Assignment 3/Mini-project

Deadline: March 20, GROUP SIZE - 2 or 3

Visual Recognition

Assignment 3a (7 marks)

Play with CNNs

- CIFAR-10 dataset
- AlexNet has 5 Conv and 3 FC layers
- Play with a medium deep network with atleast 2 conv and 2 FC layers.
- Metrics: training time and classification performance
- Compare ReLU vs tanh vs sigmoid
- With and without momentum, adaptive learning rates
- Finally: what would be your recommended architecture..

Assignment 3b (3 marks)

CNN as a feature extracter

- Pick your favourite object recognition dataset, other than CIFAR/MNIST

- Use Alexnet/any deep NN as a feature extracter (extract last layer as features), use any model on top and report the classification accuracies

- Report accuracies for Bike vs Horses dataset also

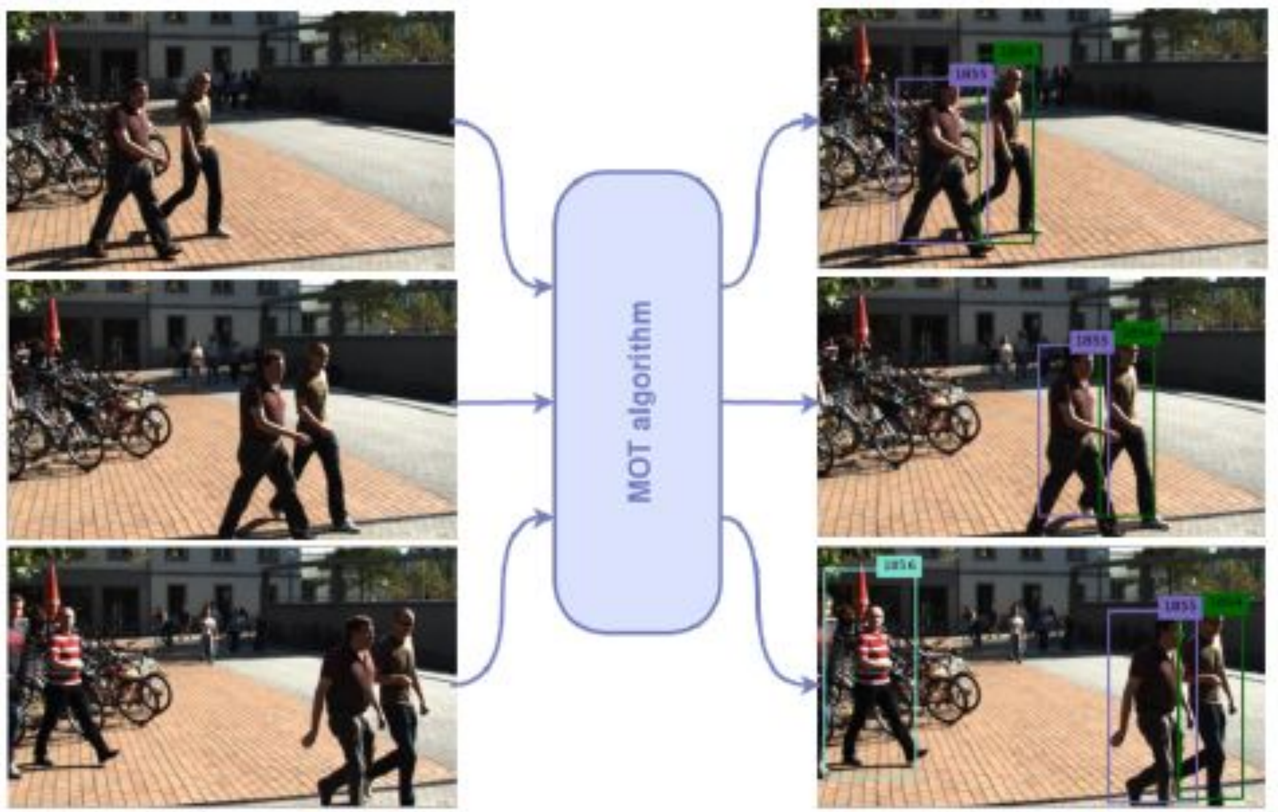
Assignment 3c: YOLO V2 (2 marks)

- Read YOLO V2 and V1 papers
- Pick 5 additional features of YOLO V2 and document it in the report

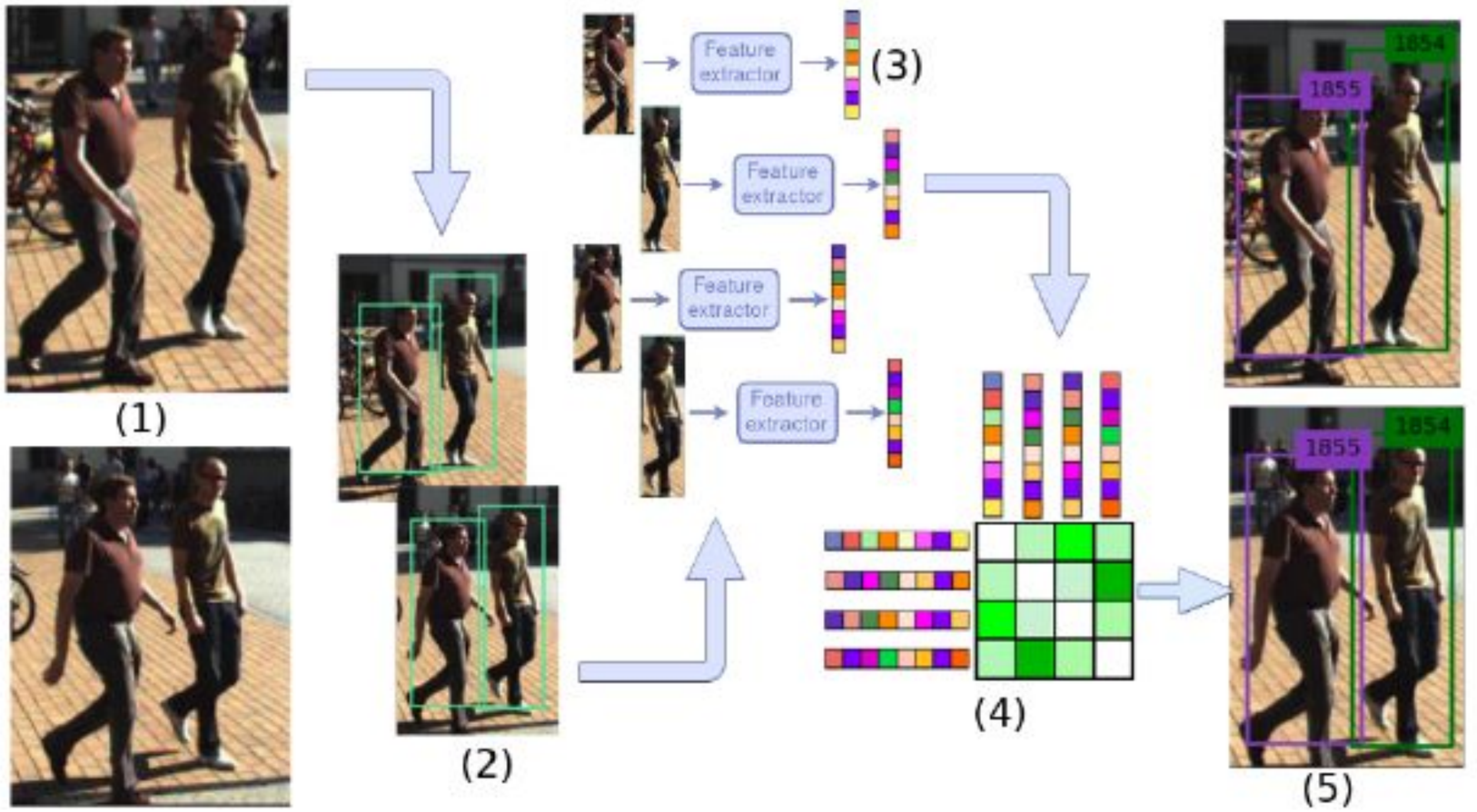
Assignment 3d: Object tracker: SORT + DeepSORT (8 marks)

- Dataset: Traffic junction videos you have collected already
- For the detections, use Faster RCNN and YOLO V2
- Implement a car counter using object tracking (SORT and DeepSORT)
- Document the results and observations
- Document the difference between SORT and DeepSORT
- Create ground truths if needed for evaluation
- Not needed to implement Faster RCNN/YOLO V2/SORT/Deep SORT from scratch

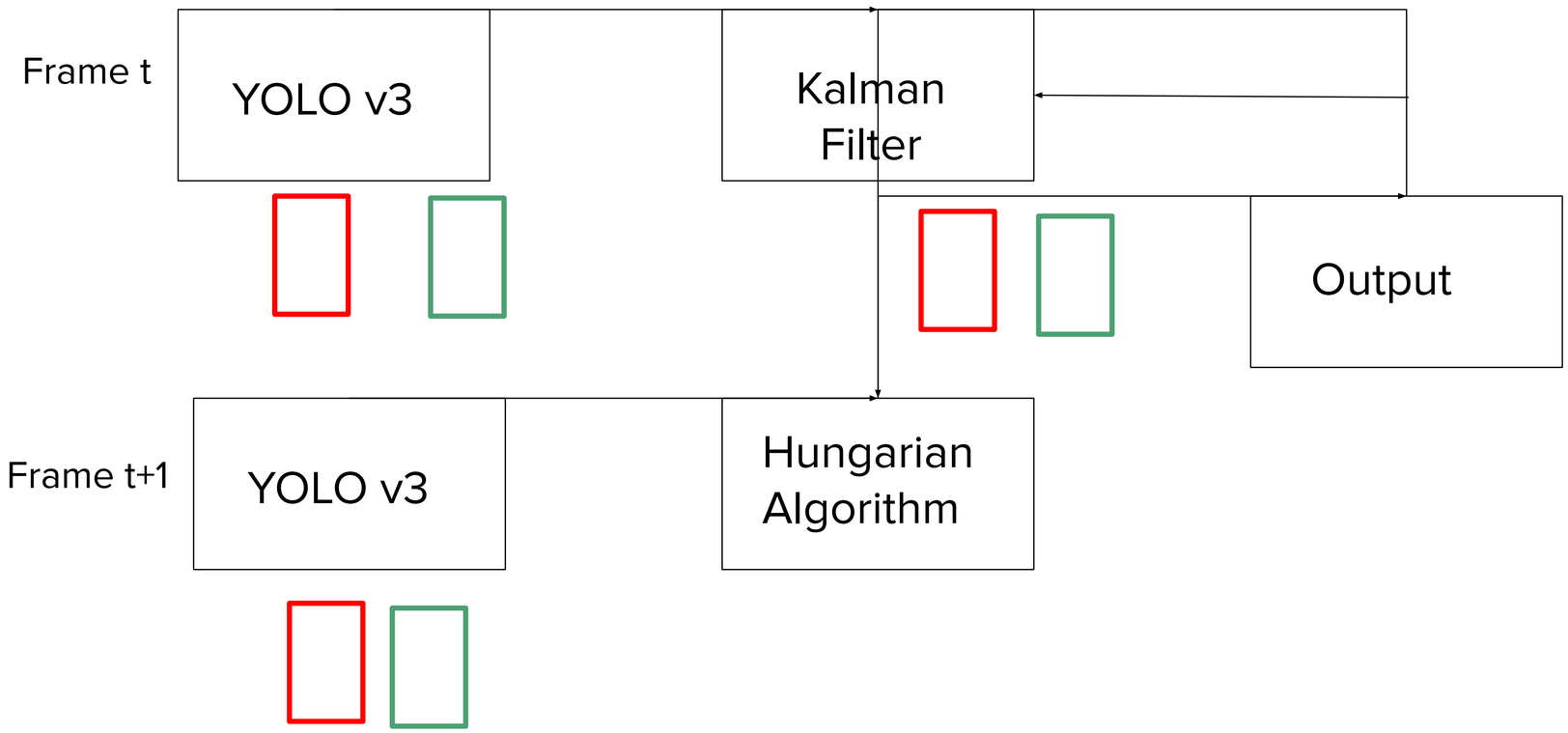
Multi-object tracking



Components of MOT tracking



Simple Online Real Time Tracker (SORT)



ICIP 2016: Alex Bewley, Zongyuan Ge, Lionel Ott “SIMPLE ONLINE AND REALTIME TRACKING”

	Clean bathroom	Sweep floors	Wash windows
Paul	$2	$3	$3
Dave	$3	$2	$3
Chris	$3	$3	$2

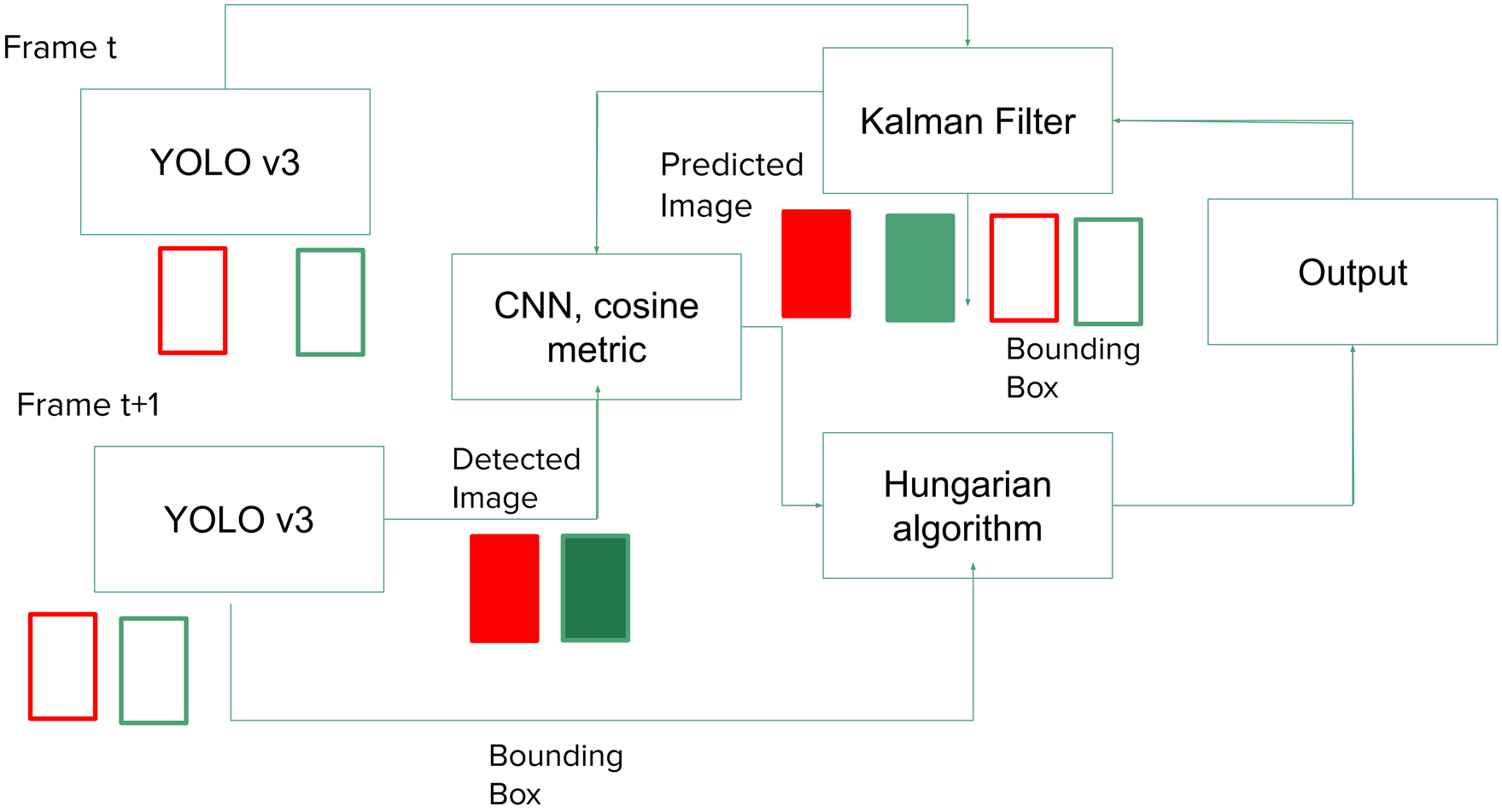
Hungarian algorithm

Hungarian Algorithm

Detection $Frame_t$/ Predicted $Frame_t$	P1	P2	P3
D1	IOU=0	IOU=0	IOU=0
D2	IOU=0.56	IOU=0	IOU=0
D3	IOU=0	IOU=0.77	IOU=0

Detection $Frame_t$/ Predicted $Frame_t$	P1	P2	P3
D1	0	0	0
D2	1	0	0
D3	0	1	0

Deep Simple Online Real Time Tracker (DeepSORT)



Wojke, Nicolai, and Alex Bewley. "Deep cosine metric learning for person re-identification." *2018 IEEE winter conference on applications of computer vision (WACV)*. IEEE, 2018.