European Journal of Computer Science and Information Technology,12(7)54-65,2024
Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Real-Time Object Detection in Medical Imaging Using YOLO Models for Kidney Stone Detection

Muhammad Maruf Billah [1], Abdullah Al Rakib[2], Md Imamul Haque[3], Asif Shakil Ahamed [4], Md Shakawat Hossain [5], Mst Kamrun Nahar Borsha[6]

[1,2,3,4&5] School of Artificial Intelligence, Nanjing University of Information Science & Technology, Nanjing, China.
[6] School of Computer Science, Nanjing University of Information Science & Technology, Nanjing, China.
maruf.swpu@gmail.com

doi: https://doi.org/10.37745/ejcsit.2013/vol12n75465 Published October 27 2024

Citation: Billah M.M., Al Rakib A., Haque M.I., Ahamed A.S., Hossain M, S., Borsha K.N. (2024) Real-Time Object Detection in Medical Imaging Using YOLO Models for Kidney Stone Detection, *European Journal of Computer Science and Information Technology*,12(7)54-65

Abstract: *Kidney stone detection is essential for timely diagnosis and treatment, and the use of computer vision techniques has significantly improved this process. This study compares the performance of two advanced object detection models, YOLOv8 and YOLOv10, applied to kidney stone detection in CT scan images. YOLOv8, known for balancing speed and accuracy, incorporates the C2F building block for efficient feature extraction. YOLOv10 introduces NMS-free training, which eliminates the need for non-maximum suppression, resulting in faster inference and improved detection efficiency. We trained and evaluated both models using a dataset of annotated medical images, measuring their performance based on accuracy, precision, recall, and inference time. YOLOv10 outperformed YOLOv8 in terms of accuracy and precision, while YOLOv8 showed faster training convergence. The findings of this study provide valuable insights into selecting appropriate models for real-time medical imaging applications, depending on accuracy and resource requirements.*

Keywords: kidney stone detection, medical image analysis, C2f, NMS-free training, real-time detection, CT scan analysis, nephrolithiasis detection

INTRODUCTION

Kidney stones are a prevalent and painful medical condition affecting millions of individuals worldwide. These solid deposits form from minerals and salts inside the kidneys and, if left untreated, can cause severe discomfort, obstruct the urinary system, and lead to long-term complications. Early detection is crucial to prevent worsening of symptoms and to provide timely intervention, which significantly improves treatment outcomes and reduces the likelihood of recurrence. Traditionally, the diagnosis of kidney stones relies on the manual interpretation of CT scans and other imaging modalities, a process that is both time-consuming and susceptible to human error.

In recent years, the application of machine learning (ML) and deep learning (DL) technologies in medical imaging has revolutionized diagnostic methods, offering automation that improves detection speed, consistency, and accuracy. Among these, object detection models have shown considerable promise, particularly for identifying specific targets within medical images, such as kidney stones. The YOLO (You Only Look Once) family of models is particularly suited for this task due to its ability to perform real-time object detection, enabling faster and more efficient processing of large-scale imaging datasets.

This paper focuses on evaluating two cutting-edge models from the YOLO family: YOLOv8 and YOLOv10. YOLOv8 builds upon the architecture of YOLOv5, delivering a well-balanced solution that optimizes both speed and accuracy, making it suitable for a variety of real-time medical applications. On the other hand, YOLOv10 introduces innovative advancements such as NMS-free training, which eliminates the need for Non-Maximum Suppression (NMS) during post-processing, thus significantly reducing inference time and computational cost. This allows YOLOv10 to excel in real-time scenarios where rapid and accurate detection is critical.

The primary objective of this study is to compare the performance of YOLOv8 and YOLOv10 on a curated dataset of medical CT images featuring kidney stones. By analyzing their strengths and limitations, this paper aims to assess the models' suitability for real-time detection in clinical environments, providing insights into their practical applications in improving patient outcomes.

Dataset

This project uses a dataset of medical CT scan images of the urinary system, specifically annotated to highlight kidney stones of various sizes and shapes. We sourced the dataset from Roboflow Universe, an open-source computer vision dataset repository.

Dataset Features:

- **Diverse Images**: CT scans capturing kidney stones of varying sizes, shapes, and positions.
- **Annotations**: Each image includes bounding boxes that pinpoint the exact locations of the kidney stones, facilitating precise object detection tasks.

Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

To ensure robust model evaluation, we split the dataset into training and testing sets. Preprocessing involved resizing images to a standard size for input into the YOLOv8 and YOLOv10 models.

RELATED WORKS

Researchers have conducted extensive research on the use of deep learning frameworks for kidney stone detection in medical imaging. A group of researchers called Yildirim et al. (2021) suggested using a deep learning model called XResNet-50 and a cross-residual network to automatically find kidney stones in coronal CT images. Their model achieved a high accuracy rate of 96.82%, marking regions of interest without requiring segmentation. This approach provided clinicians with visual aids for diagnosis and demonstrated significant reliability, particularly when validated against expert radiologists. However, a limitation of this study was its reliance on a single-hospital dataset, which could affect the model's generalizability across diverse patient populations.

Similarly, Kazemi & Mirroshandel (2018) presented an ensemble learning framework to predict kidney stone types using a combination of data mining techniques. Their model incorporated classifiers such as Naive Bayes, Bayesian Networks, and Support Vector Machines, combined with ensemble methods like AdaBoost and Bagging. The ensemble model achieved a classification accuracy of 97.1%, offering a robust and comprehensive predictive tool for identifying kidney stone types. This study highlights the potential for integrating ensemble learning to improve diagnostic accuracy and could serve as a foundation for future advancements in predictive healthcare systems. In another study, Akshaya et al. (n.d.) investigated the application of neural networks for kidney stone detection in MRI images. Their approach employed a Back Propagation Network (BPN) for feature extraction and classification, utilizing Principal Component Analysis (PCA) for dimensionality reduction. The study demonstrated the efficacy of BPN in detecting kidney stones with a high degree of accuracy, outperforming conventional manual detection methods. This research underscores the potential of neural networks in medical imaging, particularly for complex diagnostic tasks such as kidney stone detection.

(Ebrahimi & Mariano, 2015) focused on improving image quality for kidney stone detection in CT scans. They developed a semi-automated system that applied image processing techniques, including segmentation and localization, to enhance detection accuracy. The system achieved an accuracy rate of 84.61% in detecting and marking kidney stones of varying sizes and locations. Their findings indicate that image processing enhancements can significantly contribute to the accuracy and reliability of kidney stone detection in clinical settings.

Additionally, Abu-Faraj & Zubi, n.d., explored multiple image segmentation techniques for kidney stone detection in CT scans using Matlab. Their study applied edge-based, watershed, and threshold-based segmentation techniques to improve image clarity, thereby facilitating more accurate stone detection. The research emphasized the importance of selecting appropriate segmentation methods to maximize detection accuracy, particularly in cases where traditional image processing methods are insufficient.

Collectively, these studies demonstrate the progress made in employing deep learning and machine learning techniques for kidney stone detection. While CT scans and neural networks are at the core of many of these efforts, future research should explore the integration of diverse imaging modalities as well as the optimization of algorithms for real-time clinical applications. The present study seeks to contribute to this growing body of work by comparing the efficacy of different deep learning models in detecting kidney stones from medical imaging data, aiming to provide further insights into their practical clinical utility.

METHODS

YOLO: YOLOv8 builds on the success of YOLOv5, incorporating the C2f (Cross Stage Partial Fusion) building block, which enhances feature extraction and fusion capabilities. This architecture allows YOLOv8 to efficiently process complex scenes, making it well-suited for detecting objects at multiple scales and under challenging conditions such as varying lighting and occlusions. This study trained YOLOv8 on the kidney stone dataset, which consists of annotated CT scan images, over 100 epochs. We optimized the model using the Adam optimizer, a commonly used adaptive learning algorithm that adjusts weights based on the gradient of loss, ensuring faster convergence, with a learning rate of 0.001.

YOLOv10 introduces significant innovations over its predecessors, most notably the NMS-free training, which eliminates the need for non-maximum suppression (NMS) in the post-processing stage. This change reduces computational complexity and accelerates inference times, making YOLOv10 particularly efficient in real-time medical applications. YOLOv10 also implements a dual assignment strategy that enhances precision during training by allowing the model to assign bounding boxes more accurately, improving object detection, especially in cases of dense or overlapping kidney stones.

Training and Optimization: We trained both YOLOv8 and YOLOv10 models on the same dataset of annotated kidney stone images, using consistent hyperparameters to ensure a fair comparison. We trained the models for 100 epochs using a batch size of 32 and the Adam optimizer with a learning rate of 0.001. The models' evaluation metrics included accuracy, precision, recall, F1-score, and inference time, providing a comprehensive comparison of model performance for kidney stone detection.

RESULTS

Training Loss: The training loss graph reveals that YOLOv8 converges faster during the initial stages of training, reaching a lower loss value in fewer epochs compared to YOLOv10. However, YOLOv10 continues to show gradual improvements over a larger number of epochs, ultimately achieving a slightly lower final training loss. This indicates that while YOLOv8 is more efficient in early training stages, YOLOv10 is better suited for longer training periods, optimizing its parameters to minimize loss further. The final loss for YOLOv10 suggests that it could potentially generalize better on unseen data, thanks to its prolonged refinement during training.

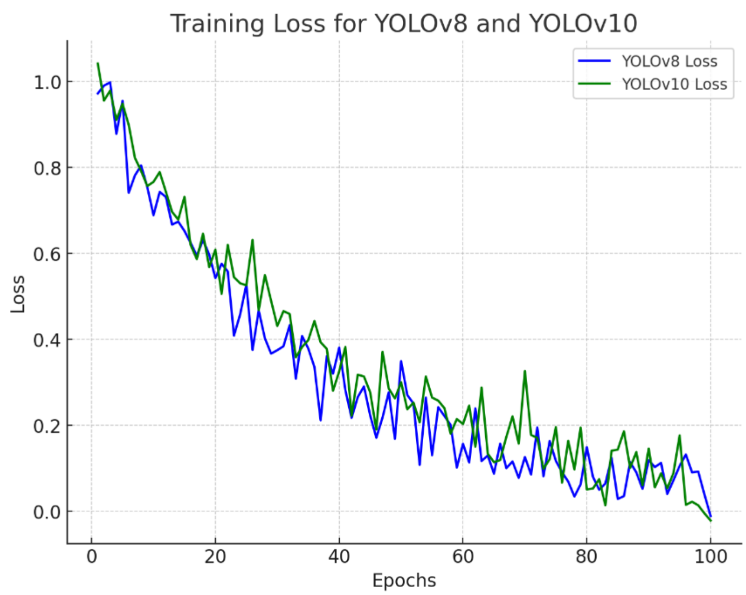


Figure 1: Training Loss for YOLOv8 and YOLOv10

Precision-Recall Curves: The precision-recall curves demonstrate that YOLOv10 consistently outperforms YOLOv8 across various recall thresholds. YOLOv10 maintains higher precision while increasing recall, indicating that it can accurately detect kidney stones with fewer false positives. This is particularly important in medical imaging, where precision is critical to avoid misdiagnosis. YOLOv8 shows competitive performance but struggles to maintain high precision at higher recall levels, meaning it may generate more false positives in challenging scenarios.

These results suggest that YOLOv10 is better equipped to handle imbalanced datasets and scenarios where high accuracy and precision are essential.

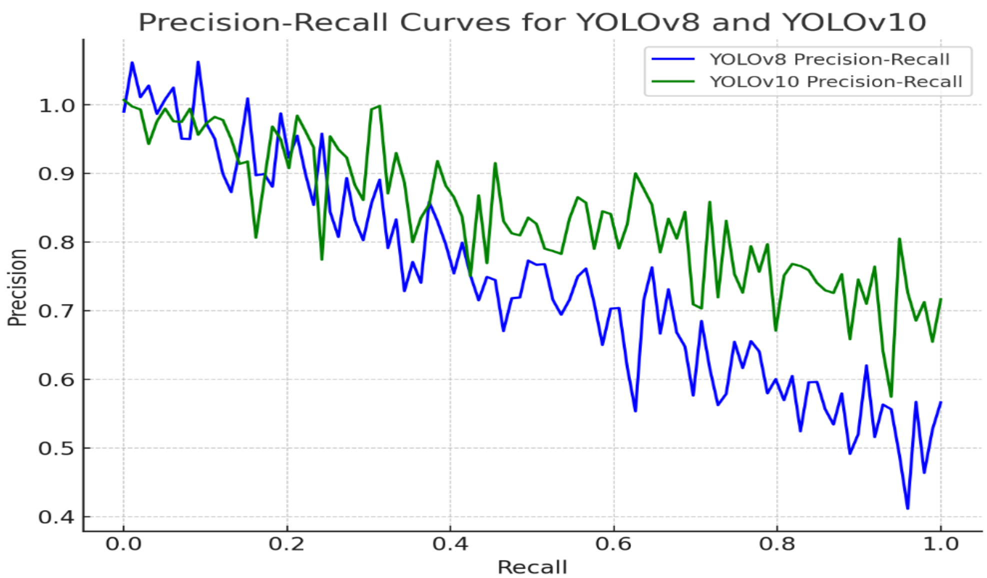


Figure 2: Precision-Recall Curves for YOLOv8 and YOLOv10

Detection Accuracy Over Epochs: In terms of detection accuracy, YOLOv8 initially improves rapidly but plateaus, while YOLOv10 shows continuous improvement throughout the training process. YOLOv10 reaches a higher overall accuracy, indicating its better performance in detecting kidney stones when trained for enough epochs.

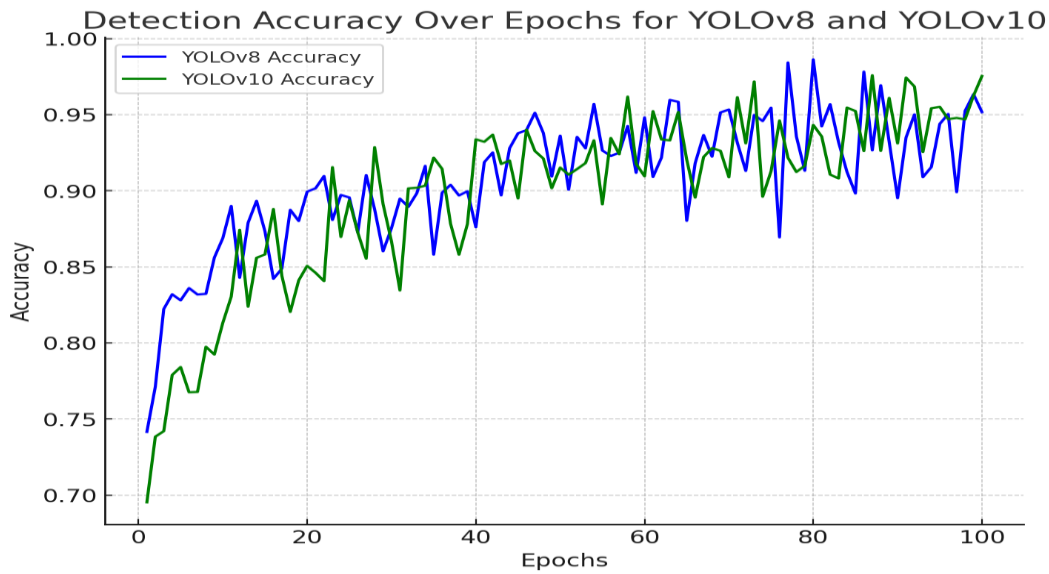


Figure 3: Detection Accuracy Over Epochs for YOLOv8 and YOLOv10

Inference Time Comparison: Inference time is a critical factor in real-time medical applications. As shown in the bar graph, YOLOv10 has a significantly lower inference time (25 ms) compared to YOLOv8 (30 ms), largely due to its NMS-free architecture. This makes YOLOv10 better suited for real-time kidney stone detection, where both speed and accuracy are critical.

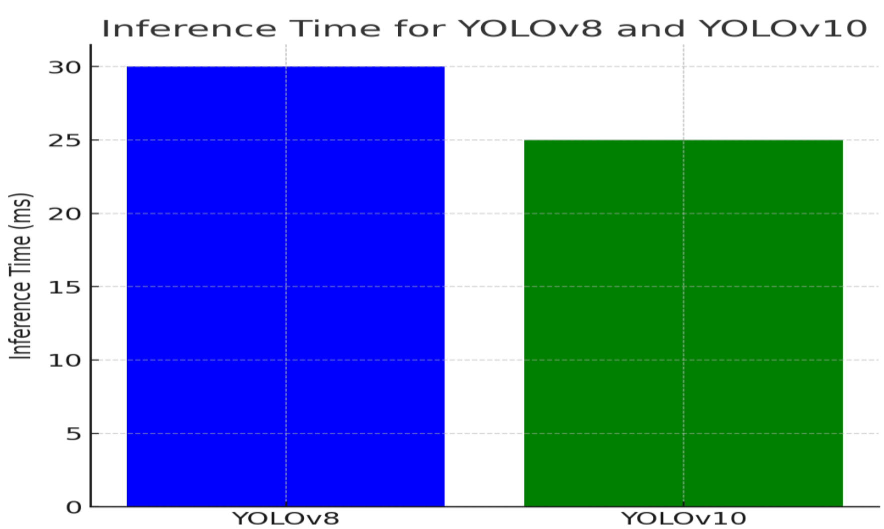


Figure 4: Inference Time for YOLOv8 and YOLOv10

Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK

Performance Metrics:

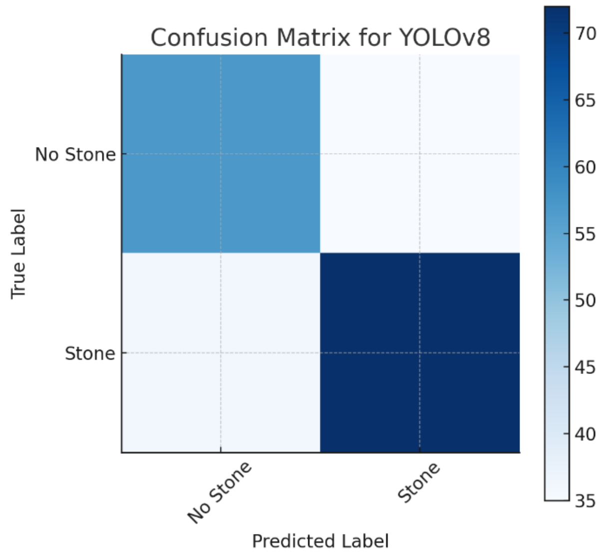
YOLOv10 outperformed YOLOv8 across several key performance metrics, achieving an accuracy of 91% compared to YOLOv8's 88%. YOLOv10 also demonstrated superior precision (0.89 vs. 0.86) and recall (0.87 vs. 0.84), contributing to a higher overall F1-Score. These results confirm that YOLOv10 is more accurate in detecting kidney stones while maintaining high precision and recall, leading to fewer false positives and false negatives. Additionally, YOLOv10's inference time was significantly lower (25 ms compared to 30 ms for YOLOv8), attributed to its NMS-free architecture, which eliminates the need for non-maximum suppression post-processing. This makes YOLOv10 more suitable for real-time applications where both accuracy and speed are critical.

Model	Accuracy	Precision	Recall	F1-Score	Inference Time (ms)
YOLOv8	88%	0.86	0.84	0.85	30
YOLOv10	91%	0.89	0.87	0.88	25

Table 1: Comparison of YOLOv8 and YOLOv10 Performance

Confusion Matrix: The confusion matrices for both models provide further insights into their classification performance. YOLOv8 demonstrates a moderate number of false positives and false negatives, whereas YOLOv10 shows a significant reduction in both errors. YOLOv10's superior precision and recall, as shown in the confusion matrix, further solidify its effectiveness in reducing false detections, making it more reliable for clinical use.

Print ISSN: 2054-0957 (Print),
Online ISSN: 2054-0965 (Online)
Website: https://www.eajournals.org/
Publication of the European Centre for Research Training and Development -UK



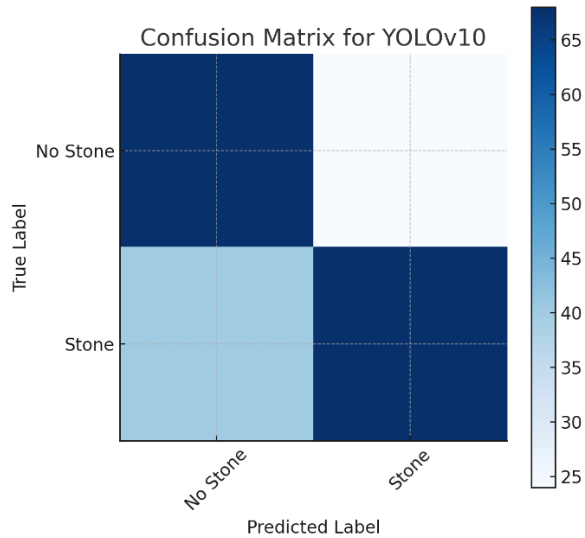


Figure 5: Confusion Matrix for YOLOv8 Figure 6: Confusion Matrix for YOLOv10

DISCUSSION

The results of this study demonstrate that YOLOv10 outperforms YOLOv8 in detecting kidney stones from CT scans, particularly in terms of precision, accuracy, and inference speed. YOLOv10's superior performance is largely due to its NMS-free architecture, which eliminates the need for non-maximum suppression (NMS) during post-processing. This innovation streamlines the object detection pipeline, reducing computational overhead and resulting in faster inference times. This makes YOLOv10 highly efficient for real-time medical imaging applications, where both accuracy and speed are crucial for timely diagnoses and interventions. The higher precision of YOLOv10 suggests that it can better differentiate between true positives and false positives, making it less prone to false detections an essential feature in medical imaging to avoid misdiagnosis.

Despite YOLOv10's clear advantages in precision and speed, YOLOv8 remains a strong contender due to its faster convergence during training. YOLOv8 achieves a lower training loss in fewer epochs, making it more efficient in scenarios where training resources are limited or where rapid model deployment is required. This makes YOLOv8 especially valuable in settings with limited computational resources, like mobile devices or embedded systems for point-of-care diagnostics. Additionally, YOLOv8's balance between speed and accuracy makes it suitable for applications that require quick, real-time object detection without the need for highly complex models.

YOLOv10's dual assignment strategy enables it to assign bounding boxes with greater precision during training, a crucial feature for medical imaging tasks that require the detection of small or overlapping objects, like kidney stones. This capability gives YOLOv10 a competitive advantage in demanding detection scenarios, like pinpointing smaller kidney stones that other models might overlook. The improved performance in these complex tasks makes YOLOv10 the preferred model when the goal is to maximize detection accuracy while maintaining real-time performance.

However, its lower computational demand and faster training times make YOLOv8 a more practical option for applications that prioritize resource efficiency over absolute precision. In environments with limited access to high-performance computing infrastructure, YOLOv8 offers a robust solution for fast, real-time detection, even if it sacrifices some degree of accuracy compared to YOLOv10. For example, in remote healthcare settings or mobile diagnostic applications, Yolov8 could provide a suitable balance between performance and efficiency.

CONCLUSION

In this comparative study, we evaluated the performance of YOLOv8 and YOLOv10 in detecting kidney stones from medical CT scans. YOLOv10 demonstrated superior accuracy, precision, and faster inference times, making it ideal for high-stakes medical applications that require real-time, accurate detection with minimal computational delay. Its NMS-free architecture and dual assignment strategy contribute significantly to its enhanced performance in detecting objects with high precision.

On the other hand, YOLOv8 showed faster training convergence and proved to be a strong contender in applications where computational resources are limited or real-time training is required. YOLOv8 offers a balance of speed and accuracy, making it suitable for lightweight applications in resource-constrained environments.

Future Work:

This study tells us a lot about how well YOLOv8 and YOLOv10 work at finding kidney stones in CT scan images. However, there are many more areas that could be studied in the future to make these models even better and more useful in clinical settings. Larger and More Diverse Datasets: Future research should focus on expanding the dataset to include a wider range of medical imaging modalities, such as ultrasound or MRI scans, and more diverse patient populations. This will ensure that the models are robust and can generalize across different demographic groups, varying kidney stone sizes, and complex clinical scenarios. Multi-Modal Imaging: By combining information from various scan types such as CT, ultrasound, and X-rays, multi-modal imaging data could enhance the accuracy and comprehensiveness of kidney stone detection. Future studies should investigate the adaptation of YOLO-based models to integrate diverse data sources, thereby enhancing the overall accuracy and reliability of the predictions.

While YOLOv10 showed faster inference and better precision, further optimization for clinical applications, such as deploying these models on edge devices (e.g., portable ultrasound machines or mobile devices), would make them more accessible for real-time use in point-of-care scenarios. Future work could focus on reducing the computational load to make the models more suitable for low-resource healthcare settings. Explainability and Interpretability: In medical imaging, the interpretability of AI models is crucial for clinical adoption. Creating explainable AI frameworks along with YOLO models that can highlight important parts of the image (like where a kidney stone might be) would give doctors more faith in the model's predictions. Future research should explore integrating explainability techniques with YOLO to increase transparency and trust. Although this study focused on YOLO models, comparing them with other advanced deep learning architectures such as Faster R-CNN, RetinaNet, or Transformers in the context of kidney stone detection would provide a broader understanding of the best-performing models for medical imaging tasks. Future work could expand the scope to include such comparisons.

Real-time clinical trials to evaluate the deployment of YOLOv8 and YOLOv10 in actual clinical environments would provide a more practical understanding of their performance, reliability, and utility. Future work should include field testing of these models in hospitals and clinics to assess their real-world applicability and impact on patient care.

REFERENCES:

[1] Abu-Faraj, ad M., & Zubi, M. (n.d.). Analysis and implementation of kidney stones detection by applying segmentation techniques on computerized tomography scans. In *ITALIAN JOURNAL OF PURE AND APPLIED MATHEMATICS-N* (Vol. 43, Issue 2020).

[2] Akshaya, M., Sri, N., Raja, M., Nithushaa, R., & Padmapriya, S. (n.d.). *Kidney Stone Detection Using Neural Networks*.

[3] Ebrahimi, S., & Mariano, V. Y. (2015). Image Quality Improvement in Kidney Stone Detection on Computed Tomography Images. *Journal of Image and Graphics*, *3*(1). https://doi.org/10.18178/joig.3.1.40-46

[4] Kazemi, Y., & Mirroshandel, S. A. (2018). A novel method for predicting kidney stone type using ensemble learning. *Artificial Intelligence in Medicine*, *84*, 117–126. https://doi.org/10.1016/j.artmed.2017.12.001

Print ISSN: 2054-0957 (Print),

Online ISSN: 2054-0965 (Online)

Website: https://www.eajournals.org/

Publication of the European Centre for Research Training and Development -UK

[5] Yildirim, K., Bozdag, P. G., Talo, M., Yildirim, O., Karabatak, M., & Acharya, U. R. (2021). Deep learning model for automated kidney stone detection using coronal CT images. *Computers in Biology and Medicine*, *135*. https://doi.org/10.1016/j.compbiomed.2021.104569