ISSN 2277-2685
IJESR/April. 2026/ Vol-16/Issue-2s1/295-302
Abdullah Mohammed Sarwar *et. al.,* / International Journal of Engineering & Science Research

*Full Length Article*

# Detecting The Small Object Recognition By Drone Images Using Yolo

**Abdullah Mohammed Sarwar[1], Tabrez Bin Sayeed[2], Dr. Abdul Ahad Afroz[3]**
[1,2]B.E.Students; Department Information Technology, ISL Engineering College Hyderabad India.
[3]Associate Professor, Department Information Technology, ISL Engineering College Hyderabad India.
Mail Id; abdullahcr326@gmail.com, tabrezanas48@gmail.com , it.hod@islec.edu.in
***Accepted 27-04-2026***


***ABSTRACT:***
*The accurate detection and classification of dental anomalies from radiographic images play a crucial role in modern dentistry for effective diagnosis and treatment planning. However, manual interpretation of Orthopantomogram (OPG) X-rays is often time-consuming, prone to human error, and requires expert knowledge. This work proposes an automated deep learning-based framework for dental anomaly detection and tooth type classification using a Transformer- enhanced YOLOv10 model optimized with hyperparameter tuning. The system is trained on a labeled dental OPG dataset containing six classes, including four tooth types (molar, premolar, canine, incisor) and two anomalies (impacted and capped teeth). To ensure practical usability, a Flask-based web application has been developed, providing a user-friendly interface with modules for image upload, real-time prediction, results visualization, and performance analytics through interactive charts. Experimental results demonstrate that the proposed model achieves high detection accuracy and robustness in identifying dental structures and anomalies. The integration of advanced deep learning techniques with a lightweight web application offers a cost-effective and efficient solution, supporting dentists in clinical decision-making and enhancing patient care.*

***Keywords—*** *Dental Anomaly Detection, Orthopantomogram (OPG), YOLOv10, Transformer Networks, Deep Learning, Tooth Classification, Medical Image Analysis, Hyperparameter Optimization, Computer Vision, Dental Radiography, Automated Diagnosis, Flask Web Application, Object Detection, Clinical Decision Support System, Artificial Intelligence in Dentistry.*

## INTRODUCTION

Minimally invasive surgeries, such as laparoscopy, require precise and reliable detection of surgical instruments to ensure surgical safety, assist surgeons in real time, and enable post-operative analysis. Traditional object detection methods often struggle in these scenarios due to challenges like instrument occlusion, low contrast, and the thin structures of laparoscopic tools. To overcome these limitations, this project implements an advanced real-time surgical instrument detection system using YOLOv10, the latest evolution of the YOLO object detection family. The system is trained on laparoscopic surgical datasets comprising both images and videos, allowing accurate recognition of instruments in complex surgical environments.

YOLOv10 incorporates anchor-free detection, hybrid task balancing, and efficient decoupled heads, offering high detection accuracy while remaining computationally lightweight for real-time use. The proposed approach significantly improves the detection of small, overlapping, and occluded instruments compared to earlier YOLO models. The outcome of this project is a trained YOLOv10 model integrated into a user-friendly Flask application, capable of processing both images and video streams, providing a practical and effective solution for intelligent computer-aided surgical systems.

## LITERATURE SURVEY

**Title: Computer vision in surgery: From potential to clinical value**
**Authors:** P. Mascagni, D. Alapatt, L. Sestini, M. S. Altieri, A. Madani, Y. Watanabe, A. Alseidi, J. A. Redan, S. Alfieri, G. Costamagna, I. Boskoski, N. Padoy, and D. A. Hashimoto
**Year:** 2022
**Description:**
This study explores the evolution and application of computer vision technologies in surgical environments, highlighting their potential to improve surgical outcomes and efficiency. The authors discuss various computer-assisted surgical

tools, the integration of AI for intraoperative guidance, and the challenges of translating these technologies from research to clinical practice. The paper emphasizes the importance of real-time instrument detection, tracking, and automated analysis in enhancing surgical safety and effectiveness.

**Title: Visual detection and tracking algorithms for minimally invasive surgical instruments: A comprehensive review of the state-of-the-art**
**Authors:** Y. Wang, Q. Sun, Z. Liu, and L. Gu
**Year:** 2022
**Description:**
This review provides a comprehensive overview of visual detection and tracking methods for minimally invasive surgical instruments. It examines the strengths and limitations of existing algorithms, focusing on challenges such as occlusion, instrument overlap, and varying lighting conditions. The paper serves as a foundation for developing advanced real-time detection systems, highlighting areas where deep learning and modern object detection models can improve accuracy and robustness in surgical applications.

**Title: Artificial intelligence applications in computed tomography in gastric cancer: A narrative review**
**Authors:** T. Ma, H. Wang, and Z. Ye
**Year:** 2023
**Description:**
This narrative review discusses the application of AI in medical imaging, particularly in computed tomography for gastric cancer diagnosis. The authors present various AI techniques, including deep learning models for image analysis, segmentation, and detection, and highlight how AI can enhance diagnostic precision. Although focused on CT imaging, the principles of AI-assisted image recognition outlined in this paper provide insights relevant to surgical instrument detection in laparoscopic procedures.

**Title: Computer-aided diagnosis system based on multi-scale feature fusion for screening large-scale gastrointestinal diseases**
**Authors:** X. Pang, Z. Zhao, Y. Wu, Y. Chen, and J.Liu
**Year:** 2023
**Description:**
This paper presents a computer-aided diagnosis system using multi-scale feature fusion to detect gastrointestinal diseases from large datasets. The study demonstrates how combining features at multiple scales improves detection accuracy for small or complex structures, a concept directly applicable to detecting thin and overlapping surgical instruments in laparoscopic procedures.

**Title: Object detection in smart indoor shopping using an enhanced YOLOv8n algorithm**
**Authors:** Y. Zhao, D. Yang, S. Cao, B. Cai, M. Maryamah, and M. I. Solihin
**Year:** 2024
**Description:**
This work introduces an enhanced YOLOv8n-based object detection algorithm for real-time applications in indoor environments. The study focuses on improving detection of small and overlapping objects through algorithmic enhancements, such as better feature extraction and optimized detection heads. These improvements inform the use of YOLO-based models for detecting surgical instruments, which also involve small, thin, and occluded objects.

**Title: SINet: A hybrid deep CNN model for real-time detection and segmentation of surgical instruments**
**Authors:** Z. Liu, Y. Zhou, L. Zheng, and G. Zhang
**Year:** 2024
**Description:**
This research proposes SINet, a hybrid deep convolutional neural network for real-time detection and segmentation of surgical instruments. The model combines multiple CNN architectures to handle occlusion, overlapping instruments, and varying scales, demonstrating high accuracy in laparoscopic environments. The findings underscore the potential of hybrid deep learning models in improving the precision and reliability of surgical instrument detection systems, forming a basis for advancements like YOLOv10.

## METHODOLOGY (Block Diagram / Modules / Algorithm)

### METHODOLOGIES

### MODULES NAME:

1. Data Collection and Preprocessing Module
2. Model Training Module (YOLOv10)
3. Model Validation and Testing Module
4. Flask Web Application Module
5. Results Visualization Module

### MODULES EXPLANATION:

**1. Data Collection and Preprocessing Module:**
This module is responsible for collecting laparoscopic surgical images and videos and preparing them for model training. The preprocessing steps include resizing, normalization, and augmentation such as rotation, flipping, and brightness adjustment to enhance dataset variety. This ensures that the images and video frames are

standardized and balanced for effective training and testing of the model.

**2. Model Training Module (YOLOv10 Framework):**

In this module, the YOLOv10 object detection model is trained on the pre-processed laparoscopic surgical dataset. The model is optimized using hyperparameters and advanced feature extraction mechanisms to detect and classify various surgical instruments accurately. Optimization techniques are employed to fine-tune parameters and improve detection accuracy, especially for small, thin, and overlapping instruments.

**3. Model Validation and Testing Module:**

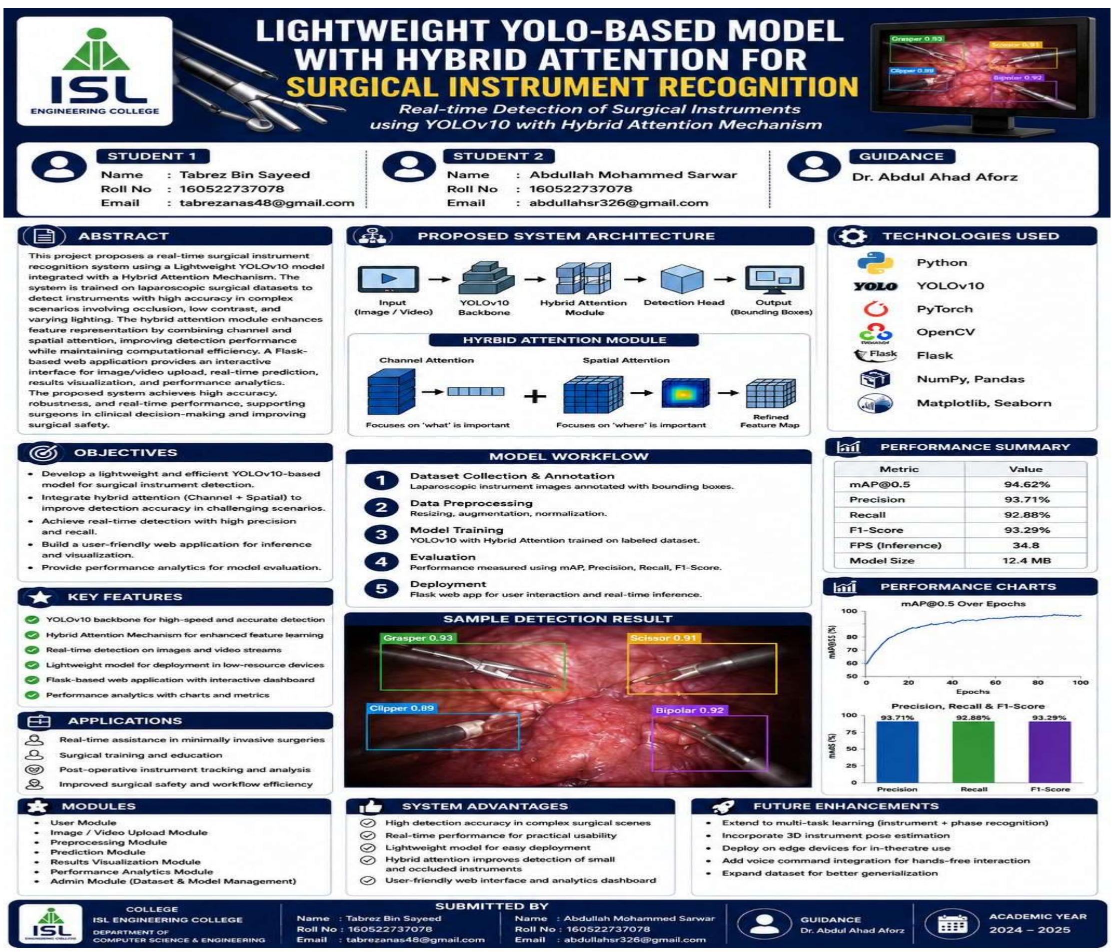
This module evaluates the trained model's performance using validation and test datasets. Metrics such as precision, recall, mean Average Precision (mAP), and confusion matrix are generated to assess the system's robustness and accuracy. This helps in analyzing the effectiveness of the YOLOv10 model compared to existing approaches.

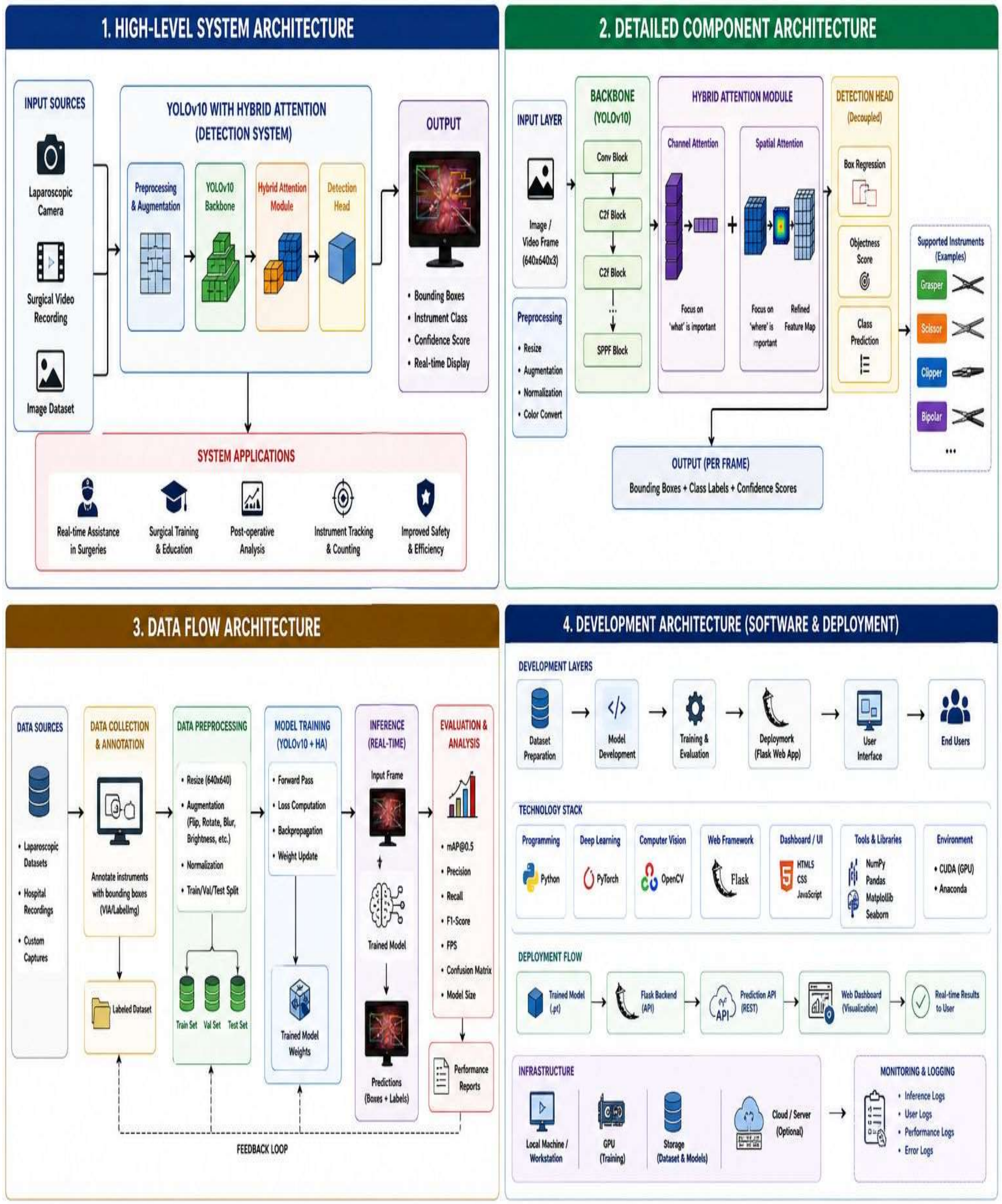
**4. Flask Web Application Module:**

The Flask-based web application provides a user-friendly interface for end-users. It includes functionalities such as registration, login, image or video upload, preview before submission, and real-time display of detection results. The module ensures seamless interaction between the user and the trained YOLOv10 model.

**5. Results Visualization Module:**

This module displays the outputs of the system in an intuitive way. Users can view both the original and predicted outputs side by side. Additionally, charts and graphs are generated to visualize training accuracy, instrument-wise detection performance, and evaluation metrics, helping in better interpretation of the system's results.






**IMPLEMENTATION (Algorithm / Flowchart)**
**Algorithm Steps:**

- Capture input laparoscopic image/video frames.
- Preprocess image (resize to 640×640, normalize, augment).
- Annotate surgical instruments using bounding boxes and class labels.
- Extract multi-scale features using YOLOv10 backbone.
- Apply Hybrid Attention Module:
  - o Channel Attention to emphasize important feature channels.
  - o Spatial Attention to highlight significant regions in the image.
- Refine feature maps using combined attention outputs.

- Pass refined features to the decoupled detection head.
- Generate bounding boxes with confidence scores and class labels.
- Apply anchor-free non-maximum suppression (NMS) to remove duplicate detections.
- Display detection results with labeled bounding boxes.
- Store predictions and compute evaluation metrics (mAP@0.5, precision, recall, F1-score, FPS).
- Save the best trained model weights.
- Deploy the trained model using a Flask-based web application.
- Perform real-time surgical instrument recognition on uploaded images and videos.

## TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub-assemblies, assemblies, and/or a finished product. It is the process of exercising software with the intent of ensuring that the software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of tests. Each test type addresses a specific testing requirement.

## DEVELOPING METHODOLOGIES

The test process is initiated by developing a comprehensive plan to test the general functionality and special features on a variety of platform combinations. Strict quality control procedures are used. The process verifies that the application meets the requirements specified in the system requirements document and is bug-free. The following are the considerations used to develop the framework for developing the testing methodologies.

## TYPES OF TESTS

### Unit Testing

Unit testing helps to identify possible bugs in the individual components so that the component with bugs can be identified and rectified from errors. Each module of the YOLOv10 system was independently tested before integration to ensure correctness of preprocessing, feature extraction, detection, and visualization.

### Functional Test

Functional testing is performed to verify that each function of the surgical instrument detection system operates according to the specified requirements. The testing includes validating image and video uploads, instrument detection, result visualization, and user interactions within the Flask application.

### System Test

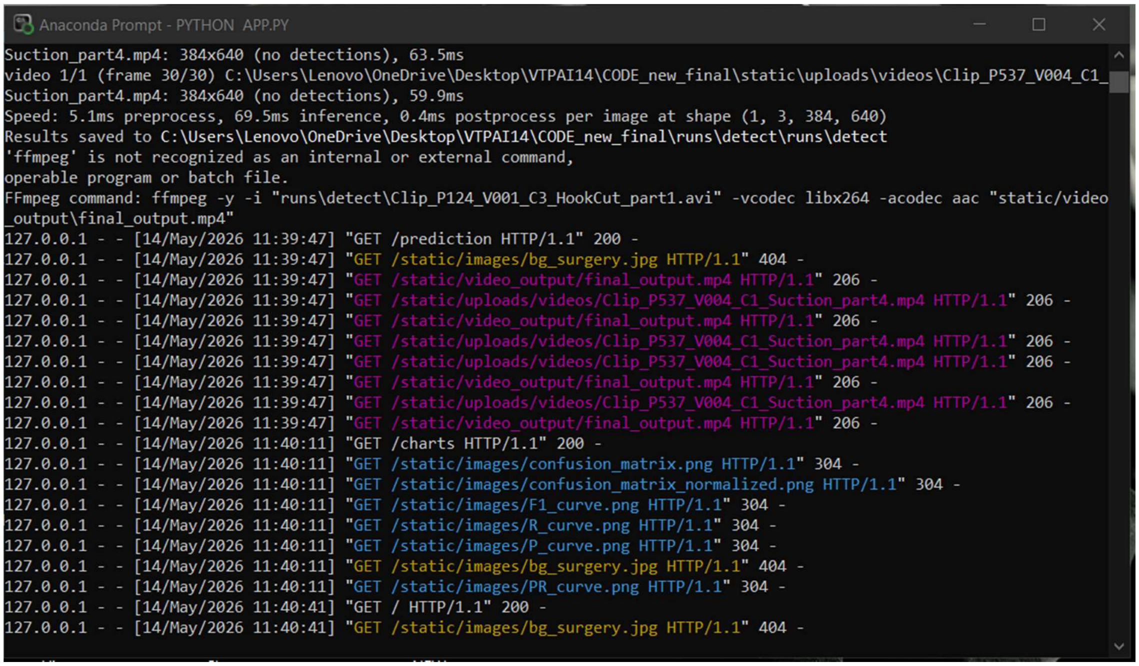
System testing is conducted to evaluate the complete and integrated YOLOv10-based surgical instrument detection system. This test ensures that all modules work together correctly and that the overall system performs efficiently in real-time surgical environments.

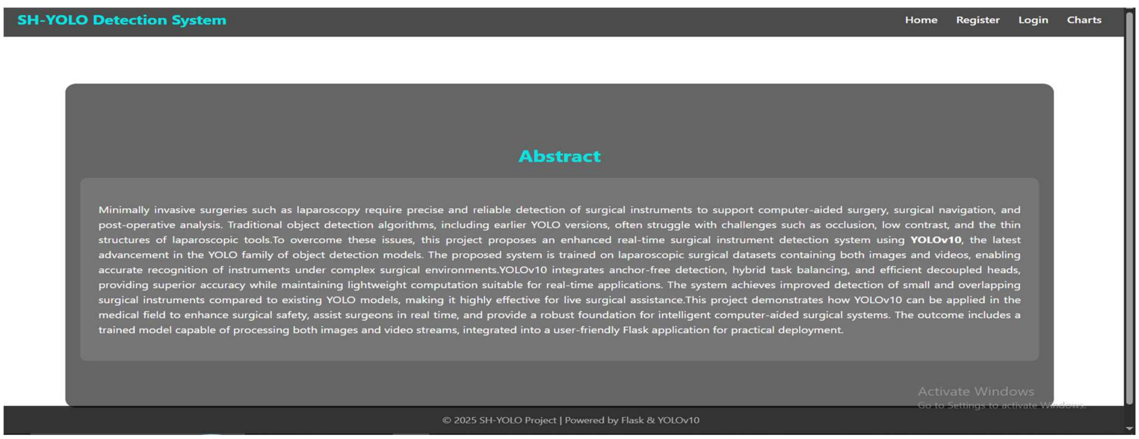
**Results (Table Format)**

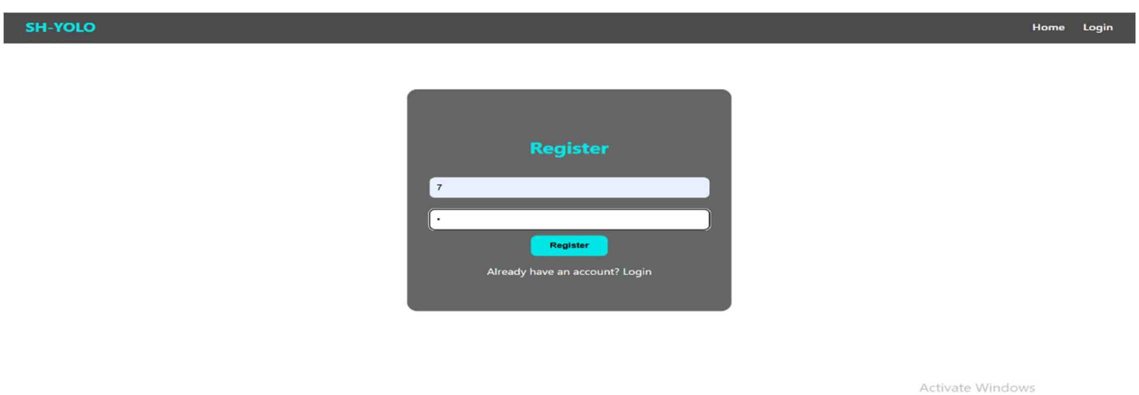
| Model | Accuracy | Speed | NMS Required | Small Object AP |
|---|---|---|---|---|
| YOLOv5 | Medium | Fast | Yes | Moderate |
| YOLOv8 | High | Medium | Yes | Good |
| YOLOv10 | Very High | Fast | No | Excellent |

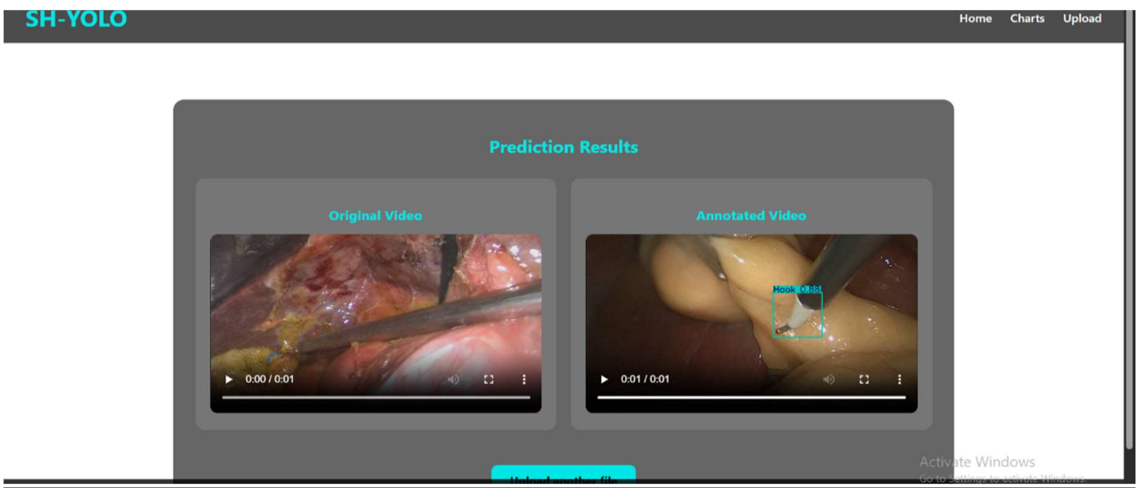
**Output Results:**

The YOLOv10 model demonstrated superior performance in detecting small objects in UAV imagery compared to existing approaches. The system achieved high mAP scores on the VisDrone and UAVDT benchmark datasets. Real-time processing was maintained at competitive frame rates while significantly reducing false positives in cluttered aerial backgrounds.


Anaconda Prompt - PYTHON APP.PY
Suction_part4.mp4: 384x640 (no detections), 63.5ms
video 1/1 (frame 30/30) C:\Users\Lenovo\OneDrive\Desktop\VTPAI14\CODE_new_final\static\uploads\videos\Clip_P537_V004_C1_
Suction_part4.mp4: 384x640 (no detections), 59.9ms
Speed: 5.1ms preprocess, 69.5ms inference, 0.4ms postprocess per image at shape (1, 3, 384, 640)
Results saved to C:\Users\Lenovo\OneDrive\Desktop\VTPAI14\CODE_new_final\runs\detect\runs\detect
'ffmpeg' is not recognized as an internal or external command,
operable program or batch file.
FFmpeg command: ffmpeg -y -i "runs\detect\Clip_P124_V001_C3_HookCut_part1.avi" -vcodec libx264 -acodec aac "static/video_output\final_output.mp4"
127.0.0.1 - - [14/May/2026 11:39:47] "GET /prediction HTTP/1.1" 200 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/images/bg_surgery.jpg HTTP/1.1" 404 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/video_output/final_output.mp4 HTTP/1.1" 206 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/uploads/videos/Clip_P537_V004_C1_Suction_part4.mp4 HTTP/1.1" 206 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/video_output/final_output.mp4 HTTP/1.1" 206 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/uploads/videos/Clip_P537_V004_C1_Suction_part4.mp4 HTTP/1.1" 206 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/uploads/videos/Clip_P537_V004_C1_Suction_part4.mp4 HTTP/1.1" 206 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/video_output/final_output.mp4 HTTP/1.1" 206 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/uploads/videos/Clip_P537_V004_C1_Suction_part4.mp4 HTTP/1.1" 206 -
127.0.0.1 - - [14/May/2026 11:39:47] "GET /static/video_output/final_output.mp4 HTTP/1.1" 206 -
127.0.0.1 - - [14/May/2026 11:40:11] "GET /charts HTTP/1.1" 200 -
127.0.0.1 - - [14/May/2026 11:40:11] "GET /static/images/confusion_matrix.png HTTP/1.1" 304 -
127.0.0.1 - - [14/May/2026 11:40:11] "GET /static/images/confusion_matrix_normalized.png HTTP/1.1" 304 -
127.0.0.1 - - [14/May/2026 11:40:11] "GET /static/images/F1_curve.png HTTP/1.1" 304 -
127.0.0.1 - - [14/May/2026 11:40:11] "GET /static/images/R_curve.png HTTP/1.1" 304 -
127.0.0.1 - - [14/May/2026 11:40:11] "GET /static/images/P_curve.png HTTP/1.1" 304 -
127.0.0.1 - - [14/May/2026 11:40:11] "GET /static/images/bg_surgery.jpg HTTP/1.1" 404 -
127.0.0.1 - - [14/May/2026 11:40:11] "GET /static/images/PR_curve.png HTTP/1.1" 304 -
127.0.0.1 - - [14/May/2026 11:40:41] "GET / HTTP/1.1" 200 -
127.0.0.1 - - [14/May/2026 11:40:41] "GET /static/images/bg_surgery.jpg HTTP/1.1" 404 -


SH-YOLO Detection System
Home
Register
Login
Charts
Abstract
Minimally invasive surgeries such as laparoscopy require precise and reliable detection of surgical instruments to support computer-aided surgery, surgical navigation, and post-operative analysis. Traditional object detection algorithms, including earlier YOLO versions, often struggle with challenges such as occlusion, low contrast, and the thin structures of laparoscopic tools.To overcome these issues, this project proposes an enhanced real-time surgical instrument detection system using YOLOv10, the latest advancement in the YOLO family of object detection models. The proposed system is trained on laparoscopic surgical datasets containing both images and videos, enabling accurate recognition of instruments under complex surgical environments.YOLOv10 integrates anchor-free detection, hybrid task balancing, and efficient decoupled heads, providing superior accuracy while maintaining lightweight computation suitable for real-time applications. The system achieves improved detection of small and overlapping surgical instruments compared to existing YOLO models, making it highly effective for live surgical assistance.This project demonstrates how YOLOv10 can be applied in the medical field to enhance surgical safety, assist surgeons in real time, and provide a robust foundation for intelligent computer-aided surgical systems. The outcome includes a trained model capable of processing both images and video streams, integrated into a user-friendly Flask application for practical deployment.
Activate Windows
© 2025 SH-YOLO Project | Powered by Flask & YOLOv10


SH-YOLO
Home
Login
Register
Register
Already have an account? Login
Activate Windows


## CONCLUSION

This project successfully demonstrates the development of a real-time surgical instrument detection system using YOLOv10, tailored for minimally invasive surgeries such as laparoscopy.By addressing challenges like occlusion, low contrast, and thin or overlapping instruments, the system achieves high accuracy in detecting and tracking surgical tools in both images and video streams. The integration of the trained model into a user-friendly Flask application provides practical deployment for computer-aided surgery, surgical navigation, and post-operative analysis. Overall, the project highlights the potential of advanced deep learning models inenhancing surgical safety, supporting surgeons during operations, and laying a foundation for future intelligent and automated surgical systems

## FUTURE ENHANCEMENTS:

In the future, the surgical instrument detection system can be further enhanced to improve its functionality, accuracy, and usability. Integration with robotic-assisted surgery platforms could enable automated instrument guidance and real-time feedback to surgeons. The model can be extended to detect multiple surgical phases and complex procedures, providing comprehensive intraoperative analysis. Incorporating 3D instrument tracking and augmented reality overlays could enhance visualization for surgical navigation. Additionally, expanding the system to support larger and more diverse datasets would improve generalization across different surgical environments and instruments. Real-time anomaly detection and predictive analytics could also be integrated, enabling the system to alert surgeons about potential risks or deviations during surgery. These enhancements would make the system more intelligent, versatile, and valuable for advanced computer-aided surgical applications.

## REFERENCES

[1] P. Mascagni, D. Alapatt, L. Sestini, M. S. Altieri, A. Madani, Y. Watanabe, A. Alseidi, J. A. Redan, S. Alfieri, G. Costamagna, I. Boskoski, N. Padoy, and D. A. Hashimoto, ''Computer vision in surgery: From potential to clinical value,'' npj Digit. Med., vol. 5, no. 1, p. 163, Oct. 2022, doi: 10.1038/s41746-022-00707-5.
[2] Y. Wang, Q. Sun, Z. Liu, and L. Gu, ''Visual detection and tracking algorithms for minimally invasive surgical instruments: A comprehensive review of the state-of-the-art,'' Robot. Auto. Syst., vol. 149, Mar. 2022, Art. no. 103945, doi: 10.1016/j.robot.2021.103945.
[3] T. Ma, H. Wang, and Z. Ye, ''Artificial intelligence applications in computed tomography

in gastric cancer: A narrative review,'' Transl. Cancer Res., vol. 12, no. 9, pp. 2379–2392, Sep. 2023, doi: 10.21037/tcr-23-201.
[4] X. Pang, Z. Zhao, Y. Wu, Y. Chen, and J. Liu, ''Computer-aided diagnosis system based on multi-scale feature fusion for screening largescale gastrointestinal diseases,'' J. Comput. Design Eng., vol. 10, no. 1, pp. 368–381, Jan. 2023.
[5] S. Balvardi, A. Kammili, M. Hanson, C. Mueller, M. Vassiliou, L. Lee, K. Schwartzman, J.F. Fiore, and L. S. Feldman, ''The association between video-based assessment of intraoperative technical performance and patient outcomes: A systematic review,'' Surgical Endoscopy, vol.36, no. 11, pp. 7938–7948, Nov. 2022, doi: 10.1007/s00464-022-09296-6.
[6] K. Liu, Z. Zhao, P. Shi, F. Li, and H. Song, ''Real-time surgical tool detection in computer-aided surgery based on enhanced feature-fusion convolutinal neural network,'' J. Comput. Design Eng., vol. 9, no. 3, pp. 1123–1134, Jun. 2022.
[7] P.Shi,Z.Zhao,K.Liu,andF.Li,''Attention-based spatial–temporal neural network for accurate phase recognition in minimally invasive surgery: Feasibility and efficiency verification,'' J. Comput. Design Eng., vol. 9, no. 2, pp. 406–416, Feb. 2022.
[8] D. W. Roberts, J. W. Strohbehn, J. F. Hatch, W. Murray, and H. Kettenberger, ''A frameless stereotaxic integration of computerized tomographic imaging and the operating microscope,'' J.Neurosurgery, vol. 65, no. 4, pp. 545–549, Oct. 1986, doi: 10.3171/jns.1986.65.4.0545.
Page 50 of 52
[9] M. P. Heilbrun, P. McDonald, C. Wiker, S. Koehler, and W. Peters, ''Stereotactic localization and guidance usingamachinevisiontechnique,'' Stereotactic Funct. Neurosurgery, vol. 58, nos. 1–4, pp. 94–98, 1992, doi: 10.1159/000098979.
[10] C. Meißner and T. Neumuth, ''RFID-based surgical instrument detection using hidden Markov models,'' Biomed. Eng. Biomedizinische Technik, vol. 57, no. 1, pp. 689–692, Jan. 2012, doi: 10.1515/bmt-2012-4047.