RMIT UNIVERSITY

Macrohard

School of Science, Engineering & Technology (SSET)

EndoDetect: Lesion Detection in Endoscopic Images using Deep Learning

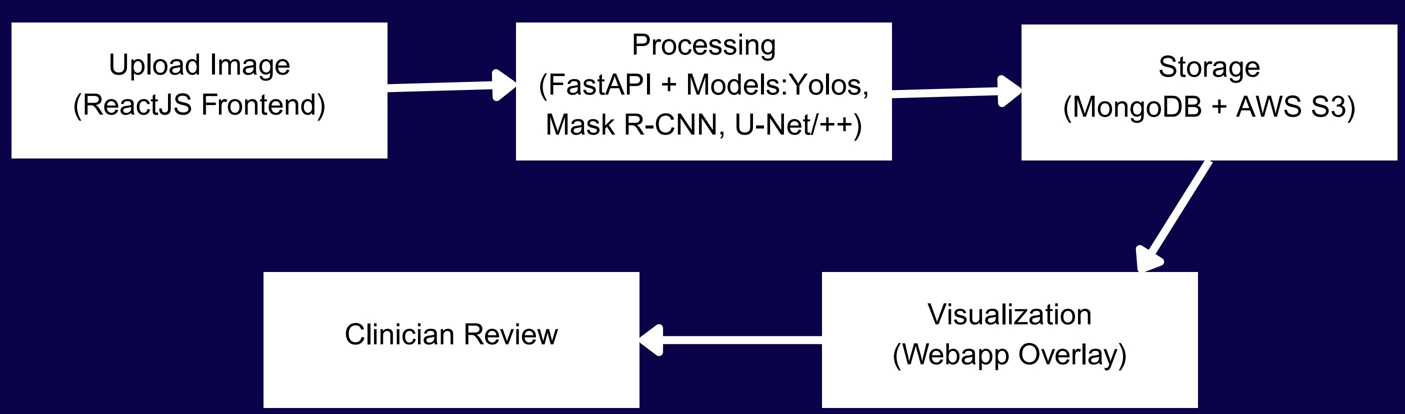
BACKGROUND AND MOTIVATION

- Colorectal cancer is the second leading cause of cancer deaths worldwide [1].
- Finding and removing polyps early during colonoscopy can greatly reduce cancer risk and deaths [2].
- Studies show that about 26% of adenomas are overlooked, and detection rates can vary widely between doctors [3].
- Motivation: Create EndoDetect, a simple web app that uses AI to highlight lesions in images, helping doctors see more clearly and reduce errors.

OBJECTIVES

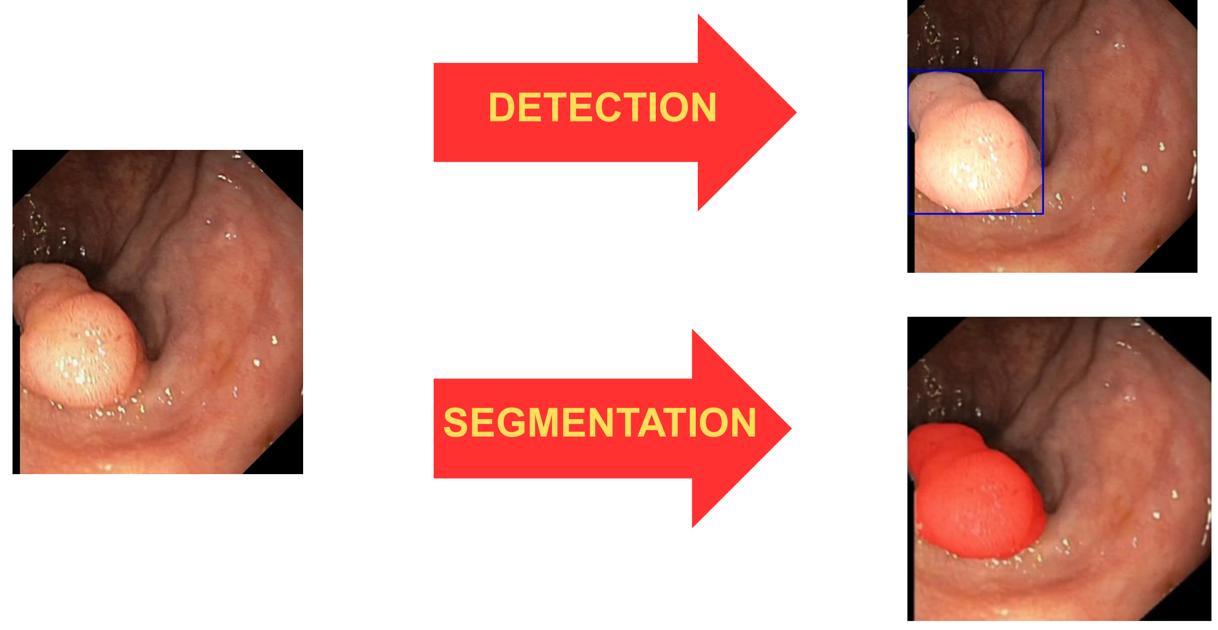
- Create a web-based platform where clinicians can upload endoscopy images and receive automated analysis.
- Evaluate detection and segmentation models to find the best trade-off between accuracy and speed.
- Provide an accessible tool for low to mid-resource hospitals without requiring high-end hardware.

METHODOLOGY



1. Image Upload (Frontend - ReactJS)
   - Clinicians interact via a web interface.
   - Supports JPEG, PNG, TIFF formats.
2. AI Model Processing (Backend - FastAPI)
   - Uploaded image sent to server for analysis.
   - Models applied:
     - YOLOv9 / YOLOv11: Detect lesion regions (bounding boxes).
     - Mask R-CNN (ResNet50-FPN): Detection + instance segmentation.
     - U-Net / U-Net++ (EfficientNet-B7): Semantic segmentation of fine boundaries.
   - Models trained/validated on Kvasir-SEG, CVC-ClinicDB, ETIS-Larib.
   - Metrics: Precision, Recall, Dice, F1, IoU, Specificity, mAP@50.
3. Data Storage & Management
   - AWS S3 stores uploaded endoscopy images
   - MongoDB stores doctor accounts (username, password), input and ouput information in JSON.
   - Ensures secure login and scalable storage of medical data and model outputs.
4. Visualization & Clinician Review
   - Processed images are displayed on the webapp.
   - Two types of overlays to help clinicians:
     - Detection: Bounding boxes around suspected lesions.
     - Segmentation: Colored masks outlining lesion boundaries.

EXAMPLE OUTPUT



EXPERIMENTS AND RESULTS

Metric	Precision	Recall	F1	mAP@0.5	mAP@0.5:0.95	Inference Speed	Complexity
YOLOv9t	0.9285	0.9209	0.9247	0.9438	0.7860	20.44 ms/img	2.13M params / 0.68G MACs
YOLO11n	0.9125	0.9033	0.9079	0.9387	0.7867	224.12 ms/img	2.62M params / 0.53G MACs

Metric	Precision	Recall	F1	IoU	Dice	Specificity	Inference Speed	Complexity
Mask R-CNN	0.8997	0.8456	0.8517	0.7864	0.8517	0.9857	38.00 ms/img (26.31 img/s)	43.92M params / 134.59G MACs
U-Net	0.9174	0.8468	0.8618	0.7944	0.8618	0.9895	46.29 ms/img (21.60 img/s)	67.10M params / 3.22G MACs
U-Net++	0.9192	0.8867	0.8938	0.8303	0.8938	0.9887	52.11 ms/img (19.19 img/s)	68.16M params / 12.26G MACs

The experiments tested different detection and segmentation models on the Kvasir-SEG and ETIS-Larib datasets, evaluating accuracy, inference speed, and computational complexity.

Detection Models: YOLOv9t vs. YOLO11n

- YOLOv9t achieved higher precision (0.9285), F1 score (0.9247), and strong mAP@0.5 (0.9438), providing more balanced detection with fewer false positives.
- YOLO11n achieved better recall (0.9033) and slightly higher mAP@0.5:0.95 (0.7867), ensuring fewer missed detections. It is also very lightweight (2.62M params / 0.53G MACs), though inference was significantly slower (224.12 ms/img) compared to YOLOv9t (20.44 ms/img).
- This makes YOLOv9t more suited for real-time detection, while YOLO11n is favorable for high-recall applications where missing objects is more costly.

Segmentation Models: Mask R-CNN, U-Net, U-Net++

- U-Net++ achieved the highest accuracy, with superior Recall (0.8867), F1 (0.8938), IoU (0.8303), and Dice (0.8938).
- Mask R-CNN provided the fastest inference (38 ms/img) but required very high computational cost (134.59G MACs).
- U-Net offered the best efficiency, combining solid accuracy with the lowest compute requirement (3.22G MACs), making it lightweight to deploy.

CONCLUSION

- EndoDetect is a multi-user web application that integrates state-of-the-art deep learning models for lesion detection and segmentation.
- It delivers accurate results while being practical for hospitals with limited computing resources.
- The system shows strong potential to support endoscopists, improve early detection, and enhance patient outcomes.

[1] World Health Organization. “Colorectal Cancer.” World Health Organization, 11 July 2023, www.who.int/news-room/fact-sheets/detail/colorectal-cancer.
[2] Zauber, Ann G., et al. “Colonoscopic Polypectomy and Long-Term Prevention of Colorectal-Cancer Deaths.” New England Journal of Medicine, vol. 366, no. 8, 23 Feb. 2012, pp. 687–696, www.ncbi.nlm.nih.gov/pmc/articles/PMC3322371/, https://doi.org/10.1056/nejmoa1100370.
[3] Zhao, Shengbing, et al. “Magnitude, Risk Factors, and Factors Associated with Adenoma Miss Rate of Tandem Colonoscopy: A Systematic Review and Meta-Analysis.” Gastroenterology, vol. 156, no. 6, May 2019, pp. 1661-1674.e11, https://doi.org/10.1053/j.gastro.2019.01.260.