Campus Sontheim

HOCHSCHULE HEILBRONN

TRAFFIC-SIGN-RECOGNITION WITH MACHINE LEARNING

❯ Ozan Anli, Rico Steinke
Computer & Robot Vision / MAS
WS2022/2023

AGENDA

- Introduction to YOLO (You Only Look Once) and its advantages for object detection
- Description of the dataset used for training and testing the model
- Explanation of the training parameters
- Results and performance evaluation of the model on the traffic sign dataset
- Discussion of potential applications and future work for traffic sign recognition using YOLO and other object detection models

INTRODUCTION TO YOLO (YOU ONLY LOOK ONCE) AND ITS ADVANTAGES FOR OBJECT DETECTION

CNN (Convolutional Neural Network)

Feature extraction of input image

- VGG16
- ImageNet
- RetineNet
- Resnet50

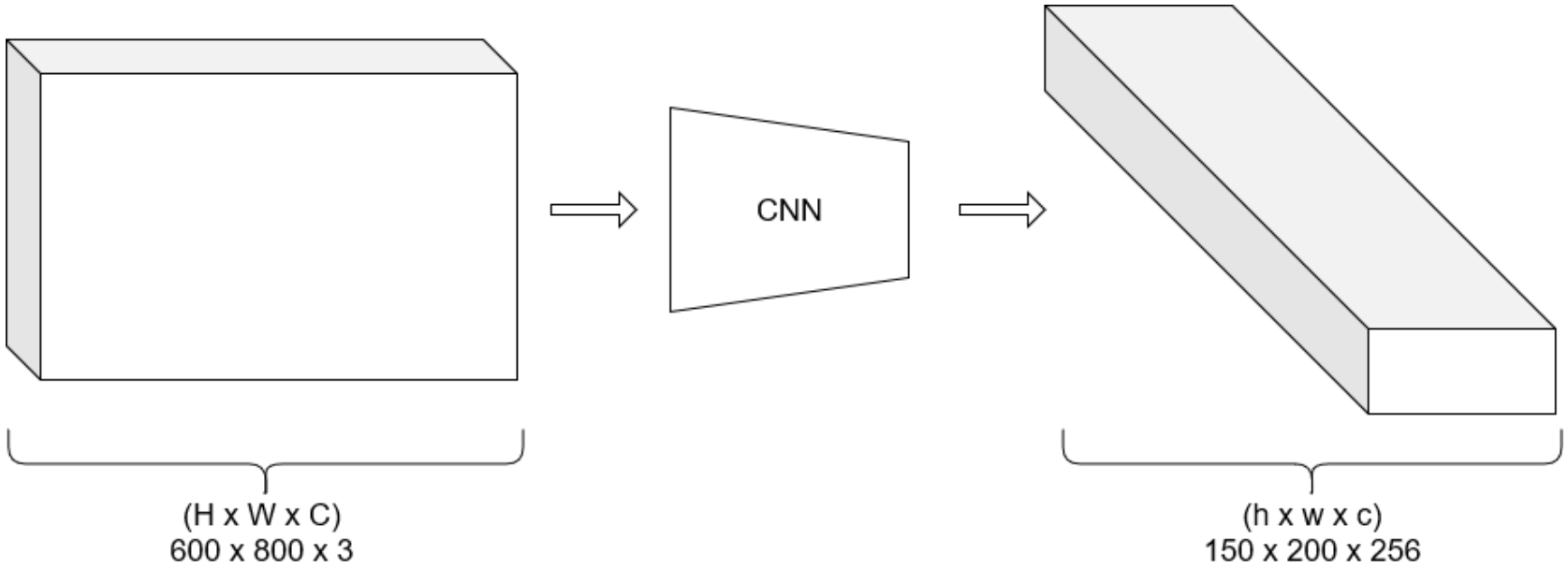


Figure 1: CNN Overview

INTRODUCTION TO YOLO (YOU ONLY LOOK ONCE) AND ITS ADVANTAGES FOR OBJECT DETECTION

FCN (Fully Connected Network)

Computes probability for each anchor box to contain an object and estimate the coordinates of the searched object. Example with 1000 anchor boxes:

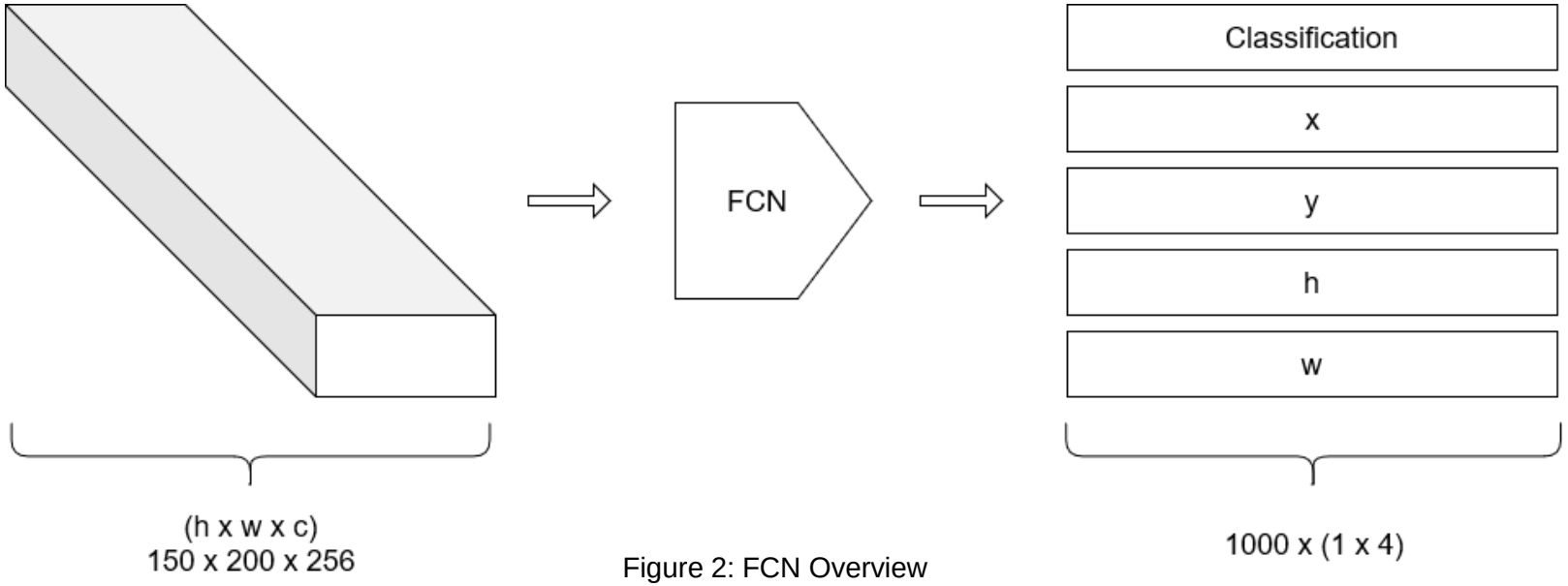


Figure 2: FCN Overview

INTRODUCTION TO YOLO (YOU ONLY LOOK ONCE) AND ITS ADVANTAGES FOR OBJECT DETECTION

NMS (Non-Maximum Suppression)

When a proposal is selected, all proposals that have an intersection over union (IoU) with the selected proposal of more than a certain threshold (usually 0.5) are removed.

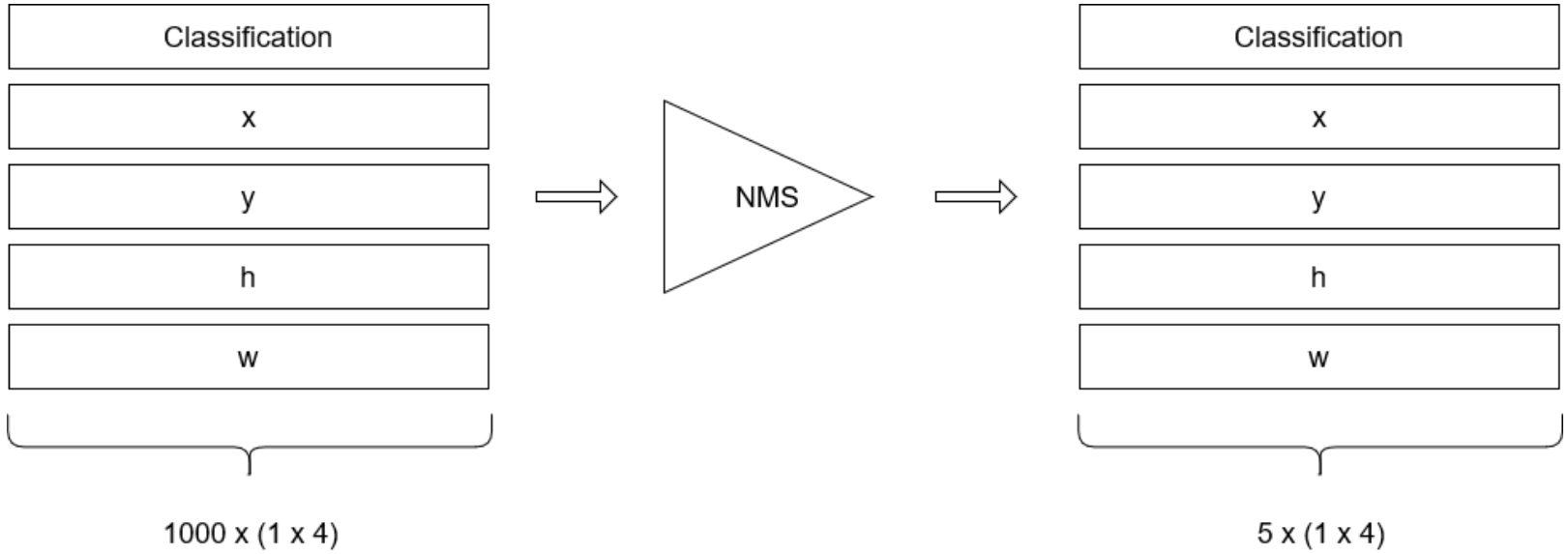


Figure 3: NMS Overview

DESCRIPTION OF THE DATASET USED FOR TRAINING AND TESTING THE MODEL

- Using 47 different traffic signs
- Each sign gets a class id for labeling
- Label-files define:
  - Class id
  - X-position (normalized)
  - Y-position (normalized)
  - Box width (normalized)
  - Box height (normalized)

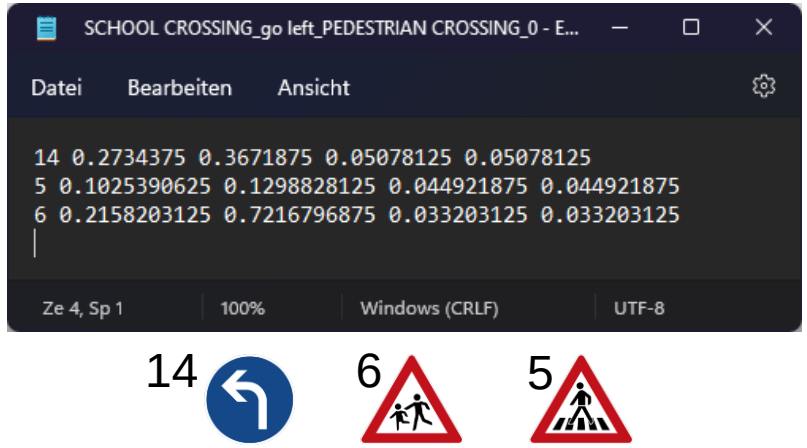


Figure 4: Example of image labeling

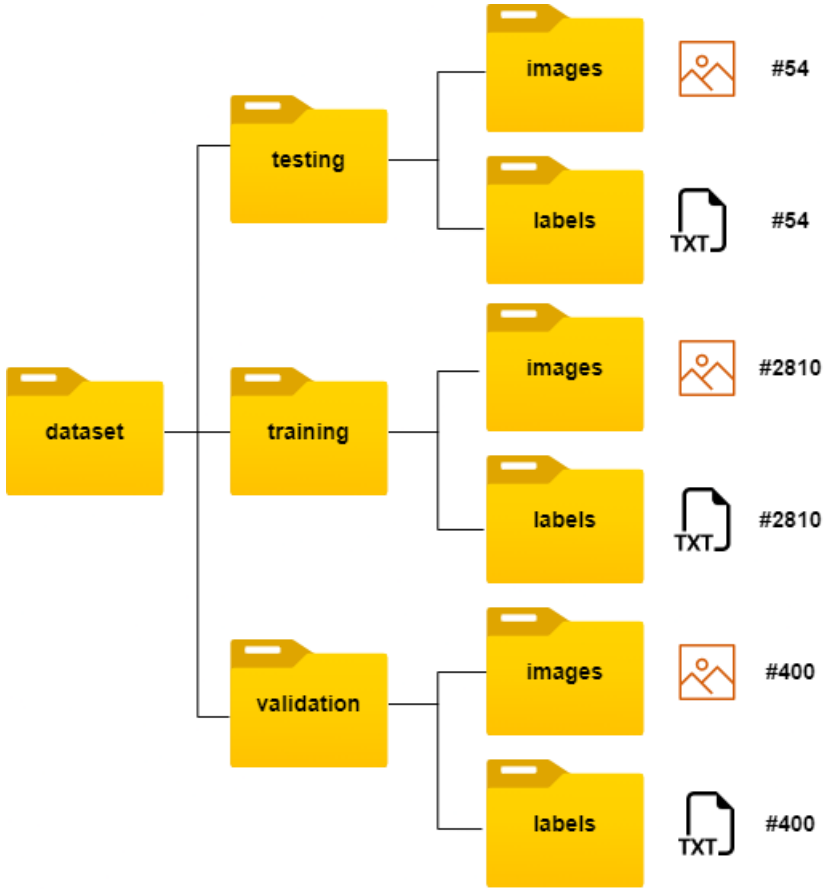


Figure 5: Folder structure of data

DESCRIPTION OF THE DATASET USED FOR TRAINING AND TESTING THE MODEL

Using AI to enhance our dataset

- Generate background images with various traffic environments with DALL●E 2
- Label traffic sign images
- Paste traffic signs in random positions and sizes on background image
- Add augmentations like weather conditions
- Create corresponding label files automatically

Figure 6: Example of AI generated images

EXPLANATION OF THE TRAINING PARAMETERS

- ≥300 Epochs
- Image size: 640x640
- Batch-Size: 4
- Transfer learning approach with pretrained weigths:
  - YoloV7 trained on Coco-Dataset

- Data Augmentation:
  - hsv_h: 0.015 [image HSV-Hue augmentation (fraction)]
  - hsv_s: 0.7 [image HSV-Saturation augmentation (fraction)]
  - hsv_v: 0.4 [image HSV-Value augmentation (fraction)]
  - degrees: 0.15 [image rotation (+/- deg)]
  - translate: 0.2 [image translation (+/- fraction)]
  - scale: 0.5 [image scale (+/- gain)]
  - shear: 0.2 [image shear (+/- deg)]
  - perspective: 0.0 [image perspective (+/- fraction), range 0-0.001]
  - flipud: 0.0 [image flip up-down (probability)]
  - fliplr: 0.0 [image flip left-right (probability)]

That makes no sense in this application! E.g. The directions of the arrows within the signs must point in the right direction.

RESULTS AND PERFORMANCE EVALUATION OF THE MODEL ON THE TRAFFIC SIGN DATASET

Precision: $\frac{TP}{TP+FP} = \frac{TP}{Positive\ results\ total}$

Recall: $\frac{TP}{TP+FN} = \frac{TP}{To\ detect\ total}$

Intersection over union (IoU): $\frac{area\ of\ overlap}{area\ of\ union}$

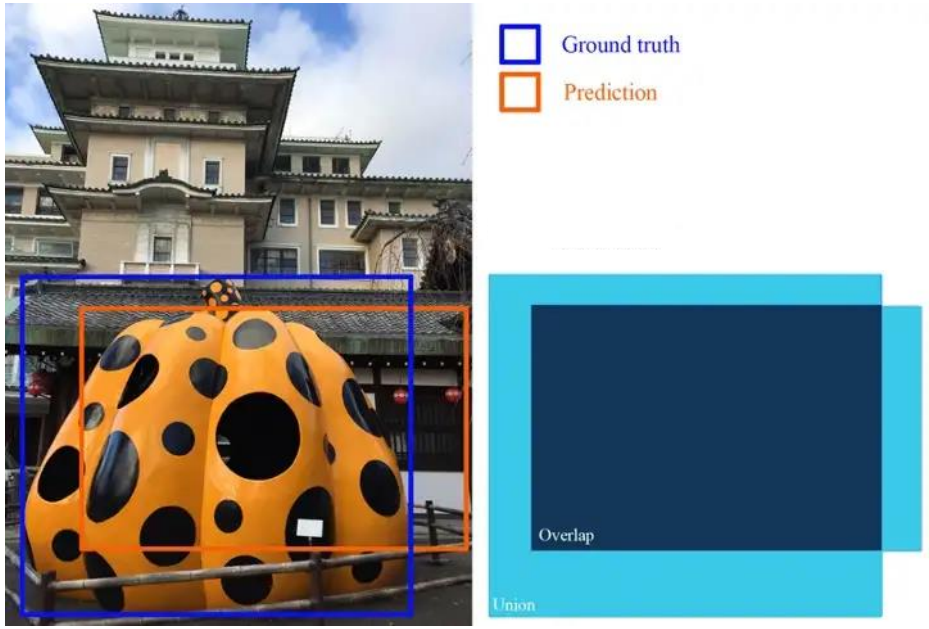


Figure 7: Example for IoU

Mean average precision (mAP): $\frac{1}{\#classes}\sum_{c\ \varepsilon\ classes}\frac{\#TP(c)}{\#TP(c)+\#FP(c)}$

Source: https://jonathan-hui.medium.com/map-mean-average-precision-for-object-detection-45c121a31173

RESULTS AND PERFORMANCE EVALUATION OF THE MODEL ON THE TRAFFIC SIGN DATASET

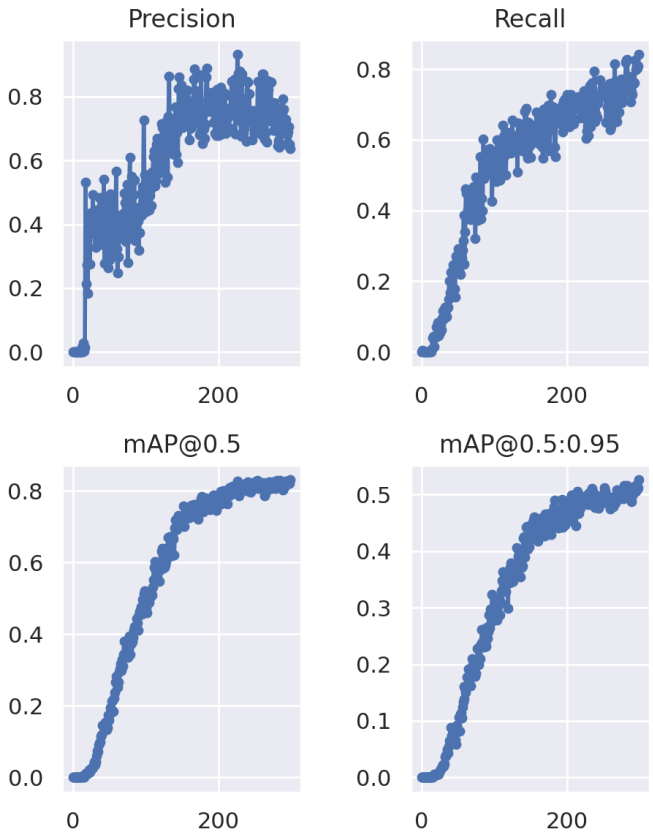


Figure 8: Example of performance evaluation

Precision: $\frac{TP}{TP+FP} = \frac{TP}{Positive\ results\ total}$

Recall: $\frac{TP}{TP+FN} = \frac{TP}{To\ detect\ total}$

Mean Average Precision (mAP): $\frac{1}{\#classes} \sum_{c\ \varepsilon\ classes} \frac{\#TP(c)}{\#TP(c)+\#FP(c)}$

mAP@0.5: IoU threshold ≥ 0.5

mAP@0.5:0.95: IoU threshold ≥0.5:0.05:0.95

RESULTS AND PERFORMANCE EVALUATION OF THE MODEL ON THE TRAFFIC SIGN DATASET

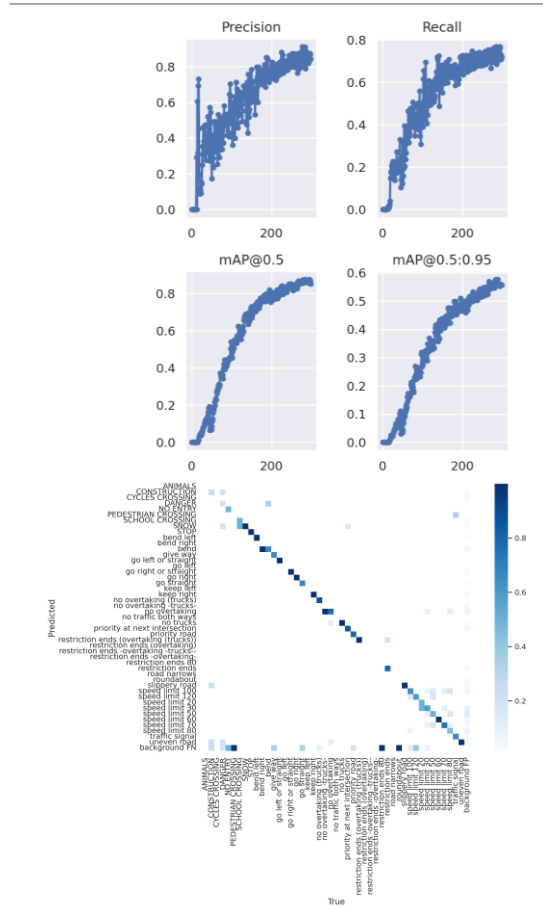


Figure 9: Performance of real-world data only dataset

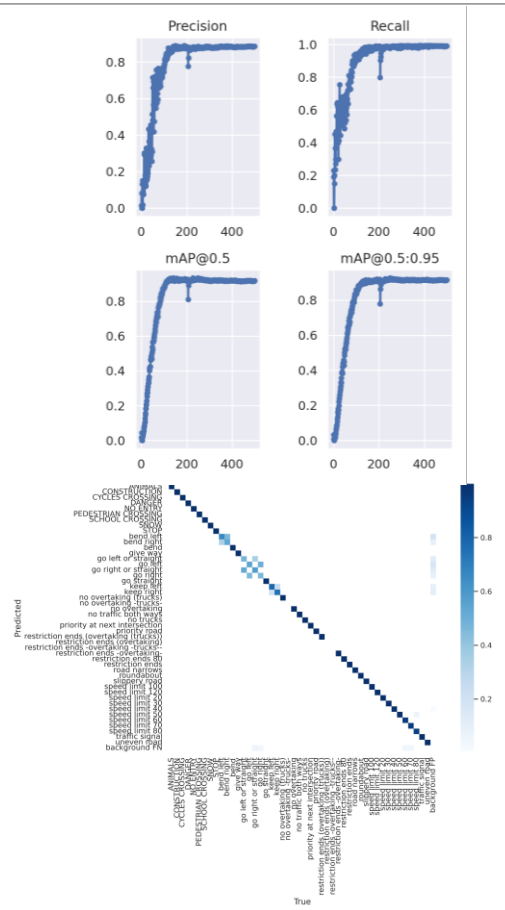


Figure 10: Performance of AI-generated data only dataset

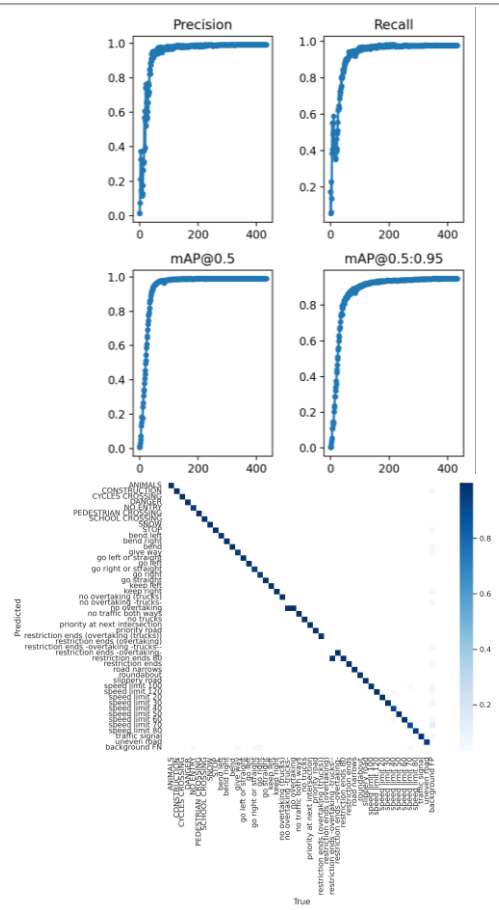
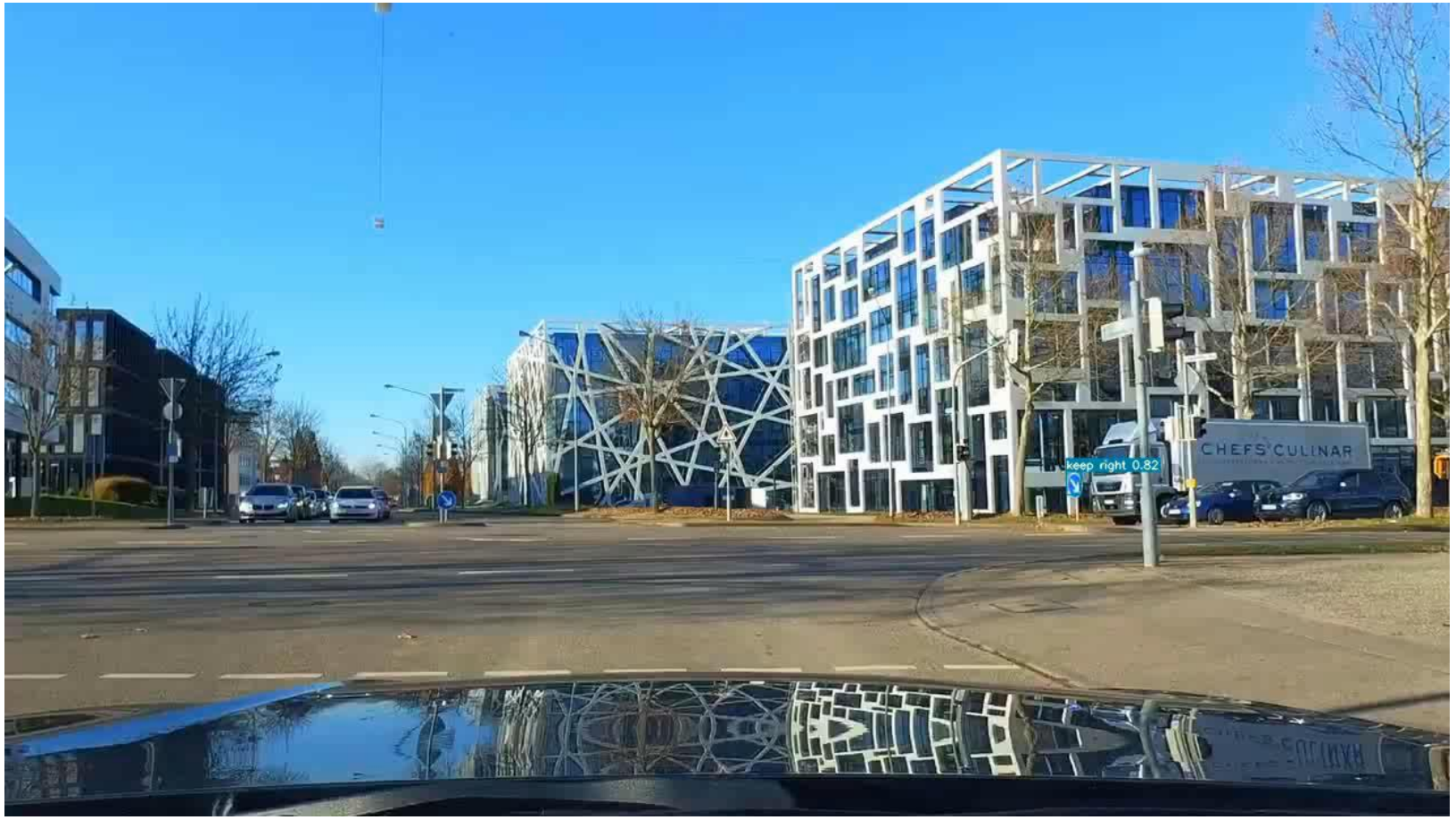


Figure 11: Performance of mixed dataset

RESULTS AND PERFORMANCE EVALUATION OF THE MODEL ON THE TRAFFIC SIGN DATASET



DISCUSSION OF POTENTIAL APPLICATIONS AND FUTURE WORK FOR TRAFFIC SIGN RECOGNITION USING YOLO AND OTHER OBJECT DETECTION MODELS

Improve detection model

- Improve hyperparameters
- Enhance dataset
- Train different object detection models and compare to YOLOv7
- Improve AI generated images
- Add the traffic signs that are not included in the dataset (more classes)

DISCUSSION OF POTENTIAL APPLICATIONS AND FUTURE WORK FOR TRAFFIC SIGN RECOGNITION USING YOLO AND OTHER OBJECT DETECTION MODELS

Potential Applications

- Modell compression with TensorRT

  [https://developer.nvidia.com/tensorrt]

- App development (Android)

  [https://developer.android.com/studio/]

- Model deployment on smartphone

  [https://www.tensorflow.org/lite/android]

→ Improve accuracy, compress model, build an app and deploy the trained model into app.

Result: traffic-sign detecting camera within smartphone!

ANY QUESTIONS?

Ozan Anli
oanli@stud.hs-heilbronn.de

Rico Steinke
rsteinke@stud.hs-heilbronn.de

INTRODUCTION TO YOLO (YOU ONLY LOOK ONCE) AND ITS ADVANTAGES FOR OBJECT DETECTION

CNN

Feature extraction of input image

- VGG16
- ImageNet
- RetineNet
- Resnet50

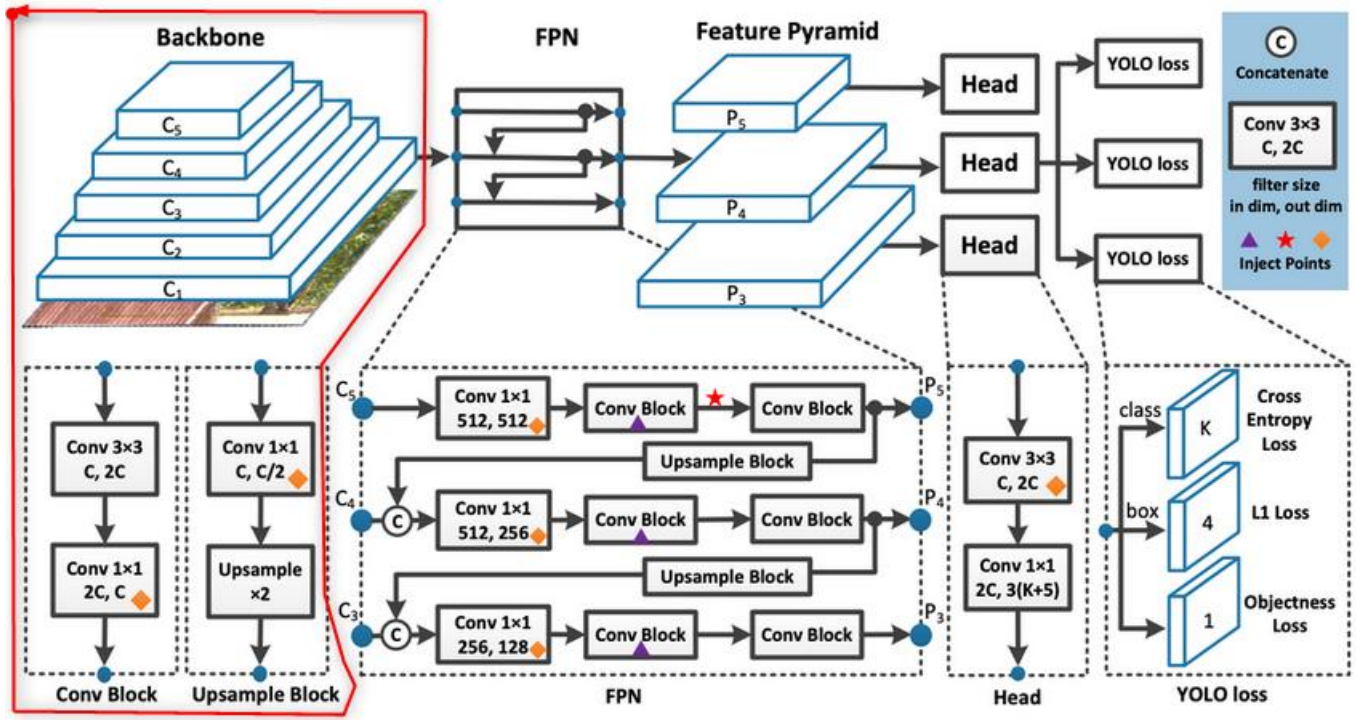


Figure 12: Detailed overview of CNN

INTRODUCTION TO YOLO (YOU ONLY LOOK ONCE) AND ITS ADVANTAGES FOR OBJECT DETECTION

Image features are combined and fused in the Neck

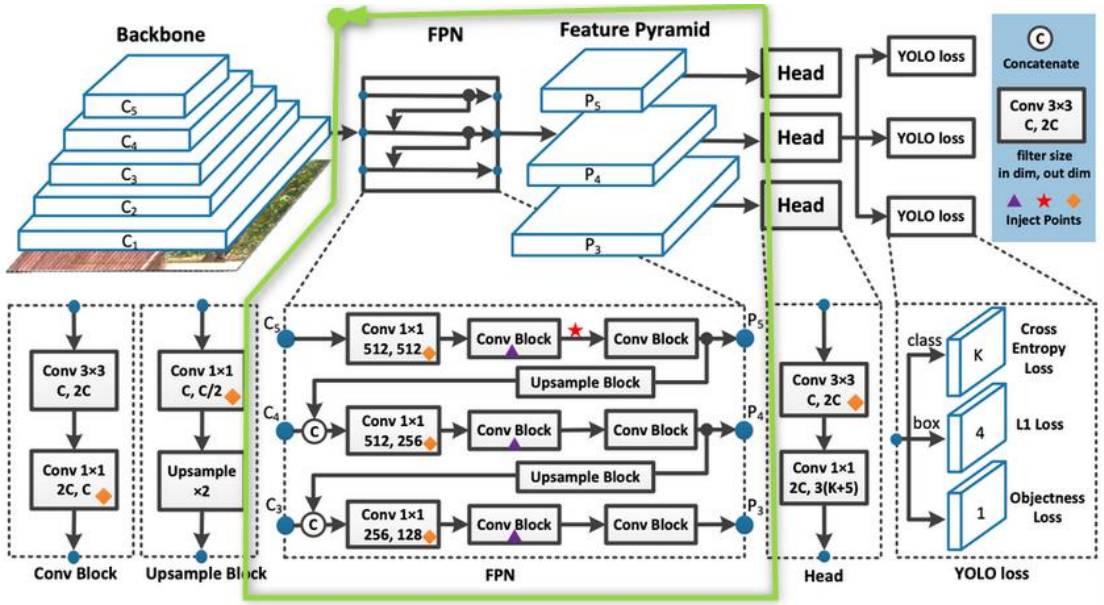


Figure 13: Detailed overview of Neck

Final layers are passed to Head, which makes prediction

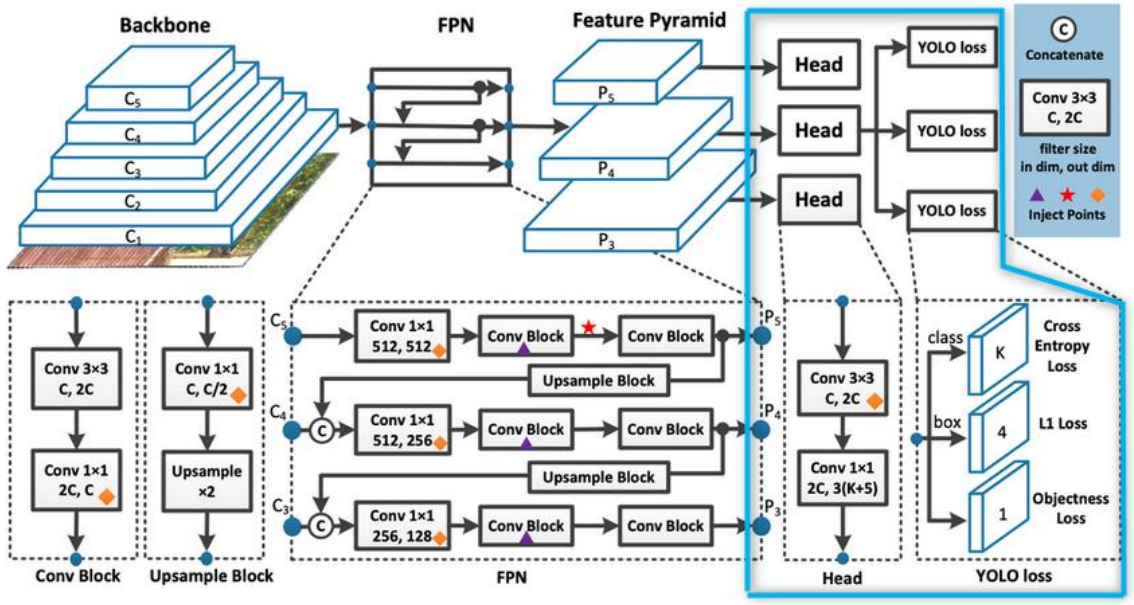


Figure 14: Detailed overview of Head

INTRODUCTION TO YOLO (YOU ONLY LOOK ONCE) AND ITS ADVANTAGES FOR OBJECT DETECTION

NMS

When a proposal is selected, all proposals that have an intersection over union (IoU) with the selected proposal of more than a certain threshold (usually 0.5) are removed.

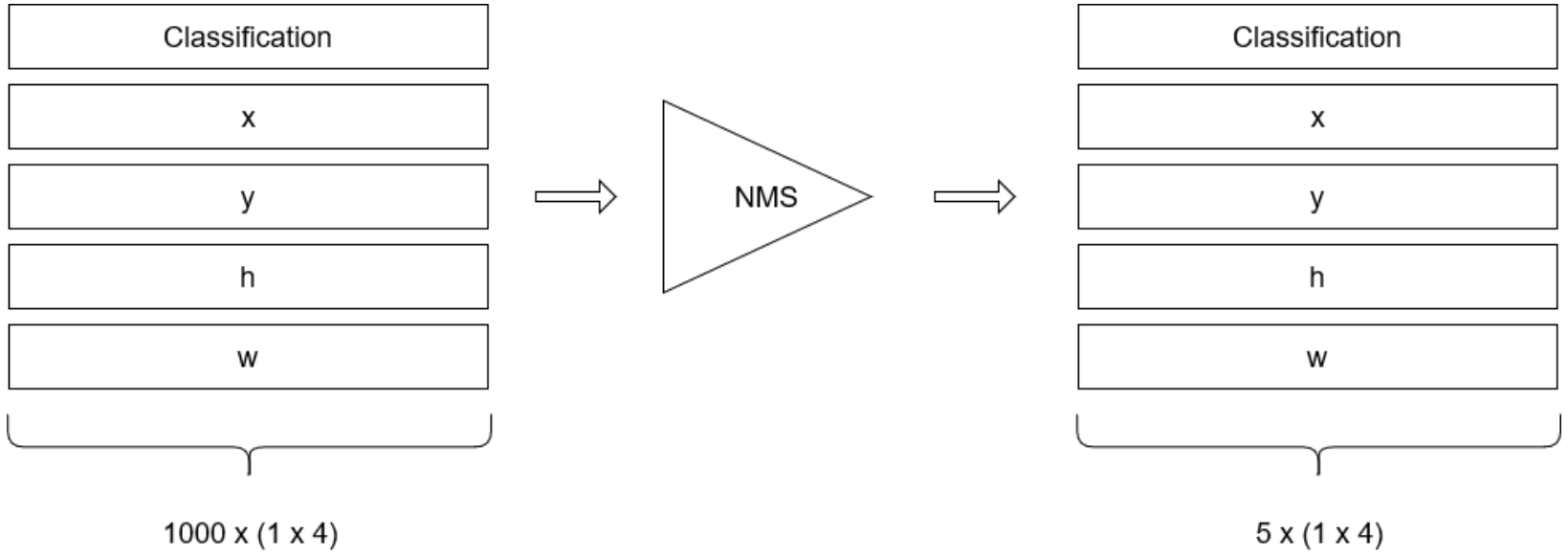


Figure 15: Overview of NMS

DESCRIPTION OF THE DATASET USED FOR TRAINING AND TESTING THE MODEL

Real world data from GTSDB database:
(GTSDB - German Traffic Sign Detection Benchmark)

Roboflow shares pictures of this category.
These images are already labeled.

- Training: 1100 images
- Validation: 108 images
- Testing: 54 images

Same data folder structure as shown as in figure 5.

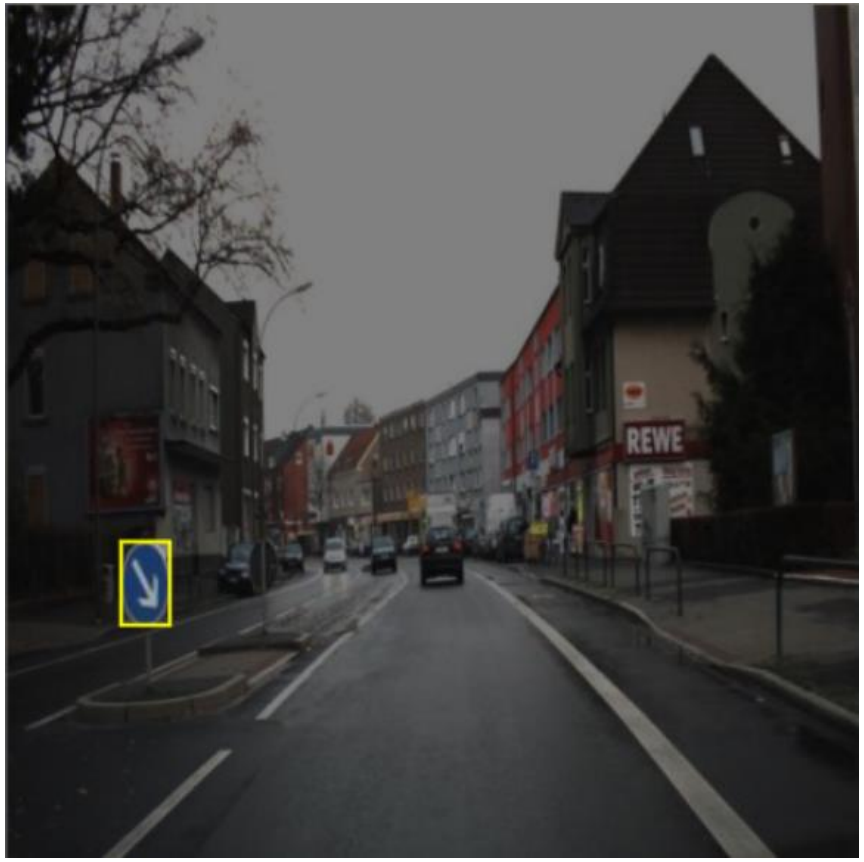


Figure 16: Example of GTSDB image

Source: https://universe.roboflow.com/mohamed-traore-2ekkp/gtsdb--german-traffic-sign-detection-benchmark