# Traffic sign detection and classification (with deep learning)

## Summary

Automatic recognition of traffic signs is important for driver assistance systems, automated driving and inventory purposes. The recognition process is usually composed of two main steps: detection of sign candidate regions and identification of signs. This is the **second part of the project** that has the objective of developing a program for the automatic **detection and classification** of a subset of traffic signs.

## Road Sign Detection Dataset

The Road Sign Detection dataset[1] is composed of 877 images of road signs of 4 different classes, namely: traffic light, stop, speed limit and crosswalk. Each image might have more than one sign, but every image has at least one.

The data has been split into stratified training and test sets (available in Moodle). The training set has 613 images, while the test set has 264. Please do not change these splits.

## General aims

To apply the theoretical knowledge about Deep Learning for Image Analysis acquired in the Computer Vision course, namely, object detection and classification techniques, using the PyTorch library.

## Specific aims

Two tasks are considered, with increasing difficulty and a different weight in the final grade.

1) **Image classification** (70%):

- implement and compare different models based on CNN architectures for classification, considering all 4 classes of the dataset
- evaluate the performance on the test set using adequate metrics and compare the results obtained by the models
- basic version (2 models): use one well-known architecture (VGG, ResNet, etc) and compare its performance when trained from scratch and using pre-trained weights + fine-tuning; consider only a multiclass problem, where the ground-truth label for each image corresponds to the class of the sign/object with the biggest area
- intermediate version (+1 model): additionally, design, implement and compare a custom architecture to solve the multiclass problem
- advanced version (+3 models): adapt the previous models to solve the original problem, i.e. multilabel classification, and compare their performance

2) **Object detection** (30%)

- implement an object detection model and evaluate its performance on the test set using adequate metrics
- basic version: consider a two-stage object detection architecture
- advanced version: additionally, consider a single-stage object detection architecture

Note that the grading of the project will not be defined by the system's performance but rather by the correctness of the adopted methodologies.

[1] https://www.kaggle.com/datasets/andrewmvd/road-sign-detection

## Project development, report and delivery

The work must be done by groups of 3 students (the same groups as in the 1st part of the project).

A short report (max. 3 pages) must be delivered, including:

- a short description of the methods/architectures implemented;
- any data processing steps performed (including data augmentation);
- relevant hyperparameters related to the training of the models;
- relevant comments about the efficacy of the used methods, describing the main problems that were encountered and any proposed solutions;
- the status of the proposed methods and the degree of fulfillment of the aims;
- an analysis of performance of the proposed methods, illustrated with some examples of the obtained results and a comparison with the ground-truth.

The report can be written in English or Portuguese and should use the template available in Moodle. Annexes may be included to show additional results that do not fit in the main report.

The code, with meaningful comments and processing examples, must be submitted in Jupyter notebook format.

The work must be submitted at the Computer Vision page in Moodle, **until the end of June 9, 2022**.

A final presentation including both parts of the project will be done by the end of the semester.

## Bibliography


- Imagenet classification with deep convolutional neural networks, A. Krizhevsky, I. Sutskever, and G. E. Hinton. Advances in neural information processing systems, pp. 1097-1105. 2012.
- Very deep convolutional networks for large-scale image recognition, K. Simonyan and A. Zisserman. arXiv preprint arXiv:1409.1556. 2014
- Deep residual learning for image recognition, K. He, X. Zhang, S. Ren and J. Sun. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 770-778. 2016.
- Faster R-CNN: Towards real-time object detection with region proposal networks. S. Ren, K. He, R. Girshick, and J. Sun. In Advances in neural information processing systems, pp. 91-99. 2015.
- You only look once: Unified, real-time object detection. J. Redmon, D. Santosh, R . Girshick, and A. Farhadi.. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788. 2016.