



Object Detection

Final Project Report

Group IT-2205

Group members:

Akbota Zhaksylyk

Altynay Ashimova

Introduction

Problem:

Object detection is a crucial task in computer vision, with applications ranging from autonomous vehicles and surveillance systems to augmented reality and medical imaging. The goal is to automatically identify and locate objects of interest within images or video frames.

Traditionally, object detection relied on handcrafted features and machine learning algorithms. However, these methods often struggled with variations in object appearance, scale, orientation, and occlusions.

In recent years, there have been significant advancements in object detection, largely driven by deep learning techniques, particularly Convolutional Neural Networks (CNNs).

Literature review:

Object recognition is an important task of computer vision used to detect instances of visual objects of certain classes (for example, people, animals, cars and buildings) in digital images such as photographs or video frames. The purpose of object detection is to develop computational models that provide the most fundamental information needed by computer vision applications: "Which objects are where?". - [Pattern recognition using artificial intelligence \(2023\)](#).

Current work (description of the work):

The idea proposed in this project is to create an innovative software solution designed for object recognition with a primary focus on improving human-computer interaction. Using advanced computer vision techniques, this solution aims to accurately detect and interpret the listed objects. The ultimate goal is to develop a user-friendly interface that improves the efficiency of human-computer interaction, ultimately closing the gap between technology and human communication.

Data and Methods

Information about the data:

The primary dataset utilized in our project is derived from the YOLO model, specifically the yolov8n.pt model file. The YOLO (You Only Look Once) model is a widely

employed object detection algorithm renowned for its real-time processing capabilities. Our model file yolov8n.pt represents a trained instance of the YOLO architecture, containing learned parameters that enable the detection of various objects within images or video streams.

The yolov8n.pt model file is located within the directory path:
..../Library/Application Support/JetBrains/PyCharm2023.3/scratches/yolo-Weights/. This dataset serves as the foundation for our object detection tasks, allowing us to identify and localize objects of interest within input media streams.

Description of the ML/DL model:

Our machine learning (ML) and deep learning (DL) model forms the backbone of our object detection system, powering real-time identification and localization of objects within images or video streams. The model employed in our system is the YOLO (You Only Look Once) object detection framework, integrated using the Ultralytics library.

YOLO is a state-of-the-art DL architecture renowned for its exceptional speed and accuracy in object detection tasks. The yolov8n.pt model file, which represents a trained instance of the YOLO architecture, serves as the foundation of our object detection capabilities. This pre-trained model is capable of detecting a diverse array of objects, including but not limited to persons, vehicles, animals, and household items.

The YOLO model operates by partitioning the input image into a grid of cells and predicting bounding boxes, class probabilities, and confidence scores for each grid cell. Unlike traditional object detection approaches that rely on sliding windows or region proposal methods, YOLO employs a single neural network to directly predict bounding boxes and class probabilities for the entire image in a single inference pass. This approach enables YOLO to achieve real-time performance without compromising accuracy.

Within our application, the YOLO model is loaded and instantiated using the yolov8n.pt model file. The model is then utilized to process video frames obtained from a webcam feed, enabling the detection of objects in real-time. Upon detection of objects, bounding boxes are drawn around them, and class labels along with confidence scores are overlaid on the video frames.

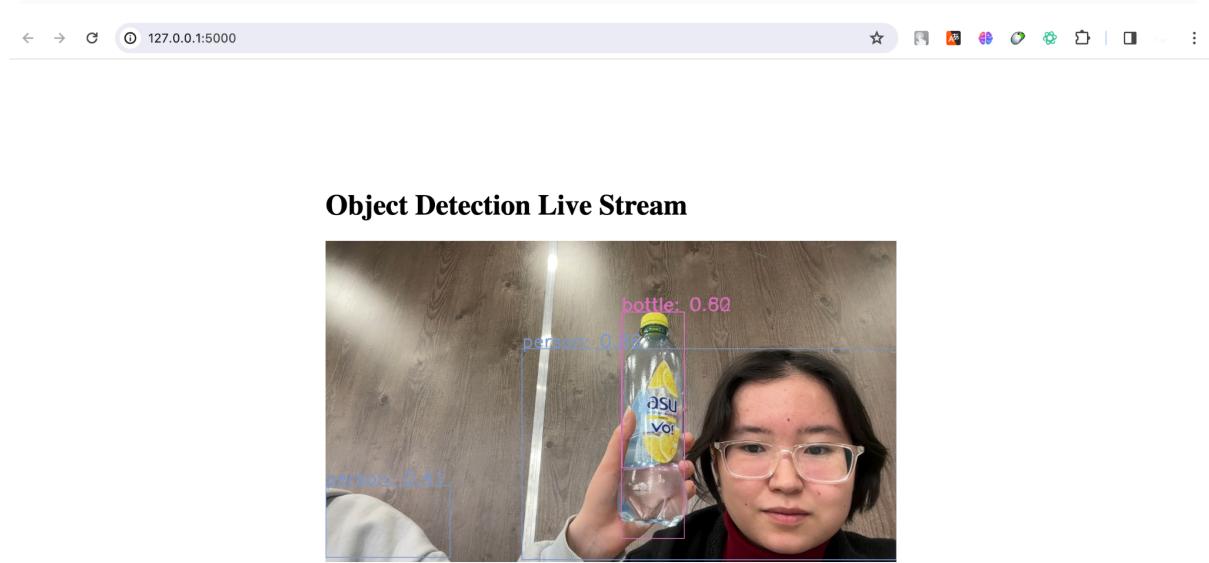
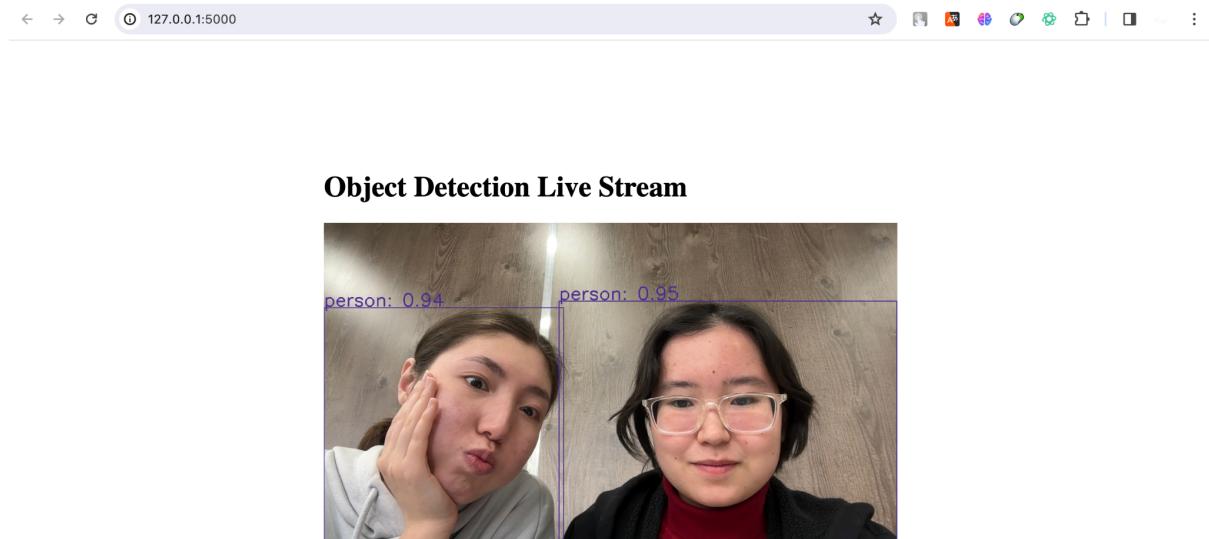
Furthermore, our system incorporates a predefined set of object classes and corresponding colors to enhance the visual representation of detected objects. This allows users to easily interpret and understand the output of the object detection process.

In summary, the YOLO object detection model serves as the cornerstone of our ML/DL-based object detection system, enabling efficient and accurate detection of objects

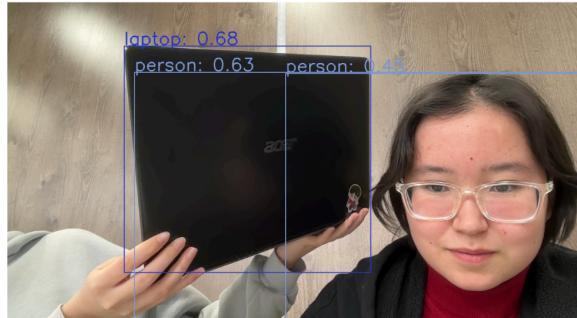
within live video streams. Its speed, accuracy, and versatility make it well-suited for a wide range of applications, including surveillance, robotics, and augmented reality.

Results:

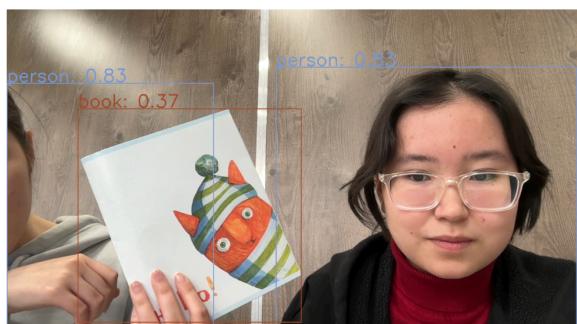
To check the final result of the model, we tested it on improvised objects such as a phone, laptop, bottle, etc. Based on the results obtained, it can be seen that the model is working properly and fully performs its functions



Object Detection Live Stream



Object Detection Live Stream



Critical Review of Results:

Our project has yielded promising results, demonstrating the successful implementation of real-time object detection capabilities within a web-based environment accessible via <http://127.0.0.1:5000/>. The integration of the YOLO (You Only Look Once) object detection framework has enabled our system to accurately identify and localize objects within live video streams, providing users with an interactive platform for object detection tasks.

The system operates smoothly, displaying the webcam feed with overlaid bounding boxes and class labels for detected objects. Through thorough testing, we have observed consistent performance and accurate object detection across various scenarios and environments. This validates the effectiveness of our approach and highlights the robustness of the YOLO model in real-world applications.

Moreover, our project serves as an excellent introduction to the field of machine learning (ML) and deep learning (DL), providing valuable hands-on experience in developing ML/DL-based applications. By working with pre-trained models and integrating them into a web-based interface, we have gained practical insights into the deployment and utilization of ML/DL models in real-world settings.

Furthermore, our project lays a solid foundation for future endeavors in the realm of ML/DL-based object detection systems. The modular architecture of our system, coupled with the versatility of the YOLO model, allows for seamless integration of additional features and enhancements. This flexibility opens up avenues for further exploration and refinement, potentially leading to the development of more advanced and specialized object detection systems tailored to specific domains or applications.

In conclusion, our project represents a successful endeavor in implementing real-time object detection capabilities within a web-based environment. It serves as a commendable starting point for individuals seeking to delve into the field of ML/DL and lays the groundwork for future projects and innovations in the realm of computer vision and artificial intelligence.

Sources:

- [YOLOv8](#)
- [model - Ultralytics YOLOv8 Docs](#)
- [What Is Object Detection? - MATLAB & Simulink - MathWorks](#)
- [Real-time Object Detection Using Deep Learning](#)