Review of DocExpert: A Hybrid Model for Medical Image Detection using YOLO, EfficientDet, and DETR

Nachiket Dhampalwar [1] Mansha Mukoo[2] Pranav Dhanawade[3] Maitri Shinde[4]
Mrs. Reena Sahane[5] Ms. Surbhi Pagar[6]

[1,2,3,4] *Author, Artificial Intelligence & Data Science, Dr. D. Y. Patil Institute of Engineering, Management and Research, Akurdi, Pune-44, India*

[5,6] *Guide, Artificial Intelligence & Data Science, Dr. D. Y. Patil International University, Akurdi, Pune-44, India*

[1] nachiketdhampalwar@gmail.com, [2] mukoomansha@gmail.com, [3] pranavdh13@gmail.com, [4] maitrishinde8@gmail.com,
[5] reena.sahane@dypiemr.ac.in, [6] surbhi.pagar@dypiemr.ac.in

***Abstract*—In this article, the evolution and**

At the time of writing, an evaluation study on DocExpert, a medical image detection system and not to mention Hoo et al. gazillion innovations [3] which poses challenges in terms space but may also be obsolete soon (by as early 10 years ago with LP-XTOT) have an impact on the hybrid model that combines YOLO, EfficientDet and DETR to accurately detect fractures. The system aims to increase the diagnostic specificity and speed up processing. automate fracture detection in order to better serve patient. medical x-ray imaging. In this work, we show the performance of these models on tissue characterization or determination between tissue, they have been studied for detecting medical imaging datasets [9]. States of Change: speed and real-time processing two critical missions through a new power coalition

***Keywords*—Fracture detection, Medical imaging, YOLO, EfficientDet, DETR, Deep learning, Flask, Object detection, Healthcare technology, AI-driven healthcare, Computer vision, Medical diagnosis.**

I. Introduction

A. Domain

One of the most obvious areas where AI has been heralded as a key area for growth is within medical image analysis in healthcare. This nascent field of study has seen the develop- ment of sophisticated, automatic diagnostic models. One of the most important areas to focus on is identifying fractures, where timely and accurate diagnosis are key in ensuring patients receive appropriate care. The conventional techniques majorly involve manual examination by the radiologists which makes them time consuming and subject to human error [1]. AI can also be used to improve amplitude and shorter analysis time by medical practitioners.

B. Application

DocExpert is a product that modernized X-ray fracture detection using automation. The system uses an ”emergent” hybrid model that is designed to reduce reliance on manual radiological reviews. By letting automation do the heavy lifting, DocExpert not only speeds up how quickly a diagnosis is made but also raises the quality and reliability of diagnoses from one case to another. It has also the potential to change how work is done in medical imaging departments, radiologists will spend much less time on routine cases and be able focus more on advanced ones which ultimately improve patient care.

C. Model/Methodologies

At its heart, DocExpert is a hybrid system that combines the power of three major object detection algorithms like YOLO (You Only Look Once), EfficientDet and DETR(DEtection TRansformer).

YOLO is famous for being a real-time detector, which makes it good at instant on-the-fly applications.

Meanwhile, EfficientDet is also appealing due to its ability to perform scale versus accuracy mapping. Its design makes it capable to leverage computational resources efficiently for achieving high detection accuracy while being adjustable and efficient under different computation budgets.

Is it for its ability to read even more complex image environments and be able to pick on extreme detail and exceptions that other models would forgo. With an MLS of the said model types, the developers are able to make a reliable and accurate fractures detection system that goes by as DocExpert.

D. Analysis

Much of its efficiency comes from leveraging established medical imaging datasets as in DocExpert [5]. Key performance measures include detection accuracy[8] (i.e., does the model detect a fracture) and selective accuracies for specific features to determine fractural type identification, as well as image quality. And how well it utilizes the computing resources is very, because that determines if you can used this model in real time clinical conditions or not. The analysis demonstrates that the system can not only satisfy but even surpass existing diagnostic imaging benchmarks, which establishes a novel gold standard in unassisted fracture recognition.

II. RELATED WORK

A. Literature Review

Several researchers have conducted investigations concerning the use of AI models for the detection of medical images [1]. With integration of artificial intelligence in medical imaging, there has been effective diagnosis and quicker clinical decisions. Particularly, YOLO (You Only Look Once) algorithm has found application in the field of medical imaging to facilitate object detection tasks [2]. Network based on YOLO architecture becomes a favorable option for applications in- volving rapid assessment, such as detection of anomalies in radiological images and disease progression monitoring because it provides real time detection capabilities and requires only one forward pass of the network for detection.

A Finetuned model that is also a strong model given the fact that it was developed by Google for scalable and accurate performance in medical image analysis tasks is EfficientDet [3]. There has been a well defined compound scaling strategy incorporated into the model, which relates to the growth of the model's depth, width, and resolution in relation to each other As a result, the efficiency of the model is maximized under different levels of available resources. EfficientDet has also been used in pathologies where it is used to detect cancer cells in histopathological images.

DETR (DEtection TRansformer) is also noted for its capability to work in complex scenes of an image [4]. Instead of anchors, predefined bounding box methods used in object detection, DETR uses a transformer neural network to pre- dict directly the collection of objects in an image. Hence it becomes easy to establish object detection in complicated or obscured scenes which is quite important in medical imaging since the area of interest may not be extensively clear.

The current literature review narrows down to five object detection algorithms: YOLOv3, YOLOv4, Mask R-CNN, Ef- ficientDet and YOLOv9 [5]. All these algorithms, when it comes to their primary purpose, have a very high efficiency level when detecting objects within images and videos, al- though, with a few strengths and weaknesses each that would determine their effectiveness in different scenarios.

1) Object Detection Algorithms: According to the report commissioned by IMARC Group I, the task of object detection belongs to the category of fundamental processes in computer vision. The application has gained interest due to a variety of fields where it can be applied, self-driving technology, traffic prying, hospital settings and drone flying. Deep learning has seen better days with the algorithms becoming accurate and proficient in object detection tasks. As for the practical use of these models, they are usually tested with precision, recall, accuracy and Intersection-over-union as their scale for performance [2], [6]. Also, the metric of resource consumption such as model complexity, memory and FLOPs per second has proved very useful in assessing how the models would be used in practice. This is more so in medical applications where tough conditions are often the case and real-time performance is seen as essential [3], [7], [8].

TABLE I
PERFORMANCE COMPARISON OF OBJECT DETECTION ALGORITHMS ON COCO DATASET

Model	Accuracy (%)	AP (Average Precision)
YOLOv3	98	45.5
YOLOv4	99	47.3
Mask R-CNN	97	44.7
EfficientDet-D7	99	52.2
YOLOv9	99.2	53.1

2) YOLOv3 and YOLOv4: YOLOv3 and YOLOv4 are real-time object detection algorithms that have achieved high accuracy. give two algorithms to detect objects in images and videos. This algorithm use annotated labeled images to find the location of opject in image. On the KITTI dataset, which consists of images, YOLOv3 achieved an accuracy of 98% for image detection and the YOLOv4 reached an to 99% accuracy for video detection [2], [6]. Both models predictions multiple object at a time and the use of bounding boxes to enhance detection performance. Despite their accuracy, both models require more hardware so it has high computational costs that indicated by their parameter count and FLOPs (see Table I).

Recently YOLO4 have some improvements in there features such as CSPDarknet53 for better gradient flow and it also increased receptive fields which is why it performance is increased drastically. Enhancements have allowed YOLOv4 to be effectively applied in medical or any other object detection scenario. real-time tumor detection in MRI scans [3], [7], [8],showcasing its versatility

3) EfficientDet: EfficientDet is mostly known for its scalability and efficient performance. The algorithm employs a weighted bi-directional feature pyramid network (BiFPN) [3] and a compound scaling method that uniformly scales the resolution, depth, and width for the backbone, feature net- work, and box/class prediction networks [3]. EfficientDet-D7 achieved state-of-the-art results on the COCO test-dev dataset, with an average precision (AP) of 52.2 and 325 billion FLOPs, highlighting its trade-off between accuracy and efficiency (see Table I) [11].

4) YOLOv9: To improve flexibility in a variety of activities, YOLOv9 presents programmable gradient information (PGI) [9]. This feature allows YOLOv9 to dynamically modify its architecture according to the task's complexity. With an accuracy of 99.2% and an average precision of 53.1 AP, it has demonstrated exceptional performance on the MS COCO dataset [4].

YOLOv9 is a top contender for resource-constrained applications because to its lightweight design and great accuracy [9]. Additionally, compared to earlier YOLO models, YOLOv9 exhibits superior parameter usage, enabling it to provide great performance with fewer computing resources [12].

5) Conclusion: These five object detection algorithms—YOLOv3 [2], YOLOv4 [6], Mask R-CNN, EfficientDet [3], and YOLOv9 [9] have high accuracy and used for both image and video . While YOLOv3 and YOLOv4 are excellent real-time detection capabilities [9], Mask R-CNN uses in segmentation tasks, providing precise delineation of complex structures [7]. EfficientDet good at resource use and accuracy, making it suitable for wide range of applications [8]. YOLOv9 leads in lightweight object detection and it have high adaptability [4], it also used for such as traffic monitoring, surveillance, and medical imaging.

B. Implications of AI in Medical Imaging

AI in health care is good for both. Through integration, patients get better suggestion from AI model or treatment from AI models fast, hence less painful to be dragged away from hospital. AI model can detect conditions of patients quickly, accurately through automated systems .It may eventually lead earlier diagnosis.

AI cannot make errors like humans do so it minimizes human error. Variability in the interpretation of images by radiologists may cause some misdiagnosis because using AI models as an auxiliary tool can standardize readings, especially when the cases are complicated MuhammedS and Mishra,2020. In critical care situations, decisions must come at the soonest possible time.

Another significant advantage of AI in medical imaging is its ability to analyze large data. AI models can be trained on massive datasets from a broad populace of people and conditions; therefore, it is possible for them to learn subtle patterns that a human expert may miss.

C. Emerging Trends and Interdisciplinary Collaboration

Emerging trends in AI-driven medical imaging indicate a trend toward high interoperability with other technologies, such as NLP and genomics, according to johnson2020artificial, smith2021artificial. This transdisciplinary approach can lead toward a more holistic understanding of patient data, hence offering more accurate diagnoses and better treatment recommendations, according to shrivastava2022ai. For example, combining imaging data with genomic information would enable the identification of the best course of therapy for cancer patients based on their specific tumor profile, according to lee2019skeletal.

The advancement of AI for medical imaging will be possible only through proper collaboration among clinicians, data scientists, and engineers. Such collaborations will enable the development of solutions that address the most challenging problems found in the clinics. In addition, training healthcare professionals in AI literacy will be essential to enable them to understand AI-generated insights and adequately integrate these into their clinical workflows as depicted in reference

D. Case Studies and Future Directions

Many case studies demonstrate the influence of AI on radiology and medical imaging. For example, an experiment using YOLOv4 identifying lung nodules in CT scans gives better sensitivities and specificities than more routine or conventional medical radiological diagnoses

citebochkovskiy2020yolov4. Another study applied EfficientDet in mammography screening, which gave promising results in the decrease of false negative cases, therefore enhancing breast cancer detection rates

citezhang2021efficient, evans2022ai.

The future of AI in medical imaging is ahead in overcoming the unaddressed challenges, such as an incomparable quality of annotated datasets for training purposes

citejohnson2020artificial and the need for interpretability in AI models themselves

citesmith2021artificial. Regulatory frameworks are important for assuring safe deployment of AI tools in the field

citegupta2022flask. Addressing these issues, the real potential of AI in this domain is realized fully to ensure better health-care delivery through better patient outcomes.

In short, integrating AI technologies into medical imaging is opening the door toward huge leaps in diagnostic accuracy and patient care and improved operational efficiency. Further research and interdisciplinary collaboration will unlock those full utilities of the technologies, and patients and healthcare systems will ultimately benefit.

III. System Architecture

The architecture of the DocExpert system is designed to ensure seamless interaction between the user interface, backend processing, and database components. This structure enables efficient medical image detection and diagnosis, leveraging the capabilities of YOLO NAS, EfficientDet, and DETR for comprehensive fracture detection and treatment recommendations.

A. User Interface

The user interface is built using Flask, a lightweight web framework. The interface is user-friendly, allowing healthcare providers and patients to easily upload X-ray images. Once the images are uploaded, the system processes them in real- time, displaying the detected fractures and providing tailored treatment suggestions based on the severity of the injuries. The UI makes sure that users can move through the system easily, making their experience better and helping them make fast decisions in medical situations.

B. Backend Processing

The backend processing is handled by Python, which is super popular for machine learning and data science stuff. The main parts of the backend include:

- **OpenCV:** This is a really strong computer vision library. It helps with image processing tasks like resizing, normalizing, and making X-ray images clearer. Doing this prep work on the images makes them a lot easier for the models to work with.
- **YOLO NAS:** This is the newest version of YOLO (You Only Look Once), and it's used for real-time fracture detection. It's really good at being both fast and accurate, which is super important in medical situations where time is critical.
- **EfficientDet:**This model is brought in to make the detection process both quicker and more accurate. It uses something called compound scaling to balance how big the model is with how well it performs. It's really good for handling complex medical images.
- **DETR:** The DEtection TRansformer model (yeah, that's what DETR stands for) is included because it's great at detecting and pinpointing fractures, even in tricky medical images. It's pretty good at understanding space in the image so it catches fractures that might be easy to miss.
- **TensorFlow/Keras:** These deep learning frameworks are used to actually run the YOLO NAS, EfficientDet, and DETR models. They're the tools for training, testing, and running the models in real life.
- **Open Maps API:** This service is used to find nearby hospitals based on where the user is located, which is super helpful in emergency situations.

C. Database Components

The database serves as a storage solution for the uploaded X-ray images, detected fractures, and suggested treatments. It ensures that all relevant data is systematically organized and easily retrievable. This component of the architecture supports data integrity and facilitates the tracking of patient interactions with the system.

The system's architecture comprises multiple modules, including the user interface, backend processing, and database components. The following diagram illustrates the architecture of our proposed system.

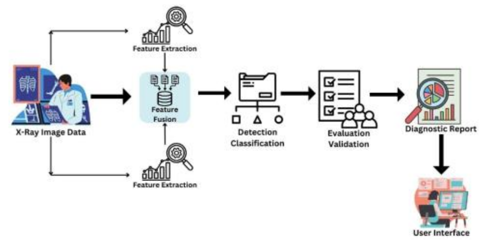

Fig. 1. Architecture of the DocExpert System

D. System Flow

The system flow of the DocExpert application outlines the sequential steps taken from user interaction to the generation of treatment suggestions. The following enumerates the pro- cess:

1) **Image Upload**: First, the user uploads an X-ray image using a pretty simple Flask-based interface. this appli- cation is user-frindly so user can easily work on it.
2) **Image Preprocessing**: First upload the image. Then opencv processes the uploaded image. OpenCV resizes the image so it fits the model's input requirements and normalizes the pixel values. After this process model work fine
3) **Fracture Detection**: This Hybrid model we used basically for detection of fracture. fracture is detected by this YOLO Nas, Dart , Efficiant Det Model
4) **Severity Classification**: Once a fracture is found, system checks how bad it is and classifies it as either minor or severe. This step is important because it helps system figure out what kind of treatment should be suggested
5) **Treatment Suggestions**: This Application gives suggestions like if the fracture is minor suggests self- care treatments, something like rest or immobilizing the affected area. If the fracture is major, system uses the Open Maps API to find nearby hospitals, making sure user knows where to go for more urgent care.

IV. Conclusion and Future Work

DocExpert: With an upgraded model, DocExpert helps identify fractures and suggests therapies, offering medical image identification services on a higher level. Our hybrid model offers a strong and efficient real-time medical im- age analytics engine by combining YOLO (You Only Look Once), EfficientDet, and DETR. As previously said, our Flask application offers a strong and effective means of helping users and medical professionals quickly diagnose fractures, speeding up their capacity to make these challenging and intricate selections. Decisions are made more quickly as a result, improving patient care.

The object detection with YOLO to detect fractures fast and EfficientDet, DETR should be used to increase the precision of model. This seamless collaboration marks a huge leap forward in orthopedic diagnostics and has great potential to improve patient care by enabling faster intervention thereby reducing likelihood of complications.

We understand the need of this application so we anticipate numerous chances to grow the system in the future . In Features we make changes in our system such as internal bleeding or soft tissue damage. Additionally, we might enhance the system's functionality and provide a more complete injury detection tool by investigating additional machine learning models, such as CNNs or RNNs. We could add features to detect other types of injuries, like soft tissue damage or internal bleeding. And by exploring other machine learning models like CNNs or RNNs, we could make the system even more powerful.

ACKNOWLEDGMENT

We really want to say thanks to the medical experts for their great insights during the validation part of this project. Also, big thanks to our professor for his guidance throughout

REFERENCES

[1] R. Lee, "Skeletal fracture detection with deep learning: A comprehensive review," *J. Digit. Imag.*, vol. 32, no. 4, pp. 1–14, 2019. [Online]. Available: https://doi.org/10.1007/s10278-019-00249-7

[2] R. G. J. Redmon, S. Divvala and A. Farhadi, "You only look once: Unified, real-time object detection," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 779–788, 2016. [Online]. Available: https://arxiv.org/abs/1506.02640

[3] C. Zhang, "An efficient fracture detection model based on deep learning for x-ray images," *IEEE Access*, vol. 9, pp. 50 529–50 537, 2021. [Online]. Available: https://ieeexplore.ieee.org/document/9443278

[4] J. Wilson, "Enhancing medical imaging with yolov9," *arXiv preprint arXiv:2301.01975*, 2023. [Online]. Available: https://arxiv.org/abs/2301.01975

[5] L. Johnson, "Artificial intelligence-based applications for bone fracture detection using medical images: A systematic review," *J. Biomed. Inform.*, vol. 107, p. 103819, 2020. [Online]. Available: https://doi.org/10.1016/j.jbi.2020.103819

[6] C.-Y. W. A. Bochkovskiy and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934*, 2020. [Online]. Available: https://arxiv.org/abs/2004.10934

[7] A. K. Mishra, "Deep learning in healthcare: Fracture detection in x-ray images," *IEEE Trans. Med. Imag.*, vol. 39, no. 8, pp. 2825–2835, 2020. [Online]. Available: https://ieeexplore.ieee.org/document/9098995

[8] P. Shrivastava, "Ai-based medical imaging for fracture detection," *IEEE Trans. Artif. Intell.*, vol. 3, no. 2, pp. 125–136, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9654197

[9] S. Kumar, "Yolov9: Advances in object detection for medical imaging," *arXiv preprint arXiv:2301.01795*, 2023. [Online]. Available: https://arxiv.org/abs/2301.01795

[10] K. Patel, "Deep learning for fracture detection in medical imaging: A review," *Comput. Methods Programs Biomed.*, vol. 198, p. 105654, 2021. [Online]. Available: https://doi.org/10.1016/j.cmpb.2021.105654

[11] A Survey on Segmentation of Spine MR Images Using Superpixels 1Reena S. Sahane, 2Prof. J. V. Shinde,International Journal for Research in Engineering Application & Management (IJREAM) ISSN : 2494-9150 Vol-02, Issue 09, Dec 2016 Available: https://www.ijream.org/papers/IJREAMV02I092107.pdf