@DocXavi

Master in Computer Vision Barcelona

[http://pagines.uab.cat/mcv/]

Xavier Giró-i-Nieto

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Department of Signal Theory and Communications
Image Processing Group

Module 6

Deep Learning for Video: Object Tracking

22nd March 2018

Object Tracking: Feature Extractor only

DNN were firstly used for feature learning to later be used by a tracker.

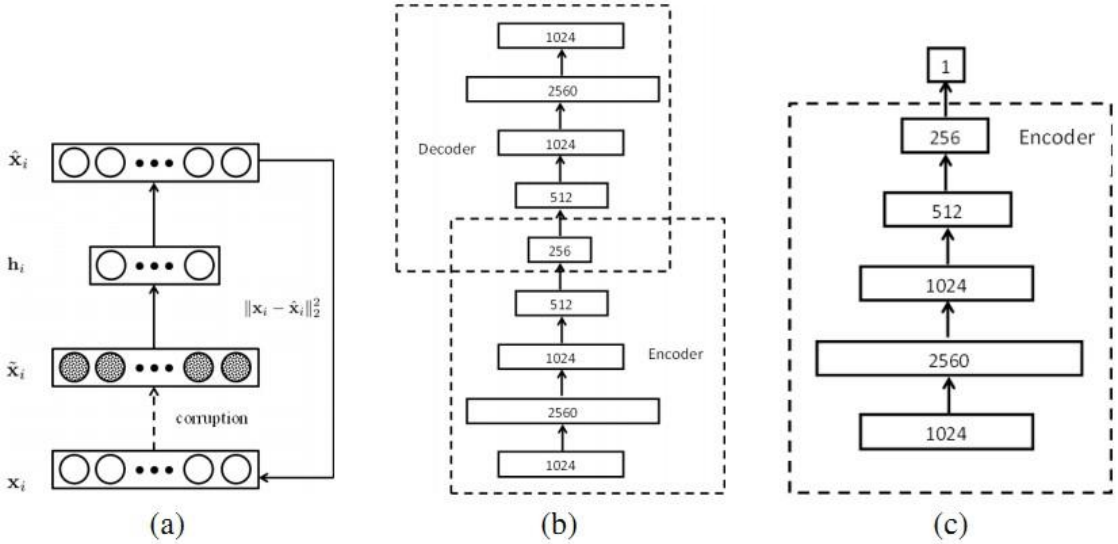
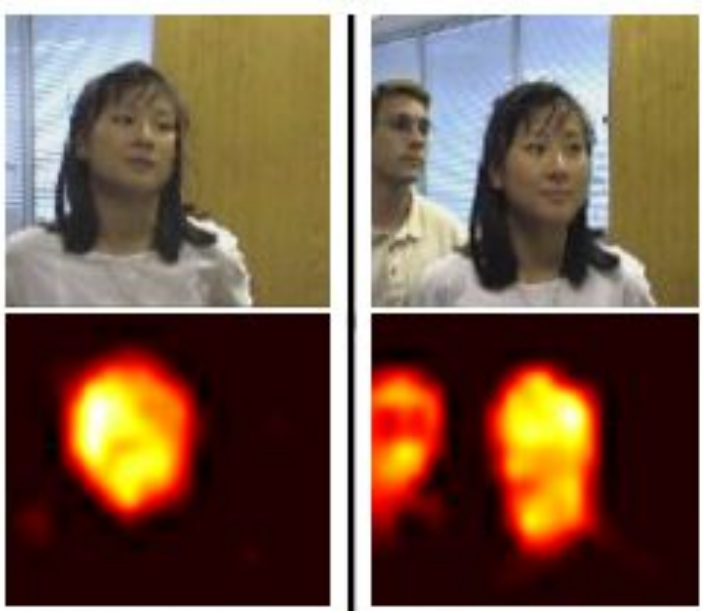


Figure 1: Some key components of the network architecture: (a) denoising autoencoder; (b) stacked denoising autoencoder; (c) network for online tracking.

Wang, Naiyan, and Dit-Yan Yeung. "Learning a deep compact image representation for visual tracking." NIPS 2013 [Project page with code]

Object Tracking: FCNT: Localization

Despite trained for image classification, feature maps in **conv5-3** enable object localization...but are not discriminative enough to discriminate between instances of the same class.

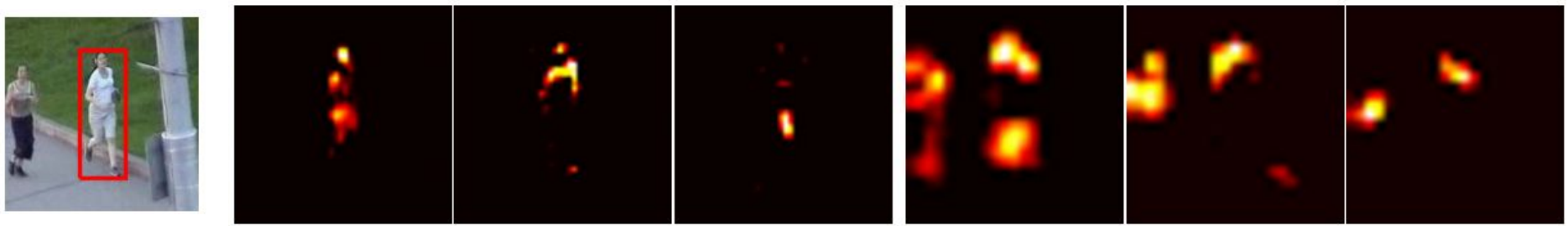


Wang, Lijun, Wanli Ouyang, Xiaogang Wang, and Huchuan Lu. "Visual Tracking with Fully Convolutional Networks." ICCV 2015. [code]

Object Tracking: FCNT: Localization

On the other hand, feature maps from **conv4-3** are more sensitive to intra-class appearance variation…

conv4-3 (specific) conv5-3 (general)



Wang, Lijun, Wanli Ouyang, Xiaogang Wang, and Huchuan Lu. "Visual Tracking with Fully Convolutional Networks." ICCV 2015. [code]

Object Tracking: FCNT: Localization

SNet=Specific Network (online update)

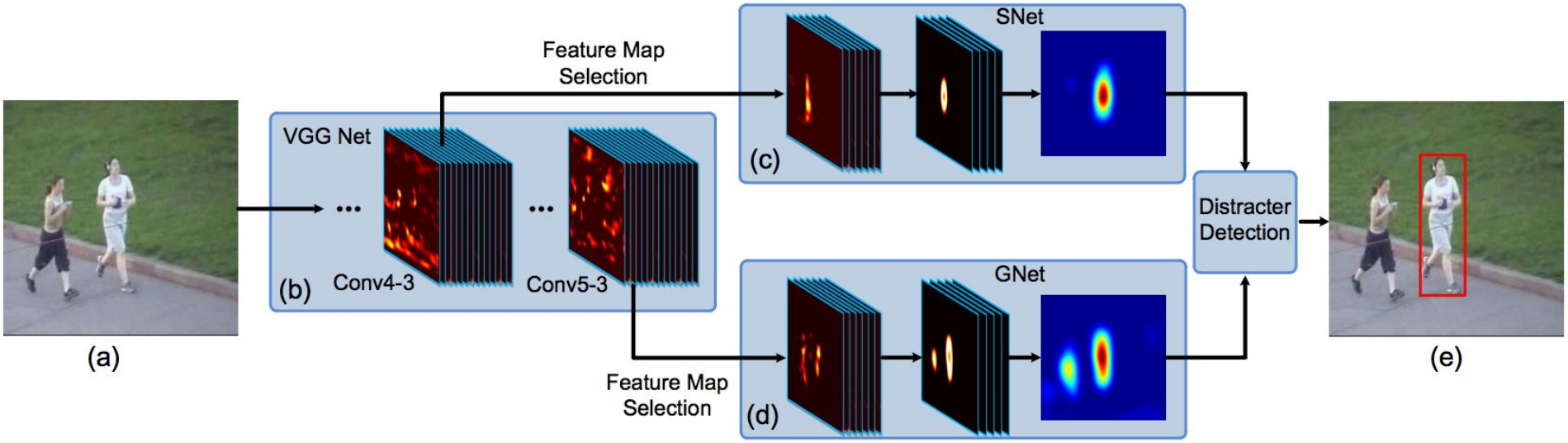
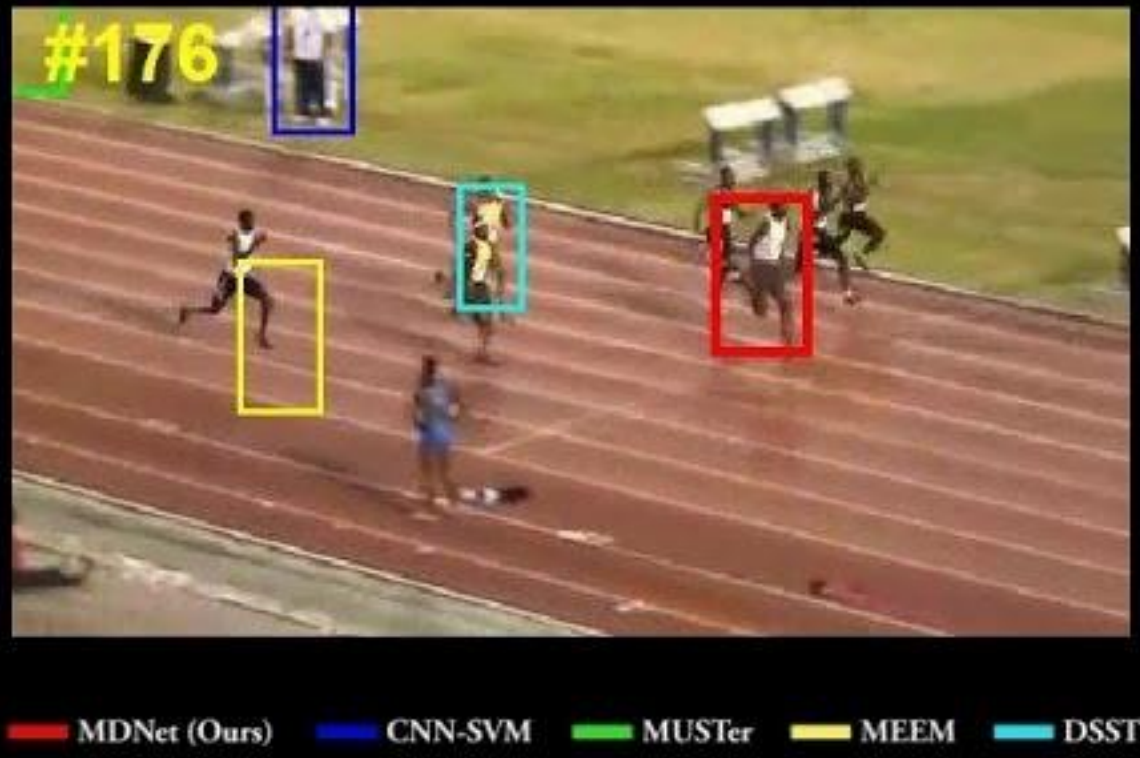


Figure 5. Pipeline of our algorithm. (a) Input ROI region. (b) VGG network. (c) SNet. (d) GNet. (e) Tracking results.

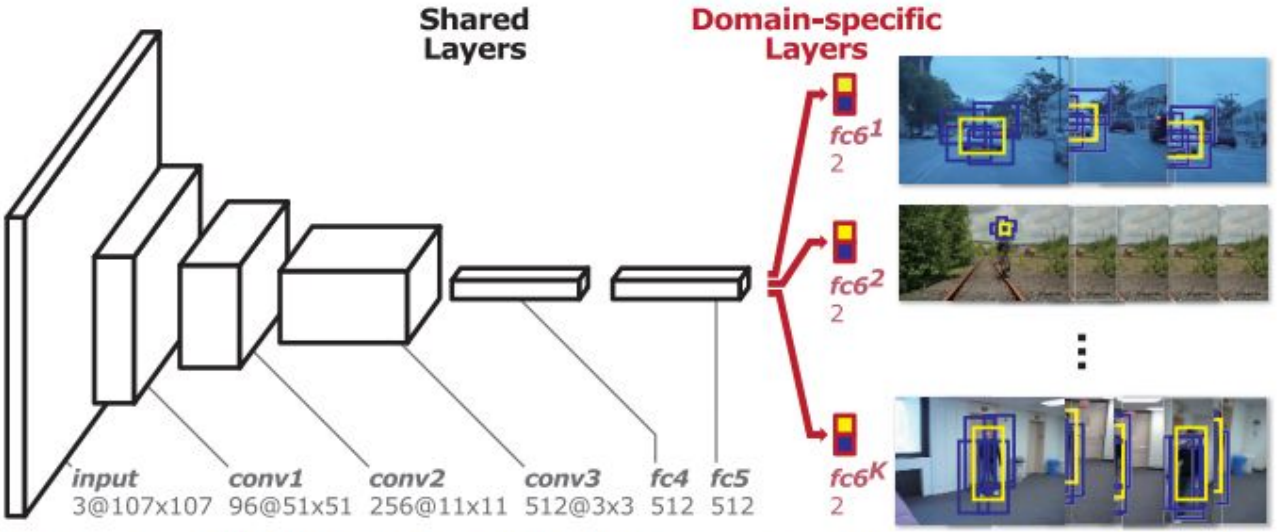
GNet=General Network (fixed)

Wang, Lijun, Wanli Ouyang, Xiaogang Wang, and Huchuan Lu. "Visual Tracking with Fully Convolutional Networks." ICCV 2015. [code]



Nam, Hyeonseob, and Bohyung Han. "Learning multi-domain convolutional neural networks for visual tracking." CVPR 2016.

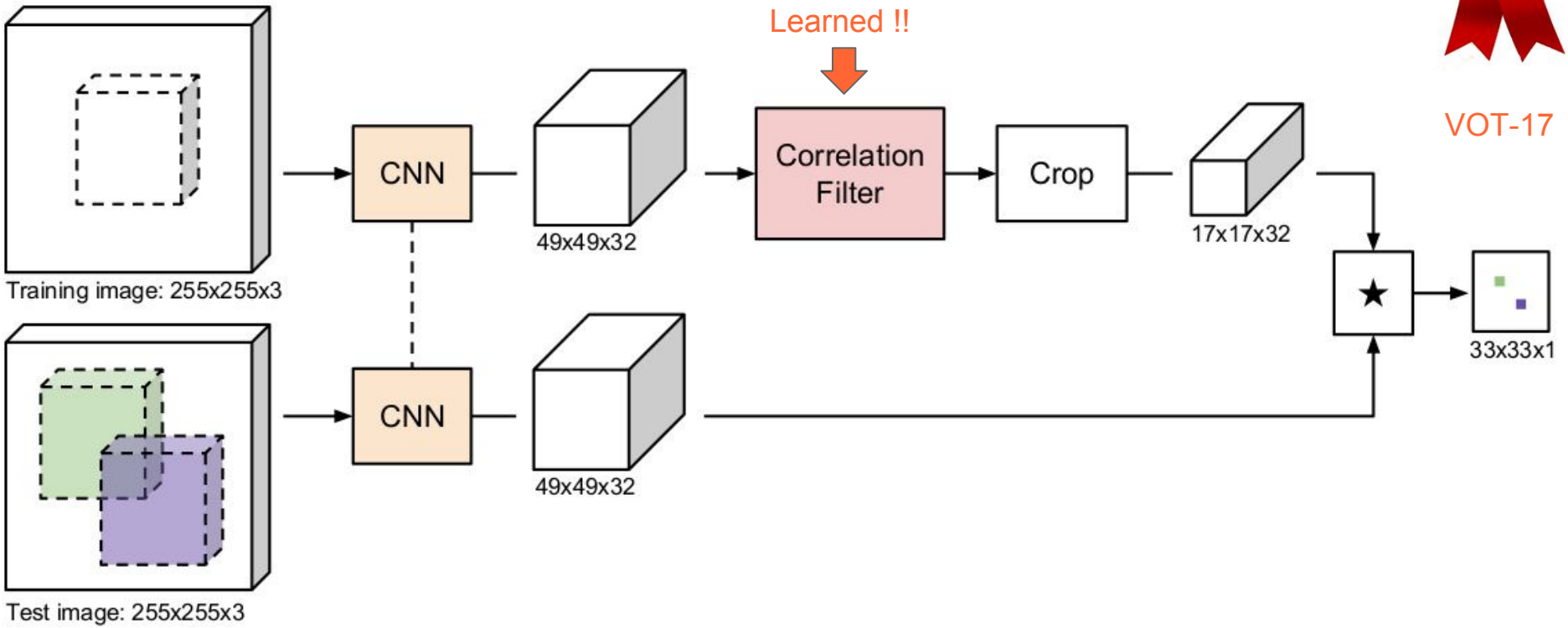
Object Tracking: MDNet



Nam, Hyeonseob, and Bohyung Han. "Learning multi-domain convolutional neural networks for visual tracking." CVPR 2016.

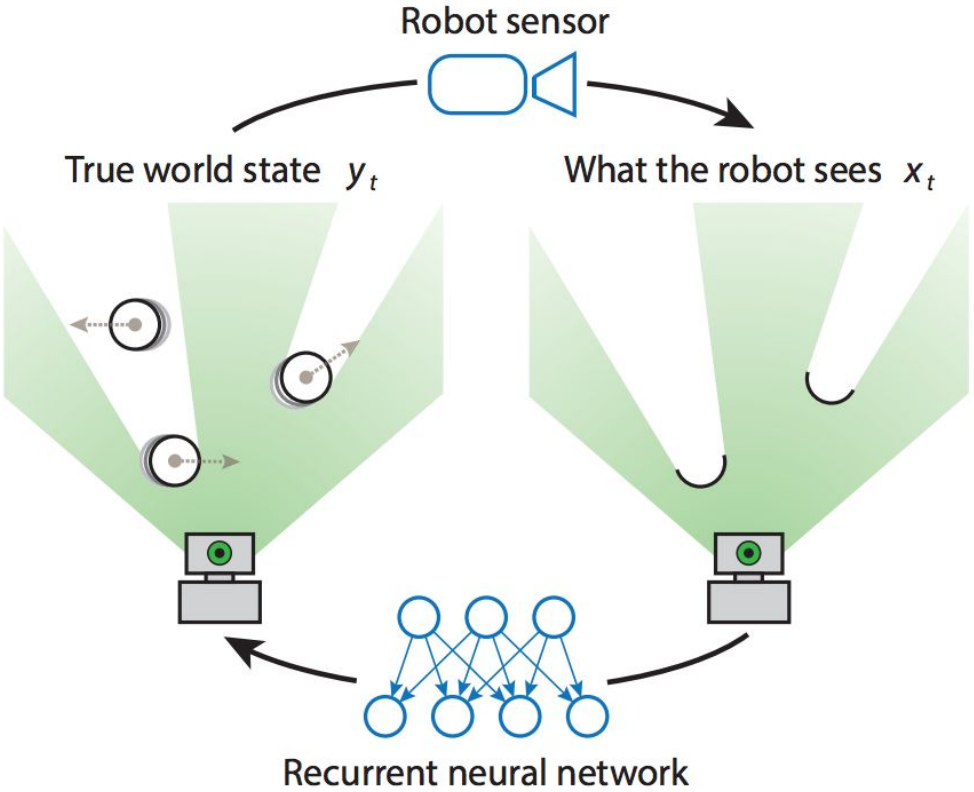
Object Tracking: CFNET

VOT-17

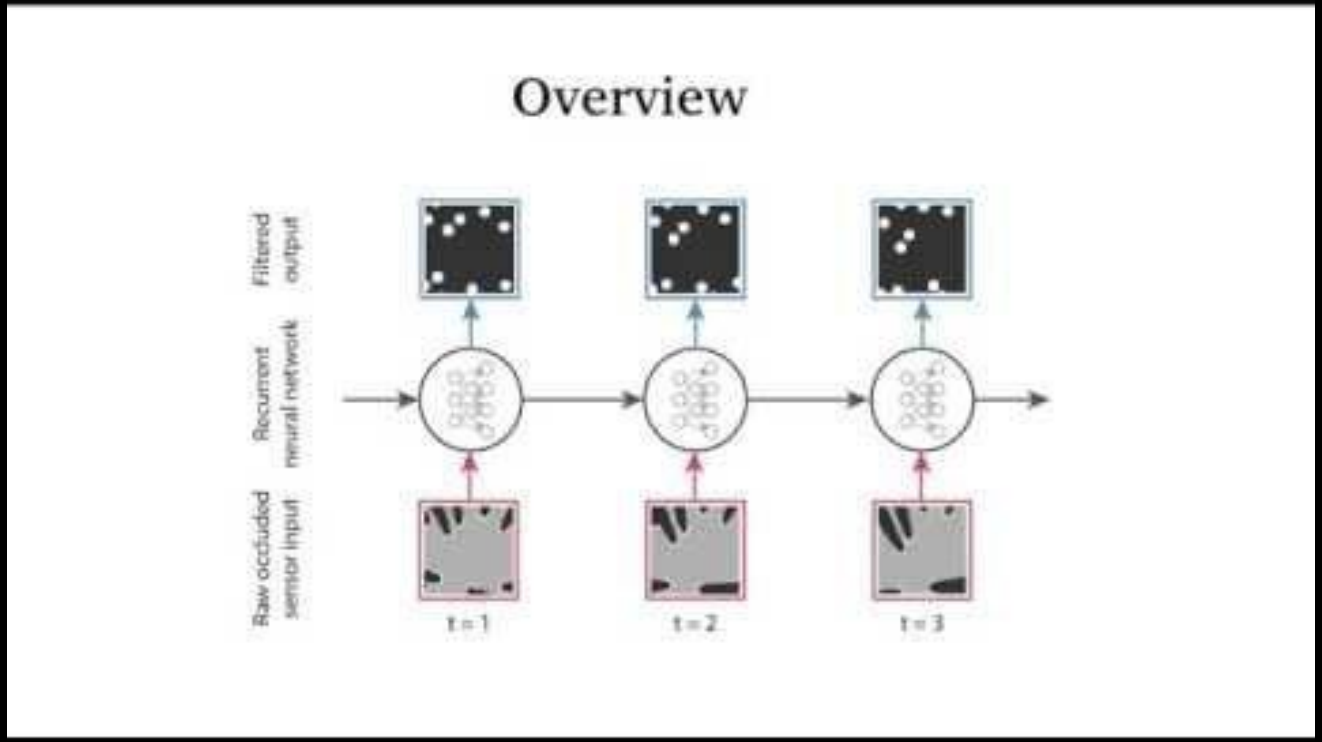


Valmadre, Jack, Luca Bertinetto, João F. Henriques, Andrea Vedaldi, and Philip HS Torr. "End-to-end representation learning for Correlation Filter based tracking." CVPR 2017

Object Tracking: RNN

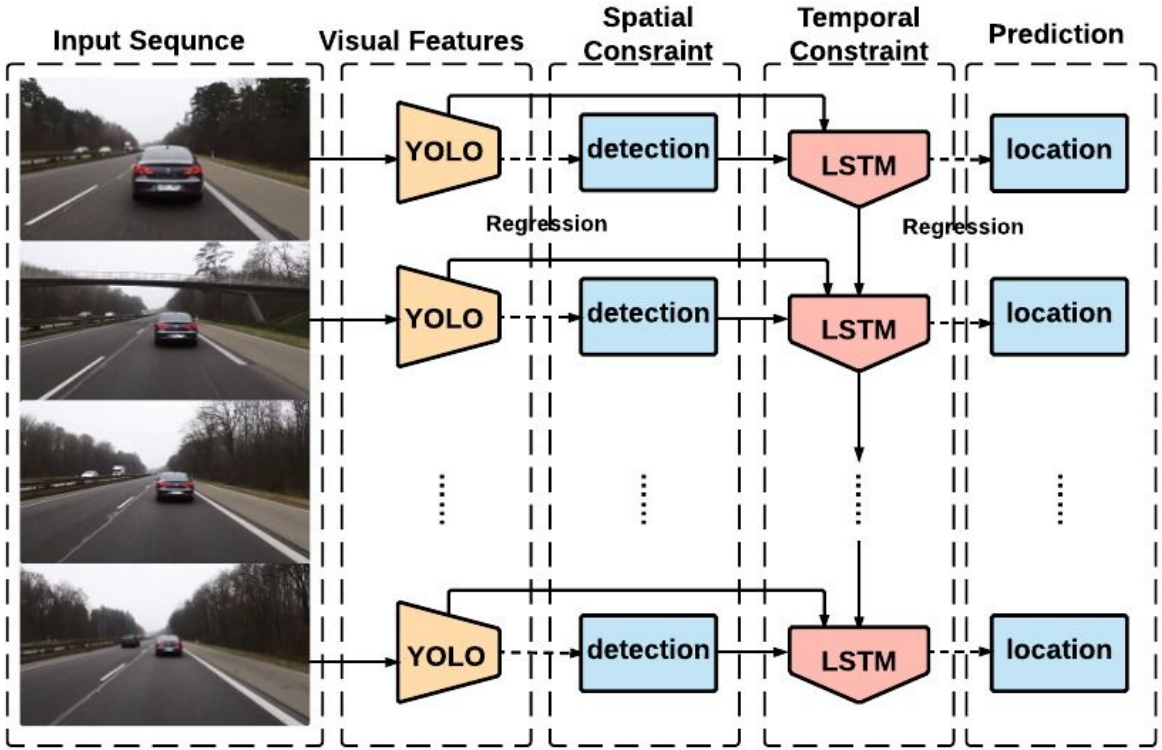


Wang, Naiyan, and Dit-Yan Yeung. "Learning a deep compact image representation for visual tracking." NIPS 2013
[Project page with code]

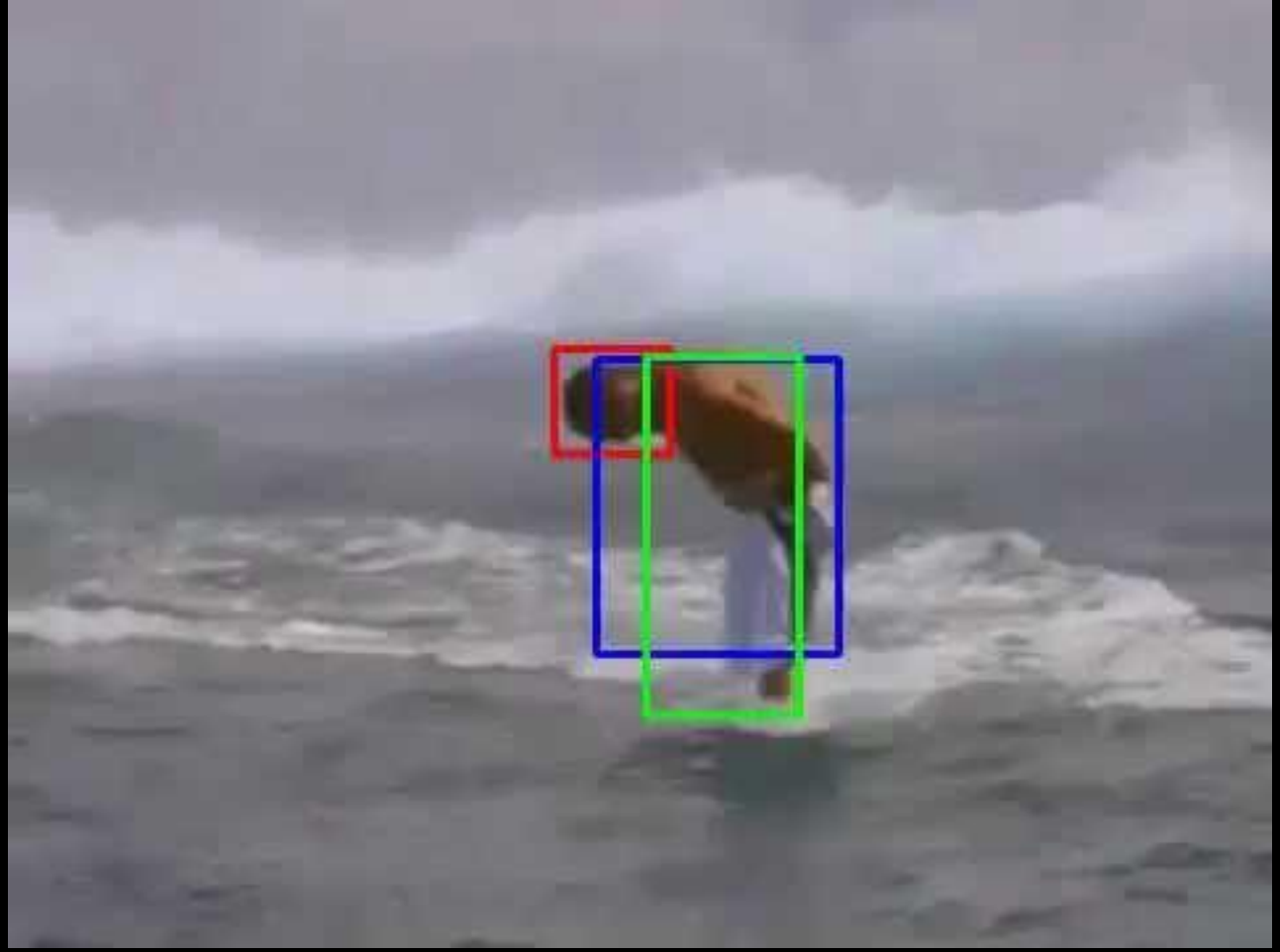


P. Ondruska and I. Posner, "Deep Tracking: Seeing Beyond Seeing Using Recurrent Neural Networks," AAAI 2016 [code]

Object Tracking: ROLO

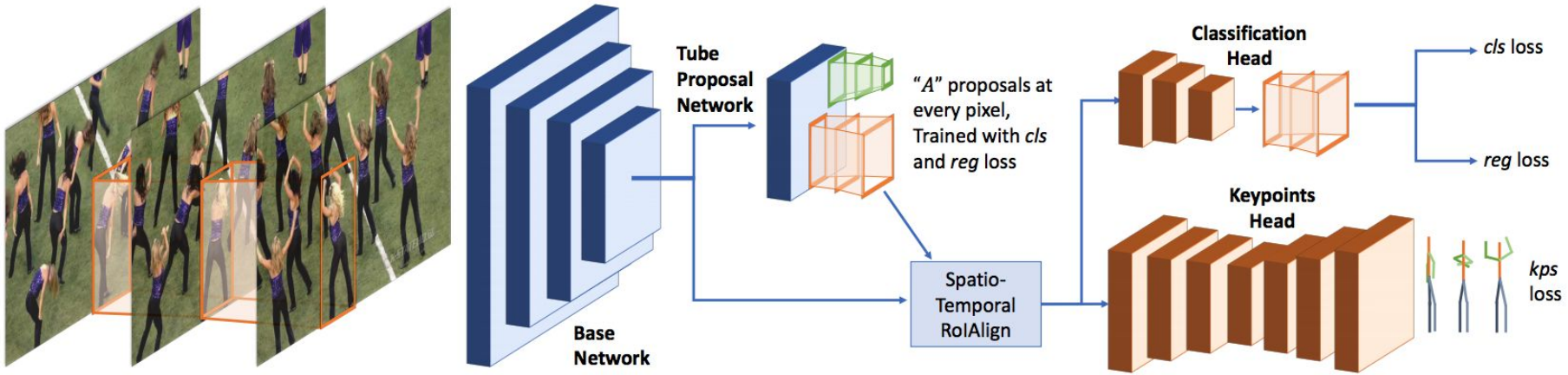


Ning, Guanghan, Zhi Zhang, Chen Huang, Zhihai He, Xiaobo Ren, and Haohong Wang. "Spatially Supervised Recurrent Convolutional Neural Networks for Visual Object Tracking." ISCAS 2017.



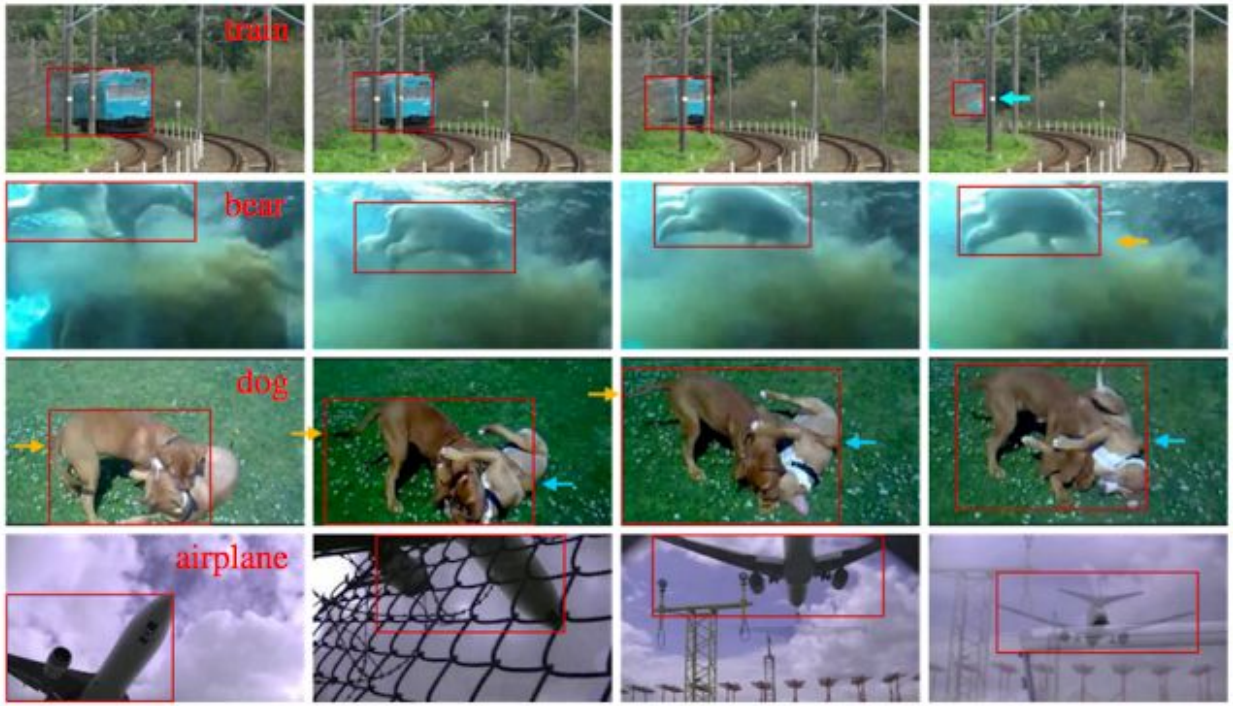
Ning, Guanghan, Zhi Zhang, Chen Huang, Zhihai He, Xiaobo Ren, and Haohong Wang. "Spatially Supervised Recurrent Convolutional Neural Networks for Visual Object Tracking." ISCAS 2017.

Keypoint tracking



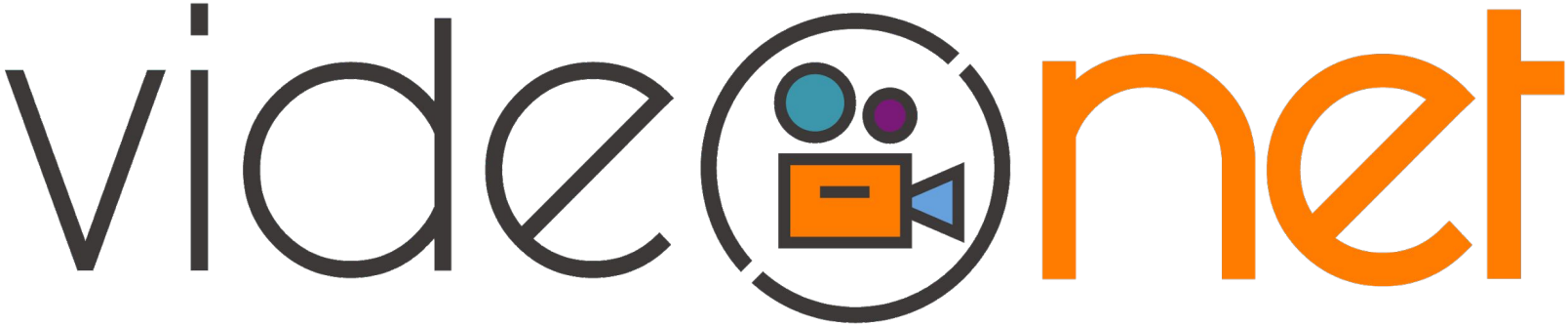
Girdhar, Rohit, Georgia Gkioxari, Lorenzo Torresani, Manohar Paluri, and Du Tran. "Detect-and-Track: Efficient Pose Estimation in Videos." CVPR 2018.

Object tracking: Datasets: YouTube-BB



Real, Esteban, Jonathon Shlens, Stefano Mazzocchi, Xin Pan, and Vincent Vanhoucke. "Youtube-boundingboxes: A large high-precision human-annotated data set for object detection in video." CVPR 2017.

Object Tracking: Challenges



Leal-Taixé, Laura, Anton Milan, Konrad Schindler, Daniel Cremers, Ian Reid, and Stefan Roth. "Tracking the trackers: an analysis of the state of the art in multiple object tracking."

Deep Learning online courses by UPC:

- [MSc course](#) (2017)
- [BSc course](#) (2018)

Next edition Autumn 2018

- [1st edition](#) (2016)
- [2nd edition](#) (2017)
- [3rd edition](#) (2018)

Summer School (late June 2018)

- [1st edition](#) (2017)
- [2nd edition](#) (2018)

Next edition Winter/Spring 2019