain. The proposed Mask R-CNN model integrates object detection and instance segmentation within a unified framework to accurately identify and segment objects in input images. The detail analysis of Precision, Recall and F1 Score of different models are for the kidney stone detection and segmentation is shown in Table 2. The graphical representation of performance metrics score like Precision, Recall &F-1 score is shown in Fig. 2.

Table 2 Show the Performance metrics Precision, Recall, F1 Score and (CA) of the proposed model.

Model	Mean average Precision (%)	Precision (%)	Recall (%)	F1 score (%)	Classification accuracy (%)
ResNet101	91	96.9	88	92	98
ResNet50	88	91	85	89	96

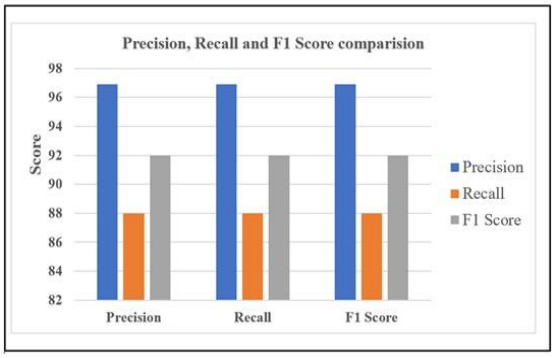

Fig.2 Graphical representation of performance metrics

Learning Curve: The learning curve depicted in Fig. 3. Presents the evolution of the proposed work's performance across successive epochs. It delineates trends in both training and validation accuracies, providing valuable insights into the model's learning trajectory. Initially, both training and validation accuracies exhibit a steady increase, indicative of effective learning and adaptation to the dataset. As epochs advance, the model further refines its performance, with training accuracy steadily approaching its peak potential. Notably, validation accuracy closely mirrors the training accuracy, highlighting successful generalization to unseen data. This learning curve offers a visual depiction of the proposed work's capacity to learn from training data and generalize effectively to validation data, crucial elements in assessing model performance and robustness.

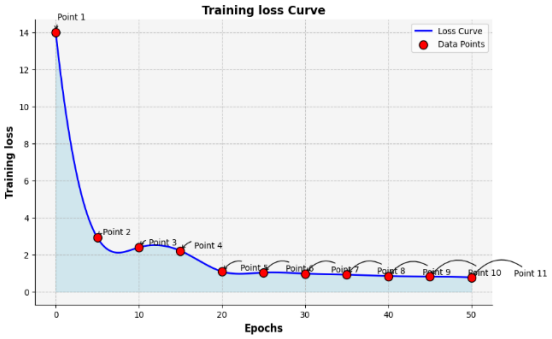

Fig. 3. The training loss progression across epochs 1 to 50. Notably, there are no significant changes in the training loss beyond epoch 21.

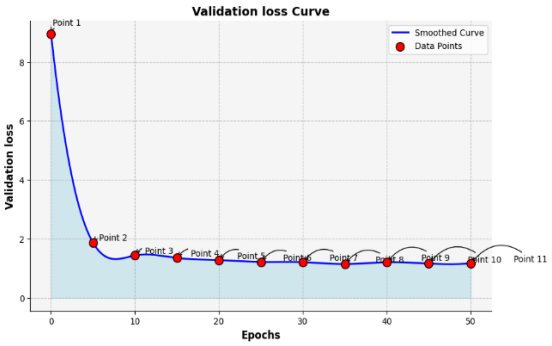

Fig. 4 Illustrates the validation loss trend over epochs 1 to 50. Notably, validation loss stabilizes after epoch 21, indicating a consistent pattern with minimal fluctuations.

The train and validation loss curve is shown in Fig. 4.

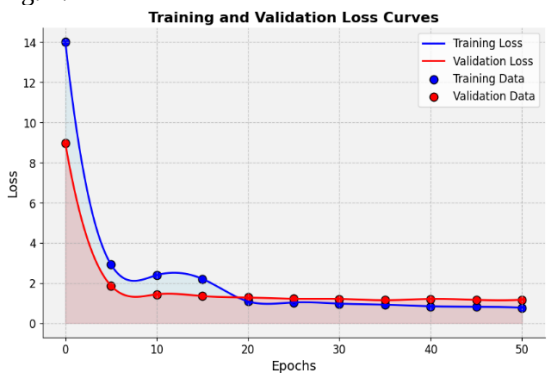

Fig. 5 The graph Illustrates the comprehensive training and validation loss trends captured throughout an experiment spanning 50 epochs, employing a batch size of 10.

B. Visual Results

Visual detection and segmentation results of the proposed model, are shown in Fig 6. In the Figure, (a) stone at left kidney, (b) multiple stones at left kidney are detected (c) Renal pelvis stone, (d) multiple stones are detected and segmented, (e) stone at left kidney upper portion of renal pelvis, (f) large size stone are detected.The Figure exhibiting the output for each class within the dataset. Each Figure depicts common attributes, comprising the bounding box (Bbox), mask, and segmentation results.

C. Discussion

The proposed method demonstrates a significant advancement in kidney stone detection and segmentation using Mask R-CNN, showcasing the superior efficiency and accuracy of deep learning techniques over traditional methods in medical image analysis. It was determined that all kidney stones were detected and segmented with a significantly high confidence factor (which is greater than 90%. while the Classification accuracy (CA) of 98% was attained with ResNet101 and 96% Classification accuracy using ResNet50%. We chose the mean Average Precision (mAP) metric to carefully evaluate how well the object detection works. Many people use this metric in evaluations like ours because it's widely accepted and understood. The mAP calculates the average precision, considering both the number of correct predictions and their order. The best mAP values were 0.88 for ResNet50 and 0.91 for ResNet101. Table 2 show the 92% F1 score for ResNet101 and for ResNet50 are 89%. Evaluation matrices precision, recall and F1 score are shown in Table 2.

The above Table 3 encapsulates a thorough examination of diverse research endeavors centered on the detection and diagnosis of kidney stones. Each entry in the table corresponds to a distinct study, revealing a spectrum of algorithms employed in the pursuit of accurate and efficient detection. Notable methodologies include traditional techniques like Artificial Neural Networks (ANN), Support Vector Machines (SVM), and advanced deep learning approaches such as Mask RCNN and Deep Convolutional Neural Networks (DCNN). The importance column succinctly outlines the unique contributions of each study, encompassing aspects like segmentation alongside detection, early detection capabilities, and the application of specific pre-processing techniques. The reported accuracy figures, ranging from 70% to 95%, shed light on the efficacy of the implemented algorithms. This compilation serves as a valuable resource for researchers and healthcare professionals, offering insights into the evolution of methodologies employed in kidney stone detection and underlining the strides made in enhancing diagnostic accuracy. The graphical representation of analysis of kidney stone detection with different models is shown in Fig. 7.

Table 4 outlines the training and validation losses associated with varying batch sizes used during model training. In the Table, the overall validation and training losses are almost equal between batch size 1 and batch size 10. Although batch number 10 having shorter training time.

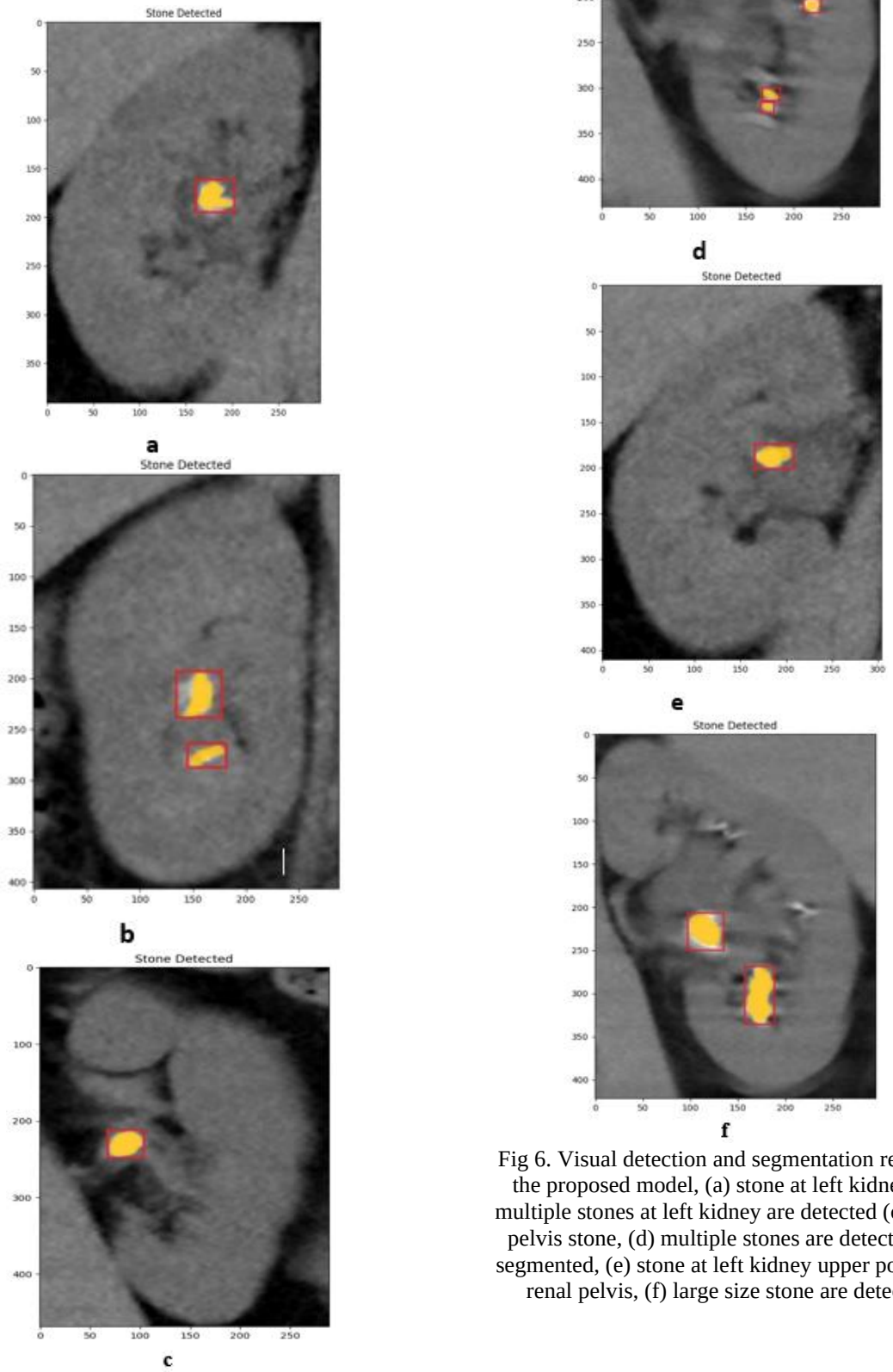

Fig 6. Visual detection and segmentation results of the proposed model, (a) stone at left kidney, (b) multiple stones at left kidney are detected (c) Renal pelvis stone, (d) multiple stones are detected and segmented, (e) stone at left kidney upper portion of renal pelvis, (f) large size stone are detected.

Table 3: Comparative Analysis of kidney stone detection of the proposed research with previous studies

S. No	Title of References	Algorithm	Importance	Accuracy
1	An automatic approach for the detection and segmentation of kidney stone in kub ct images using mask r-cnn	Mask RCNN	Segmentation has been performed alongside detection.	96%
2	Diagnosis of Kidney Stone Disease Artificial Neural Network for [28]	LVQ, RBF and feedforward architecture with back propagation algorithm	Early detection of kidney stones and best model of diagnosis.	92%
3	Urinary Stone Detection in CT Images Using DCNN [29]	Deep Convolutional Neural Networks	CT scan and kidney stone detection	95%
4	Chronic Kidney Disease (CKD) Prediction using CNN [30]	Mathematical concepts	SVM, Random Forest, XGBoost, Neural networks, Naive Bayes Classifier	90%
5	Kidney Stone Detection from ultrasound images by using Canny edge detection and CNN classification [31]	Central Neural Networks	Pre - processing, noise filtering and segmentation	70-85%
6	Artificial Neural Network for kidney Stone Detection [32]	Cuckoo Search Algorithm	generate ROI,using of sobel filtering and edge detection method	94.61%
7	Early detection of Kidney Stones in Ultrasound Images Using Median Filter and Rank Filter[33]	Median filtering algorithm is proposed	median filter and rank filter	82.2%
8	Improving the accuracy of predicting disease [34]	clustering Neural Network	Noise filter, SOM cluster	93%

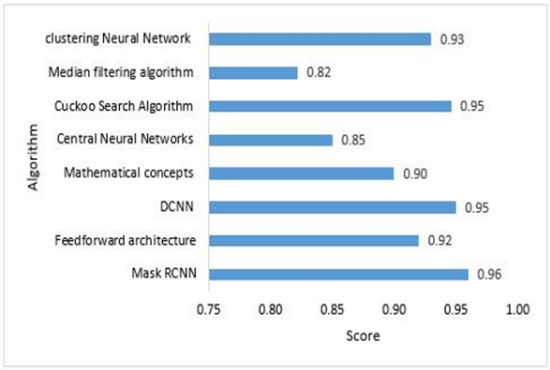

Fig 7. Graphical representation of analysis of kidney stone detection with different models

IV. CONCLUSION

In this paper, an automatic approach for kidney stone detection and segmentation in CT KUB images using Mask R-CNN was presented. Mask R-CNN with back bone architectures ResNet101 and ResNet50 was utilized to achieve kidney stone detection and segmentation with high accuracy and accelerate processing speed. The proposed model is trained on data processed and prepared from individual CT KUB images. Experimental findings indicate that the proposed model can accurately detect and segment kidney stones with a high degree of precision, achieving a mean average precision of 91%. This approach has the capacity to enhance both the efficiency and effectiveness of kidney stone disease management. The research findings suggest that the Mask R-CNN model can serve as a foundation for developing an end-to-end system proficient in automatically detecting and segmenting kidney stones when given KUB CT images as input. A Mask R-CNN-based detection system demonstrates excellent accuracy for kidney stone identification along with region segmentation however it presents multiple operational constraints. Running the model needs large processing power especially when executed with GPU acceleration capabilities either for training or inference steps. There is sufficient potential for further exploration in this area. In future research, we will try to make the model better by using more information about different kinds of kidney problems. Additionaly, improving real-time clinical workflows through the implementation of the developed model while testing its performance in actual diagnosis environments remains an important objective

Table 4: Batch Sizes Comparison Against Validation and Training Losses

Batch size	Validation Loss	Train Loss	Epochs
1	9.9673	13.014	1
	1.8796	1.9152	5
	1.4471	1.3307	10
	1.3613	1.216	15
	1.2895	1.0977	20
	1.2259	1.0229	25
	1.2187	0.9644	30
	1.1546	0.9103	35
	1.2176	0.8516	40
	1.1718	0.8219	45
	1.1811	0.7823	50
5	11.29	12.775	1
	4.469	4.9746	5
	3.0899	2.7309	10
	2.0535	1.9974	15
	1.6614	1.6909	20
	1.5781	1.4612	25
	1.5517	1.3665	30
	1.3778	1.3693	35
	1.4525	1.2955	40
	1.3676	1.2143	45
	1.3668	1.1865	50
10	10.633	11.519	1
	1.9503	1.8724	5
	1.4069	1.3511	10
	1.2847	1.1707	15
	1.2601	1.1045	20
	1.1916	1.0186	25
	1.2302	0.969	30
	1.1932	0.9195	35
	1.1225	0.8686	40
	1.1439	0.8296	45
	1.1556	0.7795	50

REFERENCES

[1] M. Calista, M. Harbaugh, M. Maksim, A. Shlykov, A. Ryan, E. Tsuchida, S. Sven, A. Holcombe, S. Jake, J. Hirschl, C. Stewart, C. Wang, and P. F. Ehrlich, "Morphometric analysis of abdominal organs and rib cage: Implication for risk of solid organ injuries in children," Journal of Trauma-Injury Infection and Critical Care, vol. 78, no. 6, pp. 1129–1133, 2015.

[2] "A comparative study of cellular diversity between the Xenopus pronephric and mouse metanephric nephron," Kidney International, vol. 103, no. 1, pp. 77–86, 2023, doi: 10.1016/j.kint.2022.07.027.

[3] M. Omar, A. Abdulwahab-Ahmed, H. Chaparala, and M. Monga, "Does stone removal help patients with recurrent urinary tract infections?," The Journal of Urology, vol. 194, no. 4, pp. 997–1001, 2015.

[4] C. D. Scales Jr, A. C. Smith, J. M. Hanley, and C. S. Saigal, "Prevalence of kidney stones in the United States," European Urology, vol. 62, no. 1, pp. 160–165, Jul. 2012.

[5] M. Goyal, J. Guo, L. Hinojosa, K. Hulsey, and I. Pedrosa, "Automated kidney segmentation by Mask R-CNN in T2-weighted Magnetic Resonance Imaging," Medical Imaging, vol. 12033, Apr. 2021.

[6] E. Kopelowitz and G. Engelhard, "Lung nodules detection and segmentation using 3D Mask-RCNN," in Proc. IEEE Int. Conf. Computer Vision Workshops (ICCVW), pp. 103–108, 2019.

[7] M. N. U. Haq, A. Irtaza, N. Nida, M. A. Shah, and L. Zubair, "Liver tumor segmentation using ResNet-based Mask R-CNN," in Proc. 2021 Int. Bhurban Conf. Applied Sciences and Technologies (IBCAST), pp. 276–281, Jan. 2021.

[8] P. Rani, A. Thakkar, A. Verma, V. Mehta, R. Chavan, V. S. Dhaka, R. K. Sharma, E. Vocaturo, and E. Zumpano, "KUB-UNet: Segmentation of organs of urinary system from a KUB X-ray image," Computer Methods and Programs in Biomedicine, vol. 224, p. 107031, 2022.

[9] J. Verma, M. Nath, P. Tripathi, and K. K. Saini, "Analysis and identification of kidney stone using Kth nearest neighbor (KNN) and support vector machine (SVM) classification techniques," Pattern Recognition and Image Analysis, vol. 27, no. 3, pp. 574–580, 2017.

[10] A. Soni and A. Rai, "Kidney stone recognition and extraction using directional emboss & SVM from computed tomography images," in Proc. IEEE 3rd Int. Conf. Multimedia Processing, Communication, and Information Technology (MPCIT), pp. 57–62, 2020.

[11] N. Thein, H. A. Nugroho, T. B. Adji, and K. Hamamoto, "An image preprocessing method for kidney stone segmentation in CT scan images," in Proc. 2018 Int. Conf. Computer Engineering, Network, and Intelligent Multimedia (CENIM), pp. 147–150, 2018.

[12] W. M. Hafizah, E. Supriyanto, and J. Yunus, "Feature extraction of kidney ultrasound images based on intensity histogram and gray level co-occurrence matrix," in Proc. 6th Asia Int. Conf. Mathematical Modelling and Computer Simulation (AMS), pp. 115–120, 2012.

[13] R. Girshick, J. Donahue, T. Darrell, and T. Malik, "Region-based convolutional networks for accurate object detection and segmentation," IEEE Trans. Pattern Anal. Mach. Intell., vol. 38, no. 1, pp. 142–158, 2015.

[14] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in Proc. IEEE Int. Conf. Computer Vision (ICCV), 2017.

[15] J. Zhang, P. Ma, T. Jiang, X. Zhao, W. Tan, J. Zhang, S. Zou, X. Huang, and M. Grzegorzek, "SEM-RCNN: A squeeze-and-excitation-based mask region convolutional

neural network for multi-class environmental microorganism detection," Applied Sciences, vol. 12, no. 19, p. 9902, 2022, doi: 10.3390/app12199902.

[16] "Face mask detection using convolutional neural network," Proc. ICBIR, 2023, doi: 10.1109/icbir57571.2023.10147730.

[17] "Deep learning based defect classification and detection in SEM images: A Mask R-CNN approach," arXiv, 2022, doi: 10.48550/arxiv.2211.02185.

[18] V. Karunakaran, "Deep learning based object detection using mask RCNN," in Proc. 6th Int. Conf. Communication and Electronics Systems (ICCES), IEEE, 2021.

[19] Y. Ge, Q. Zhang, Y. Sun, Y. Shen, and X. Wang, "Grayscale medical image segmentation method based on 2D&3D object detection with deep learning," BMC Medical Imaging, vol. 22, no. 1, p. 33, 2022.

[20] T. Hou and J. Li, "Application of mask R-CNN for building detection in UAV remote sensing images," Heliyon, vol. 10, no. 19, 2024.

[21] P. Kiatphaisansophon, D. Wanvarie, and N. Cooharojananone, "Efficient text bounding box identification using Mask R-CNN: Case of Thai documents," IEEE Access, 2024.

[22] J. Wang, M. Wu, Y. Guo, H. Wu, and Z. Wan, "Evaluating fairness of Mask R-CNN for kidney infection detection based on renal scintigraphy," in Proc. 2024 IEEE Int. Conf. Big Data (BigData), pp. 4637–4642, Dec. 2024.

[23] D. Tian, Y. Han, B. Wang, T. Guan, H. Gu, and W. Wei, "Review of object instance segmentation based on deep learning," Journal of Electronic Imaging, vol. 31, no. 4, p. 041205, 2022.

[24] W. Zhang, Y. Wang, G. Shen, C. Li, M. Li, and Y. Guo, "Tobacco leaf segmentation based on improved MASK RCNN algorithm and SAM model," IEEE Access, vol. 11, pp. 103102–103114, 2023.

[25] "Kidney stone detection," GitHub, [Online]. Available: https://github.com/muhammedtalo/Kidney_stone_detection