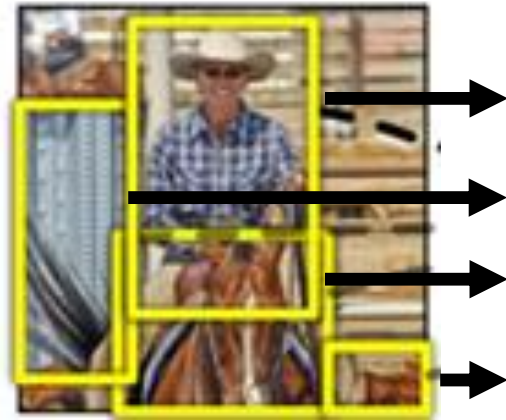


Cons

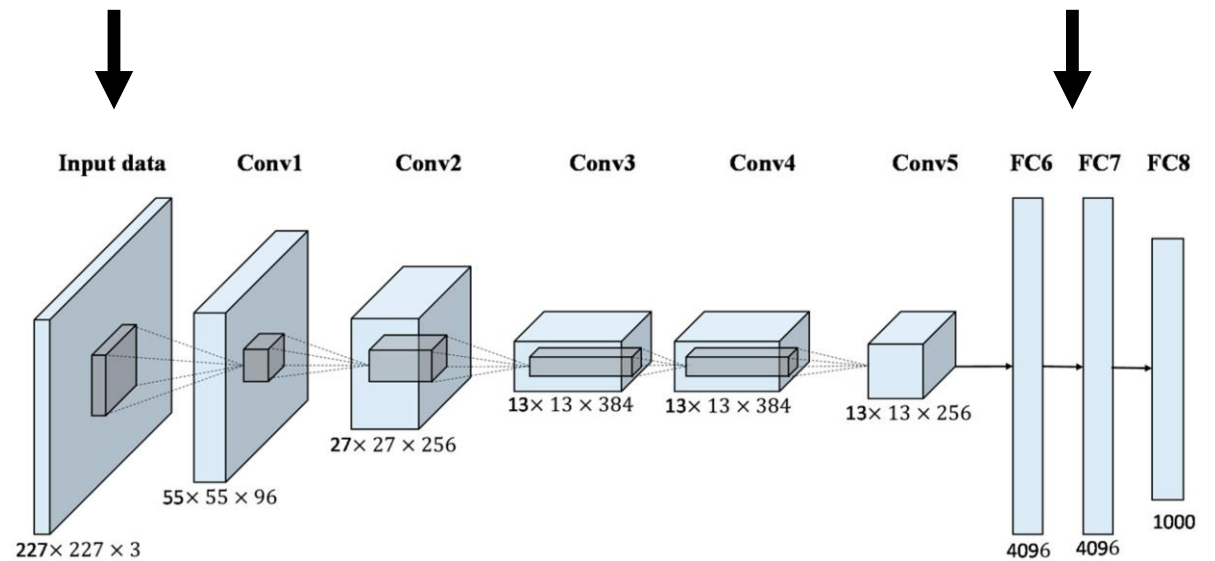
Presented by Hsin-Ping Huang

CNN requires a fixed-size input image

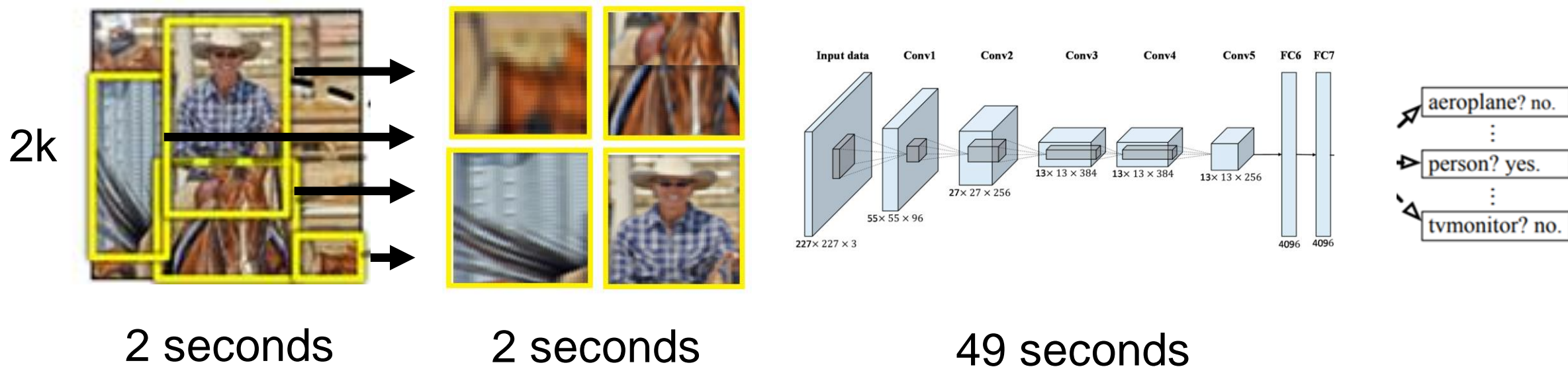
Independent forward passes



Cropping
Warping

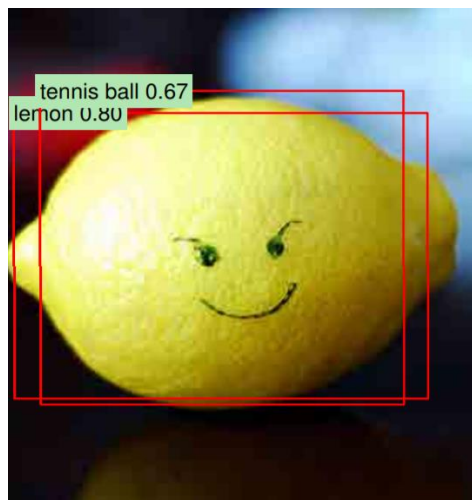


Cons 1: slow due to independent forward passes



	DPM v5	Regionlets	R-CNN
Test time	13 sec	2.2 sec	53 sec

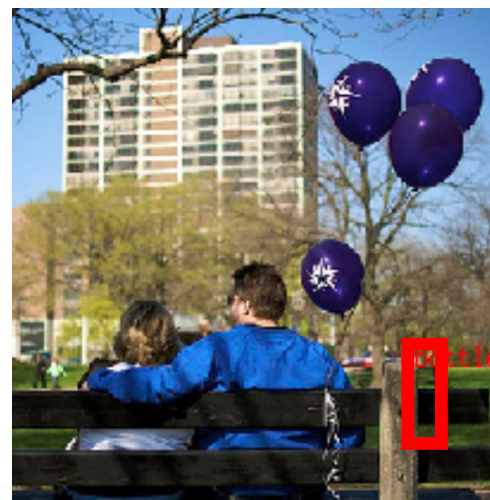
Cons 2: cropping/warping cause information loss



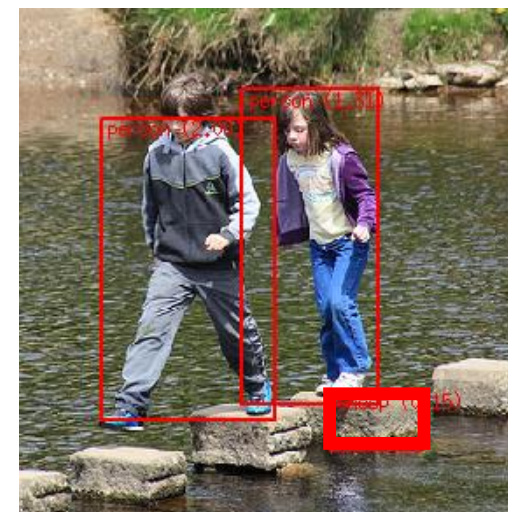
lemon



tennis ball



bottle



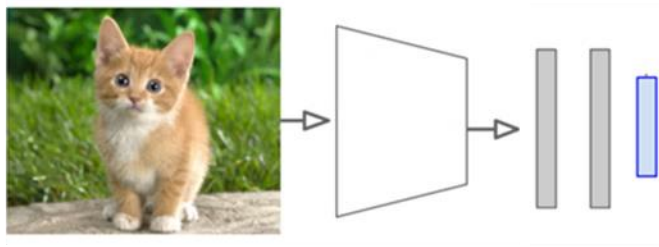
sheep

Confusion with a similar category

False positives on background

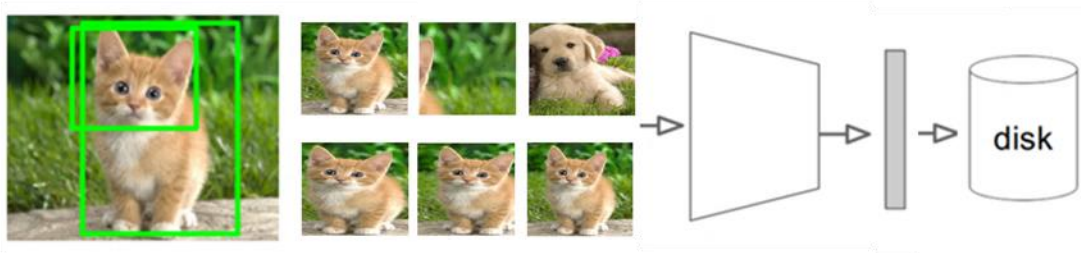
Cons 3: training is expensive in time and space

Fine-tuning



18 hours

Feature extraction



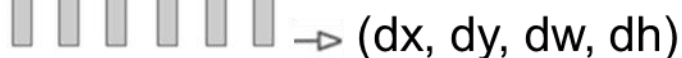
63 hours

200G storage

SVM



Bbox regressor

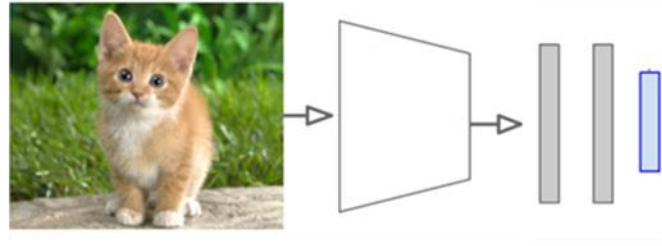


3 hours

total: 84 hours

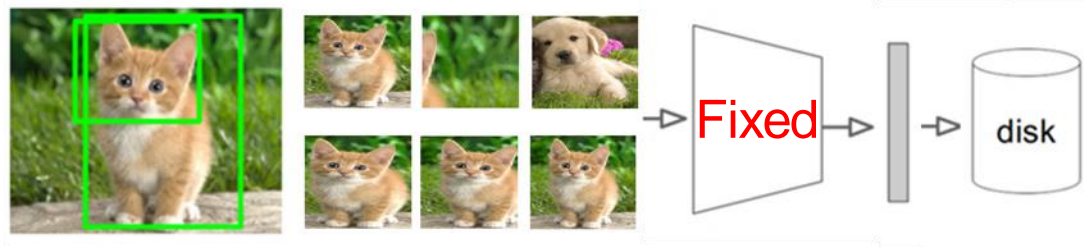
Cons 4: SVMs and regressors are post-hoc

Fine-tuning



log loss

Feature extraction

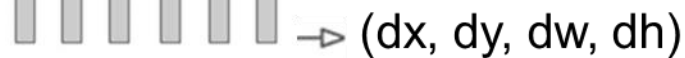


SVM



hinge loss

Bbox regressor



squared loss