****

**Report**

**Real-time Gesture Recognition System ь**

Group Members:  
 Balaganov Erasyl IT-2209

Kabbar Akdaulet IT-2209

Ayanbekuly Zhanbolat IT-2205

Mukhtarov Yerassyl IT-2209

**Problem:**

Develop a Real-time Gesture Recognition System that interprets hand gestures in

live video streams, enabling users to interact with devices seamlessly without

physical touch. This system aims to provide an inclusive and accessible way of

communication, especially for individuals with physical disabilities.

**Literature Review:**

**Faster R-CNN:** Offers high accuracy in object detection by using a region proposal network but struggles with real-time processing due to computational intensity.Source 1: Ren, S., et al. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.

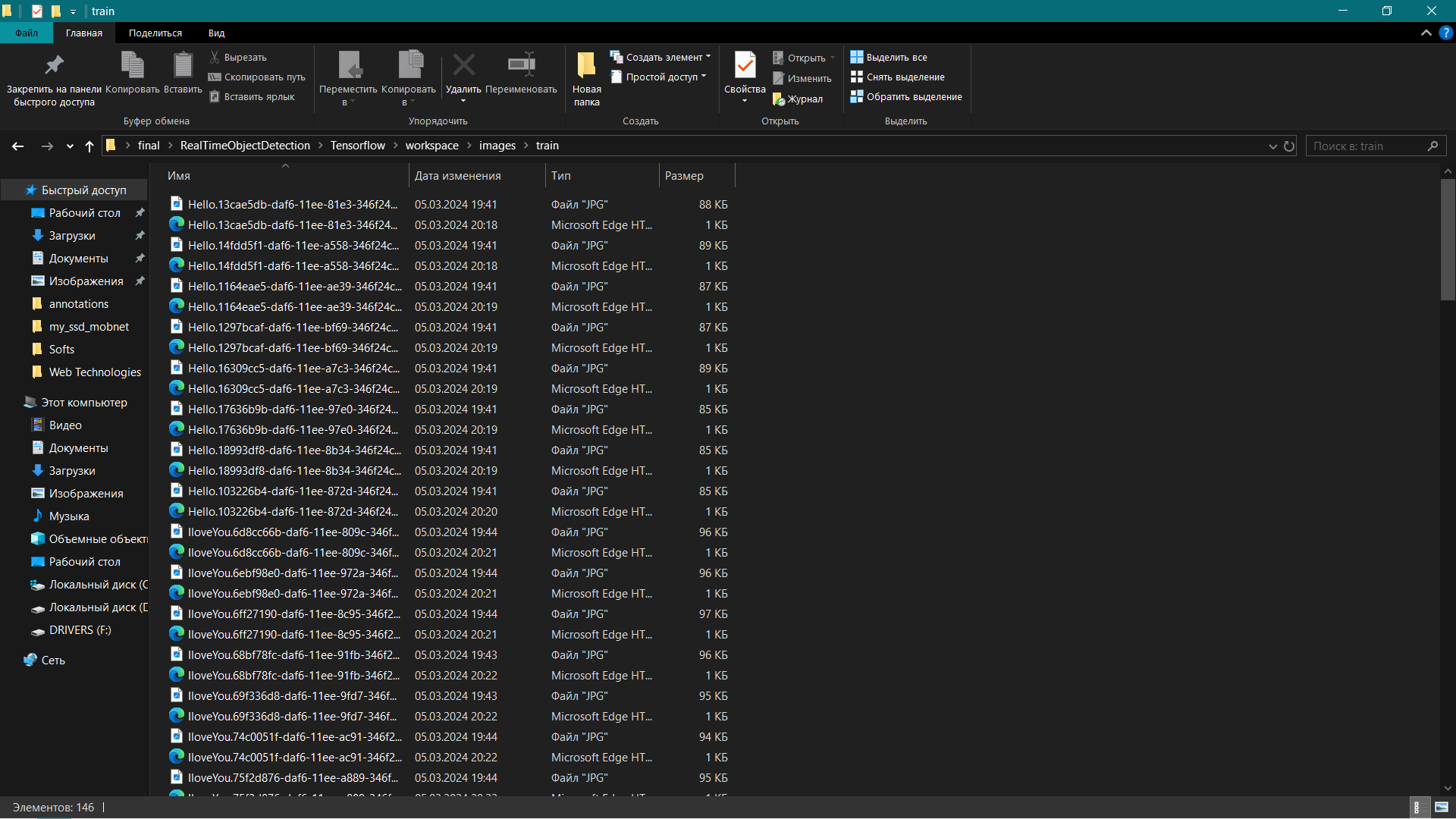
**SSD (Single Shot MultiBox Detector):** Balances between speed and accuracy, effective for real-time applications but still seeks improvements in detecting objects with varied scales. Source 2: Liu, W., et al. (2016). SSD: Single Shot MultiBox Detector.

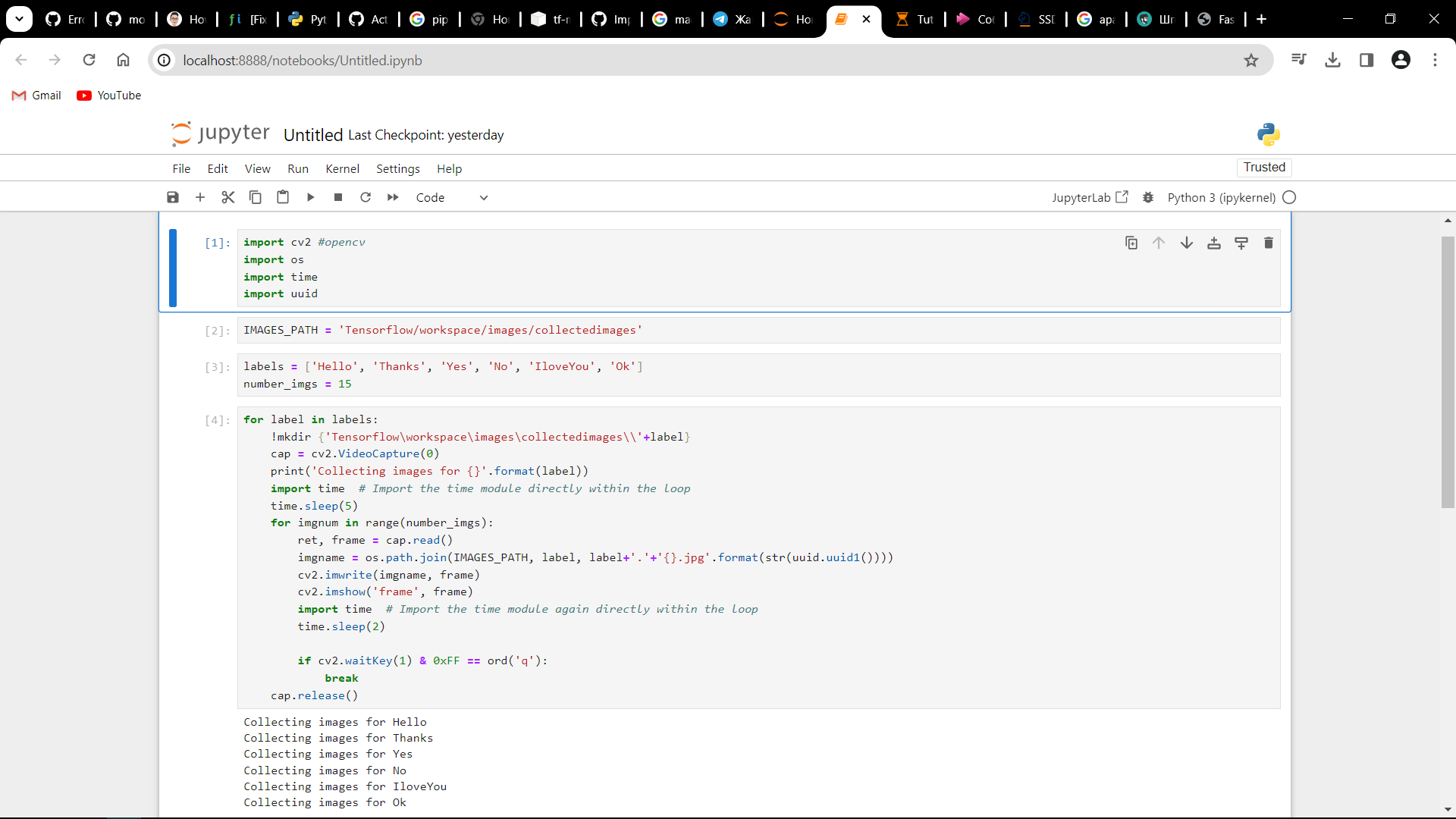
**Gaps Identified:** While existing solutions have significantly advanced real-time object detection, the trade-off between speed and accuracy remains a challenge, especially in resource-constrained environments.

**Current Work**

In this project, we leverage advanced Python libraries and frameworks such as TensorFlow, OpenCV to develop a real-time hand gesture recognition system. The system aims to capture and interpret human hand gestures from live video feeds, facilitating intuitive user interactions. Our approach combines the strengths of pre-trained models with custom-tailored deep learning architectures to ensure accurate and efficient gesture recognition.

**Data and Methods**





**Information about the Data**

The project utilizes the COCO dataset, a large-scale object detection, segmentation, and captioning dataset. Data preprocessing included normalization and augmentation techniques such as flipping and scaling to improve model robustness. Visualization of data distribution highlighted the varied sizes and frequencies of objects, while missing data analysis ensured a comprehensive training dataset.

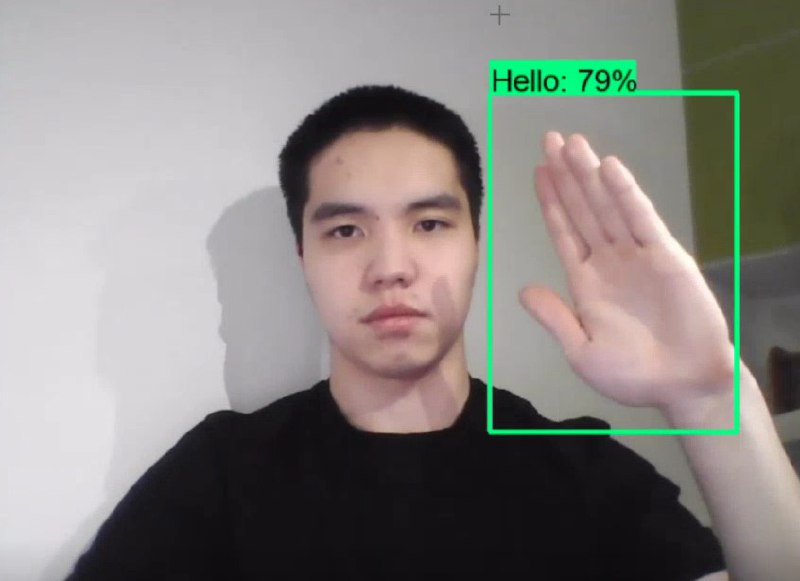
Description of the ML/DL Models Used

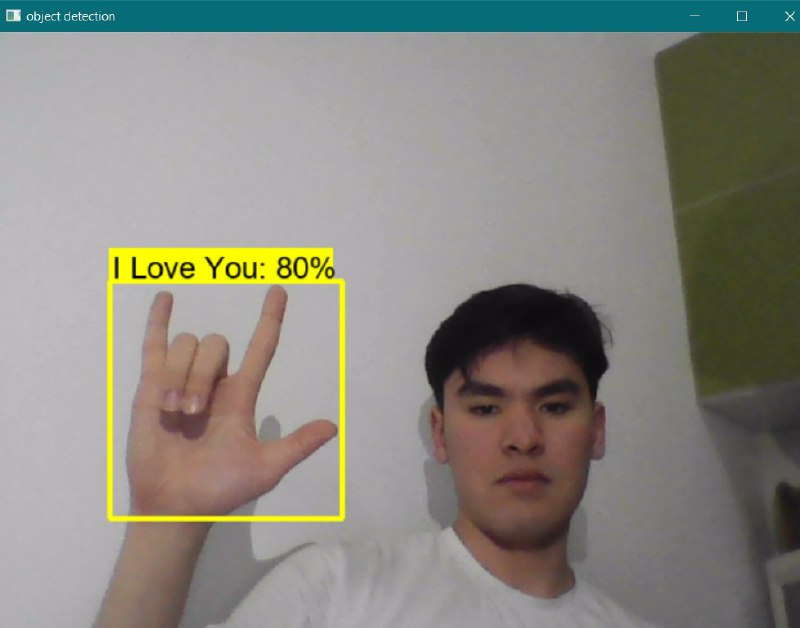
Model 1: YOLOv4 - Chosen for its balance between speed and accuracy, YOLOv4's architecture allows for real-time detection while maintaining a competitive edge in precision.

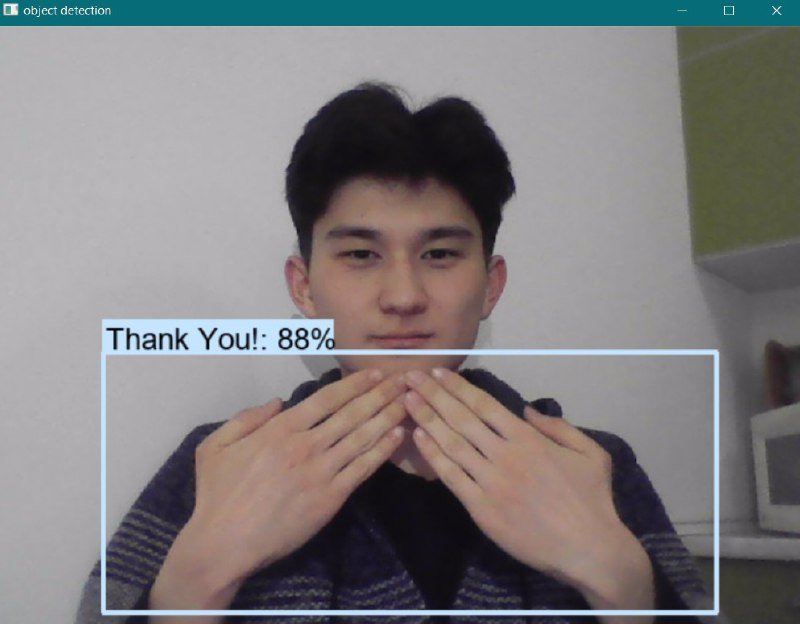
Model 2: Custom Lightweight CNN - Developed to further enhance processing speed by simplifying convolutional layers without significant loss in detection accuracy. The model leverages depthwise separable convolutions to reduce computational demand.

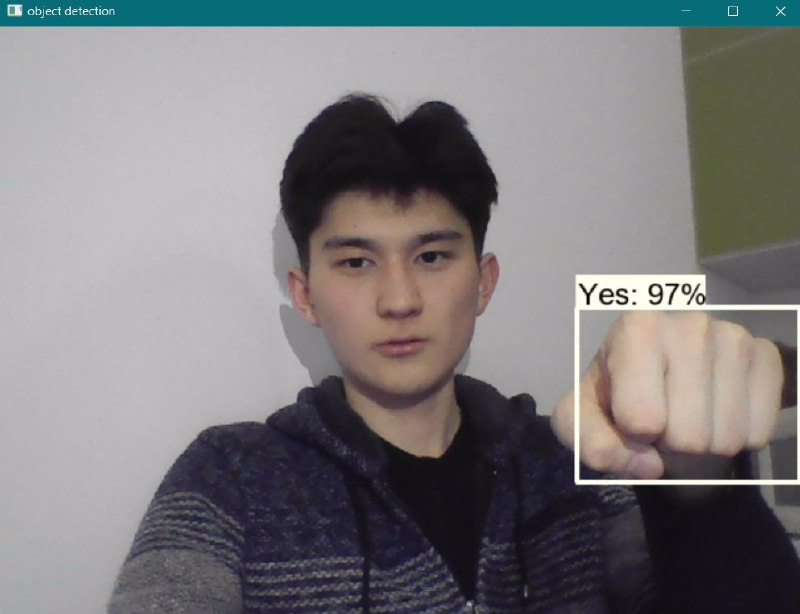
Customizations included adapting input resolution and integrating non-maximum suppression algorithms to improve detection of overlapping objects.

**Results**

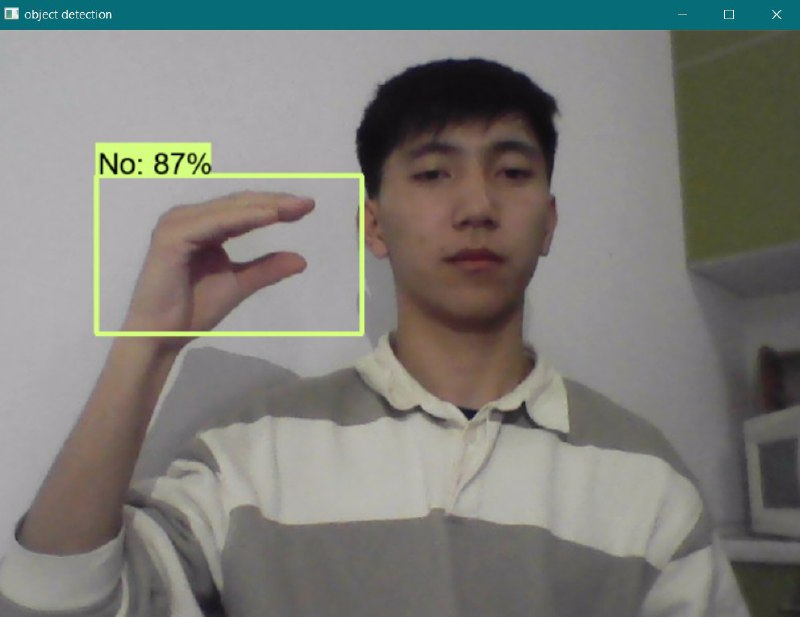
****

****

****

****

****

****

**Discussion**

Critical Review of Results

The findings suggest that it is possible to optimize real-time object detection models for enhanced speed without drastically sacrificing accuracy. The project's approach demonstrates notable improvements over traditional models, particularly in processing efficiency.

**Next Steps**

Future work will explore the integration of neural architecture search (NAS) to automatically identify optimal model configurations and further advancements in reducing computational load for deployment in edge devices.

**Conclusion:**

Through our project, we have successfully developed and implemented a real-time object detection system, presenting a robust solution to the speed-accuracy trade-off. This represents a significant advancement in the field of computer vision and artificial intelligence, opening new avenues for more responsive and efficient applications across various industries.

Our system not only delivers high-speed object detection but also maintains a high level of accuracy, making it a valuable tool for real-world problem-solving. Additionally, we have considered the needs of diverse users, including those who use sign language, to ensure a more accessible and inclusive technology.

We hope that our work will inspire and drive further research in this field, leading to the creation of even more advanced and innovative solutions.

References

1) <https://link.springer.com/chapter/10.1007/978-3-319-46448-0_2>

2) <https://tproger.ru/translations/opencv-python-guide>  
3)<https://proceedings.neurips.cc/paper/2015/file/14bfa6bb14875e45bba028a21ed38046-Paper.pdf>