Anastasia Churina

UNIVERSITY OF ARTIFICIAL INTELLIGENCE

Determining the number of people in the video

Date of presentation

31.08.2022

Beginning of study

01.10.2021

Problem definition

- Train the neural network to find and recognize people in the video;
- Display the total number of found people;
- Maximize the accuracy of detecting objects.

DataBase

Volume of database

8774 marked up images;

Source of data collection

free sources (Internet);

Data collection tools

the main part is an already marked base,

additional part - marked up using the makesense.ai service;

Difficulties encountered

inaccuracies found in the marking

of the main part of the database,

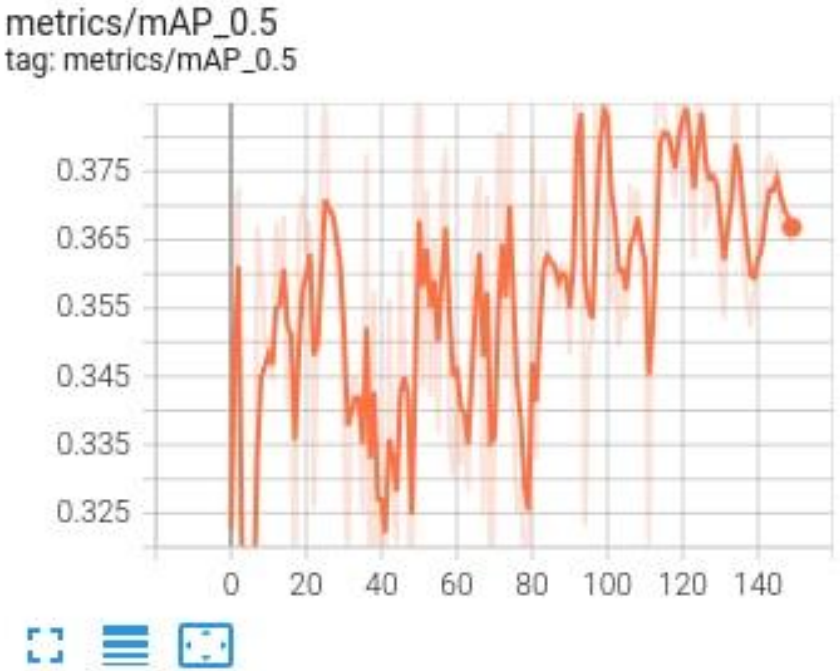
edits had to be performed in Roboflow.

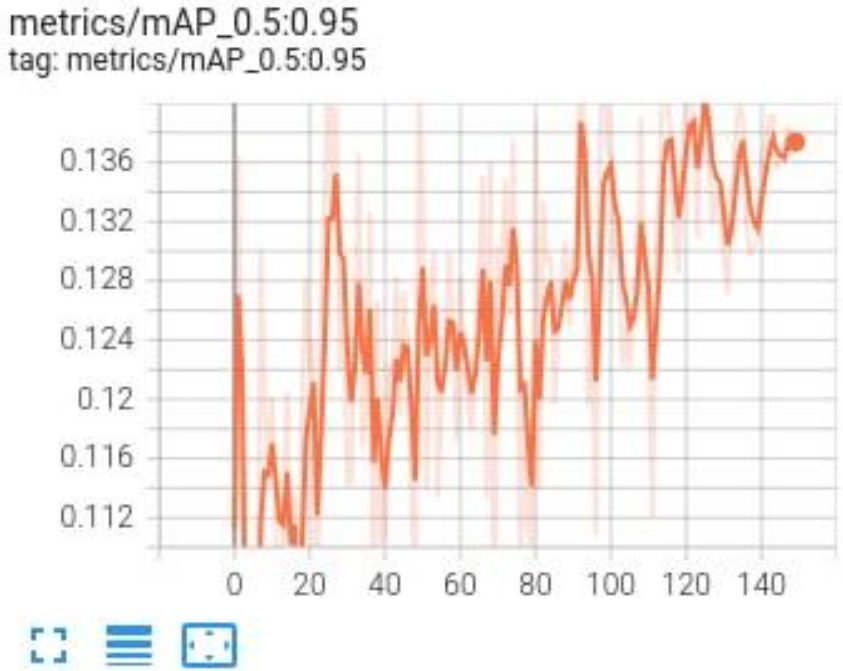
roboflow

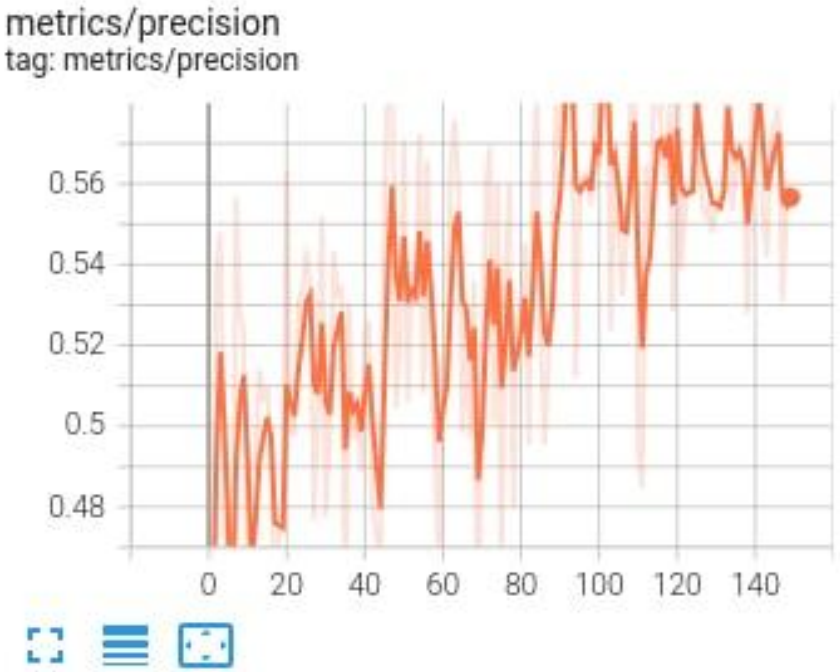
7063 Training Set
850 Validation Set
861 Test Set

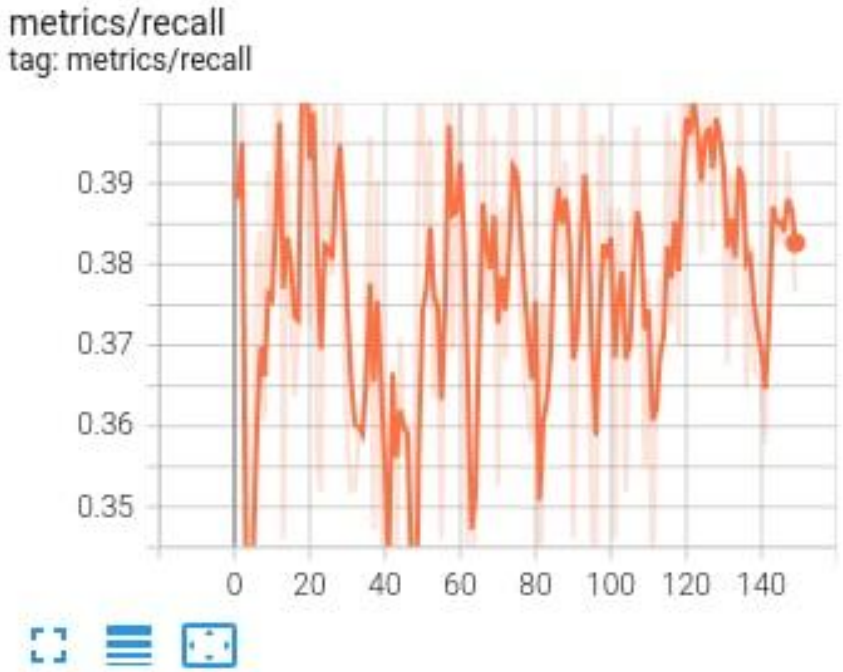
Problems with initial database.

The first attempt at training of the network resulted in low metrics and sawtooth learning curves.







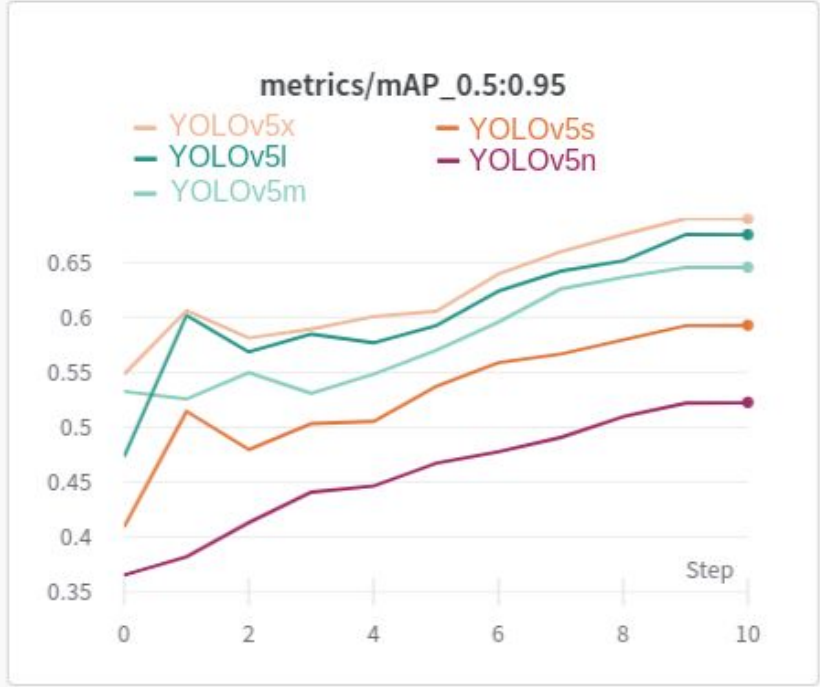


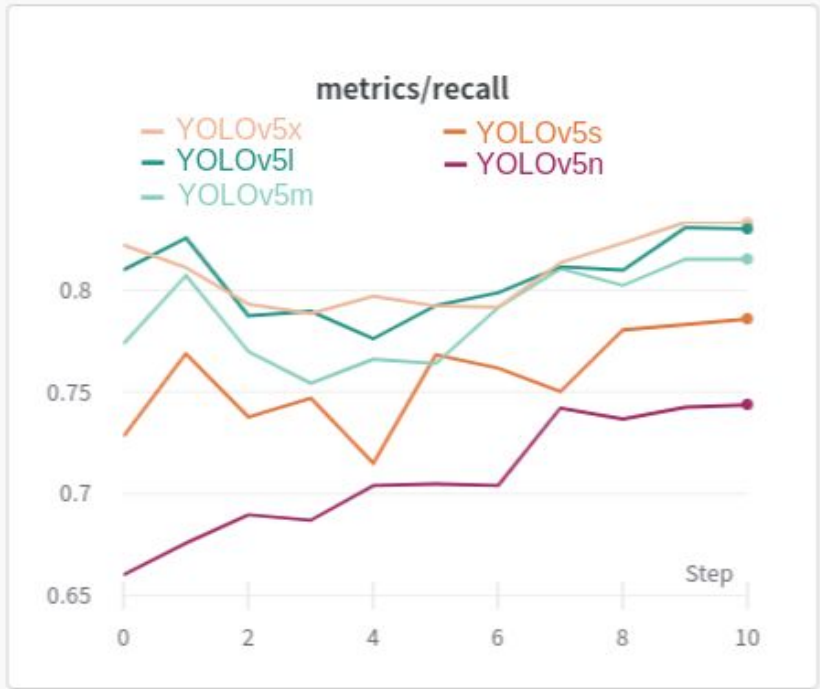
Architecture selection

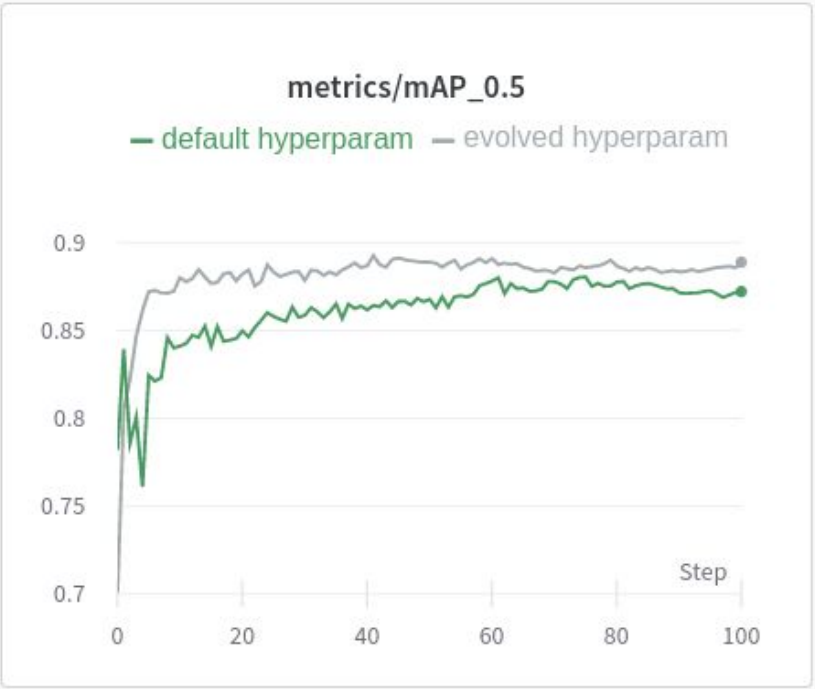
To solve our task YOLOv5 was selected.

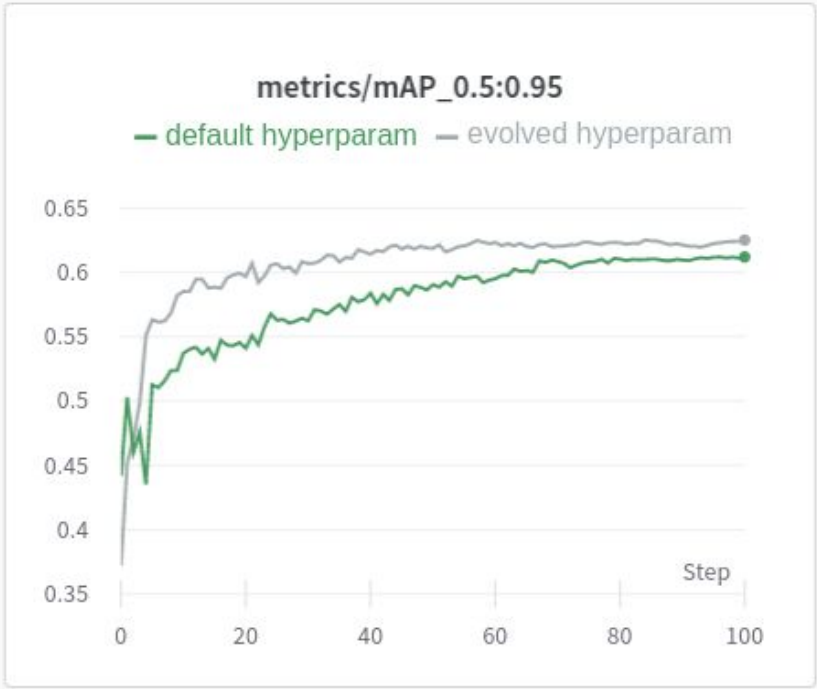
It has architectures of varying complexity:

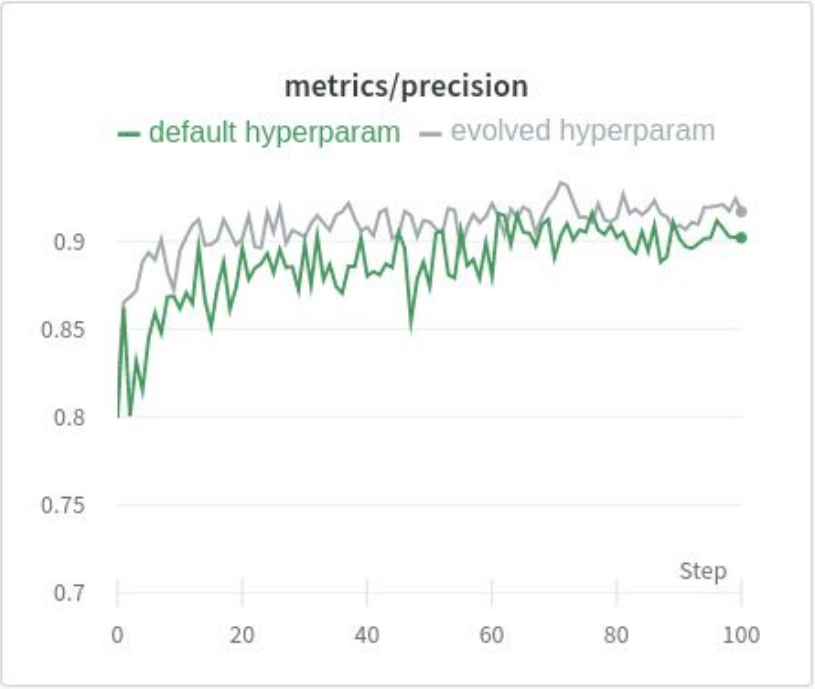
- YOLOv5n - nano;
- YOLOv5s - small;
- YOLOv5m - medium;
- YOLOv5l - large;
- YOLOv5x - XL.

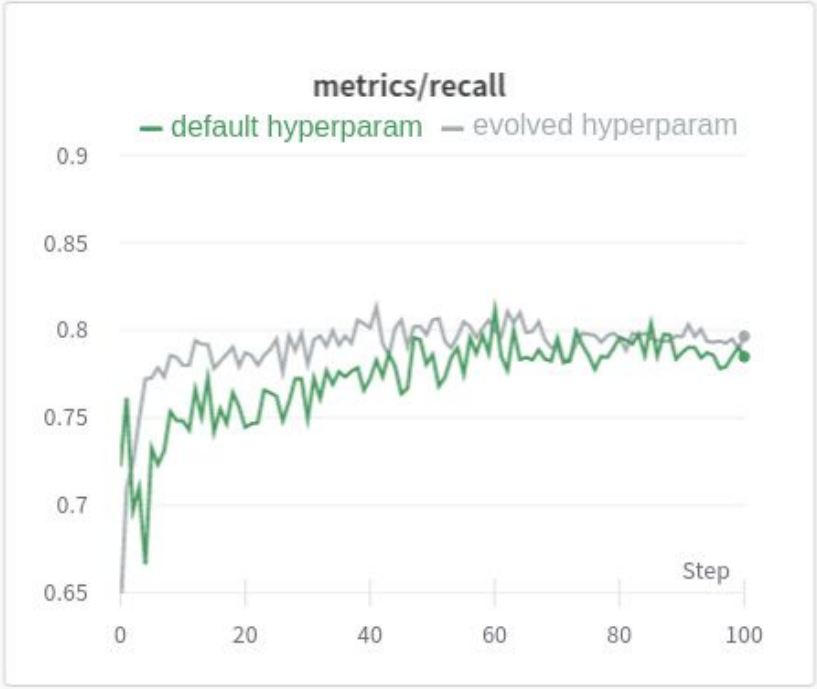








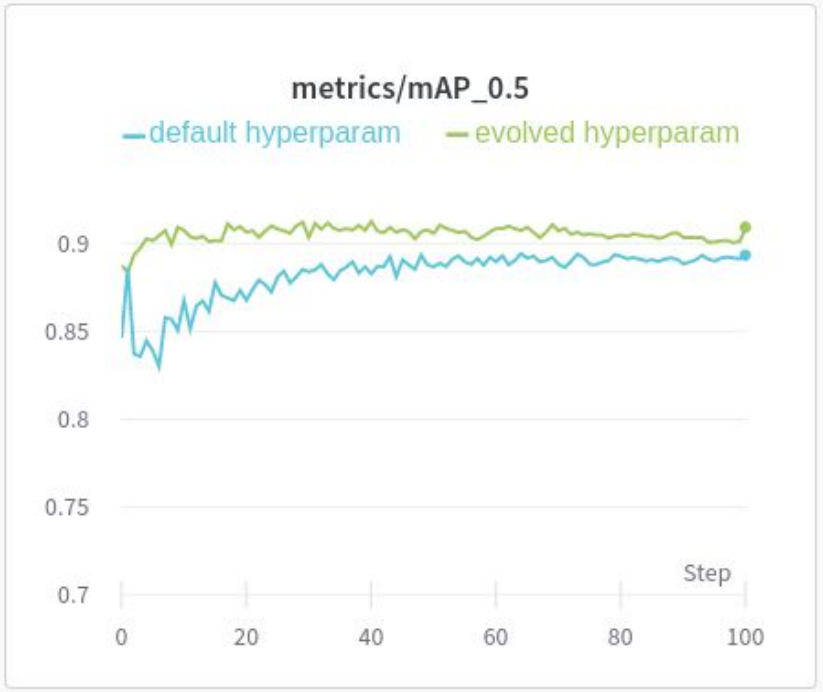


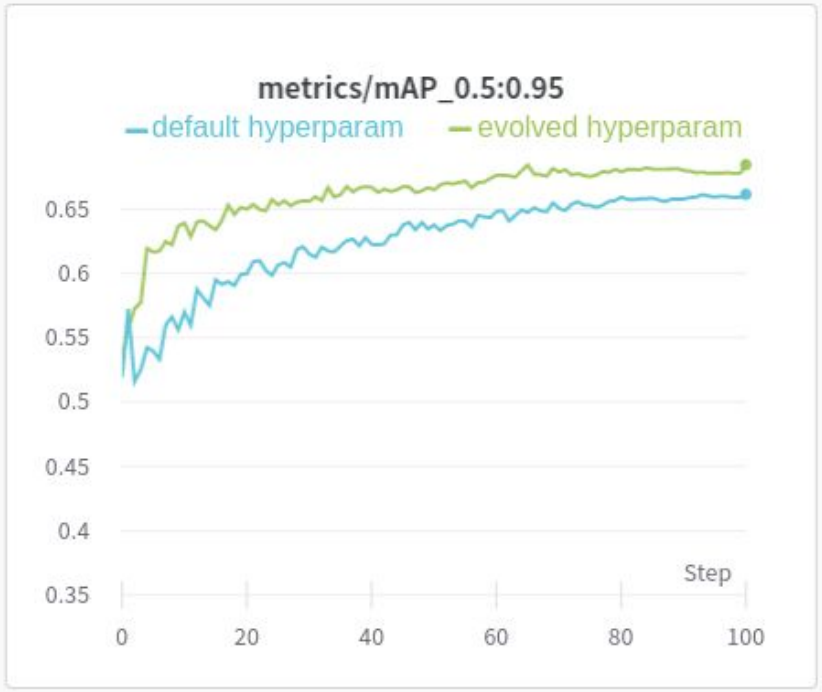


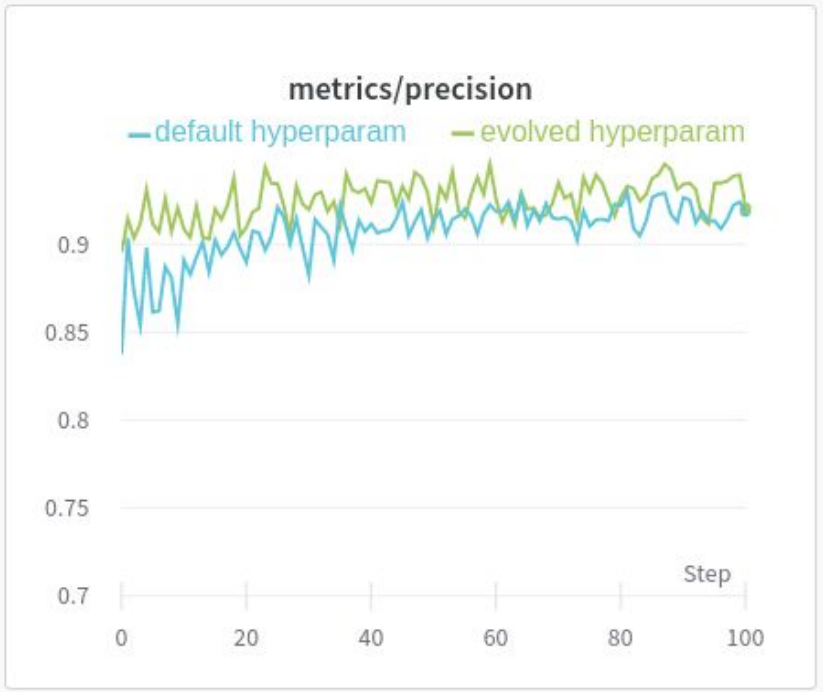
Selection of hyperparameters YOLOv5s

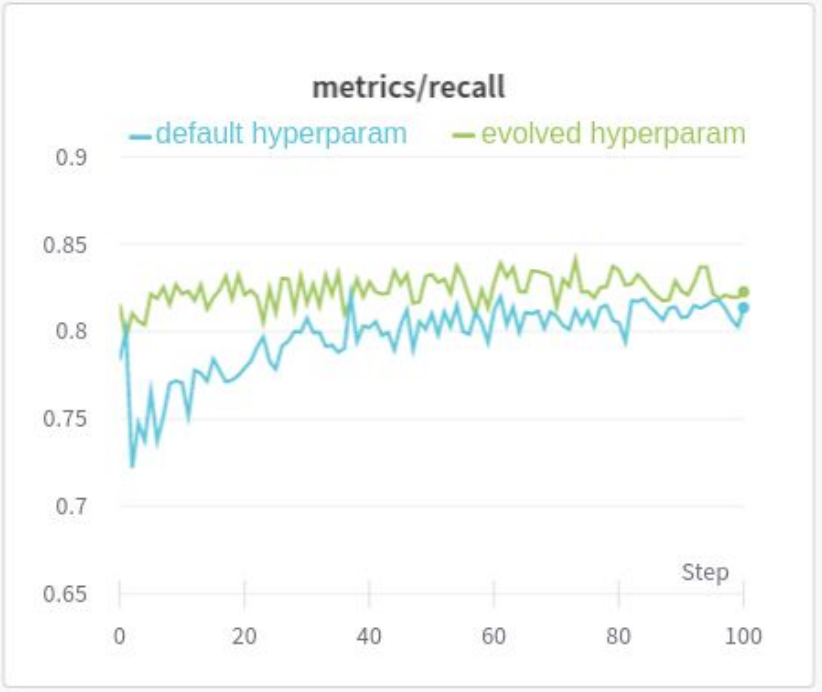
The selection of hyperparameters was implemented using a genetic algorithm.

The graphs show the learning curves of the YOLOv5s network with default hyperparameters and hyperparameters obtained using a genetic algorithm.









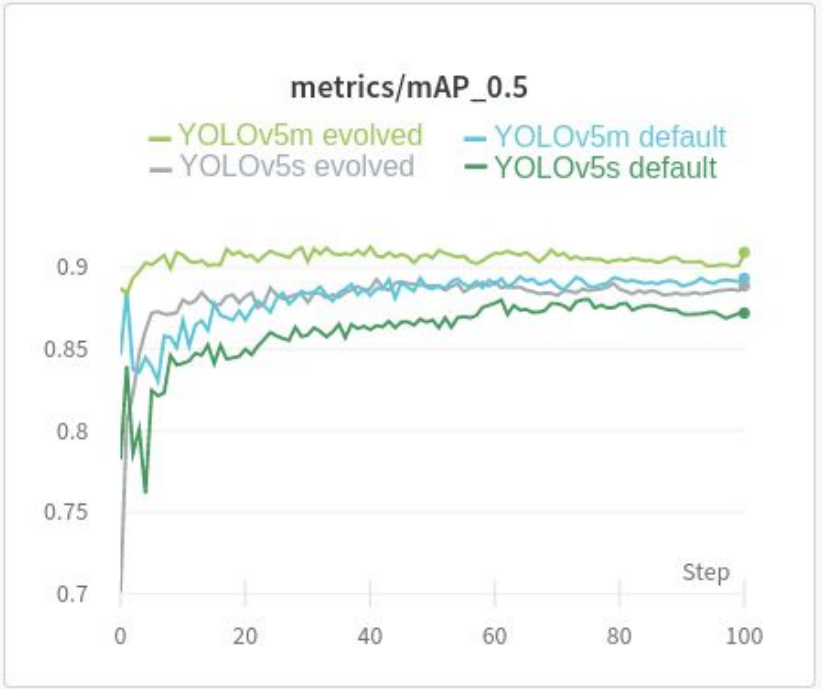
Selection of hyperparameters YOLOv5m

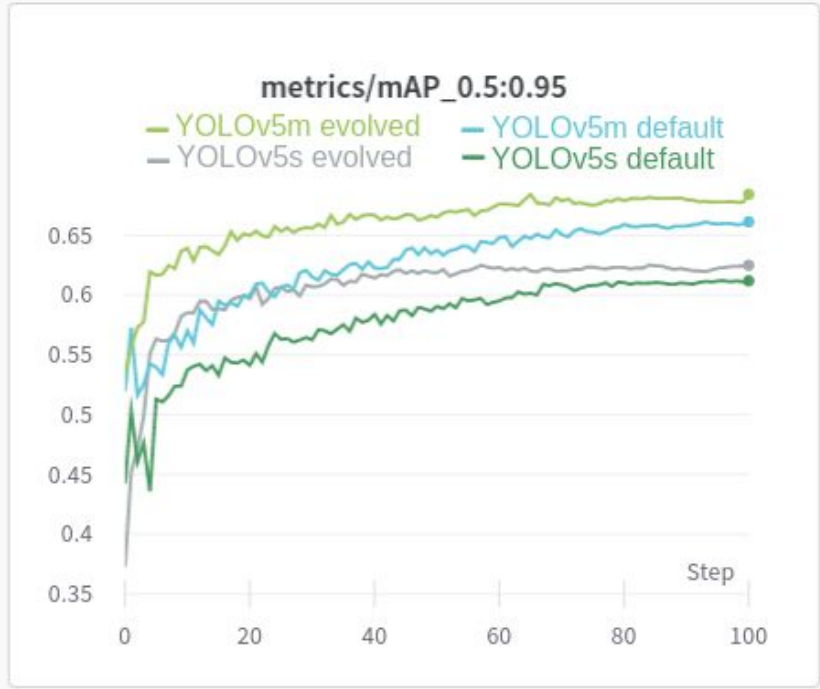
The selection of hyperparameters was implemented using a genetic algorithm.

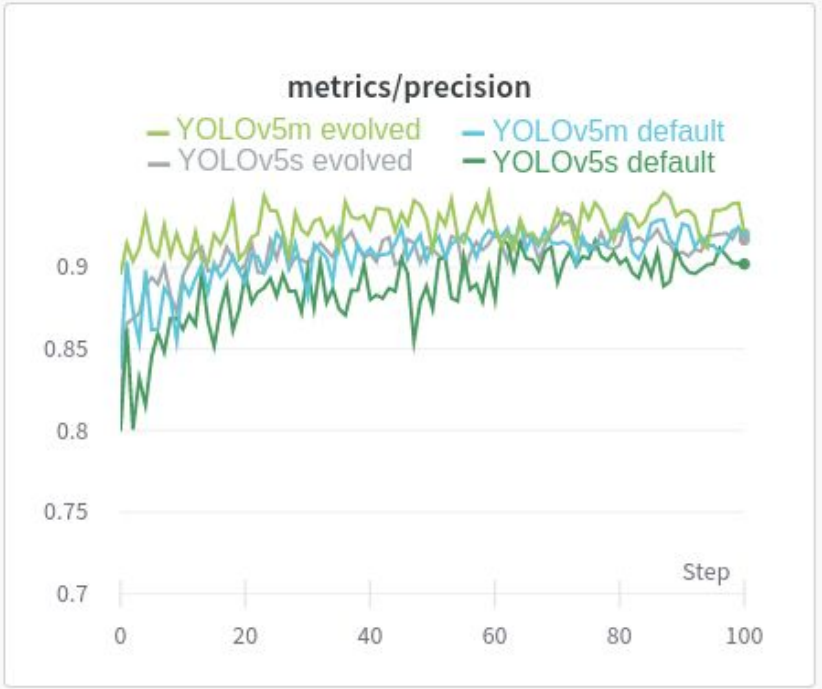
The graphs show the learning curves of the YOLOv5m network with default hyperparameters and hyperparameters obtained using a genetic algorithm.

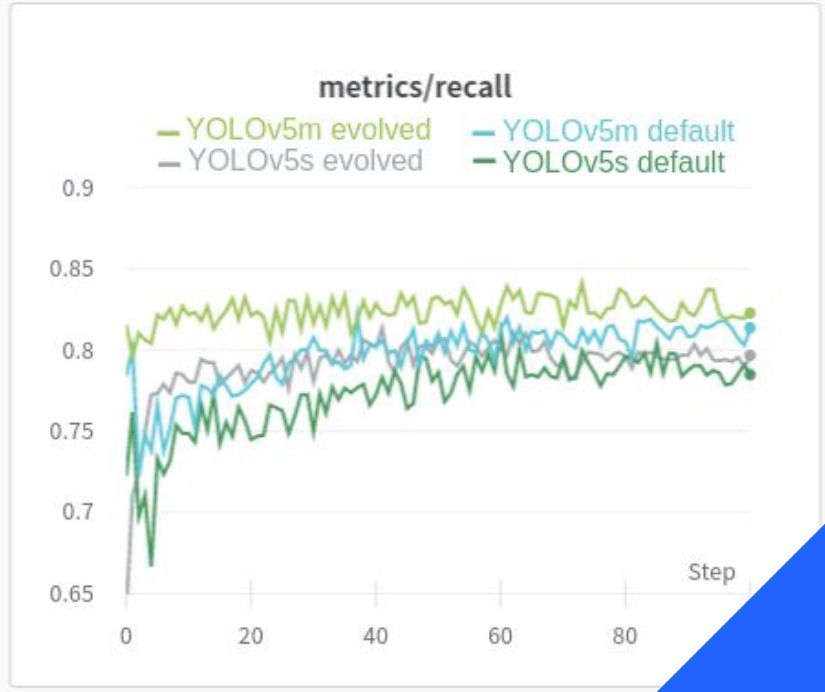
YOLOv5s
YOLOv5m

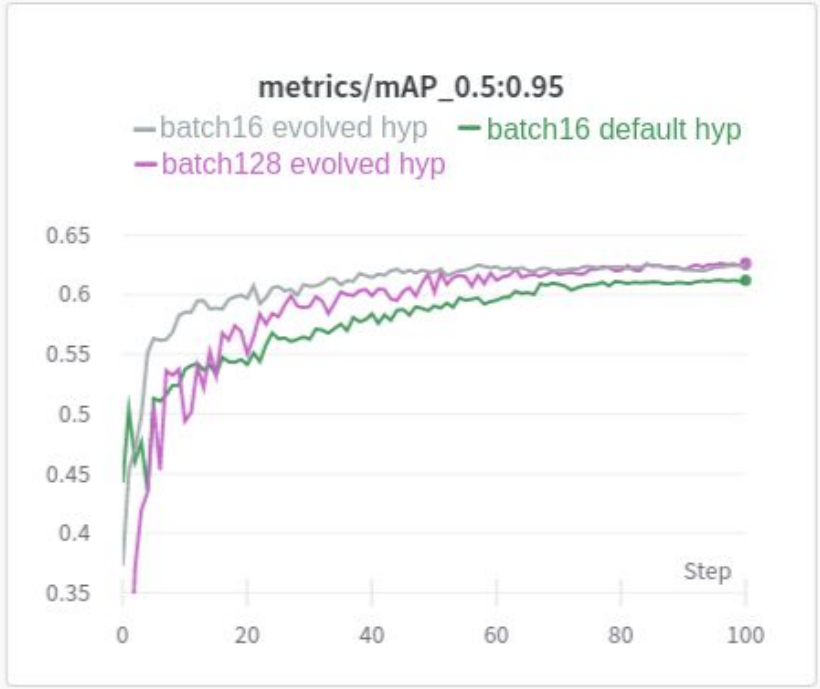
networks with default hyperparameters and hyperparameters obtained using a genetic algorithm.

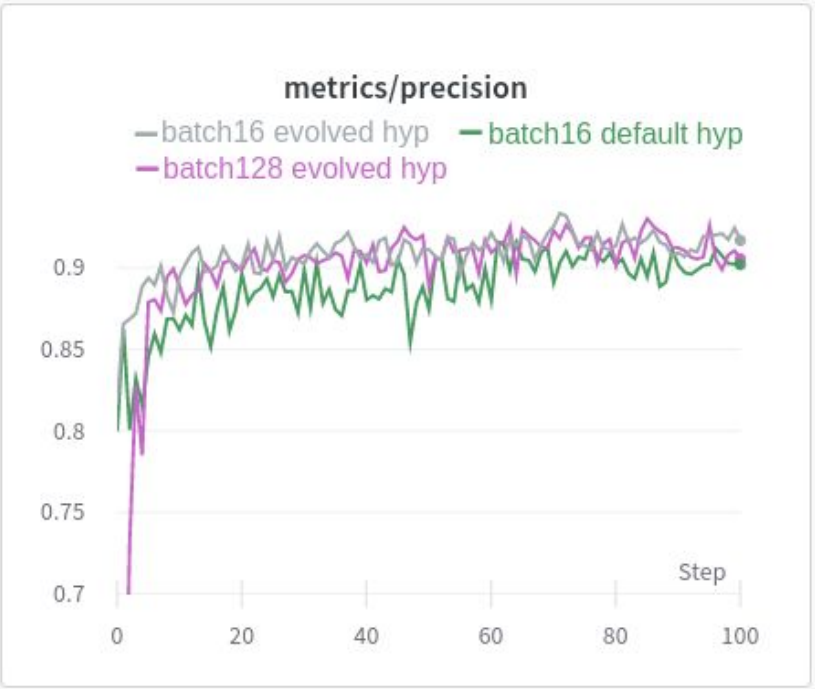


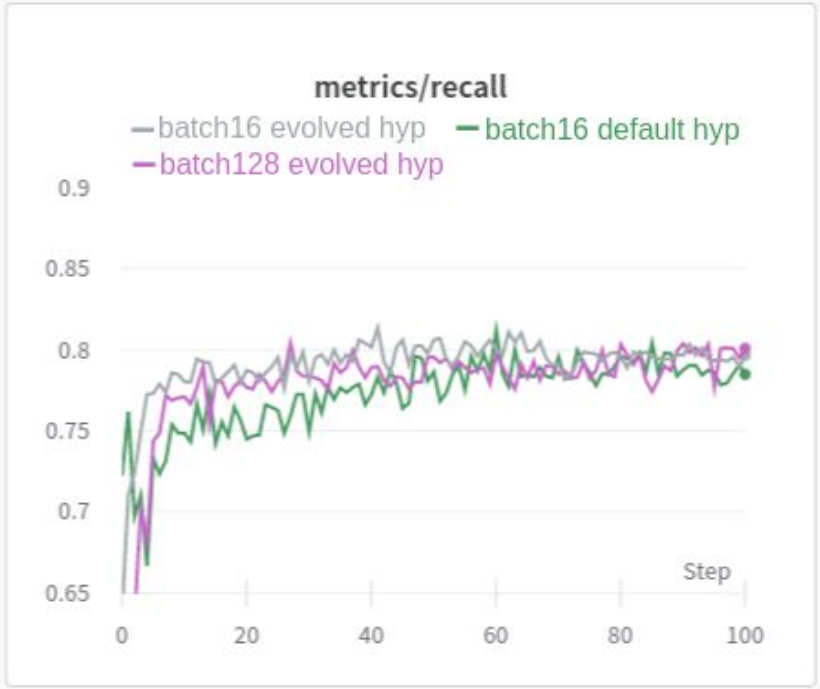






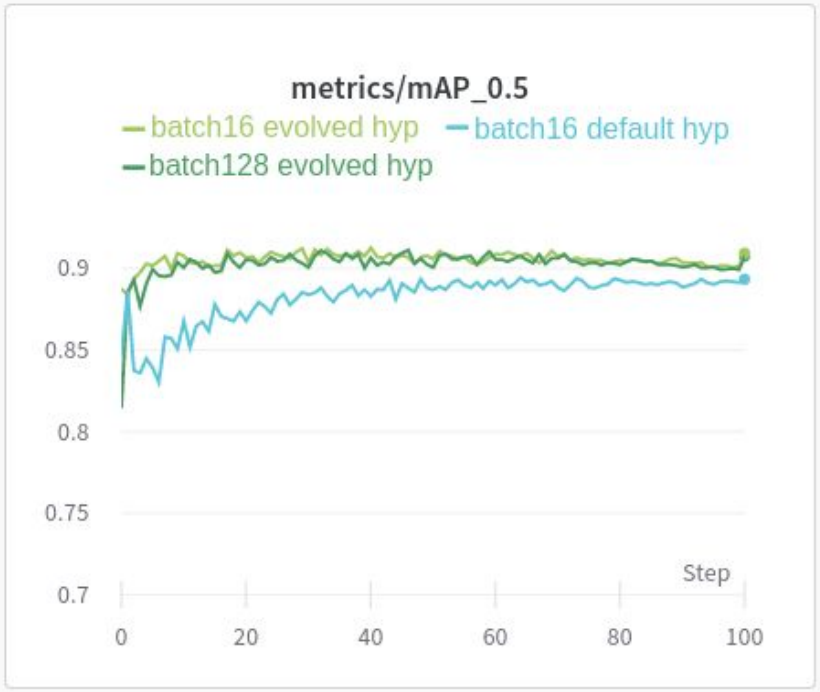


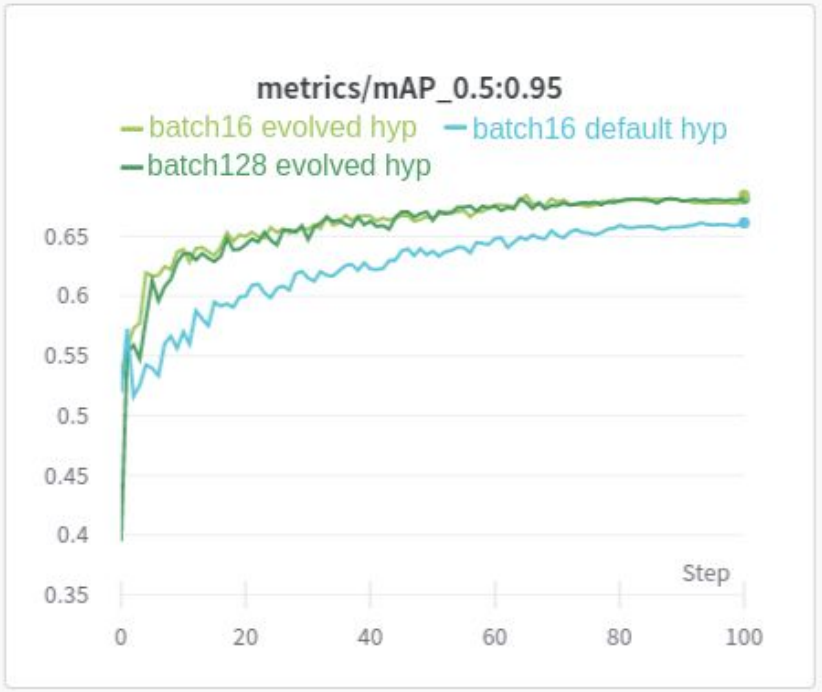


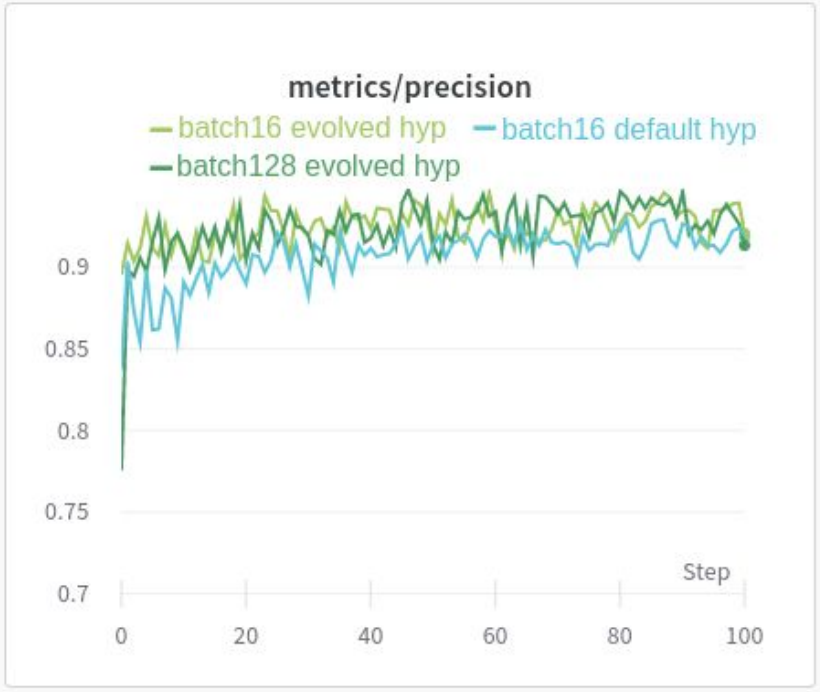


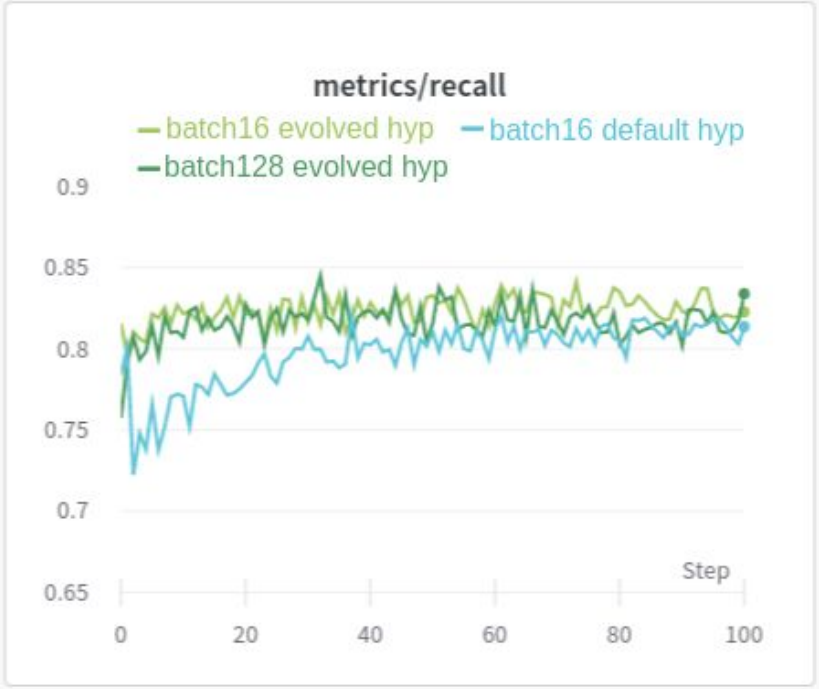
Experiment with batch size YOLOv5s

Increasing the batch size from 16 to 128 did not improve the accuracy of the neural network, but reduced the training time from 110 to 75 minutes.









Experiment with batch size YOLOv5m

Increasing the batch size from 16 to 128 did not improve the accuracy of the neural network, but reduced the training time from 200 to 180 minutes.

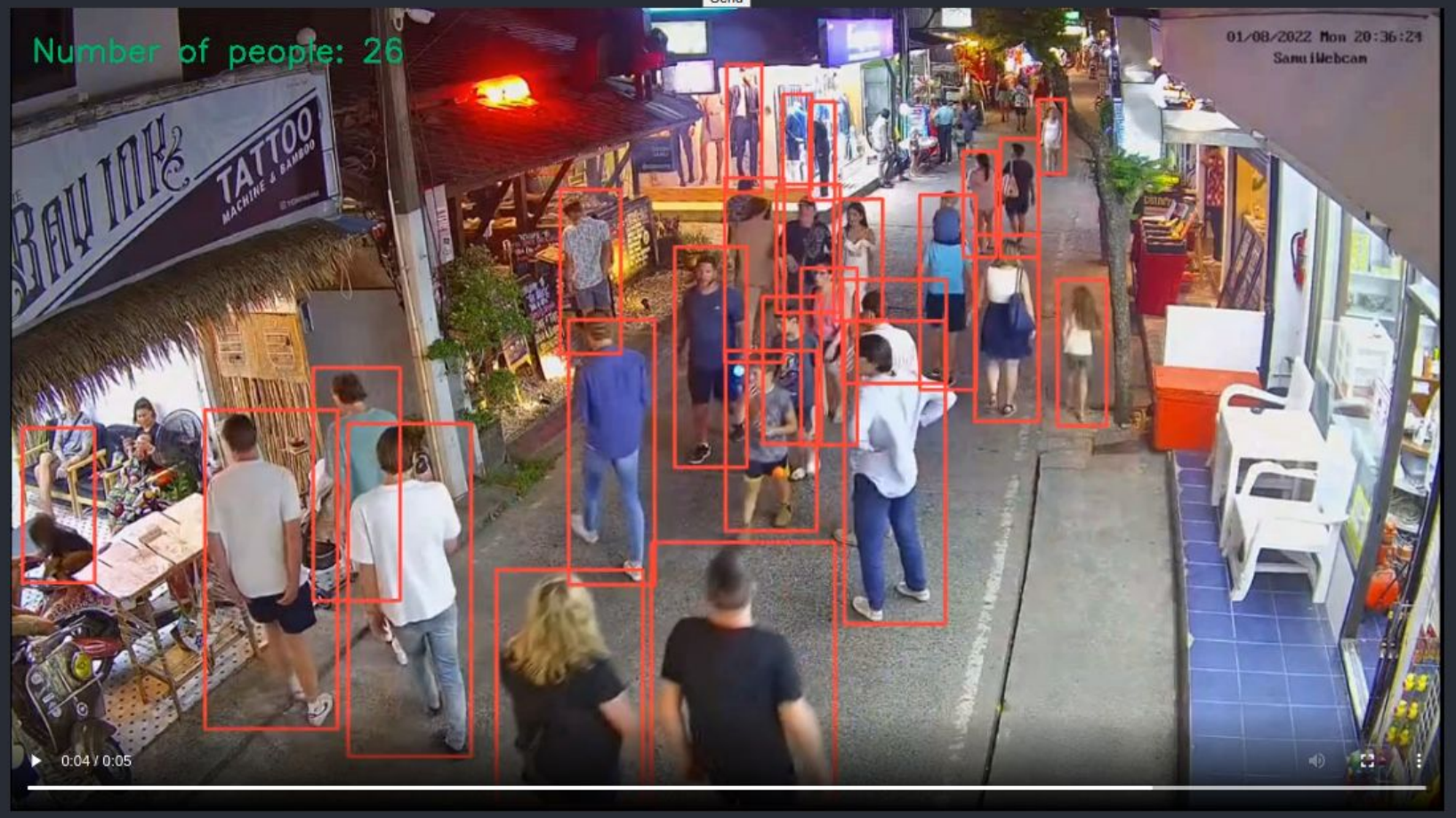
Detection example

YOLO - Detect and count people

Download

Choose File 5sec_Samui.mp4

Send



Production integration

YOLOv5s

Flask

Docker

Recap

The maximum improvement in NN metrics was achieved by improving the quality of the training base (from 0.57 to 0.90).
The selection of hyperparameters also gave some improvement in the metrics (from 0.90 to 0.92).

When training neural networks, a high-quality training base is a paramount factor.
When solving specific practical problems, it is necessary to collect a database that is as relevant to the task as possible (for example, if you need to calculate traffic in a store, take video for marking from surveillance cameras in this store).

UNIVERSITY
OF ARTIFICIAL
INTELLIGENCE

Conclusion

Achieved goals:

- the neural network is trained to detect people;
- the number of people is displayed on the screen;
- 92% accuracy achieved;
- the network is integrated with the Flask application and placed in a Docker container.

The result obtained can be used to count the number of people in stores, in queues, at airports, train stations, etc.