Wrocław University
of Science and Technology

Sign language recognition program

Contractors:
Ryan Oliver Roberts
Rafał Cielenkiewicz

Supervisor : Doctor Eng. Mateusz
Cholewiński

Wrocław 2024

Contents

1 Project Overview

1.1 Project Description

The Sign Language Recognition project is a groundbreaking initiative aiming to develop an advanced system for the recognition and interpretation of sign language gestures. At the core of the project is the state-of-the-art YOLOv3 (You Only Look Once) object detection algorithm. The main goal is to create an efficient and accurate tool that will significantly advance the field of sign language interpretation.

1.2 Case Studies

While previous projects have focused on sign language recognition, this endeavor stands out by exploiting the capabilities of YOLOv3. By integrating this state-of-the-art object detection algorithm, we aim to improve accuracy and achieve real-time performance. The use of YOLOv3 brings a novel approach to the field, promising a significant step forward in sign language recognition and interpretation.

`https://www.pjreddie.com/media/files/papers/YOLOv3.pdf`
`https://www.youtube.com/watch?v=MPU2HistivI`

1.3 Impact on Lives

The significance of our code and analogous projects extends beyond the technological sphere. Through the precise real-time interpretation of sign language, we empower individuals with hearing impairments to communicate and integrate more seamlessly. Aligned with broader societal objectives, this project contributes to enhancing accessibility and dismantling communication barriers.

2 Project overview

2.1 Introduction

Sign language, a timeless form of communication, has traversed centuries and is ingrained in diverse communities globally. While pinpointing its exact origin remains elusive, many societies have embraced various forms of gesture and sign communication throughout history. Notably, modern sign languages like American Sign Language (ASL) and Polish Sign Language (PJM) have evolved into fully-fledged modes of communication for the Deaf community.

The World Federation of the Deaf estimates that there are over 70 million Deaf individuals worldwide. It's crucial to recognize that not all within this demographic use sign language, and diverse cultures may boast unique sign language systems.

While we lack direct connections with individuals who communicate through sign language, this project has afforded us profound insights into this distinctive form of expression. Sign language serves as a pivotal means of communication for the Deaf community, prompting our pioneering effort in the Sign Language Recognition project. Our goal is to harness the capabilities of YOLOv3, renowned for its accuracy and real-time object detection speed, to elevate the field of sign language recognition.

In the realm of computer vision, YOLOv3 stands as a cutting-edge object detection algorithm tailored for real-time applications. Its groundbreaking methodology involves dividing the input image into a grid, enabling simultaneous prediction of bounding boxes and class probabilities across the entire image in a single pass.

Initially, our plan was to take our own photos for each sign language symbol. However, this turned out to be more challenging than anticipated. After capturing 20 photos for each symbol, we realized it wasn't enough to train our system effectively. Consequently, we decided to change our approach and started downloading images from the internet. This adjustment allowed us to access a larger and more diverse dataset, improving the overall quality and performance of our project.

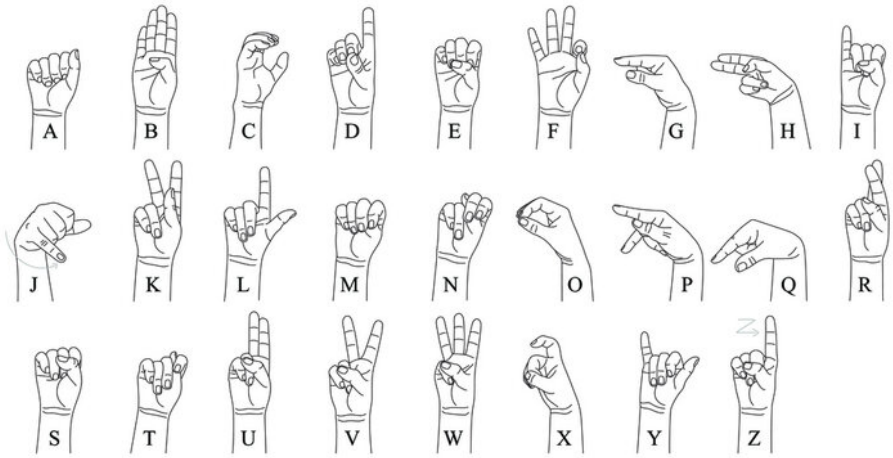

Figure 2.1: American Sign Language

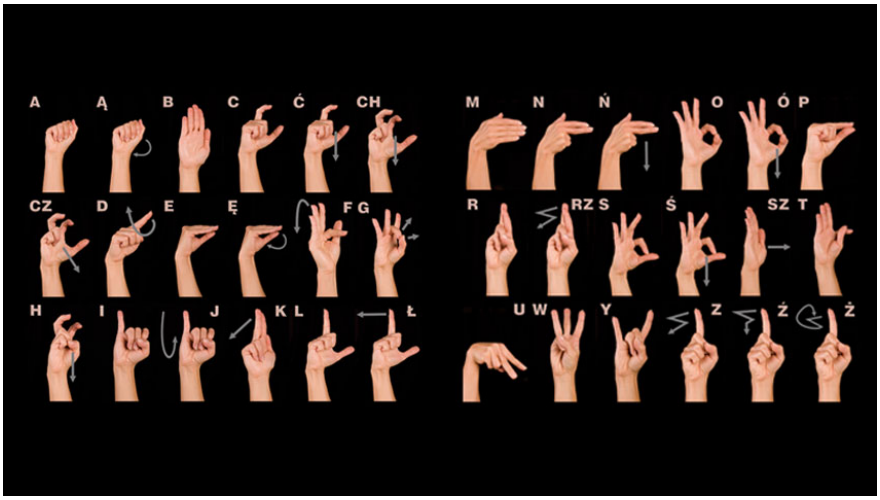

Figure 2.2: Polish Sign Language

2.2 Objectives

- **The project has well-defined and big goals that guide how it grows. These goals are carefully crafted to tackle important parts of understanding sign language, using new and advanced technology.**
  - Develop a robust sign language gesture detection system: Using the capabilities of YOLOv3, our aim is to develop a robust system capable of accurately detecting and recognising a wide range of sign language gestures.
  - Implementing sign language interpretation logic: In addition to detection, the project aims to implement interpretation logic that can translate recognised gestures.
  - Ensuring real-time performance: Real-time performance is a critical aspect of the project as it has a direct impact on the usability of the system

3 YOLO-v3 Model

3.1 Introduction

YOLO-v3 (You Only Look Once) is a state of the art real time objection algorithm, Created by Joseph Redmon and Al Farhadi, this model is faster and processes better than other competition. It excels in single pass detection, real time detection, high accuracy, versatility and objectness score. This makes it an ideal solution for the implementation of computer vision in our project

3.2 Database

A well made database is essential for the robustness of the implementation of Computer Vision. There is an inbuilt database present for public usage made by Microsoft with common objects present in the world. This collection is perfectly suitable for our project since our project uses a robot that moves in common spaces and the ability to detect this will be a great fortune for the entirety of the project.

3.2.1 Personal Database

The initial round of training included the painstaking process of taking multiple photos of each class and then annotating them for a response.

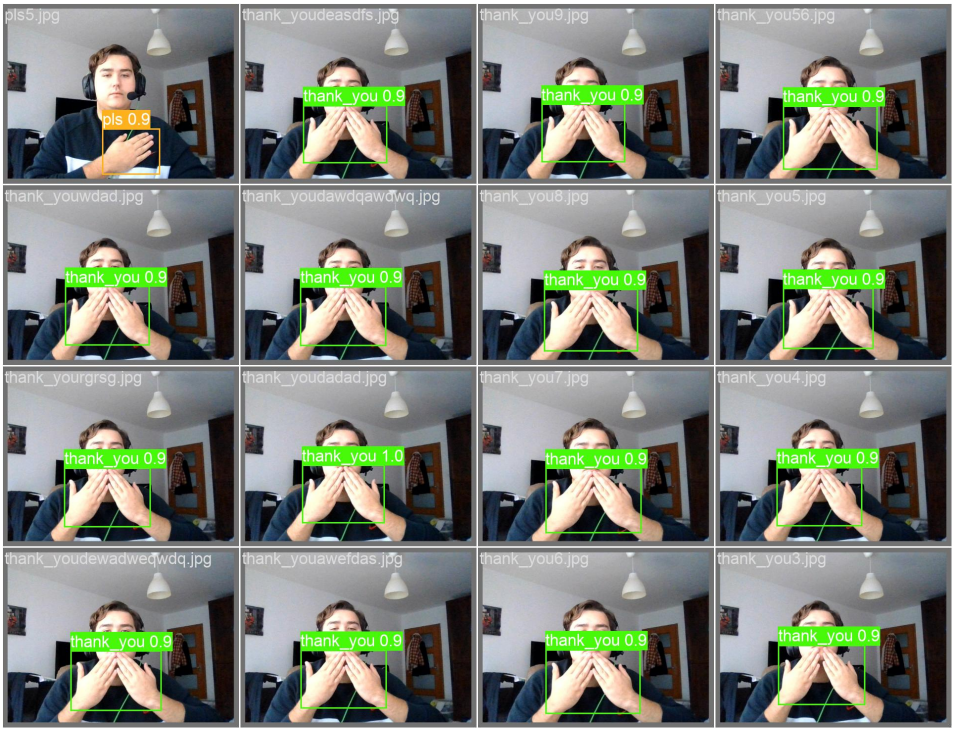

Fig. 1. Public Database

3.2.2 Public Database

There is a pre made dataset present in the dataset website ”Roboflow” which was non annotated neither trained and was a composite of 1000 pictures in ASL depicting different letters.

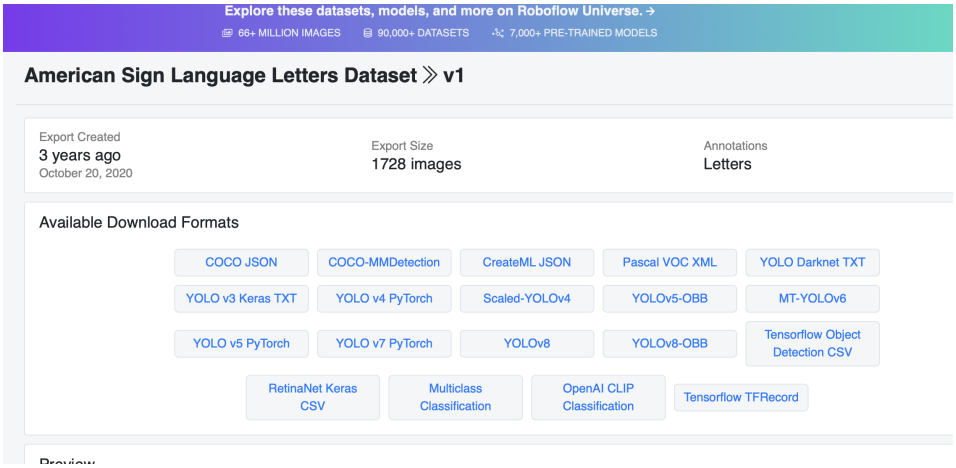

Fig. 2. Public Database

3.2.3 Annotation

To annotate, the label.mg software was used to create classes and then annotate the images present in the dataset. This software lets us create bounding boxes for each of the classes and helps us get a more fine-grain dataset with real usage value.

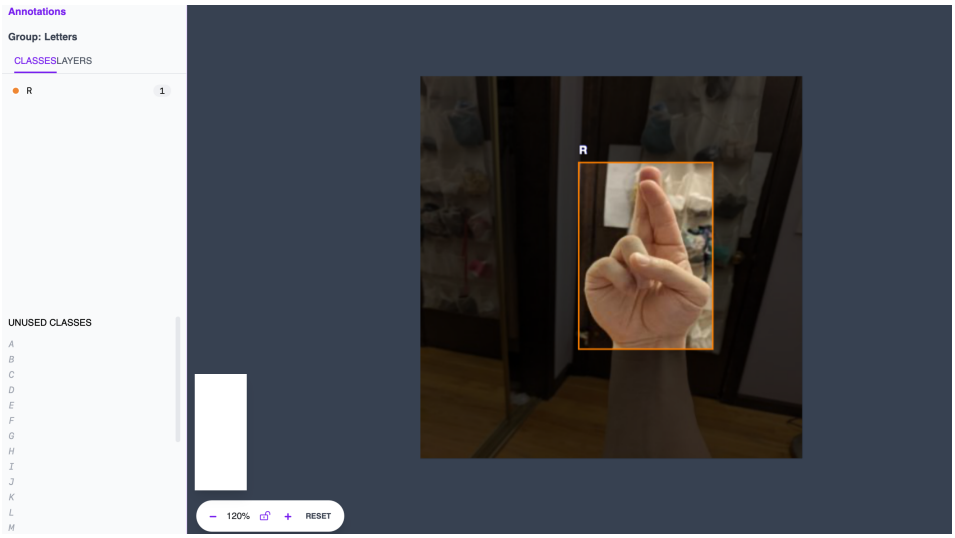

Fig. 3. Annotation

4 Learning and Training

4.1 Introduction

A database alone cannot act in the field of Computer Vision. It needs to be trained and deployed in a meaningful manner for optimal success. With the help of some training tools we are able to train a model that is suited for our needs and will work in a smooth and possible manner.

4.2 Training

To train the model, the usage of the 'Ultralytics' library was chosen for this purpose, the library consists of YOLO in which we were able to run a small neural network for processing, after splitting the images into test, train and validation. we ran 100 EPOCH's to see the outcome.

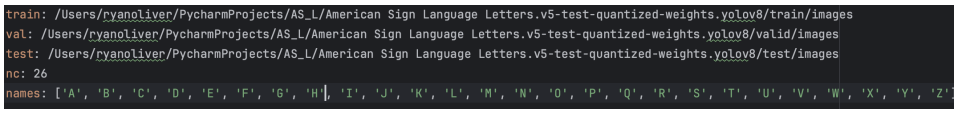

Fig. 4. data.yaml

In the figure above here is the YAML file with all the classes and the 3 segregation of test, train and validation files.

5 Conclusion

The Sign Language Recognition project represents a significant step forward in leveraging cutting-edge technology, specifically the YOLOv3 object detection algorithm, to address the unique challenges faced by the Deaf community. Through meticulous planning, adaptation, and training, our team has worked towards achieving several key objectives.

The project's primary goal is the development of a robust sign language gesture detection system. By utilizing the capabilities of YOLOv3, we aim to accurately identify and recognize a diverse range of sign language gestures. This involves not only the implementation of sophisticated object detection but also the integration of an interpretation logic that can translate recognized gestures.

One of the crucial aspects of our project is ensuring real-time performance. The ability to provide instantaneous feedback and interpretation is vital for the usability of the system, especially in real-world scenarios where effective communication is time-sensitive.

The choice of YOLOv3 as the underlying model brings advantages such as single-pass detection, real-time processing, high accuracy, and versatility. This model, coupled with a well-constructed database, has allowed us to train our system effectively. The training process involved not only a personal database created through meticulous image capture and annotation but also the utilization of a publicly available dataset to enhance the model's versatility.

Despite initial challenges in collecting a sufficiently large and diverse dataset, our decision to incorporate images from the internet significantly improved the quality and performance of the project. This adaptation reflects our commitment to overcoming obstacles and optimizing the system for real-world applications.

In conclusion, the Sign Language Recognition project goes beyond the realms of technology. Its impact extends to empowering individuals with hearing impairments, breaking down communication barriers, and contributing to the broader societal goal of enhancing accessibility. As we move forward, continuous refinement, testing, and user feedback will be crucial for further improving the system's accuracy and effectiveness in real-world scenarios. The journey undertaken in this project exemplifies the fusion of technology and social impact, demonstrating the potential of AI to positively transform lives.

6 Acknowledgement

We would like to give our sincere gratitude to our professor Dr.Mateusz Cholewiński for his help and guidance during the entirety of the project and giving out valuable feedback to make the project a success.