Bridging the communication gap

AI-Enabled Sign Language Understanding and Learning

Zewen Yang

Agenda

1. Introduction
2. Sign Language Recognition with Object Detection
3. Sign Language Understanding with Action Detection
4. Sign Language Learning with Gesture Analysis
5. Discussion

Agenda

1. Introduction
2. Sign Language Recognition with Object Detection
3. Sign Language Understanding with Action Detection
4. Sign Language Learning with Gesture Analysis
5. Discussion

Introduction

Background:

- Over **5%** of the world's population – or **430 million** people – require to address their disabling hearing loss

Introduction

Background:

- Over **5%** of the world's population – or **430 million** people – require to address their disabling hearing loss
- WHO: **1 in 4** people projected to have hearing problems by 2050

Introduction

Background:

- Over **5%** of the world's population – or **430 million** people – require to address their disabling hearing loss
- WHO: **1 in 4** people projected to have hearing problems by 2050
- More than **300** different sign languages in use around the world

Introduction

Motivation:

- Bridge communication gaps and enable more effective interactions

Introduction

Motivation:

- Bridge communication gaps and enable more effective interactions
- Develop tools for independent living and mobility

Introduction

Motivation:

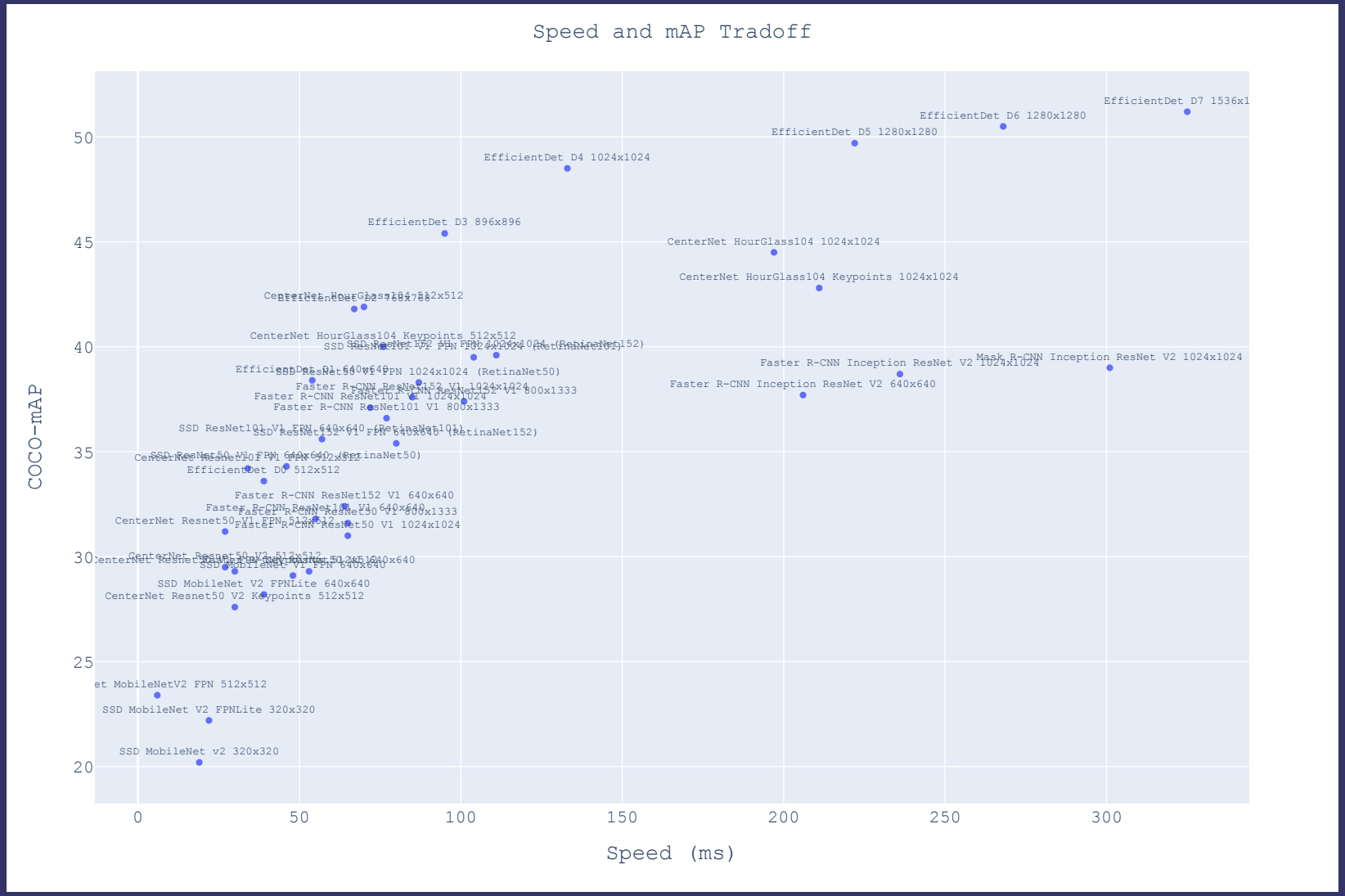
- Bridge communication gaps and enable more effective interactions
- Develop tools for independent living and mobility
- Enable more seamless interactions with AI technology

Agenda

1. Introduction
2. **Sign Language Recognition with Object Detection**
3. Sign Language Understanding with Action Detection
4. Sign Language Learning with Gesture Analysis
5. Discussion

Sign Language Recognition with Object Detection

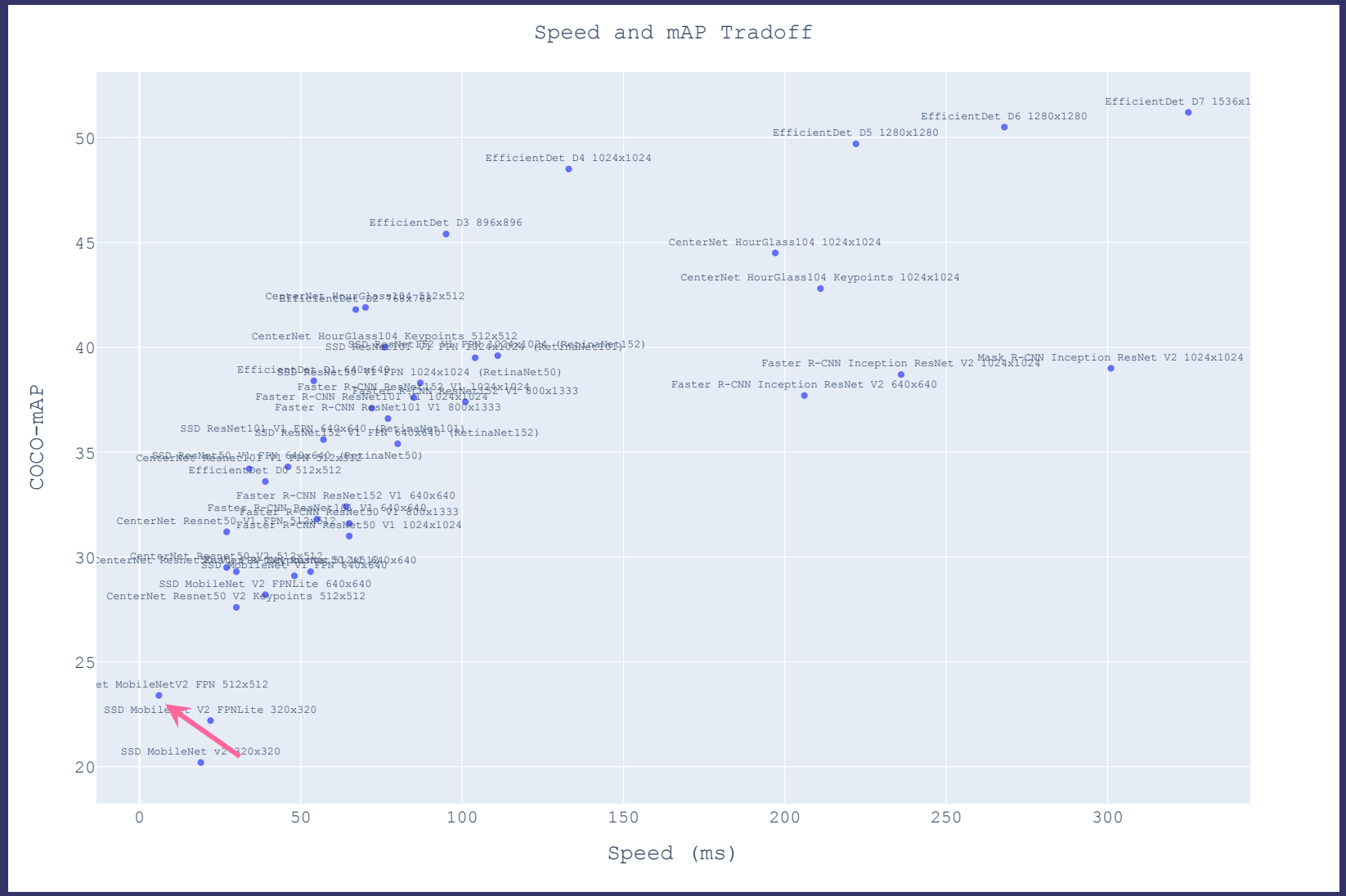
TensorFlow2 Detection Model Zoo:



Sign Language Recognition with Object Detection

TensorFlow2 Detection Model Zoo:

CenterNet MobileNetV2 FPN 512*512



Sign Language Recognition with Object Detection

Config:

- 20 images for each class
- Epochs = 90
- Retrain times = 2

Evaluation:

- mean Average Precision (mAP): 0.04

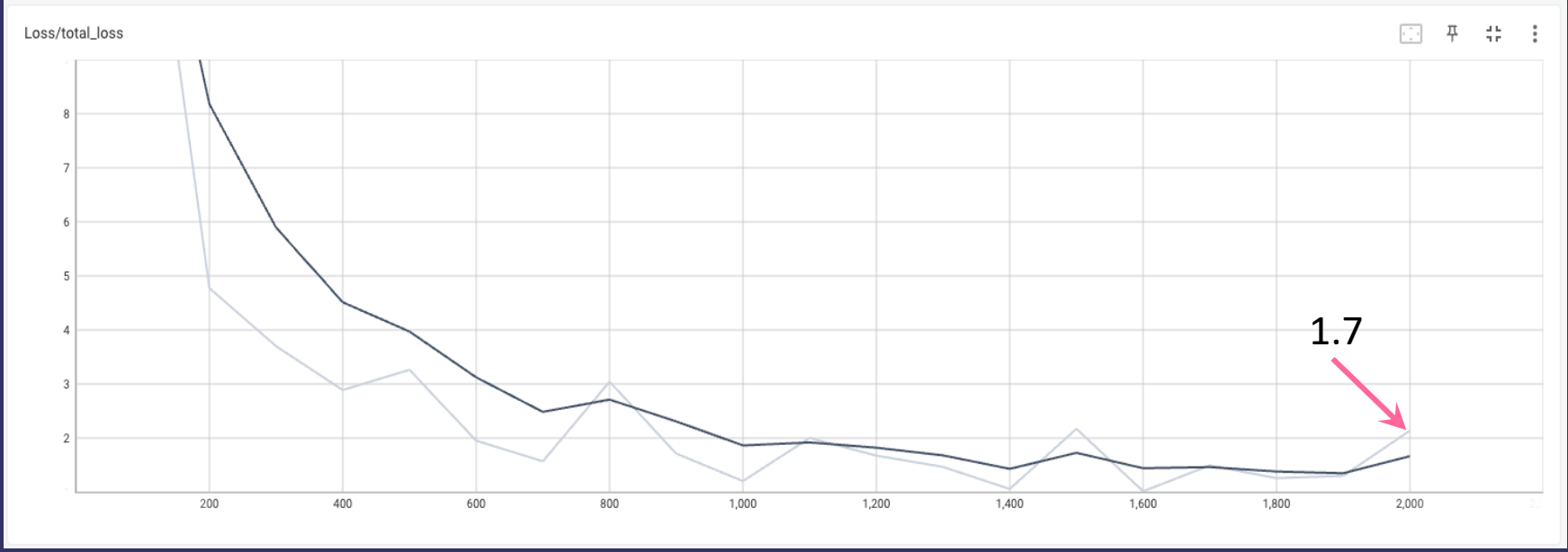
Sign Language Recognition with Object Detection

Config:

- 20 images for each class
- Epochs = 90
- Retrain times = 2

Evaluation:

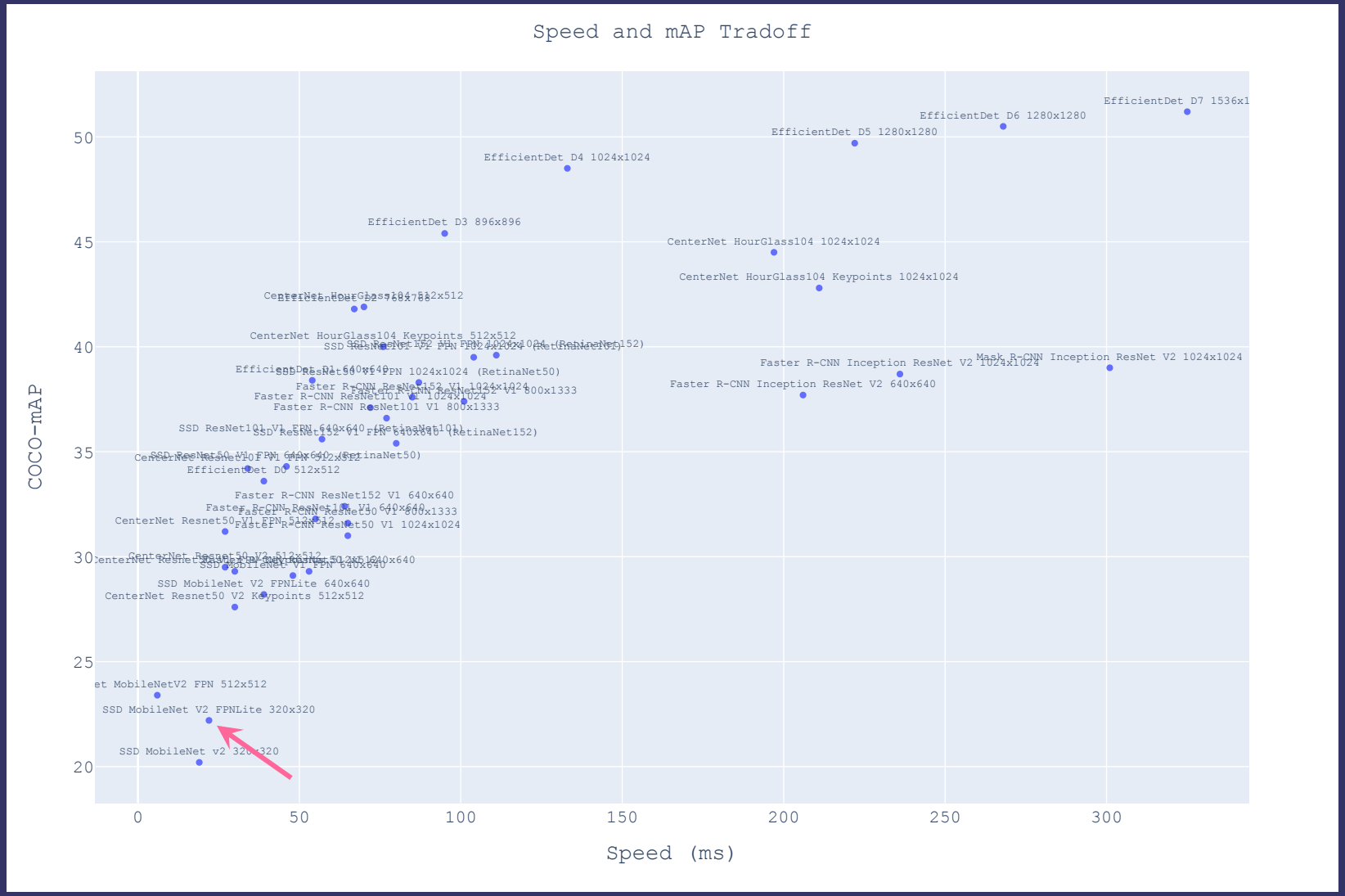
- mean Average Precision (mAP): 0.04



Sign Language Recognition with Object Detection

TensorFlow2 Detection Model Zoo:

SSD MobileNet V2 FPNLite 320*320



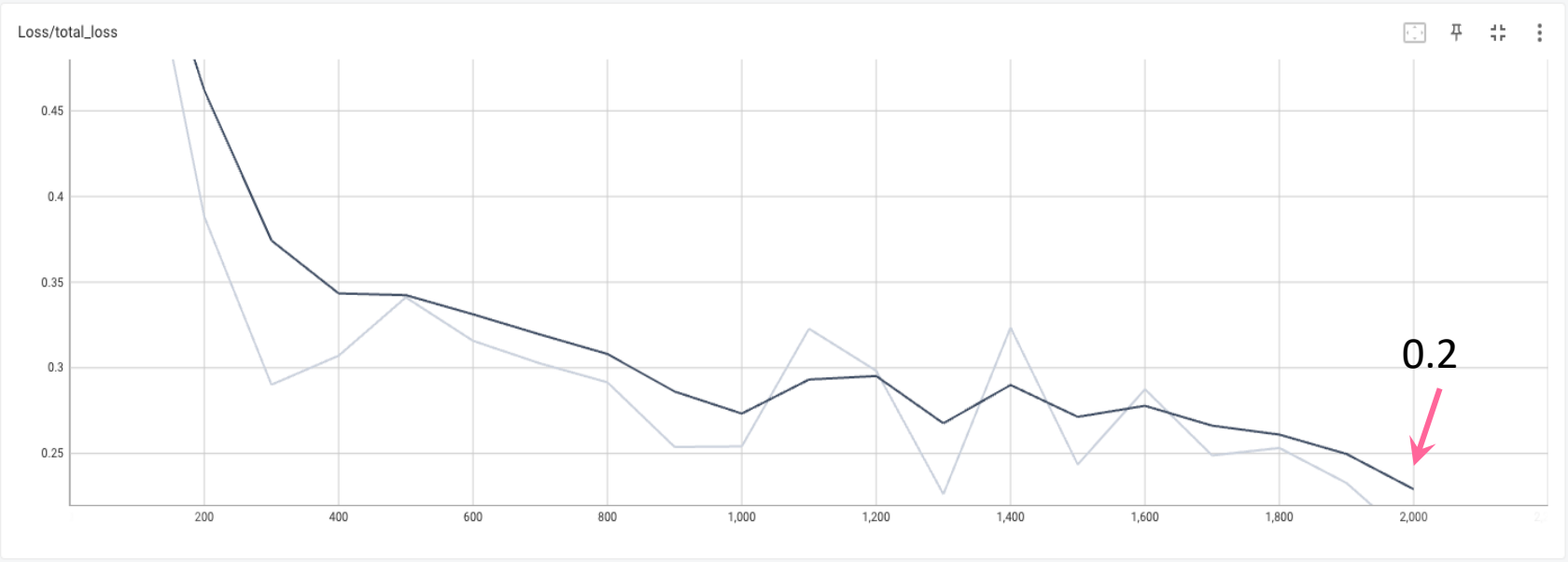
Sign Language Recognition with Object Detection

Config:

- 20 images for each class
- Epochs = 90
- Retrain times = 2

Evaluation:

- mAP: 0.77



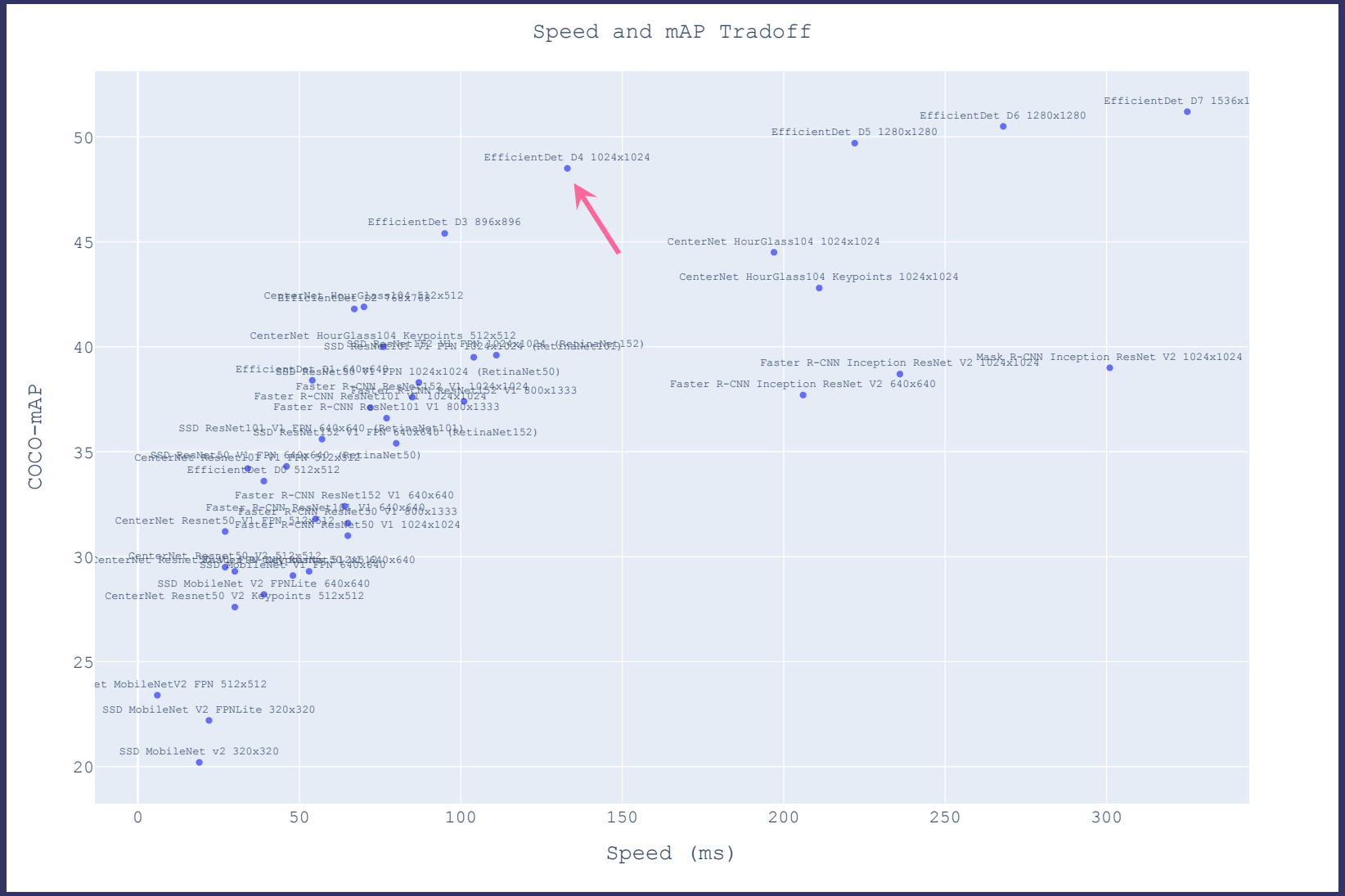
Sign Language Recognition with Object Detection

Live result:

Sign Language Recognition with Object Detection

TensorFlow2 Detection Model Zoo:

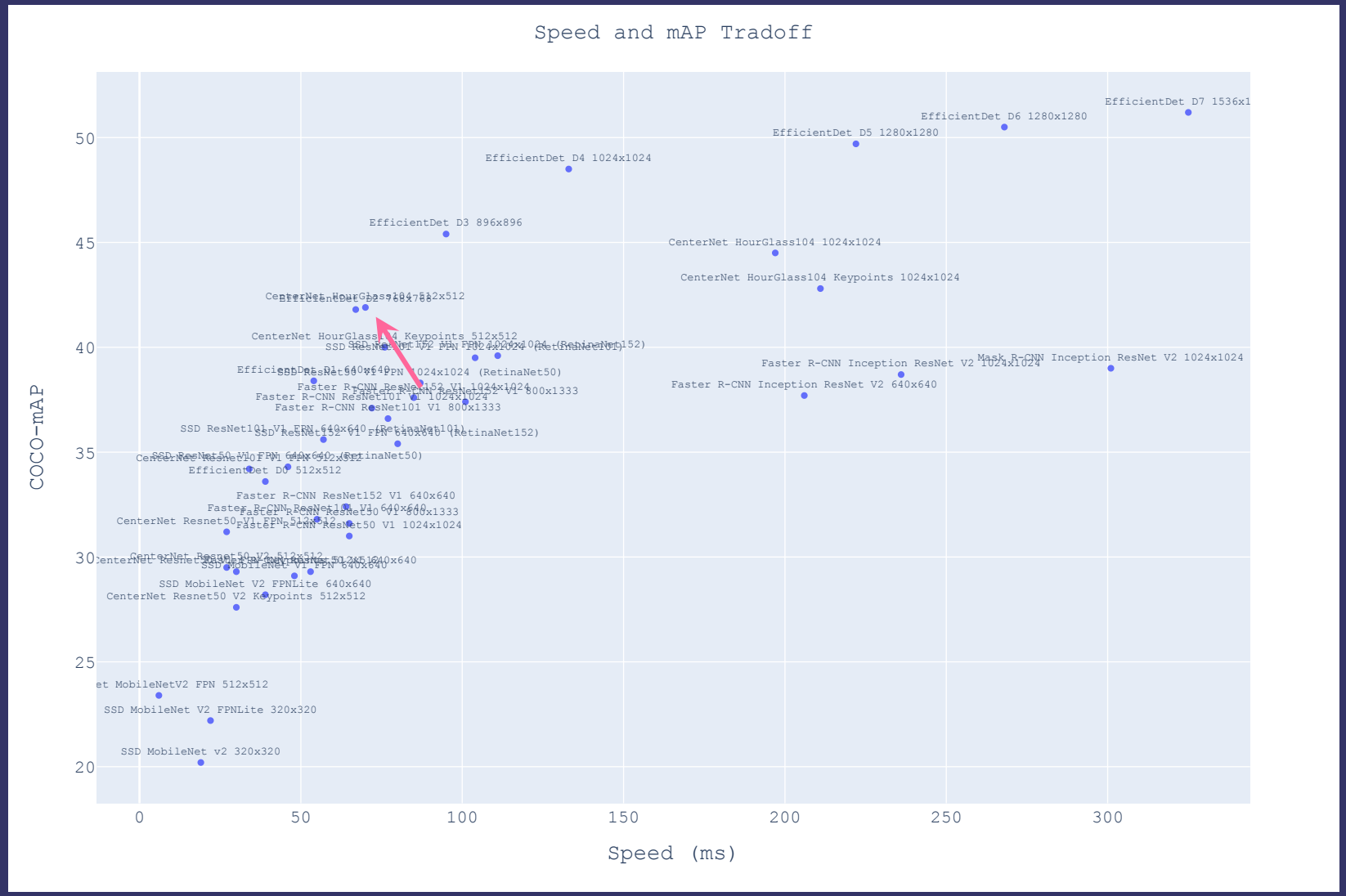
EfficientDet D4 1024*1024



Sign Language Recognition with Object Detection

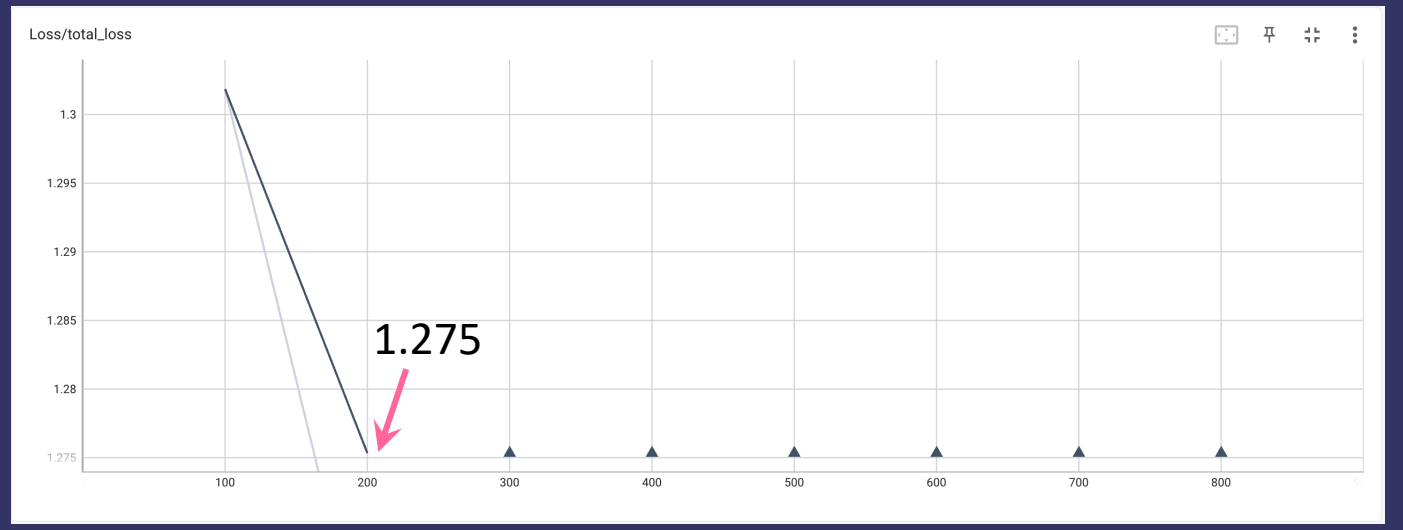
TensorFlow2 Detection Model Zoo:

EfficientDet D3 512*512



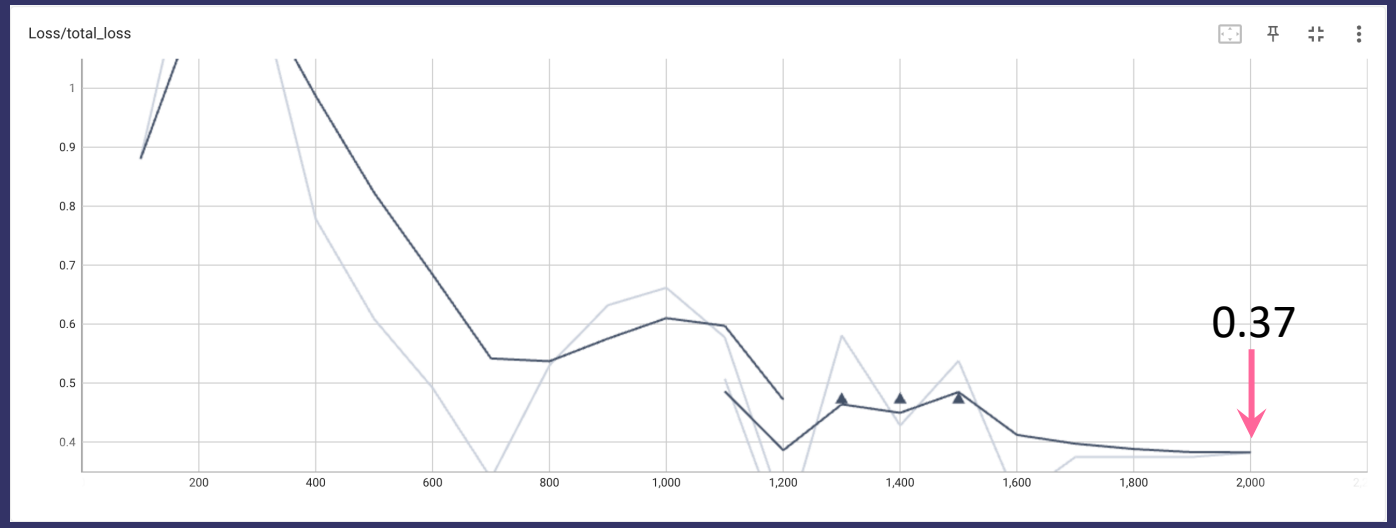
Sign Language Recognition with Object Detection

EfficientDet D4 1024*1024:



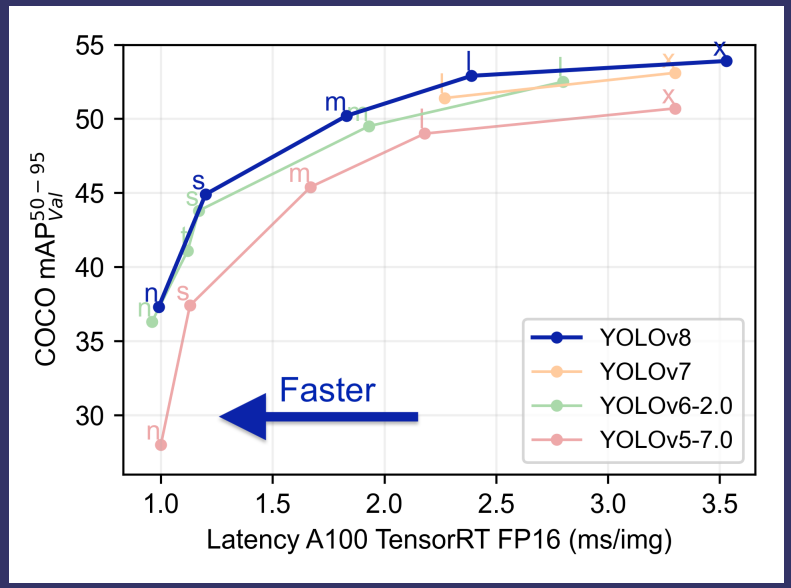
EfficientDet D3 512*512:

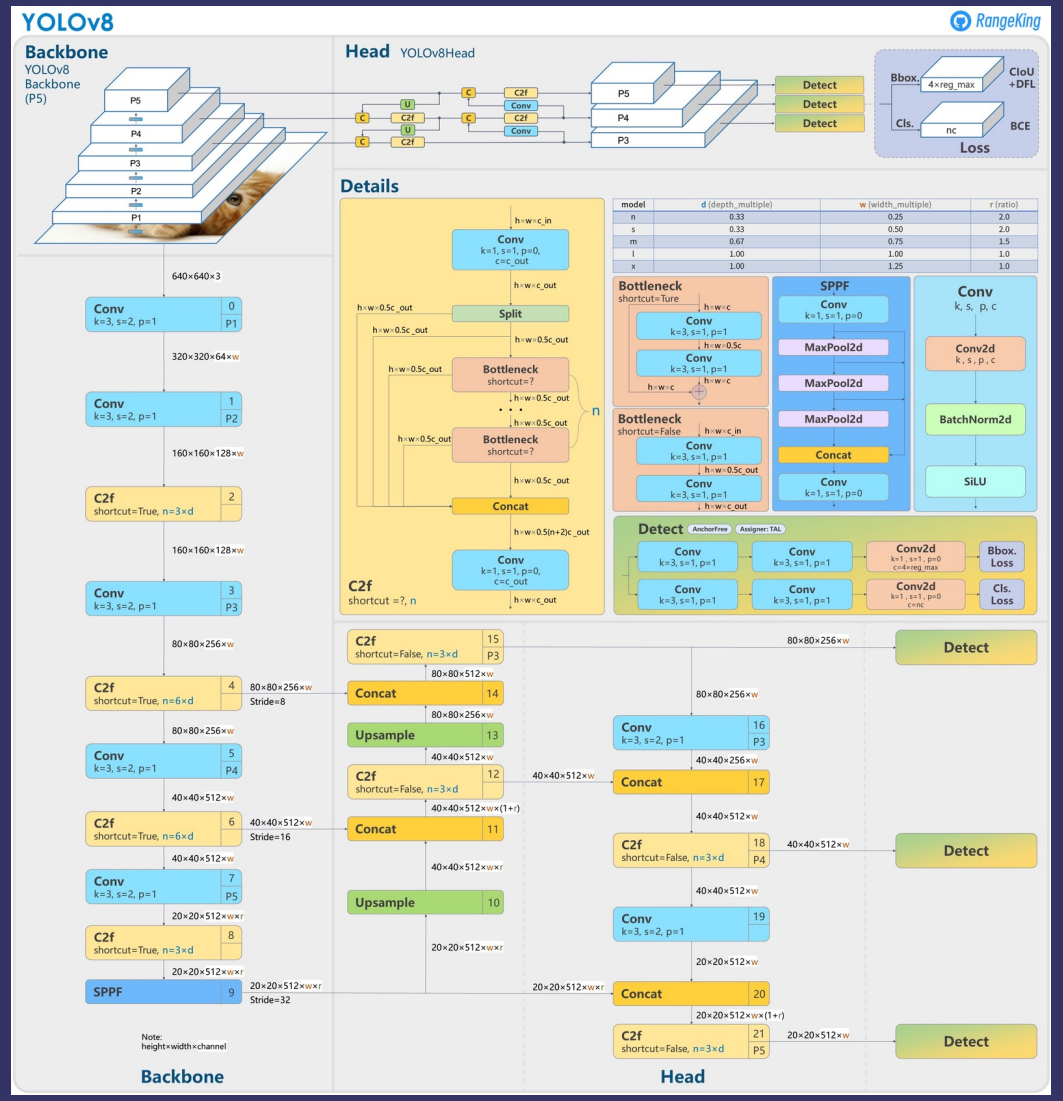
mAP :0.01



Sign Language Recognition with Object Detection

State-of-the-Art: YOLOv8





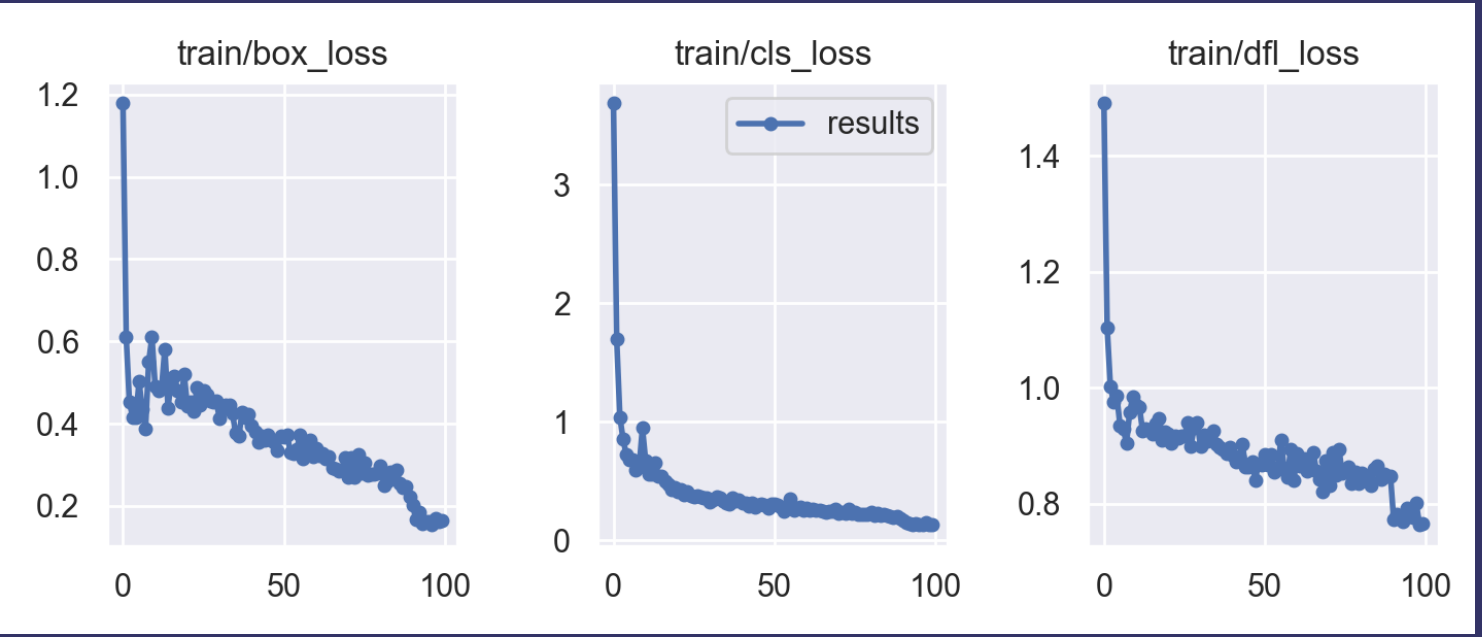
Sign Language Recognition with Object Detection

Config:

- 20 images for each class
- Epochs = 100
- Retrain times = 0

Evaluation:

- mAP: 0.995



Sign Language Recognition with Object Detection

Live Result:

Sign Language Recognition with Object Detection

What’s next?

Sign Language Recognition with Object Detection

What's next?

More photos for myself

Sign Language Recognition with Object Detection

What's next?

More photos for myself

Stop collecting data

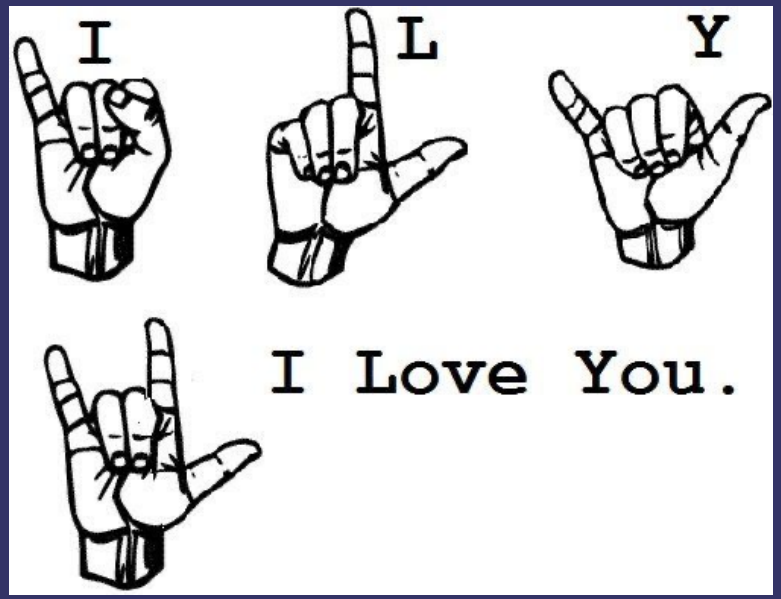
Sign Language Recognition with Object Detection

Question

Sign language is a dynamic form of communication.

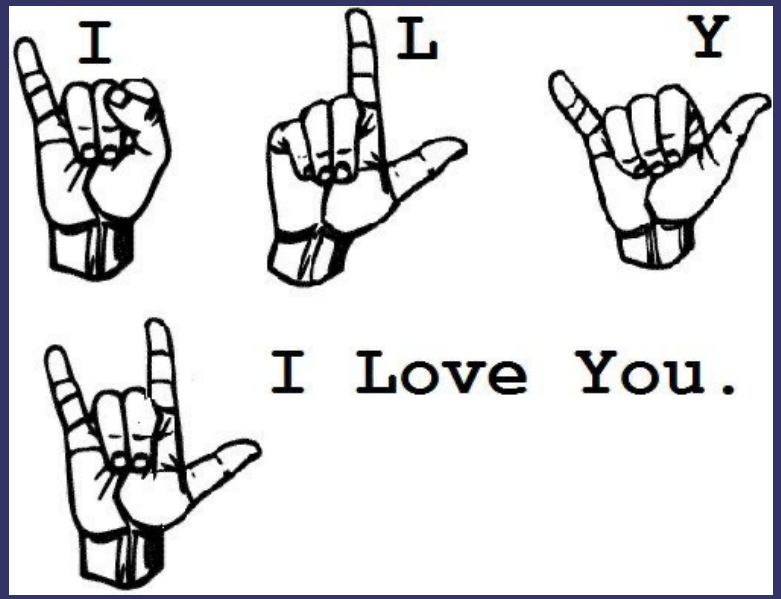
Sign Language Recognition with Object Detection

Static From:



Sign Language Recognition with Object Detection

Static From:



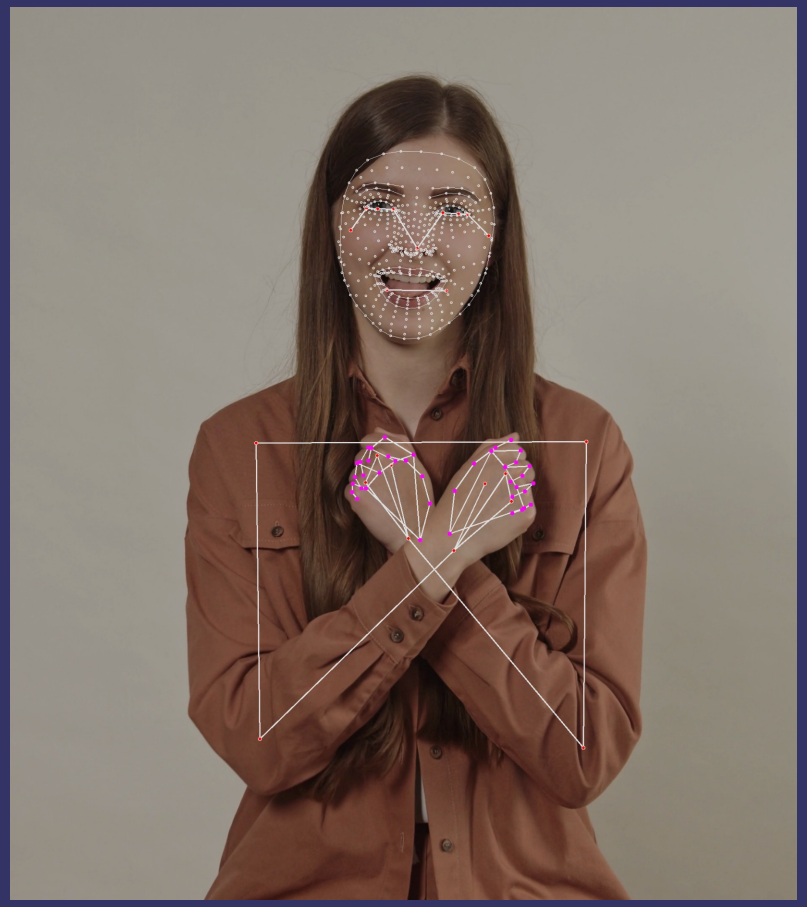
Action From:

Agenda

1. Introduction
2. Sign Language Recognition with Object Detection
3. **Sign Language Understanding with Action Detection**
4. Sign Language Learning with Gesture Analysis
5. Discussion

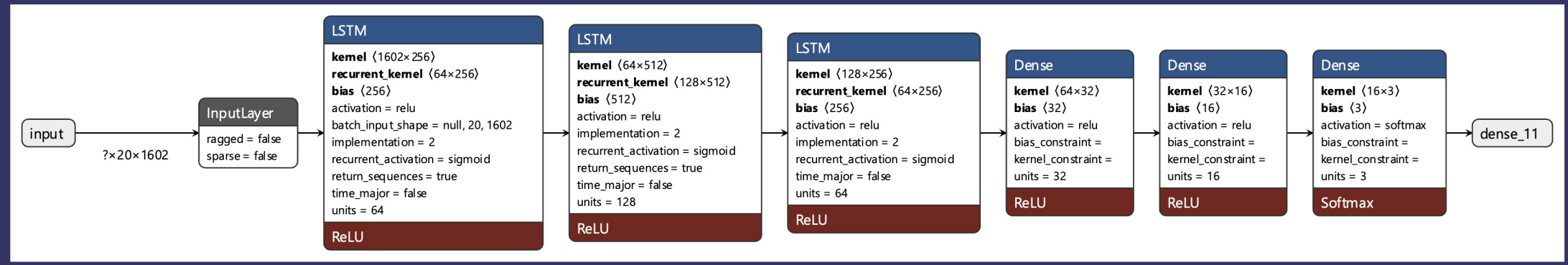
Sign Language Recognition with Action Detection

MediaPipe



Sign Language Recognition with Action Detection

Sign Language Recognition with Action Detection



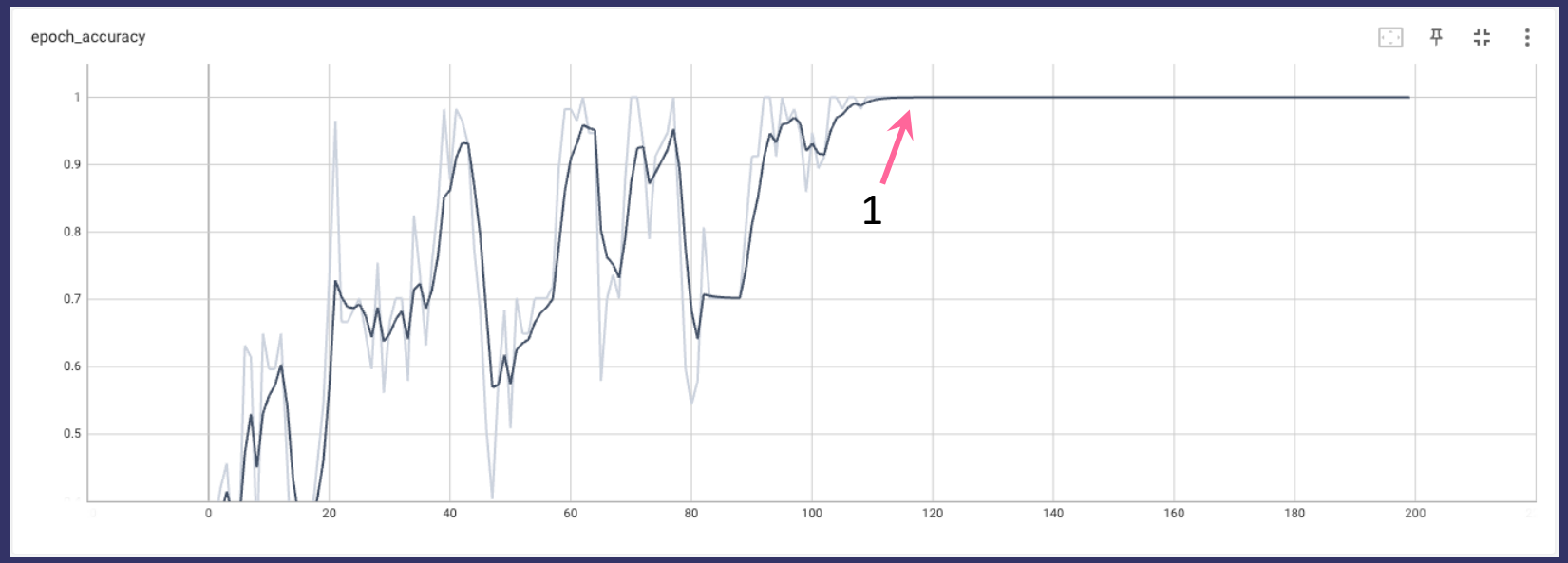
Sign Language Recognition with Action Detection

Config:

- 20 images for each class
- Epochs = 100
- Retrain times = 0

Evaluation:

- Accuracy: 1



Sign Language Recognition with Action Detection

Live Result:

Sign Language Recognition with Action Detection

What's next?

More photos for myself

Stop collecting data

Sign Language Recognition with Action Detection

Question

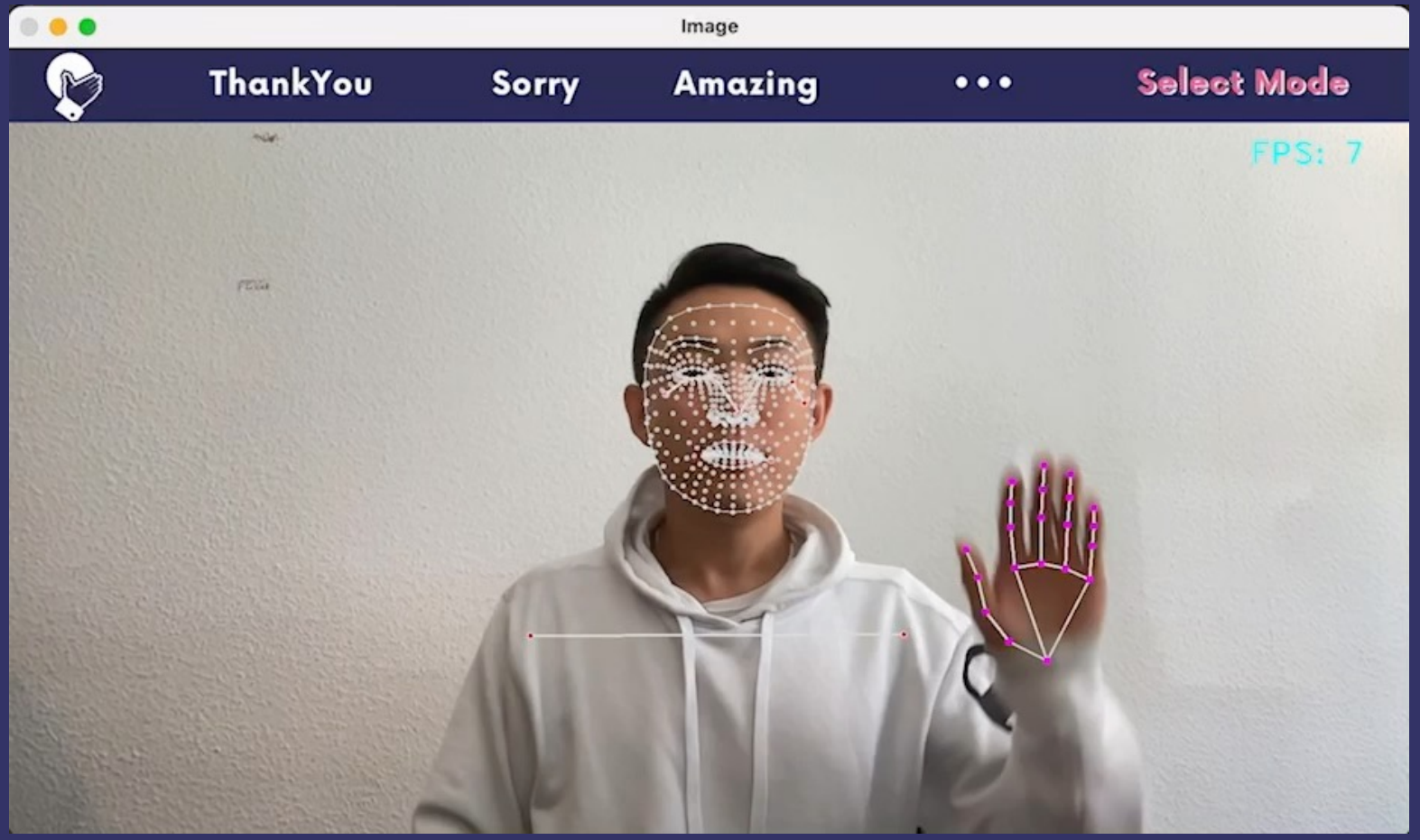
Effective communication is a two-way process.

Agenda

1. Introduction
2. Sign Language Recognition with Object Detection
3. Sign Language Understanding with Action Detection
4. **Sign Language Learning with Gesture Analysis**
5. Discussion

Sign Language Recognition with Gesture Analysis

Gesture Analysis with Expert Evaluation:



Agenda

1. Introduction
2. Sign Language Recognition with Object Detection
3. Sign Language Understanding with Action Detection
4. Sign Language Learning with Gesture Analysis
5. **Discussion**

Discussion

What could be tried:

- **Collect more data**
- **Train models longer**
- **Parameters turning**

Discussion

What could be tried:

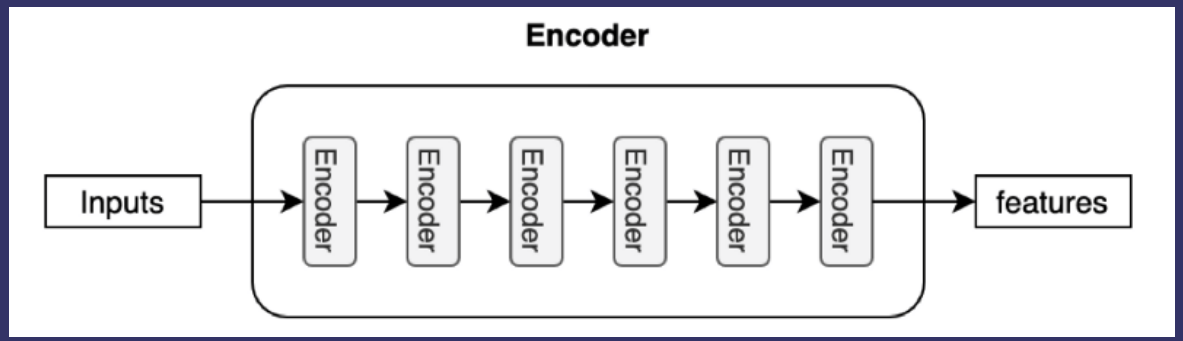
- Collect more data
- Train models longer
- Parameters turning
- Different models
  - Detectron2
  - Transformer encoder

Detectron2

Discussion

What could be tried:

- Collect more data
- Train models longer
- Parameters turning
- Different models
  - Detectron2
  - Transformer encoder



Discussion

Limitations:

- **Categories**
- **Detection Speed**
- **Gesture analysis requires human expert**

Bridging the communication gap

Thanks for listening!