Improving Cascade of Classifiers by Sliding Window Alignment

Karel Zimmermann, David Hurych, Tomáš Svoboda
Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Cybernetics

Overview

Overview

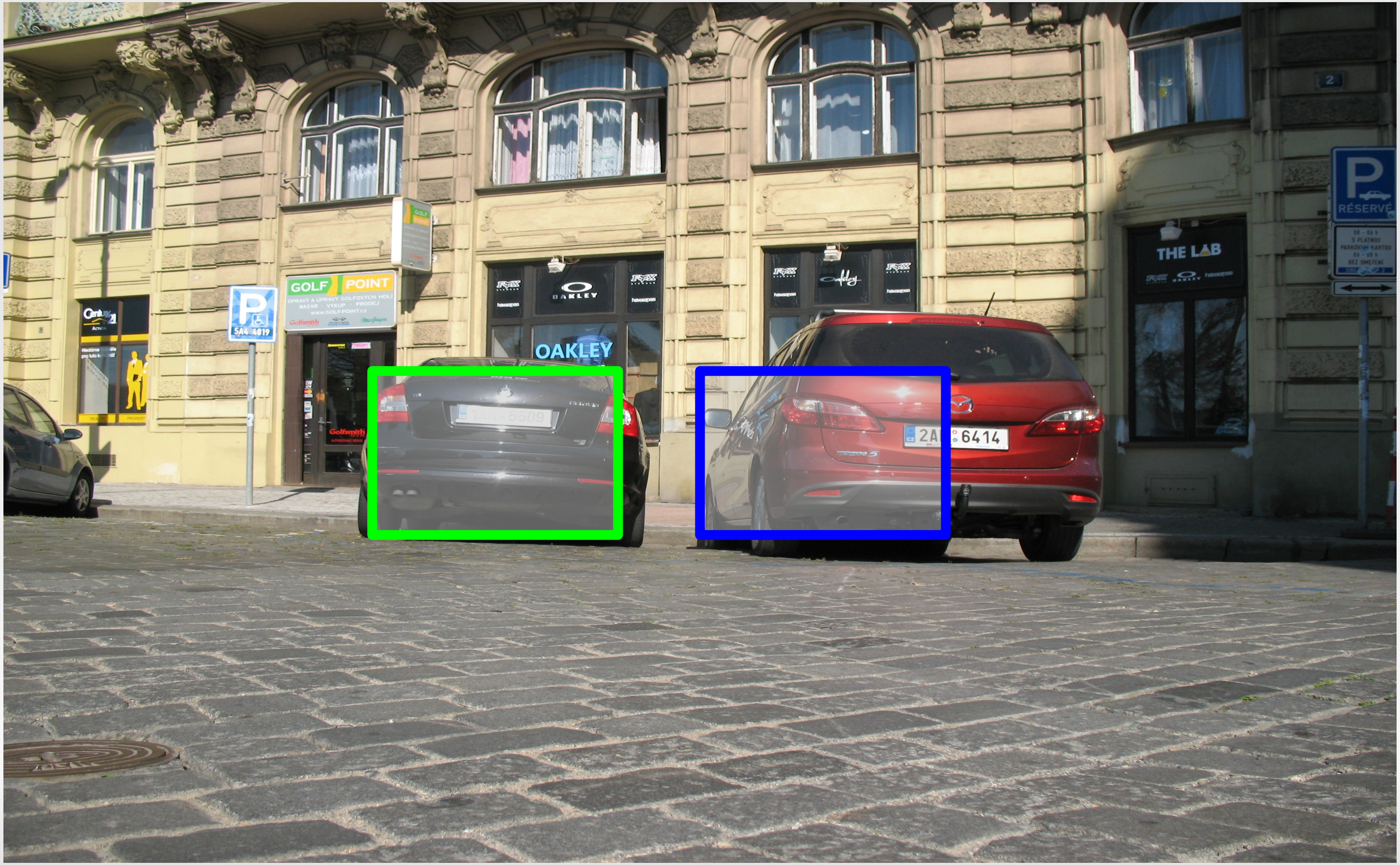
Overview

Overview

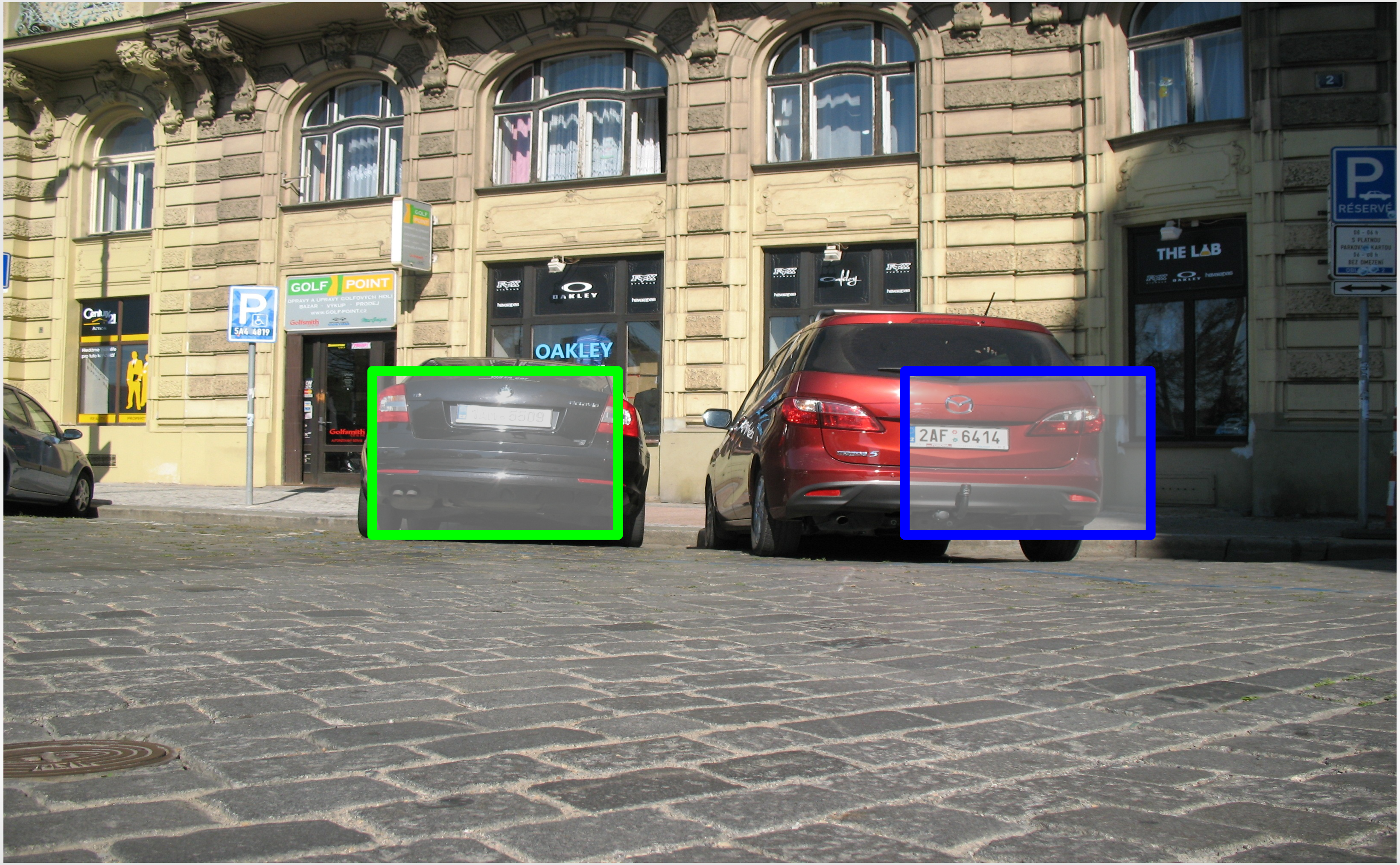
Overview

Overview

Overview

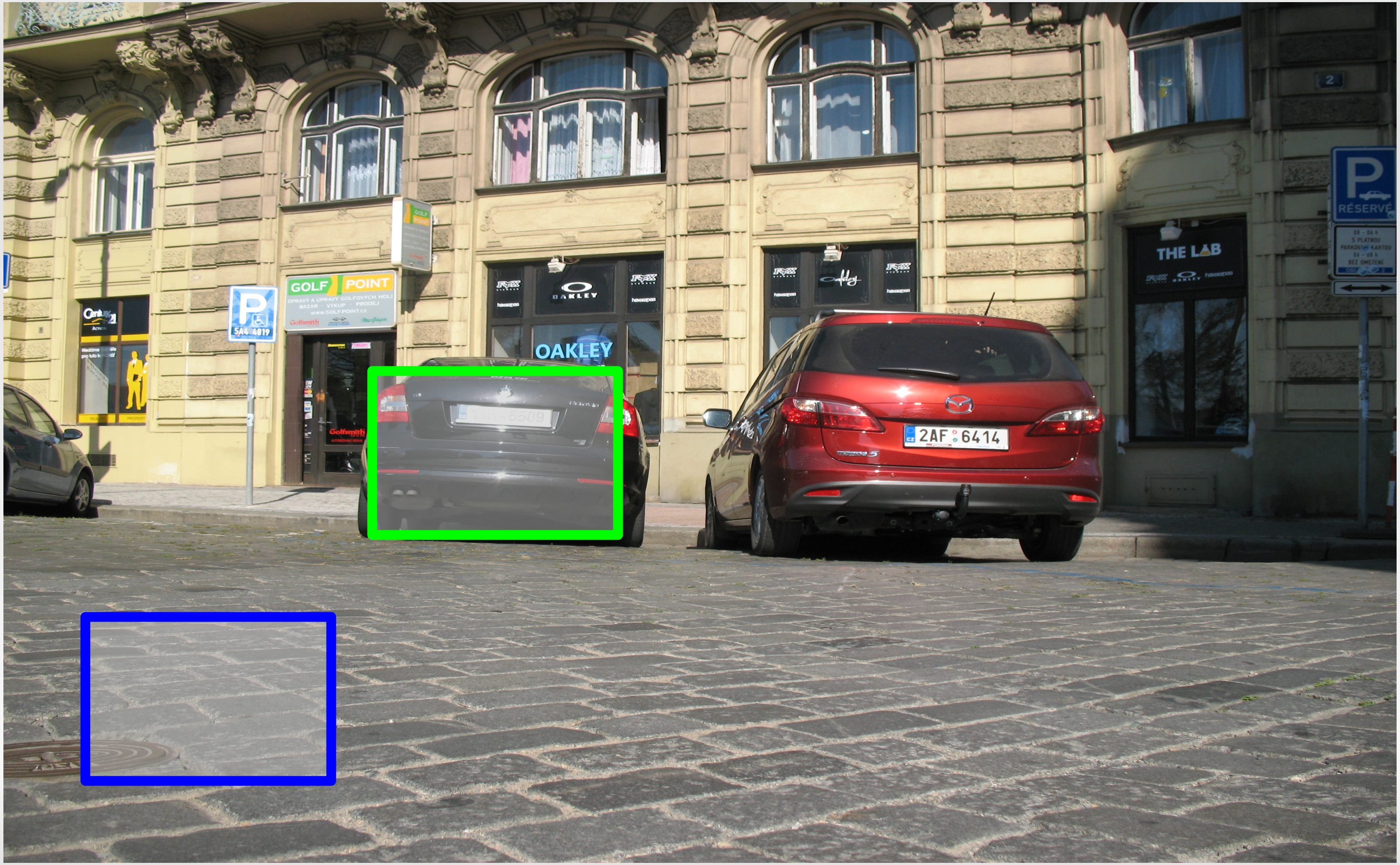


Overview

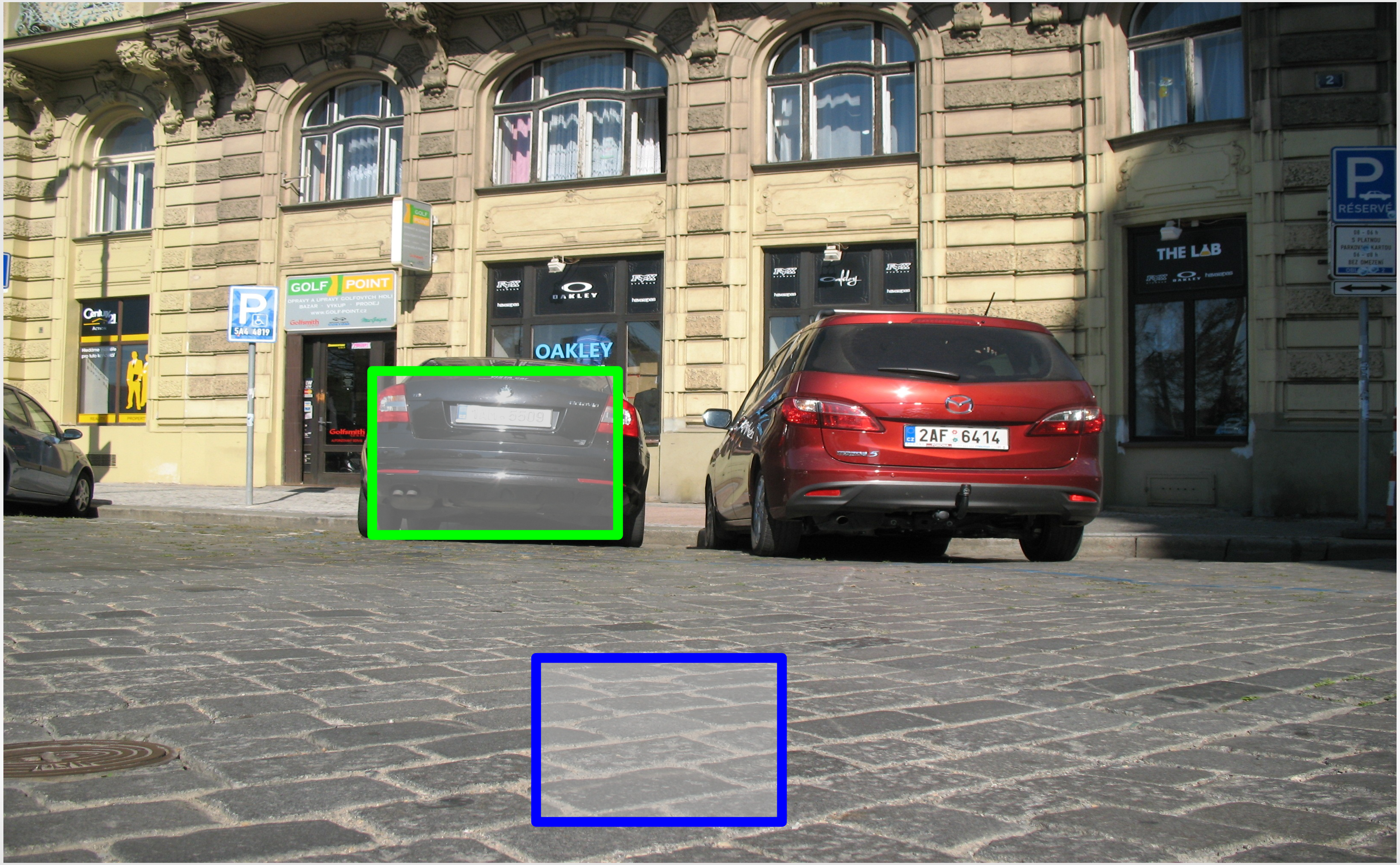


Overview

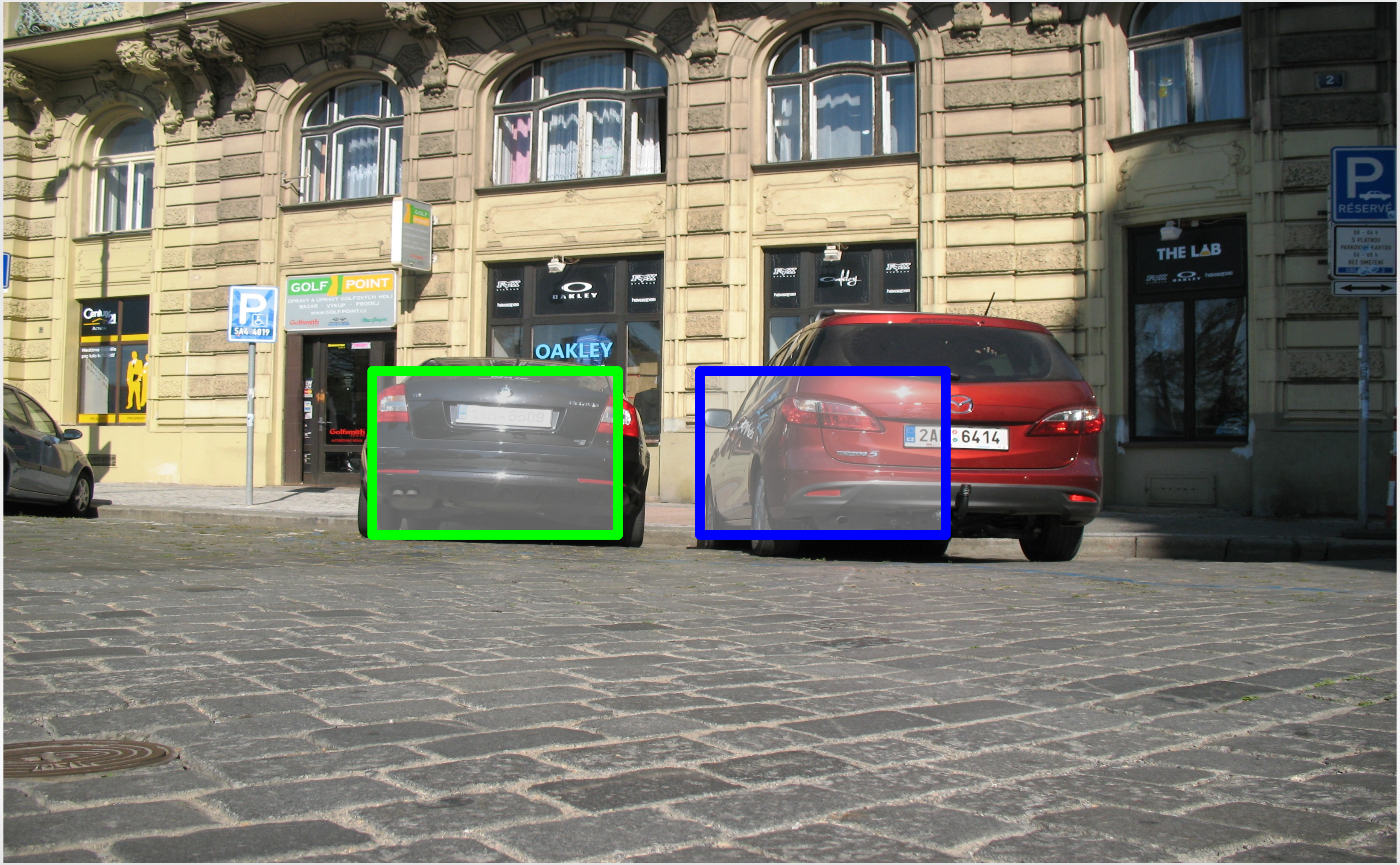
Overview



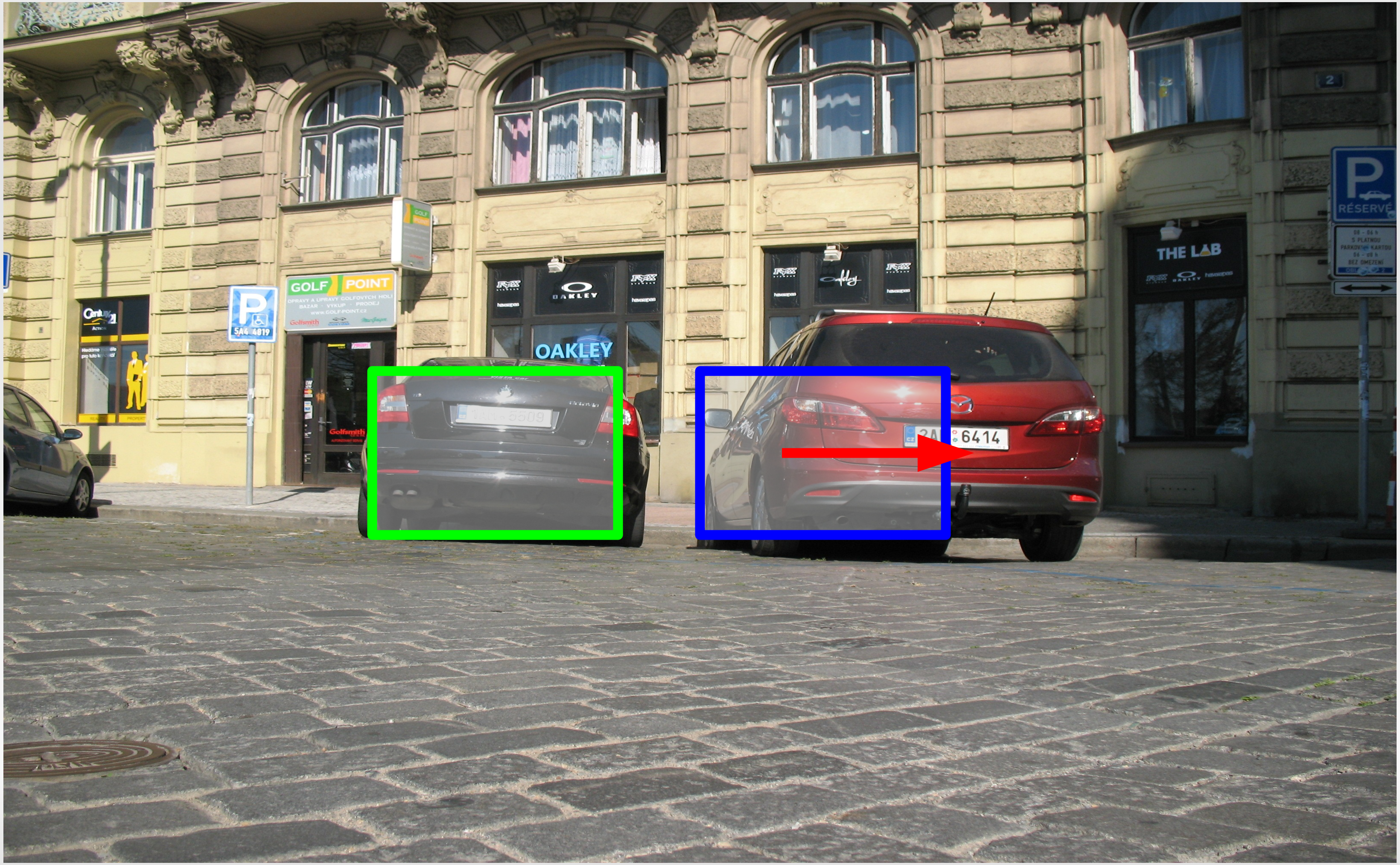
Overview



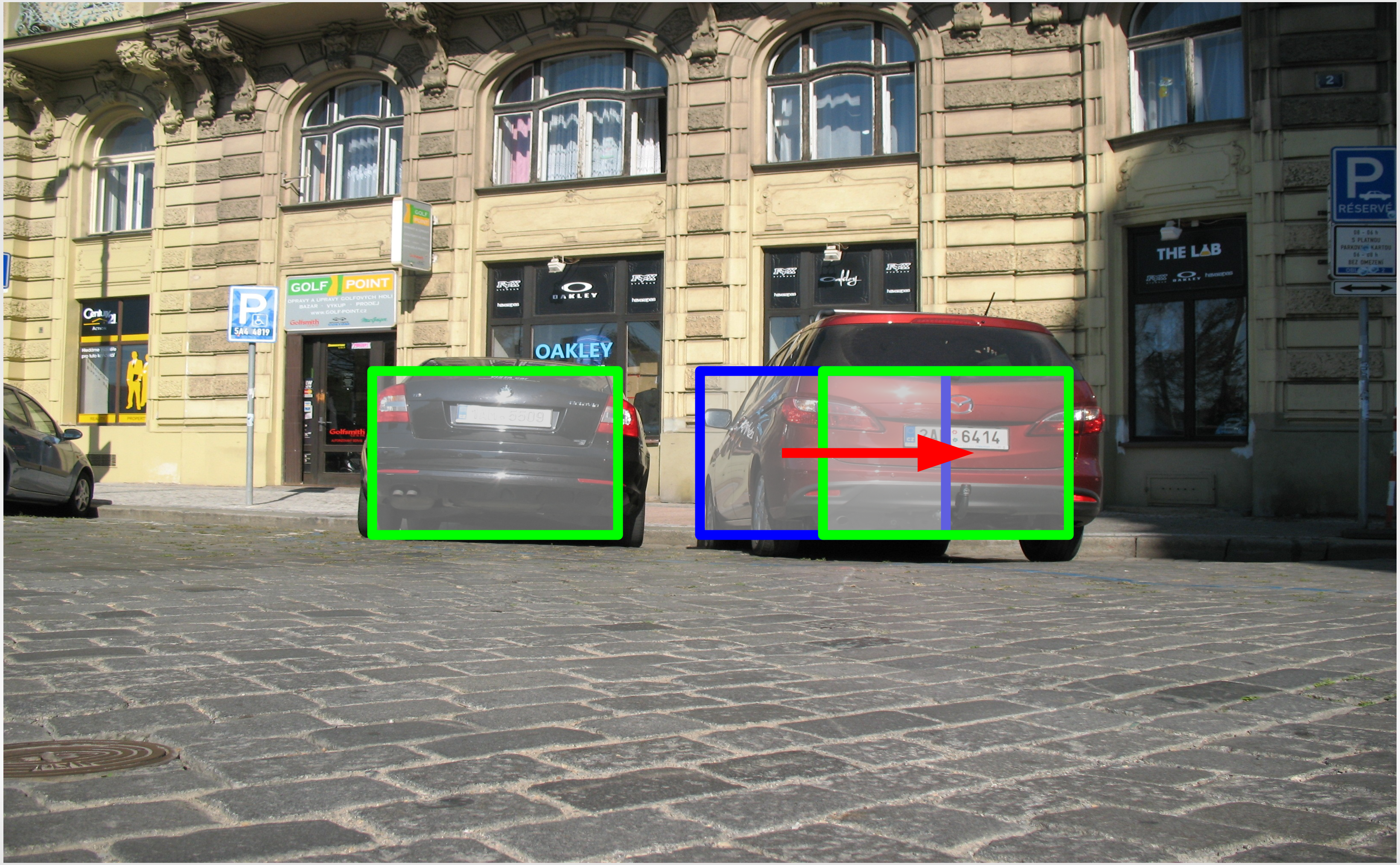
Overview



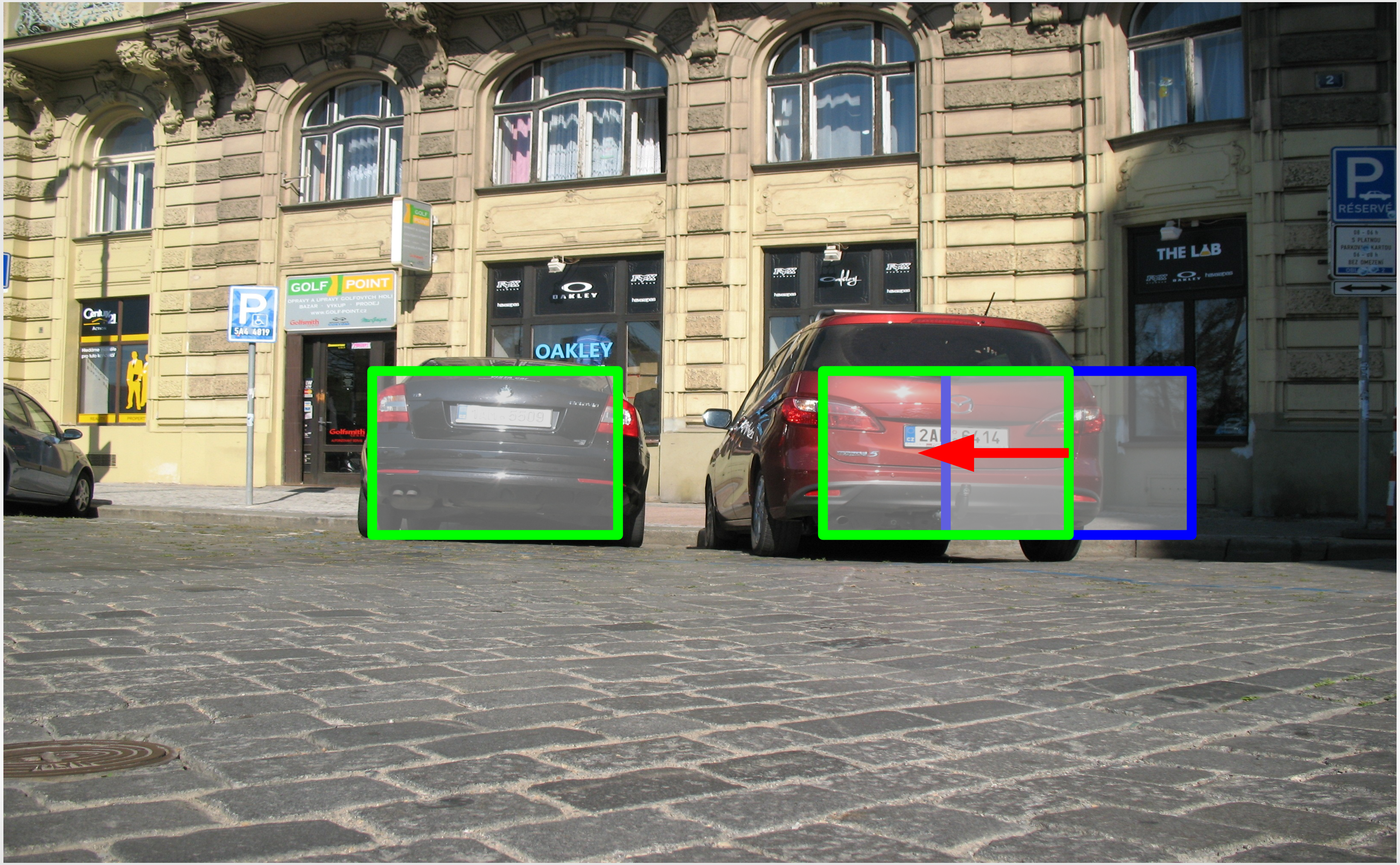
Overview



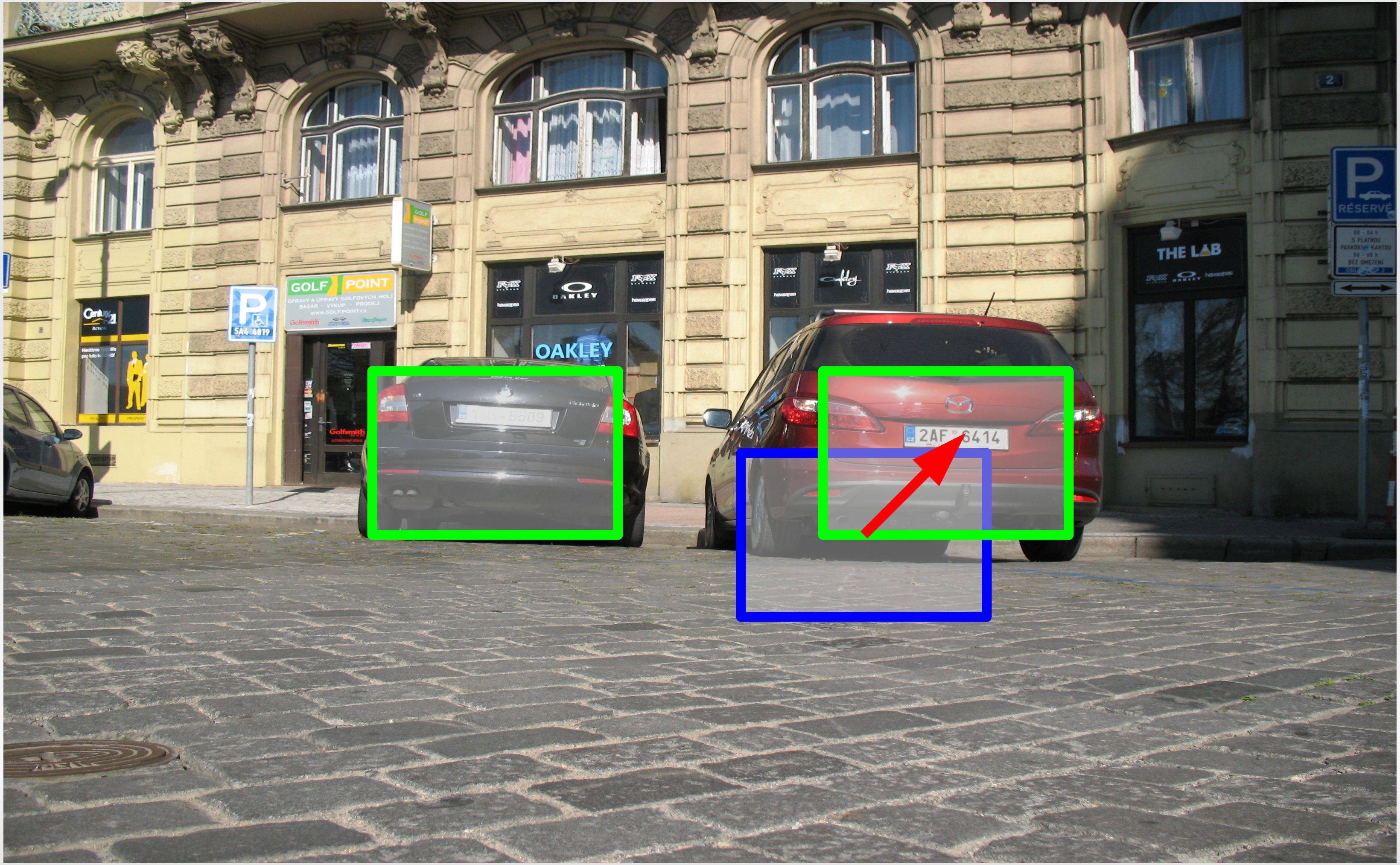
Overview



Overview

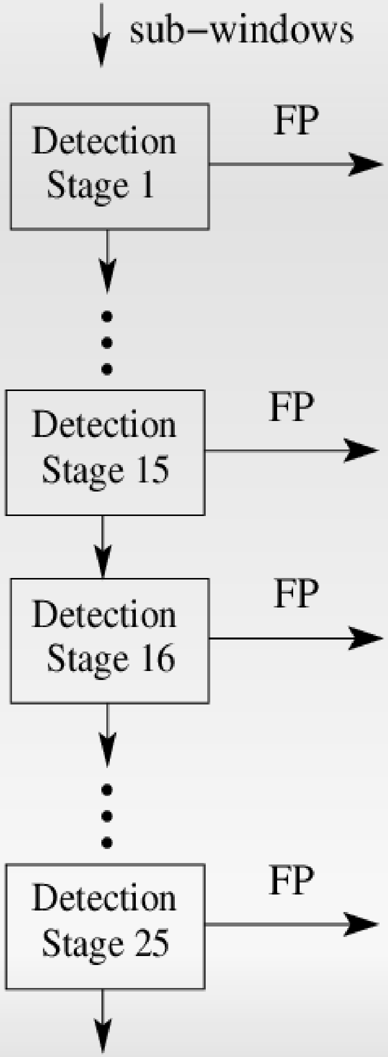


Overview



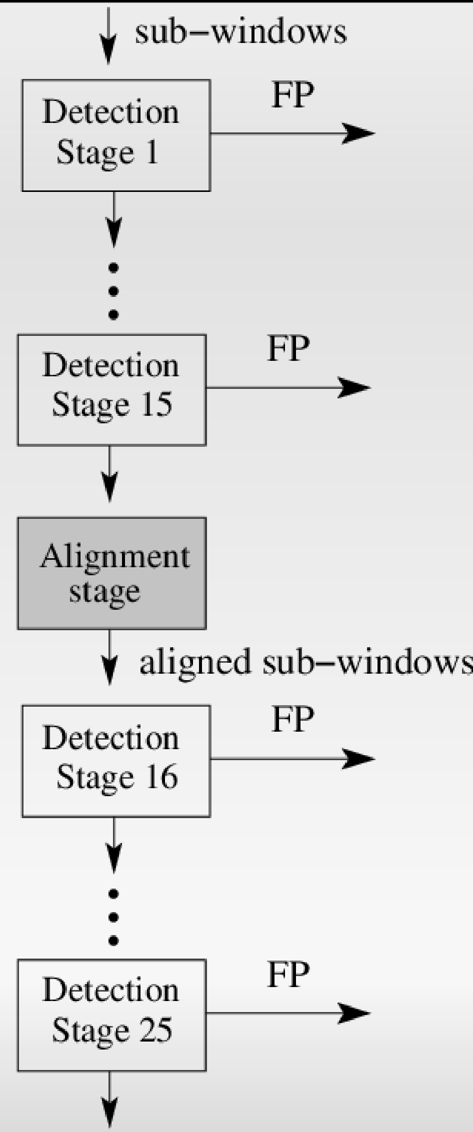
Overview

- Cascade of Gentle-Boost classifiers (on Haar features).
- Each Detection stage rejects 40% of background sub-windows.
- Each detection stage preserve 99.5% positive samples



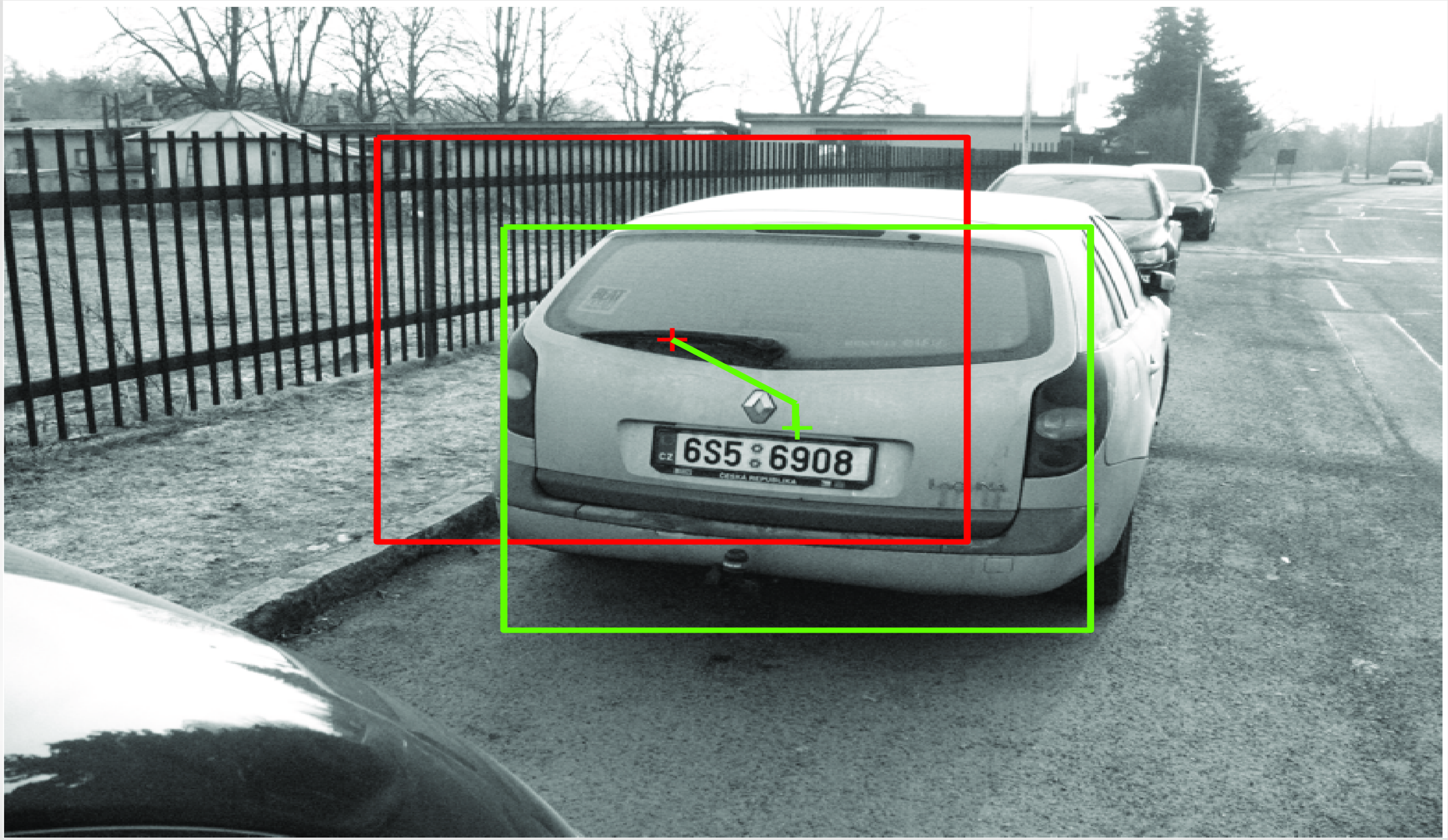
Overview

- Sliding window detection with a cascade extended by alignment.
- The **alignment** is invoked iff the sub-window survives up to a certain detection stage.
- Well aligned sub-window is more likely to be detected.



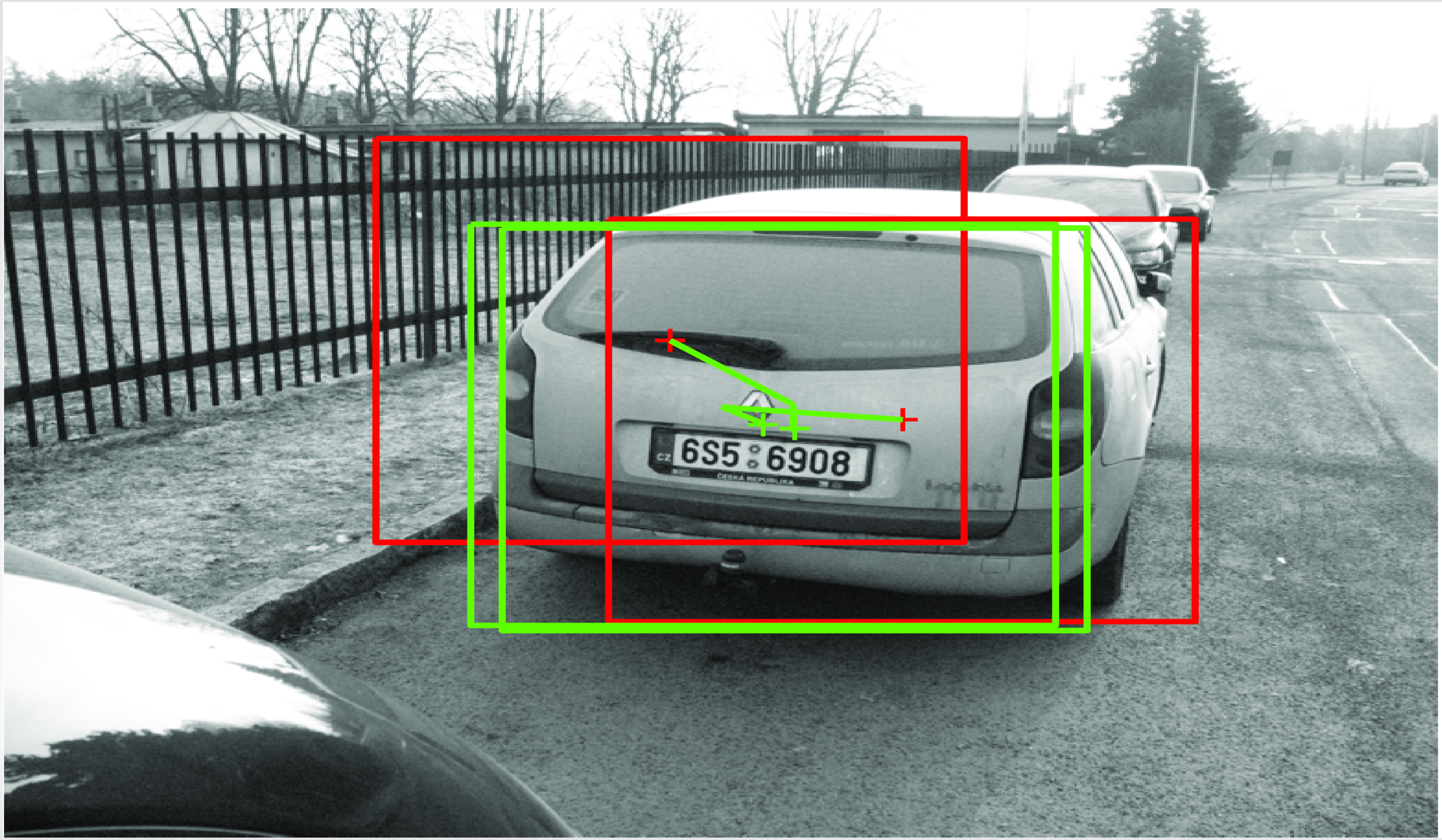
Detection with aligned cascade

- Alignment method assigns translation to sub-windows



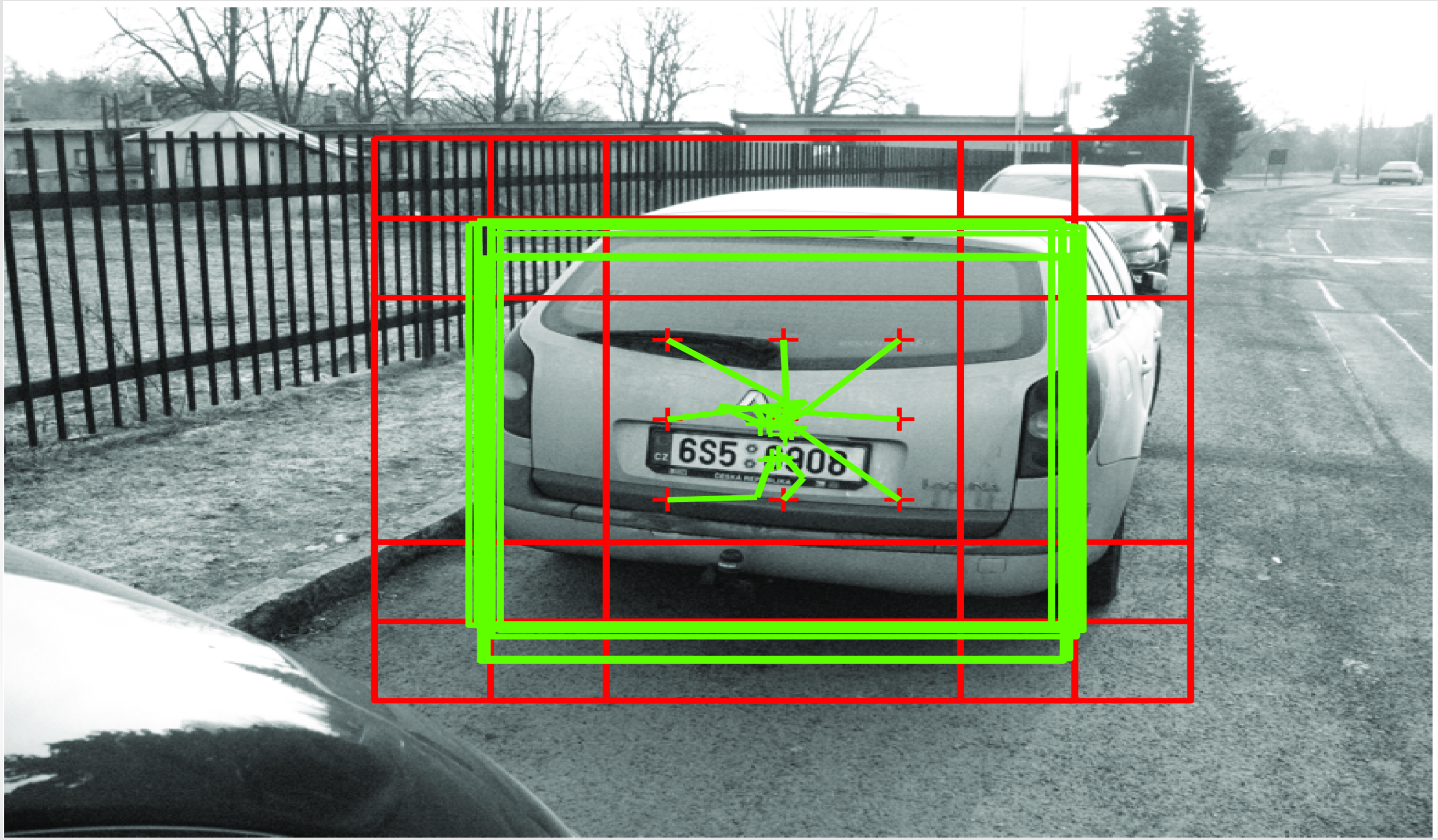
Detection with aligned cascade

- Sub-windows close to the object are aligned on the object



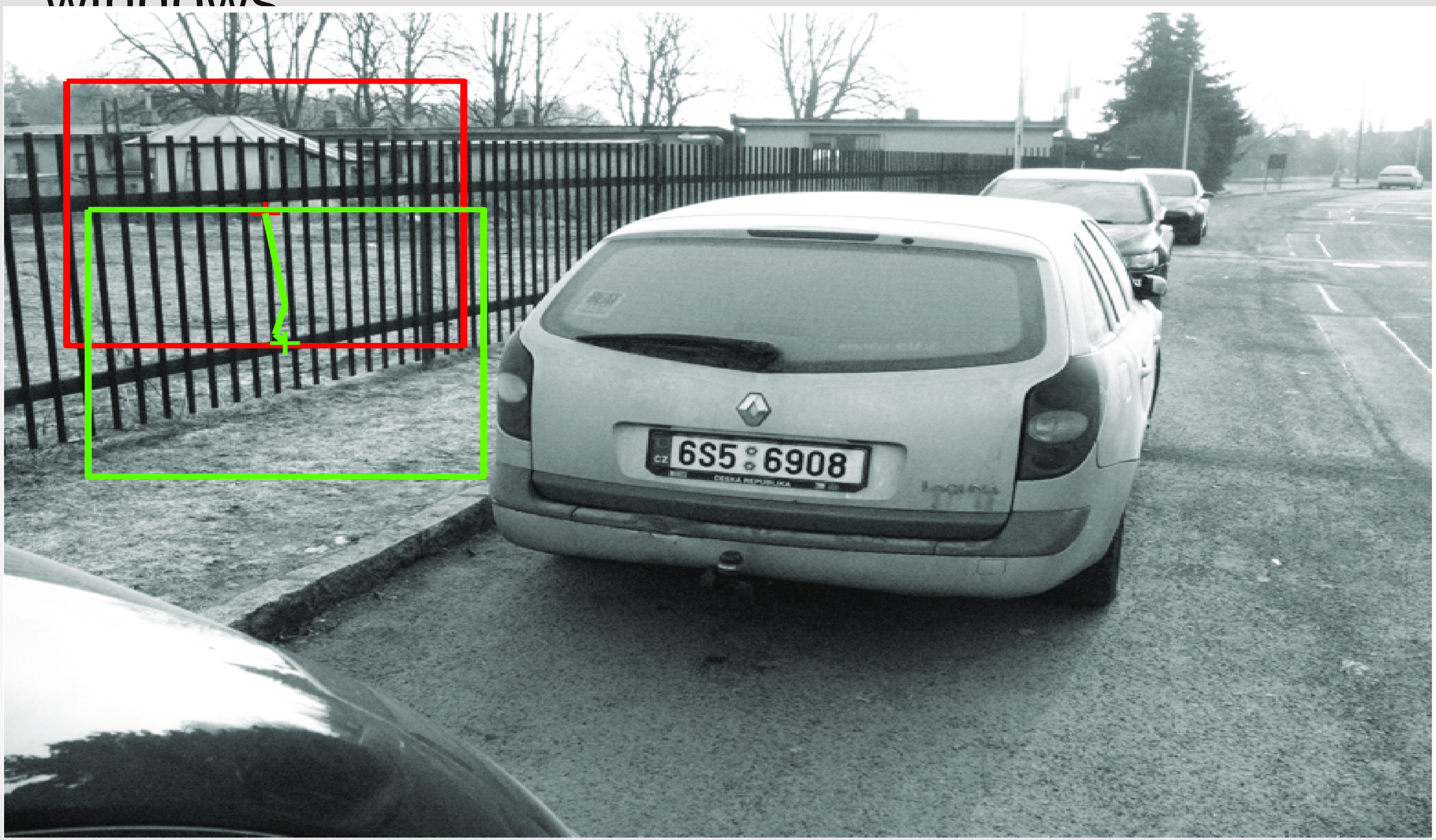
Detection with aligned cascade

- Sub-windows close to the object are aligned on the object



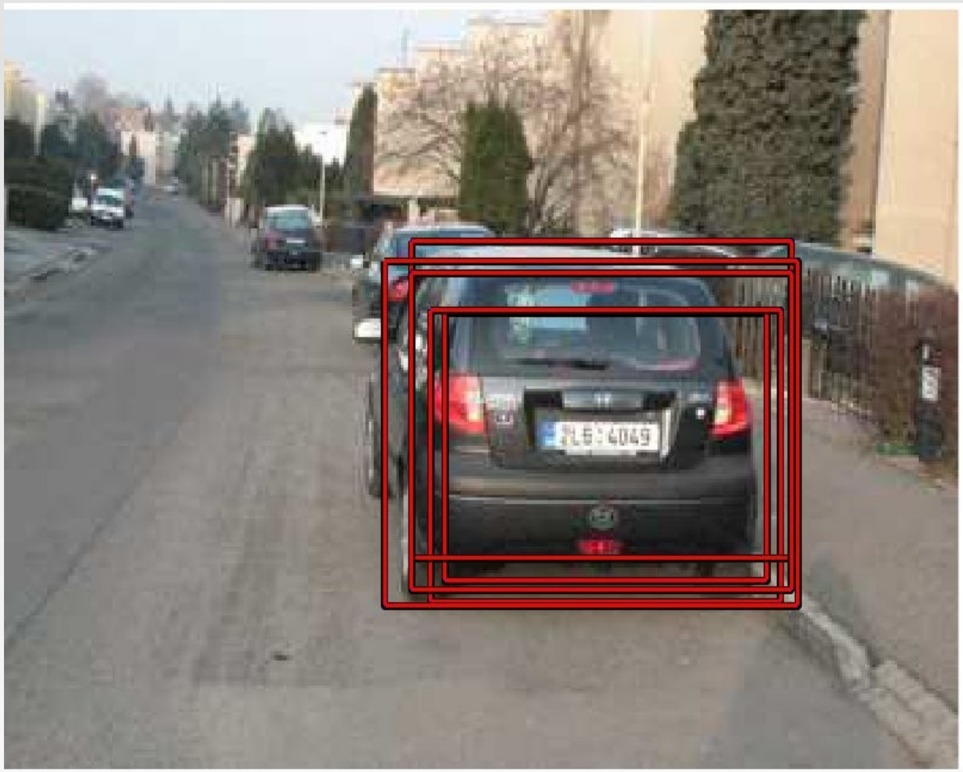
Detection with aligned cascade

- Alignment method *randomly* shifts background sub-windows

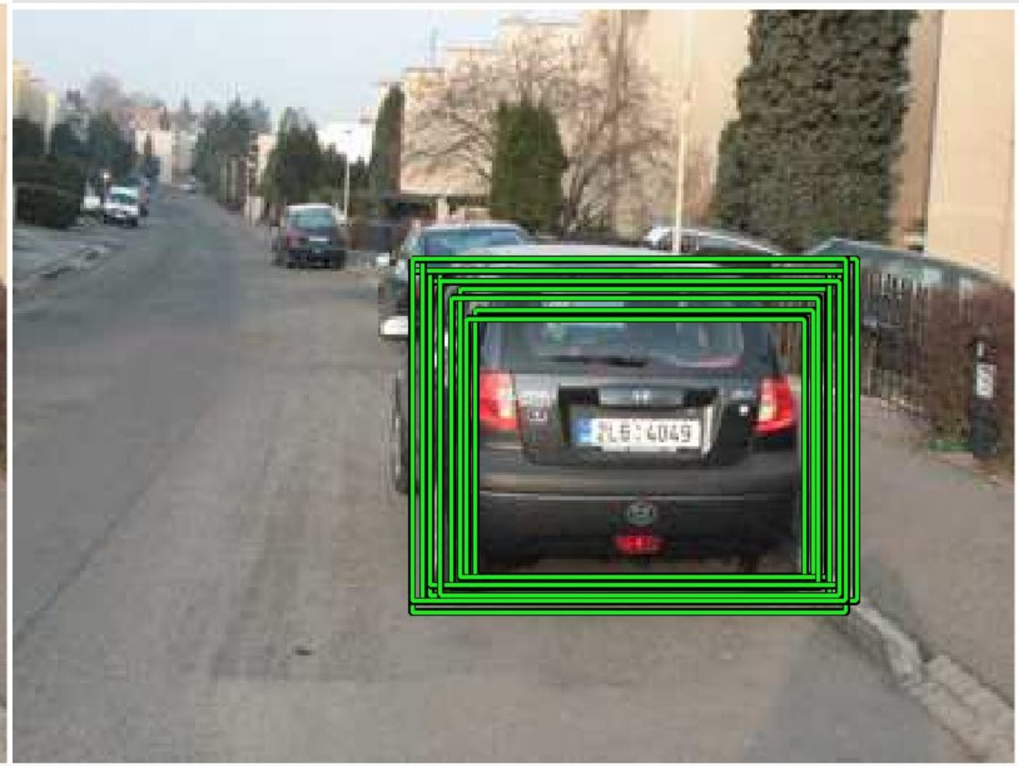


Detection with aligned cascade

Detector

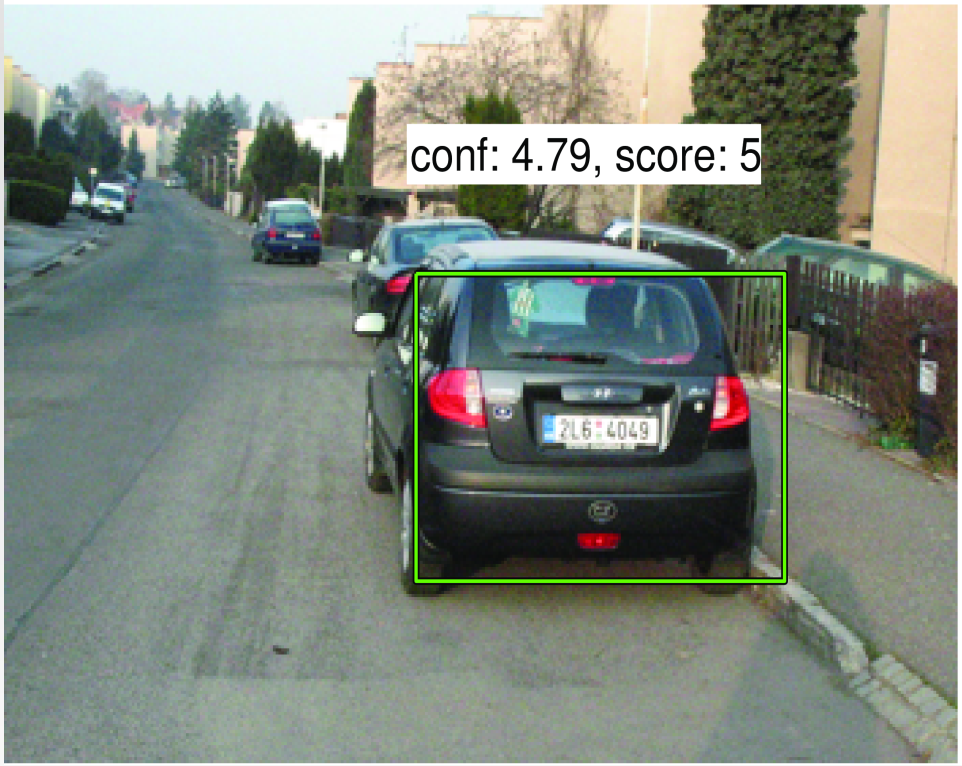


Detector + Alignment

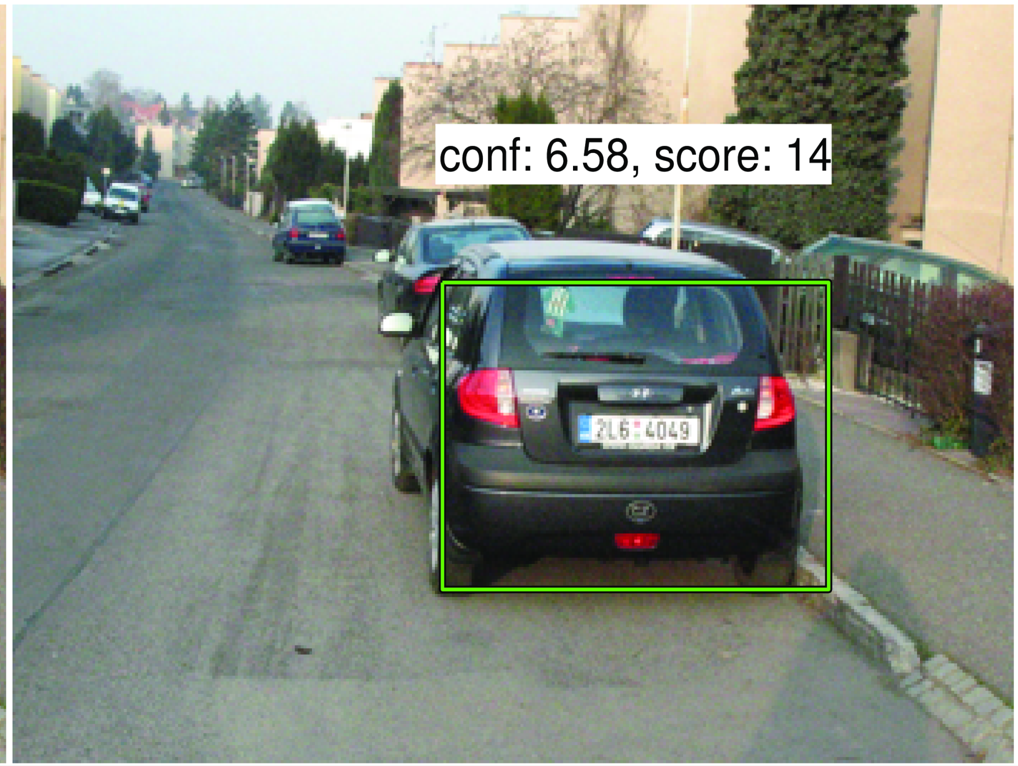


Detection with aligned cascade

Detector



Detector + Alignment



Sequential linear predictor

- Two different alignment methods were studied:
  - Linear predictor
  - Fern

Alignment with linear predictor

- Linear predictor maps features (f) from the evaluated sub-window to local displacement (t)

$$\boldsymbol{t} = H \cdot \boldsymbol{f}$$

where **f** is absolute value of Haar-like features.

Alignment with linear predictor

- Linear predictor maps features (f) from the evaluated sub-window to local translation (t)

$$\boldsymbol{t} = H \cdot \boldsymbol{f}$$

where **f** is absolute value of Haar-like features.

- Linear regression function (H) learned by the Least-Squares method on the training set.

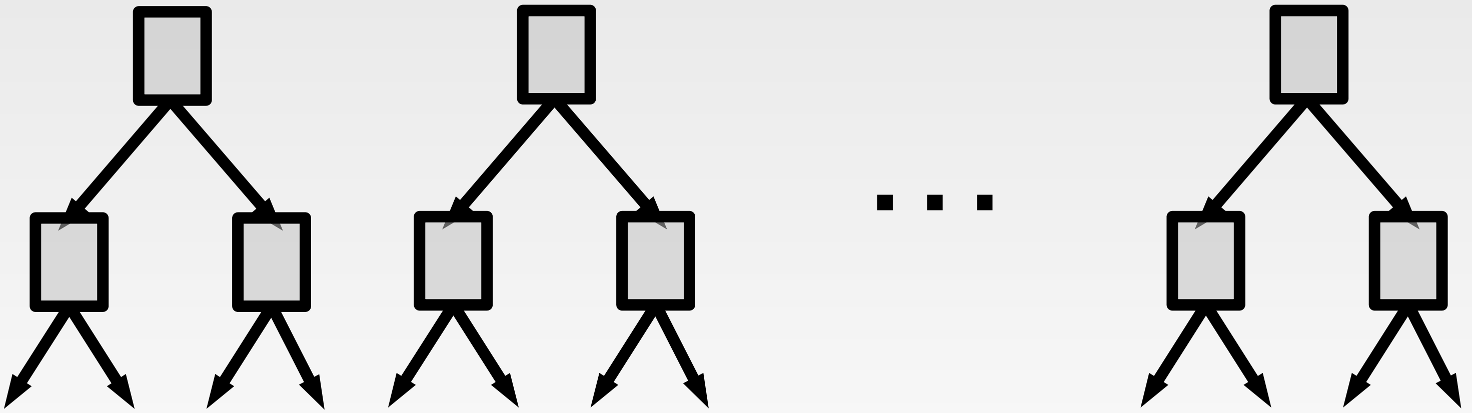
Alignment with linear predictor

Alignment with linear predictor

- Single linear predictor has a low accuracy.
- Train a sequence of linear predictors.
- Each predictor is trained on the range of translation errors of its predecessor.

Alignment with ferns

- Ferns: forest of random binary decisions trees.



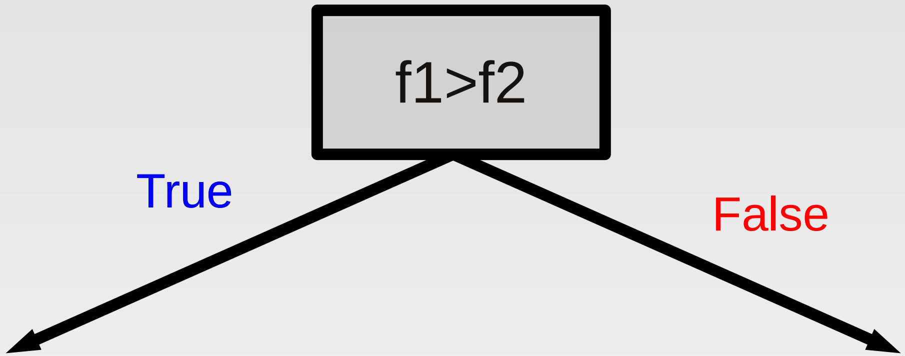
Alignment with ferns

- Each node forms a simple binary condition:

f1>f2

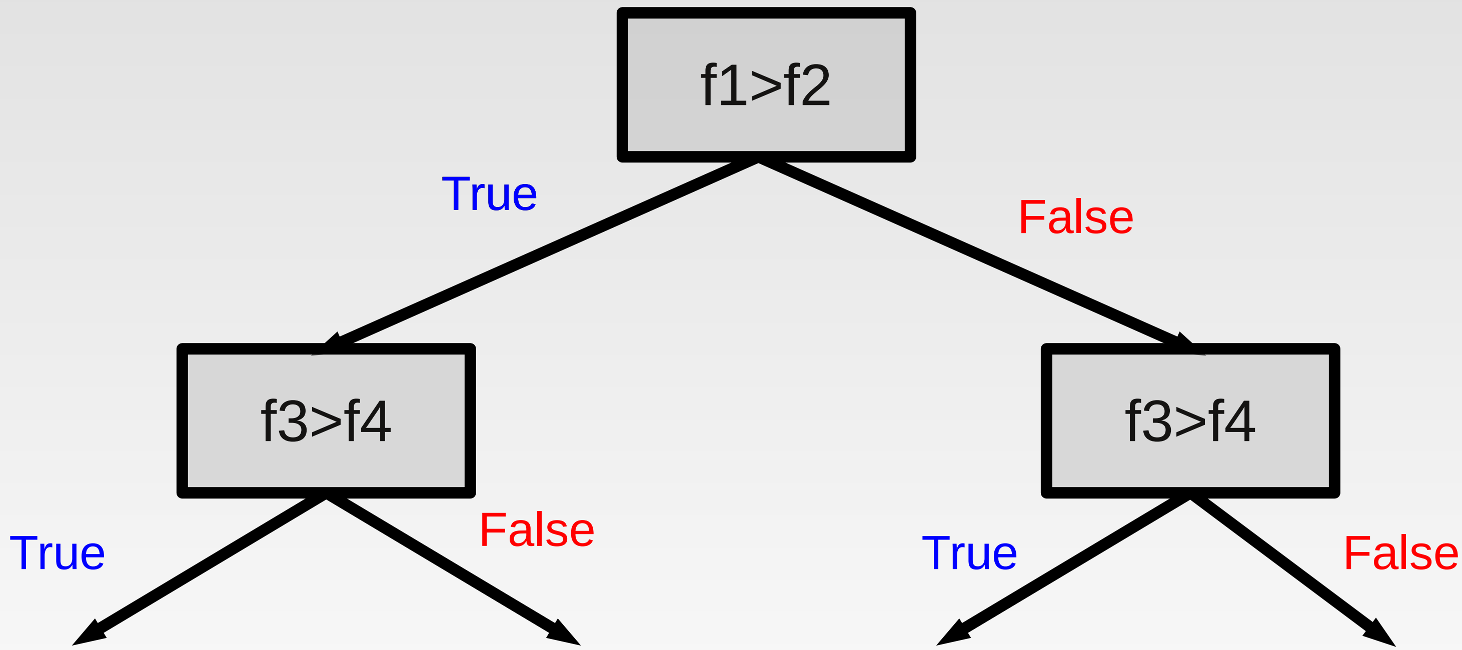
Alignment with ferns

- Result of the condition determines the direction in which the evaluated sub-window continues.



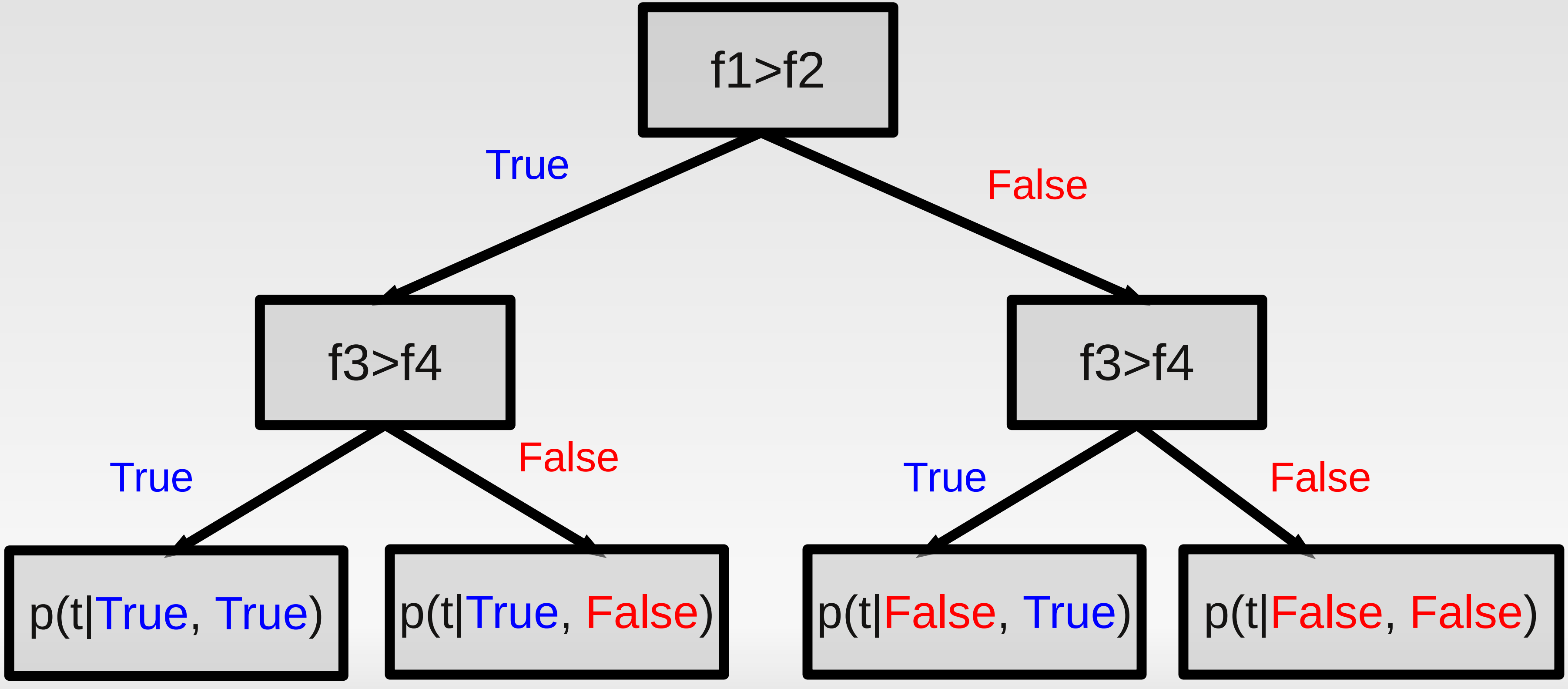
Alignment with ferns

- Then a different features are compared



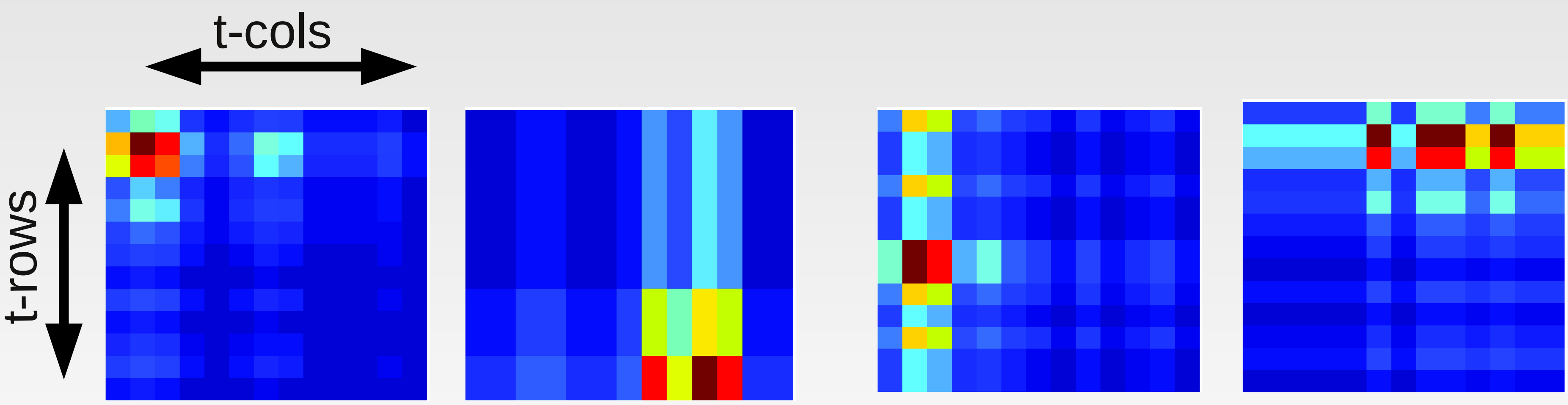
Alignment with ferns

- Leaves contain conditional probability of discretized translations.



Alignment with ferns

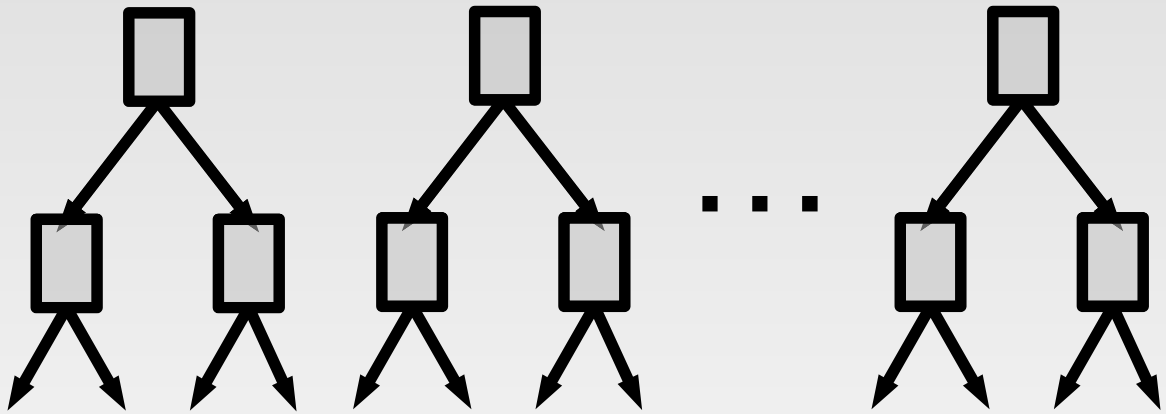
- Probability is learned off-line from the training set.



p(t|True, True) p(t|True, False) p(t|False, True) p(t|False, False)

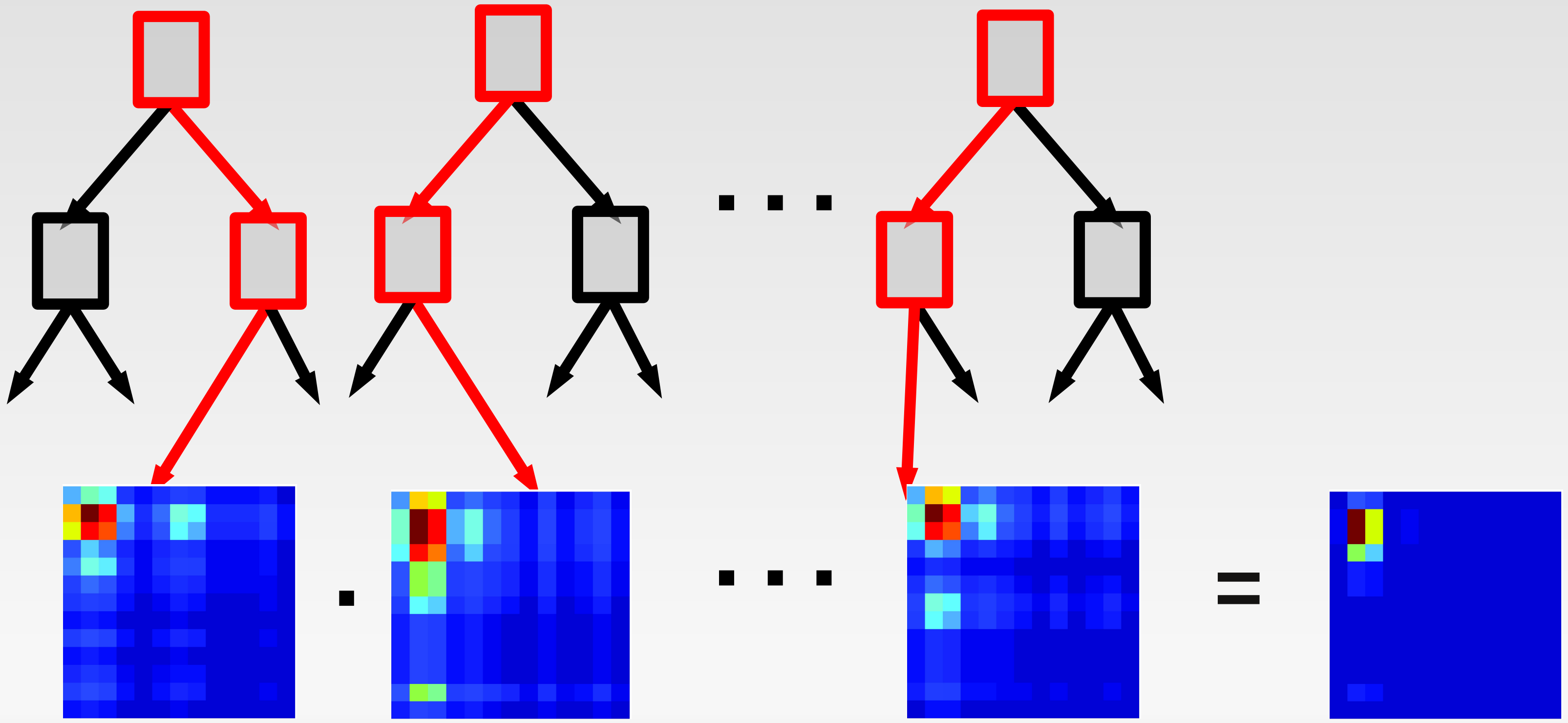
Alignment with ferns

- 50 trees, each with depth 11 (i.e. with 2048 leaves)

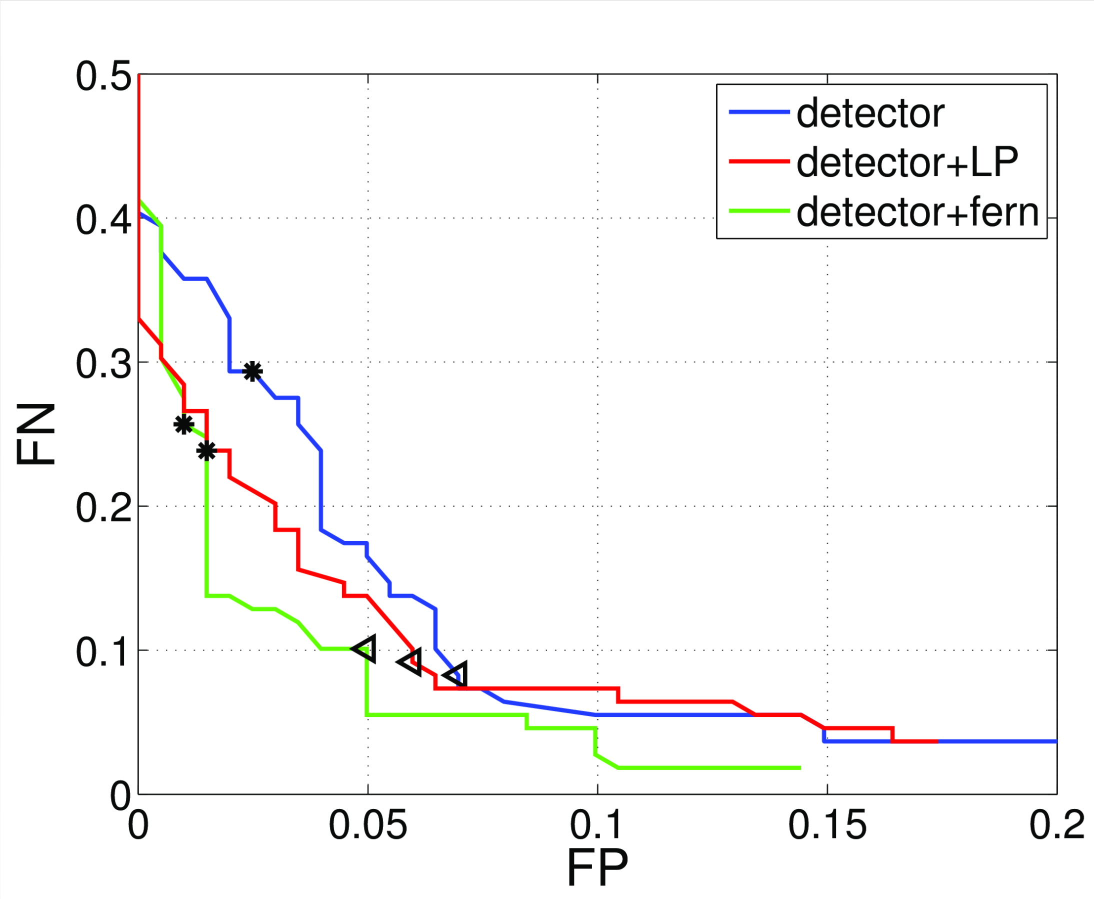


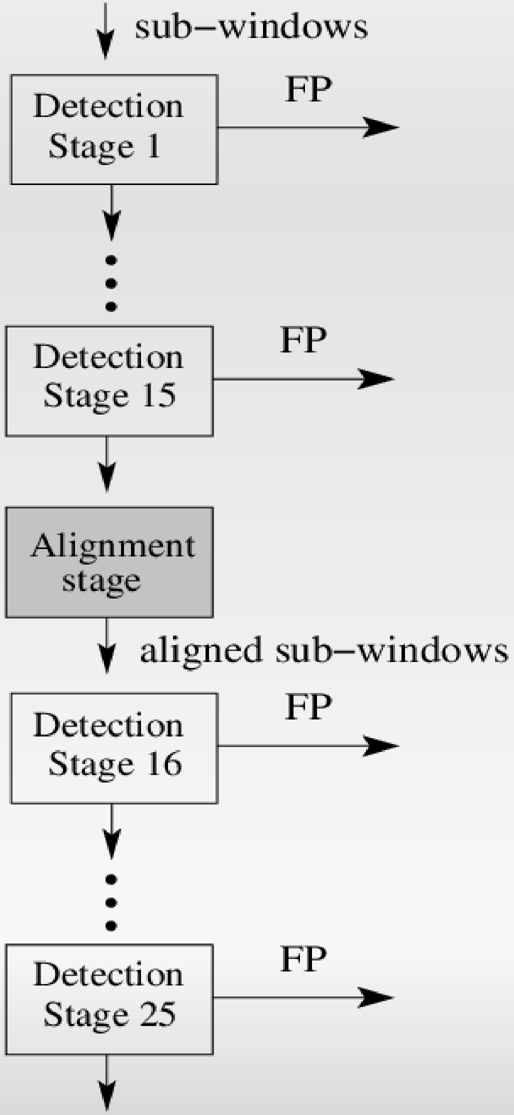
Alignment with ferns

- Final probability estimated as multiplication

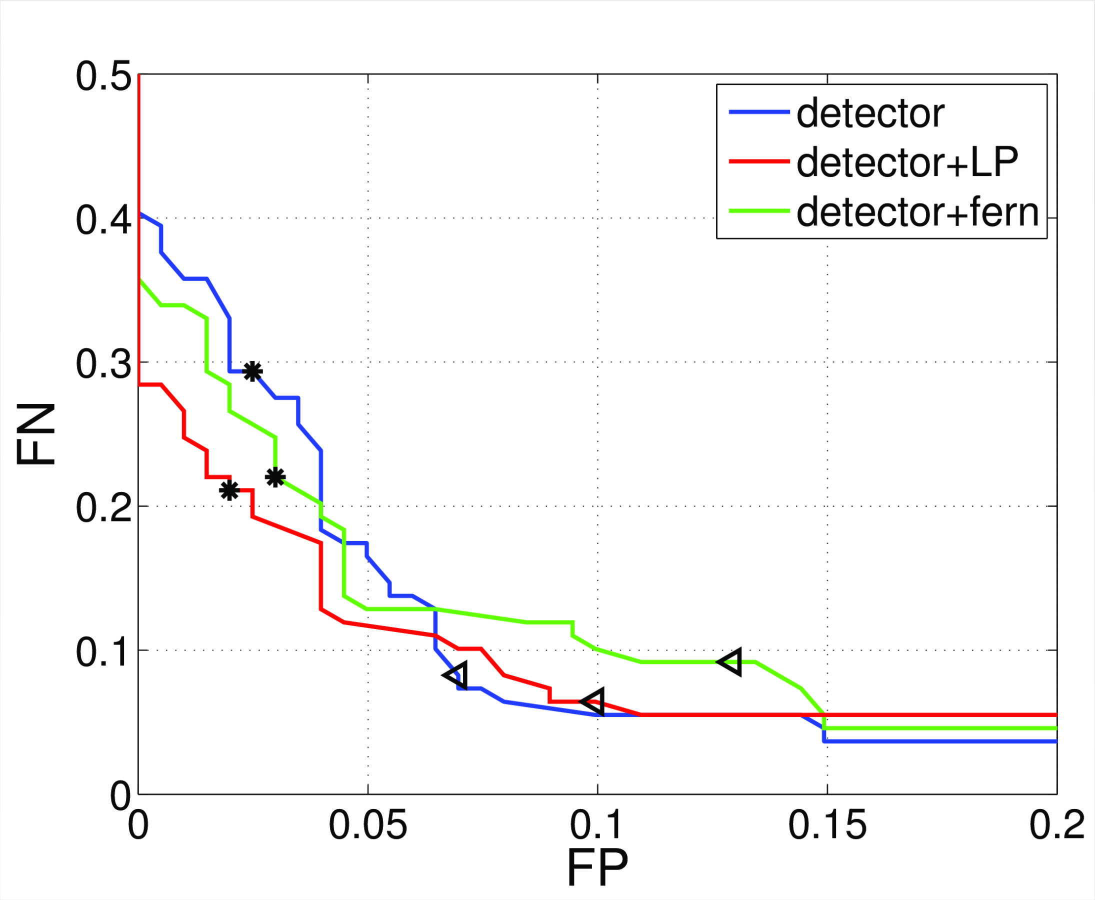


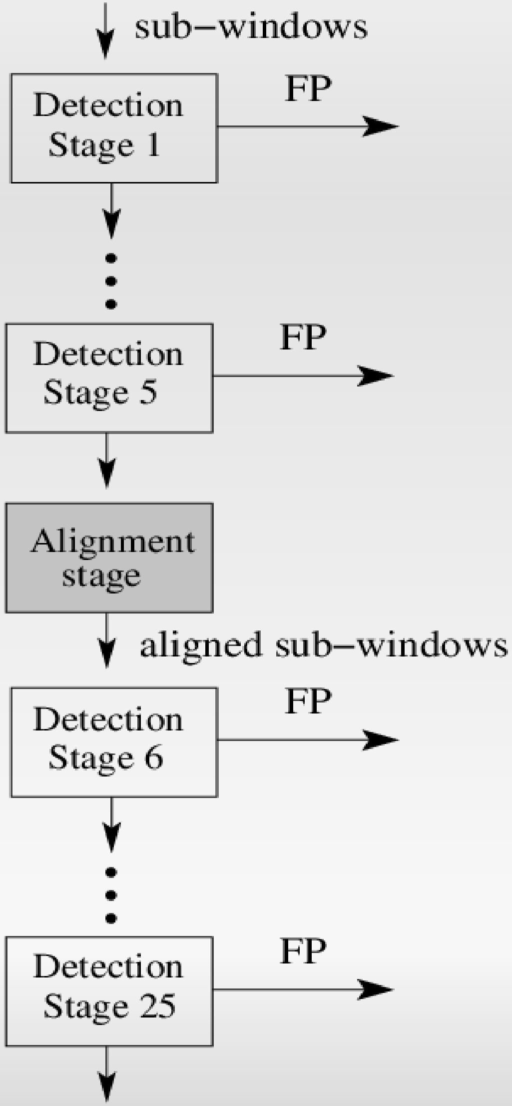
Experiments (alignment stage 15)



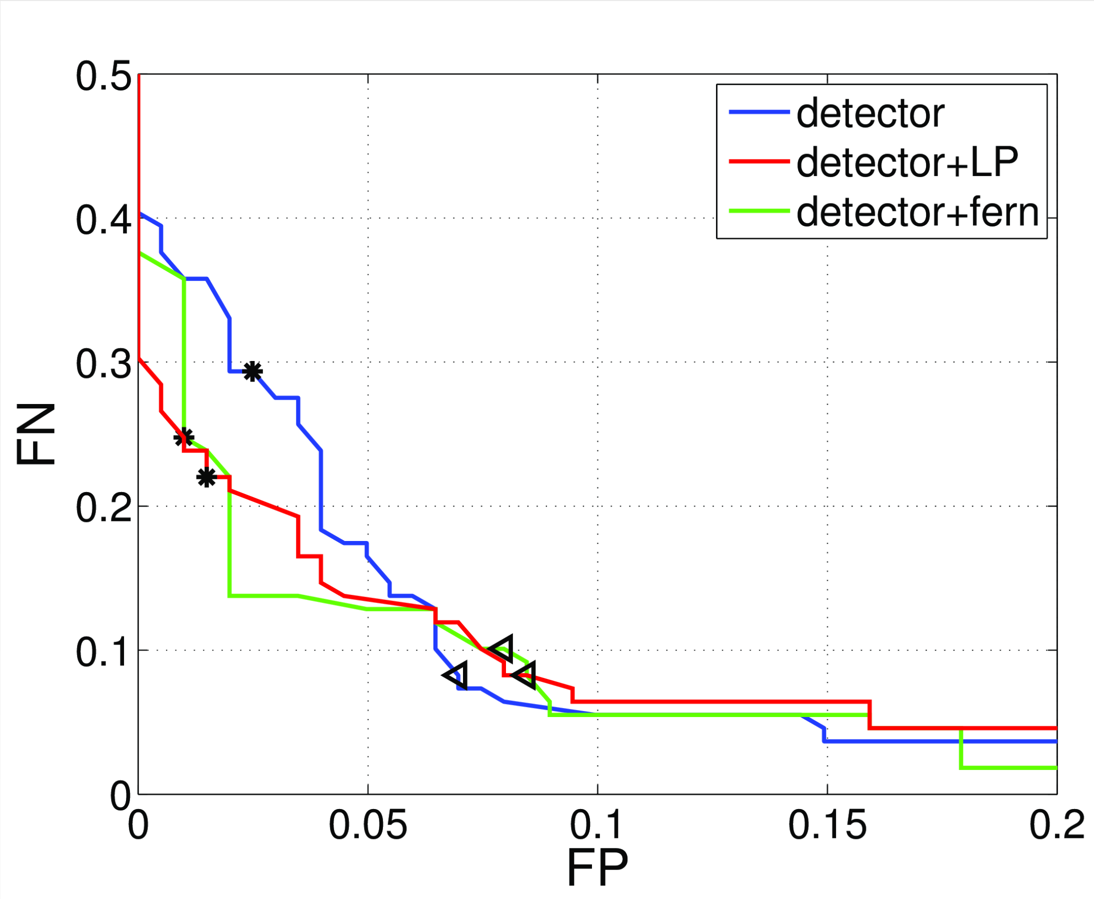


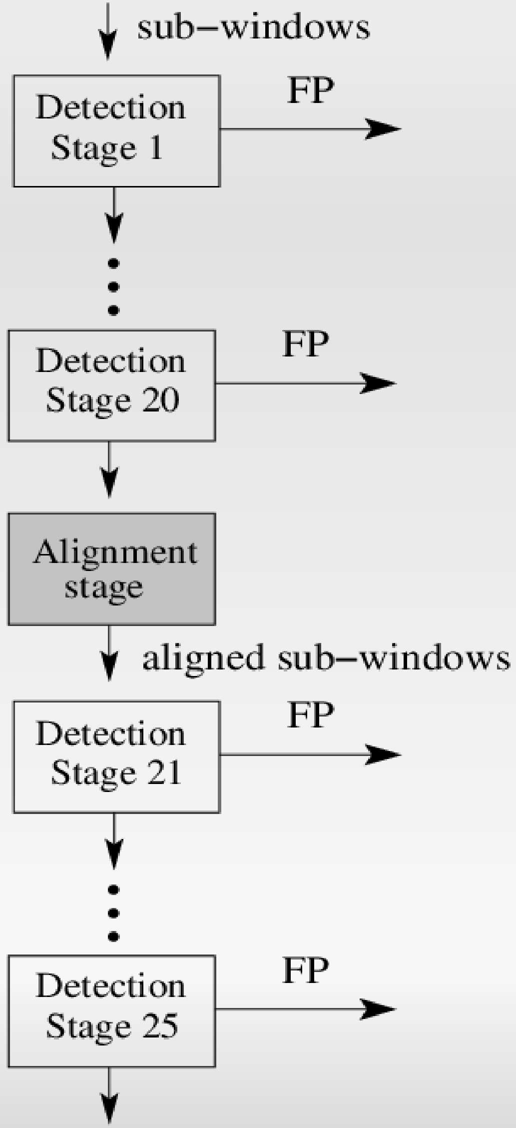
Experiments (alignment stage 5)



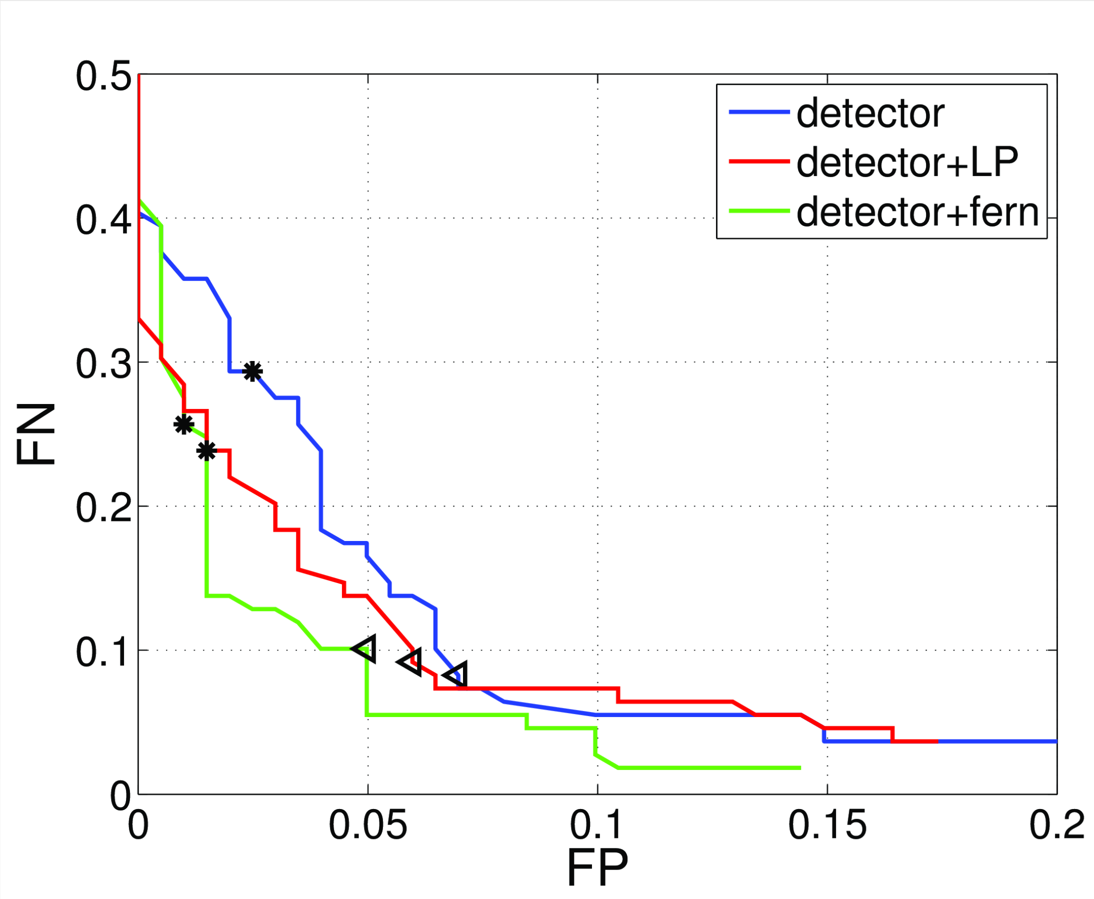


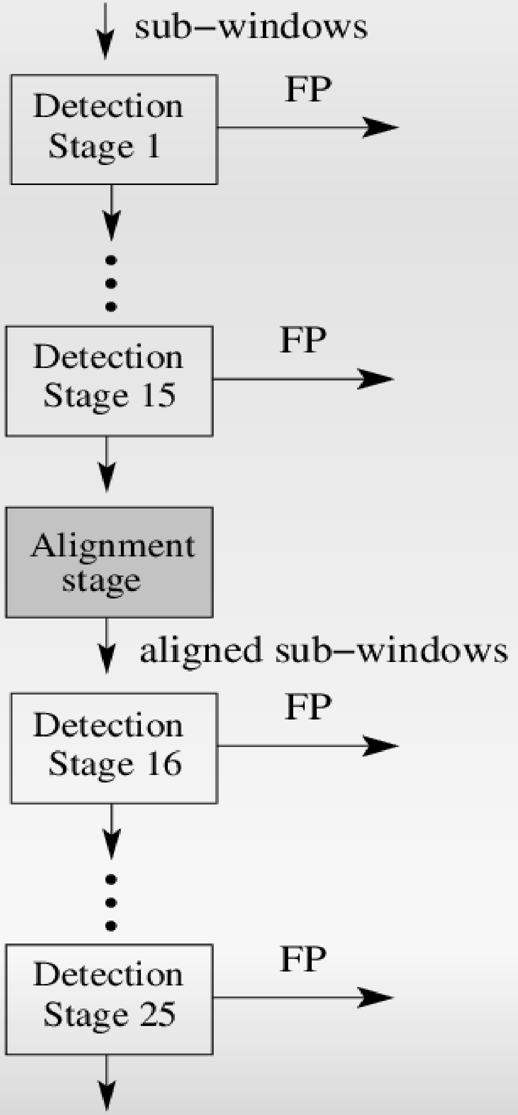
Experiments (alignment stage 20)





Experiments (alignment stage 15)





Implementation details

- Cascade of Gentle-Boost detectors trained on Haar features.
- Alignment methods use absolute value of Haar features.
- Less than 0.05% sub-windows survives up to stage 15,
- Ferns: Using the same condition at each level yields speed up (trees-> hash tables).

Future work

- Affine or perspective alignment
- Local gradient-based maximization of the detection function.