CSE463: Neural Networks “Computer Vision and Image Processing: Recent Advances and Trends”

Deep Learning

by:
Hossam El Din Hassan Abd El Munim
حسام الدين حسن عبد المنعم
Computer & Systems Engineering Dept.,
Ain Shams University,
1 El-Sarayat Street, Abbassia, Cairo 11517

Essential Books

1. **Ian Goodfellow and Yoshua Bengio and Aaron Courville, Deep Learning, MIT Press, 2016.**

2. **I Charu C. Aggarwal, Neural Networks and Deep Learning, Springer, 2018.**

3. **I Michael Nielsen, Neural Networks and Deep Learning. http://neuralnetworksanddeeplearning.com/, 2016**

4. Rudolf Kruse, C. Borgelt, F. Klawonn, C. Moewes, M. Steinbrecher, P. Held, Computational Intelligence: A Methodological Introduction, Springer, 2014. (available online)
5. S. Haykin, Neural Networks and Learning Machines, 2009.

Course Contents

- **Basics**: Biological Neuron, Idea of computational units, McCulloch{Pitts unit and Thresholding logic, Linear Perceptron, Perceptron Learning Algorithm, Linear Separability. Convergence theorem for Perceptron Learning Algorithm.
- **Feedforward Networks**: Multilayer Perceptron, Gradient Descent, Backpropagation, Empirical Risk Minimization, regularization, auto-encoders.
- **Deep Neural Networks**: Difficulty of training deep neural networks, Greedy layer-wise training.
- **Better Training of Neural Networks**: Newer optimization methods for neural networks (Adagrad, adadelta, rmsprop, adam, NAG), second order methods for training, Saddle point problem in neural networks, Regularization methods (dropout, drop connect, batch normalization).
- **Recurrent Neural Networks**: Back propagation through time, Long Short Term Memory, Gated Recurrent Units, Bidirectional LSTMs, Bidirectional RNNs
- **Convolutional Neural Networks**: LeNet, AlexNet.
- **Generative models**: Restrictive Boltzmann Machines (RBMs), Introduction to MCMC and Gibbs Sampling, gradient computations in RBMs, Deep Boltzmann Machines.

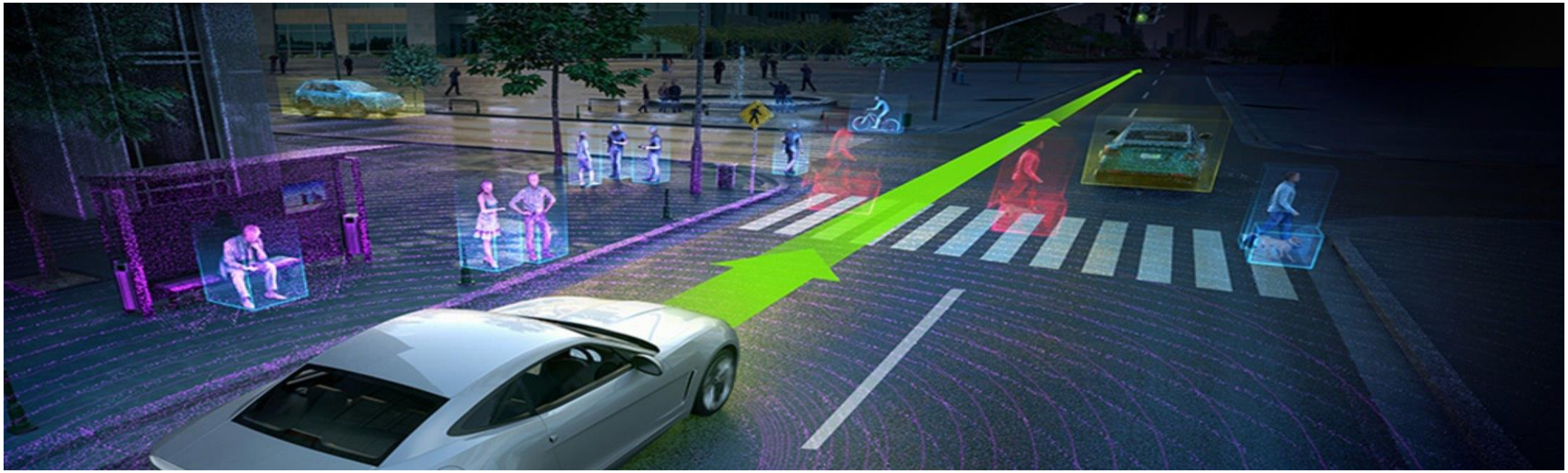
Grading System

- **Final examination 90**
- **Midterm examination 15**
- **Attendance 5**
- **Quizzes (TWO) 5**
- **Project 10**

- **Warnings:**
  - **A quiz may be given without being informed before.**
  - **Copying assignment is prohibited.**
  - **Delay of submission influences on marks.**
  - **No Plagiarism!**

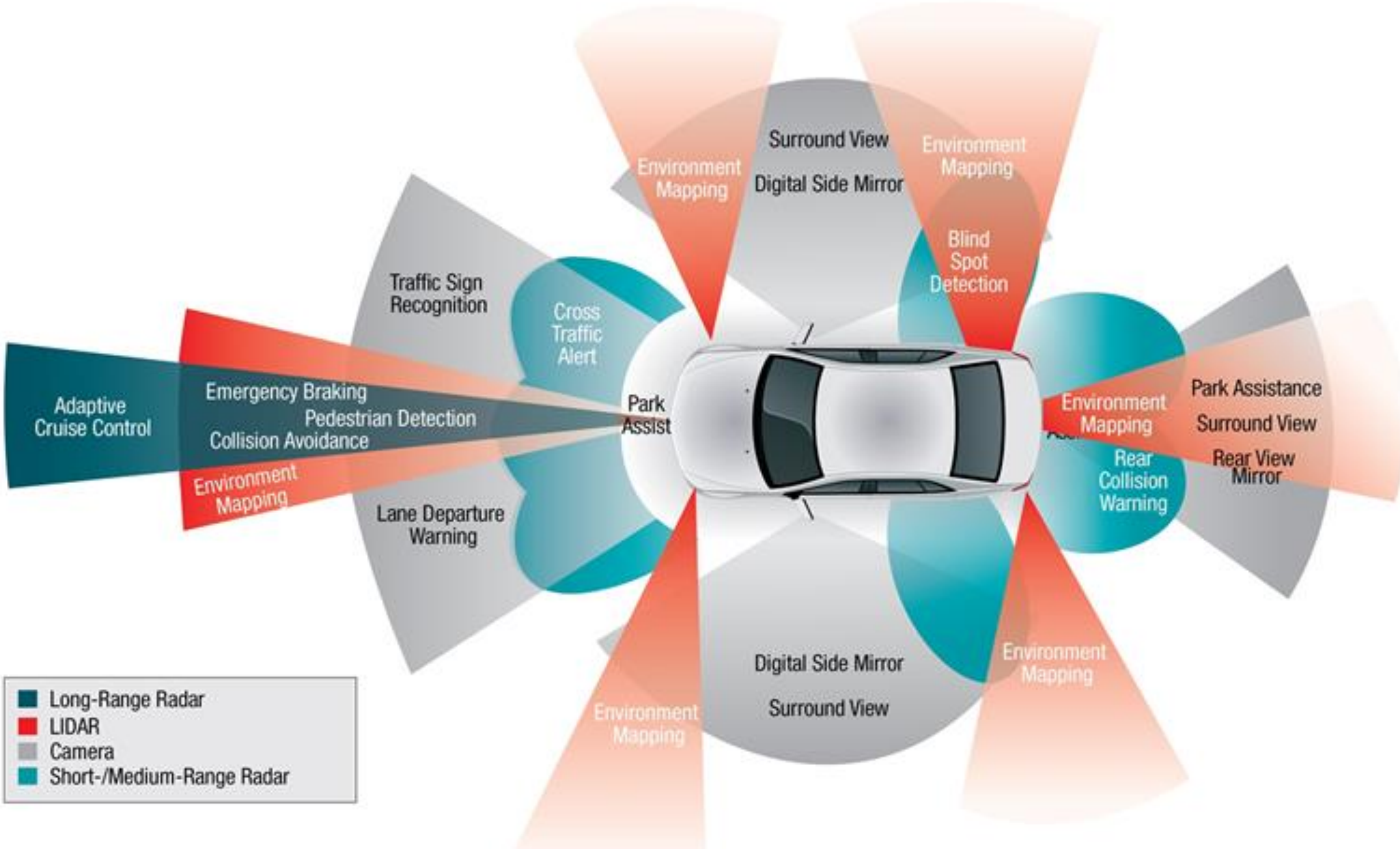
Autonomous Driving

Photo: www.shellypalmer.com



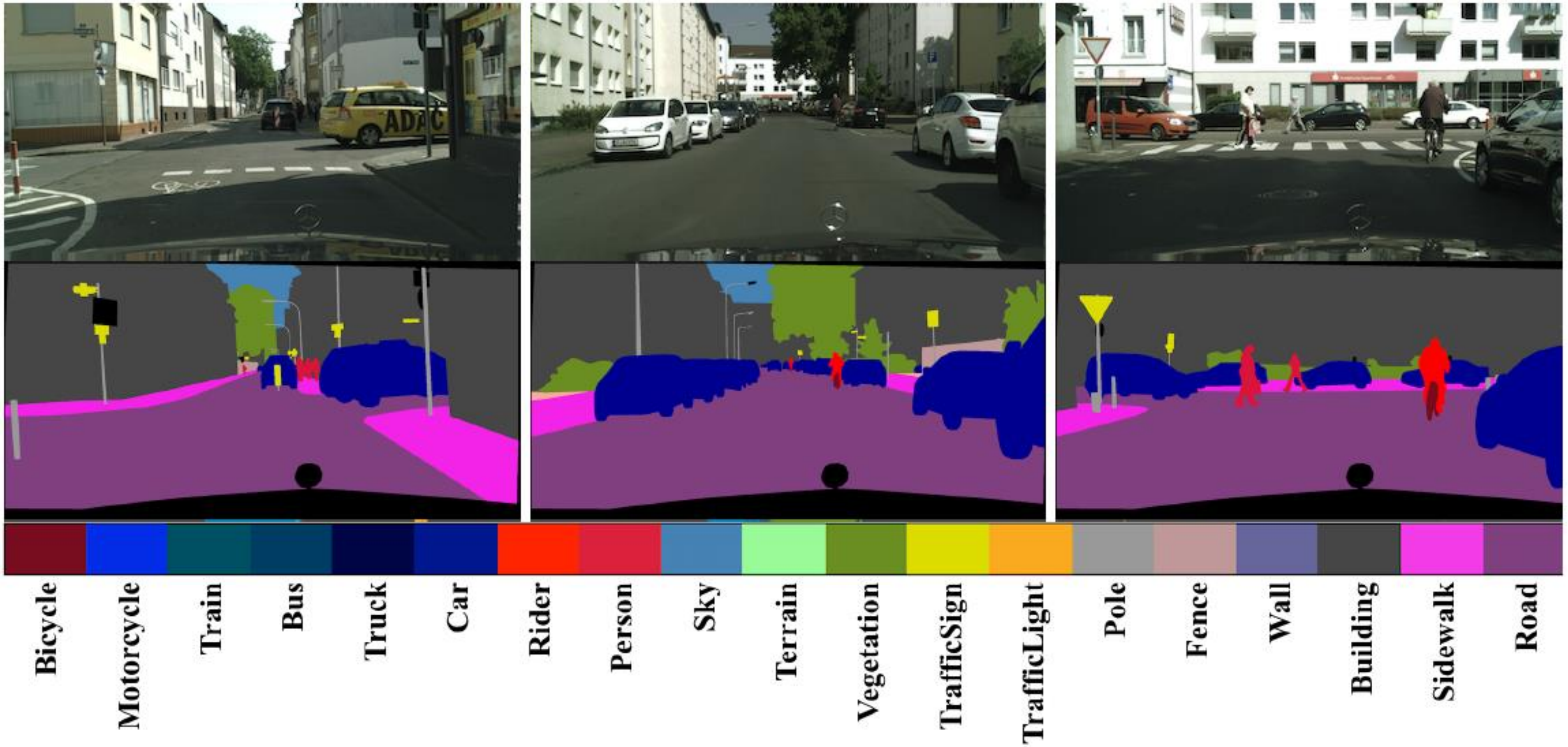
https://emerj.com/ai-sector-overviews/how-self-driving-cars-work/

Sensors Scheme and Functionality



https://test.neurohive.io/en/state-of-the-art/self-driving-cars/

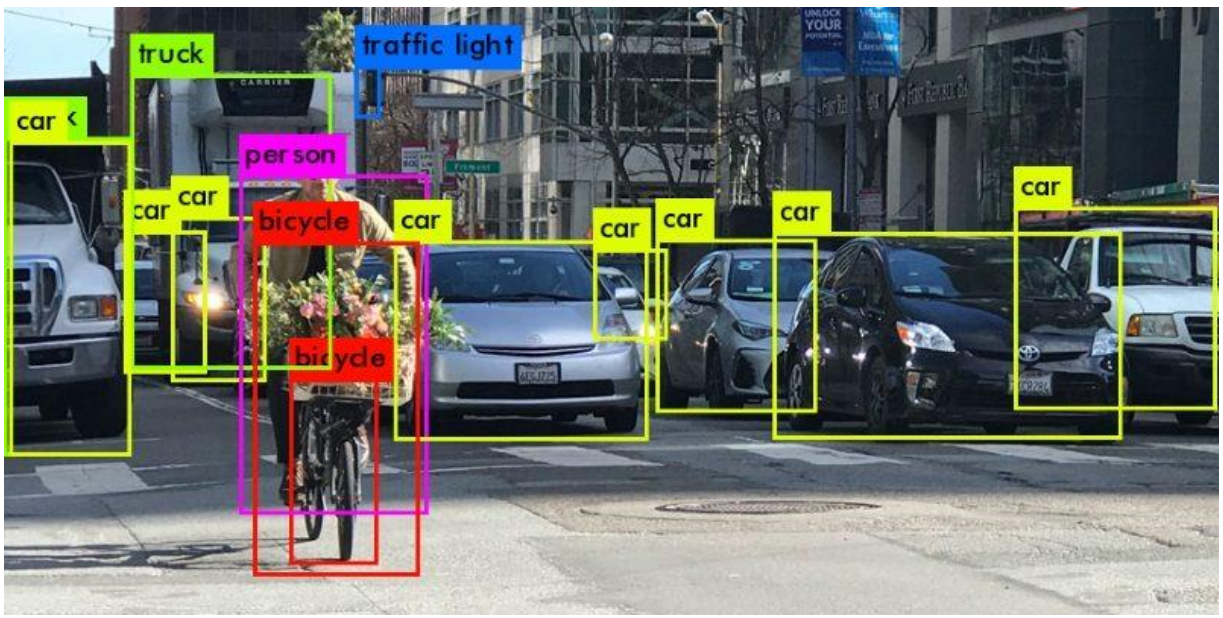
Role of Computer Vision and Image Processing



Taha Emara, Hossam E. Abd El Munim, and Hazem M. Abbas, "LiteSeg: A Novel Lightweight ConvNet for Semantic Segmentation", International Conference on Digital Image Computing: Technqiues and Applications (DICTA), Perth, Australia, Dec, 2019

What is Computer Vision?

- Deals with the development of the **<u>theoretical</u>** and **<u>algorithmic</u>** basis by which useful information about the 3D world can be automatically extracted and analyzed from a **<u>single</u>** or **<u>multiple</u>** 2D images of the world.



https://azati.ai/image-detection-recognition-and-classification-with-machine-learning/

He et al, "Mask R-CNN", ICCV 2017

Computer Vision, Also Known As …

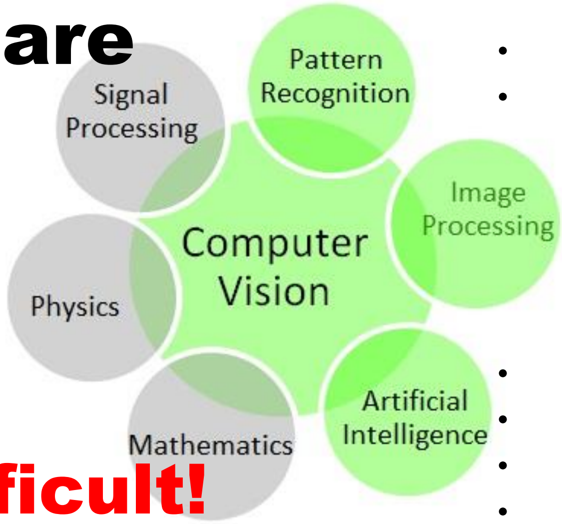
- Image Analysis
- Scene Analysis
- Image Understanding

Applications

- Industrial inspection/quality control
- Surveillance and security
- Face recognition
- Gesture recognition

Some Related Fields are

- Image Processing
- Computer Graphics
- Pattern Recognition
- Robotics
- Artificial Intelligence

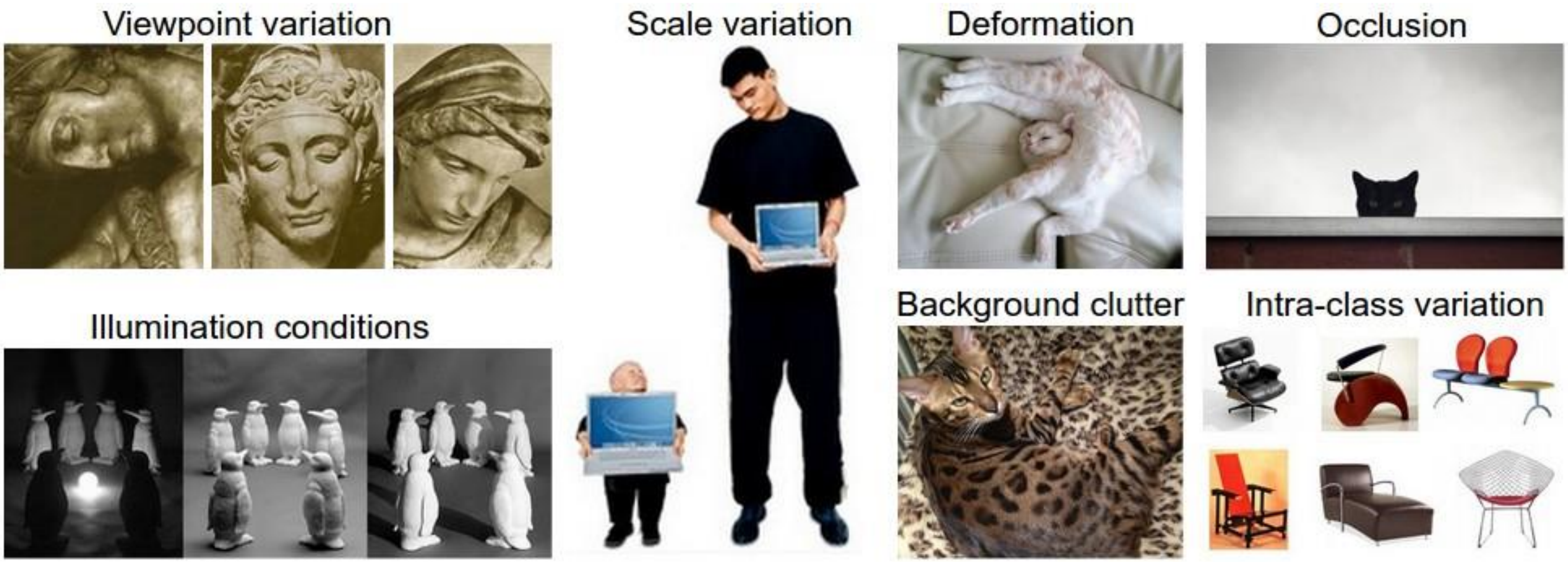


- Space applications
- Medical image analysis
- Autonomous vehicles
- Virtual reality and much more ……

Computer Vision is Difficult!

- **Mapping 2D information to the real world is not unique.**
- **It is a computationally expensive process.**
- **Pose, illumination, scale variations always represent big challenges.**

Visual Recognition/Image Classification Challenges



http://cs231n.github.io/classification/

Top IEEE/Springer/IET Journals

Rank	Publisher	Journal Details
10	IEEE	IEEE Transactions on Pattern Analysis and Machine Intelligence ISSN:0162-8828 , Monthly
44	IEEE	IEEE Transactions on Image Processing ISSN:1057-7149 , Monthly
51	IEEE	IEEE Transactions on Information Forensics and Security ISSN:1556-6013 , Monthly
63	IEEE	IEEE Transactions on Robotics ISSN:1552-3098 , Bimonthly
68	IEEE	IEEE Transactions on Medical Imaging ISSN:0278-0062 , Monthly
83	IEEE	IEEE Transactions on Circuits and Systems for Video Technology ISSN:1051-8215 , Monthly
104	IEEE	IEEE Robotics and Automation Magazine ISSN:1070-9932 , Quarterly
139	IEEE	IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing ISSN:1939-1404 , Monthly
156	IEEE	IEEE Geoscience and Remote Sensing Letters ISSN:1545-598X , Monthly
193	IEEE	IEEE Transactions on Audio, Speech and Language Processing ISSN:1558-7916 , Monthly

Rank	Publisher	Journal De
11	Springer	International Journal of Computer Vision ISSN:0920-5691 , Monthly
258	Springer	Journal of Real-Time Image Processing ISSN:1861-8200 , Quarterly
261	Springer	Machine Vision and Applications ISSN:0932-8092 , Bimonthly
265	Springer	Journal of Mathematical Imaging and Vision ISSN:0924-9907 , Monthly
269	Springer	Eurasip Journal on Advances in Signal Processing ISSN:1687-6180 , Irregular

Rank	Publisher	Journal De
431	IET The Institution of Engineering and Technology	IET Signal Processing ISSN:1751-9675 , Bimonthly
494	IET The Institution of Engineering and Technology	IET Image Processing ISSN:1751-9659 , Monthly
533	IET The Institution of Engineering and Technology	IET Computer Vision ISSN:1751-9632 , Bimonthly

Top Conferences

Hindex		Publisher	Conference Details
1	158	IEEE	CVPR : IEEE Conference on Computer Vision and Pattern Recognition, CVPR Jun 15, 2019 - Jun 21, 2019 - Long Beach , United States http://cvpr2019.thecvf.com/
5	89	IEEE	ICCV : IEEE International Conference on Computer Vision Oct 27, 2019 - Nov 3, 2019 - Seoul , South Korea http://iccv2019.thecvf.com/
36	50	IEEE	IROS : IEEE/RSJ International Conference on Intelligent Robots and Systems Nov 3, 2019 - Nov 8, 2019 - Macao , China http://www.iros2019.org/
94	34	IEEE	ICIP : IEEE International Conference on Image Processing Sep 22, 2019 - Sep 25, 2019 - Taipei , Taiwan http://2019.ieeeicip.org/
137	28	IEEE	FG : IEEE International Conference on Automatic Face & Gesture Recognition May 14, 2019 - May 18, 2019 - Lille , France http://fg2019.org/
146	27	IEEE	BTAS : IEEE International Conference on Biometrics: Theory Applications and Systems (BTAS) Sep 10, 2018 - Sep 13, 2018 - Los Angeles , United States https://www.isi.edu/events/btas2018/home
170	25	IEEE	ICME : IEEE International Conference on Multimedia and Expo Jul 23, 2018 - Jul 23, 2018 - San Diego , United States http://www.icme2018.org/important_dates

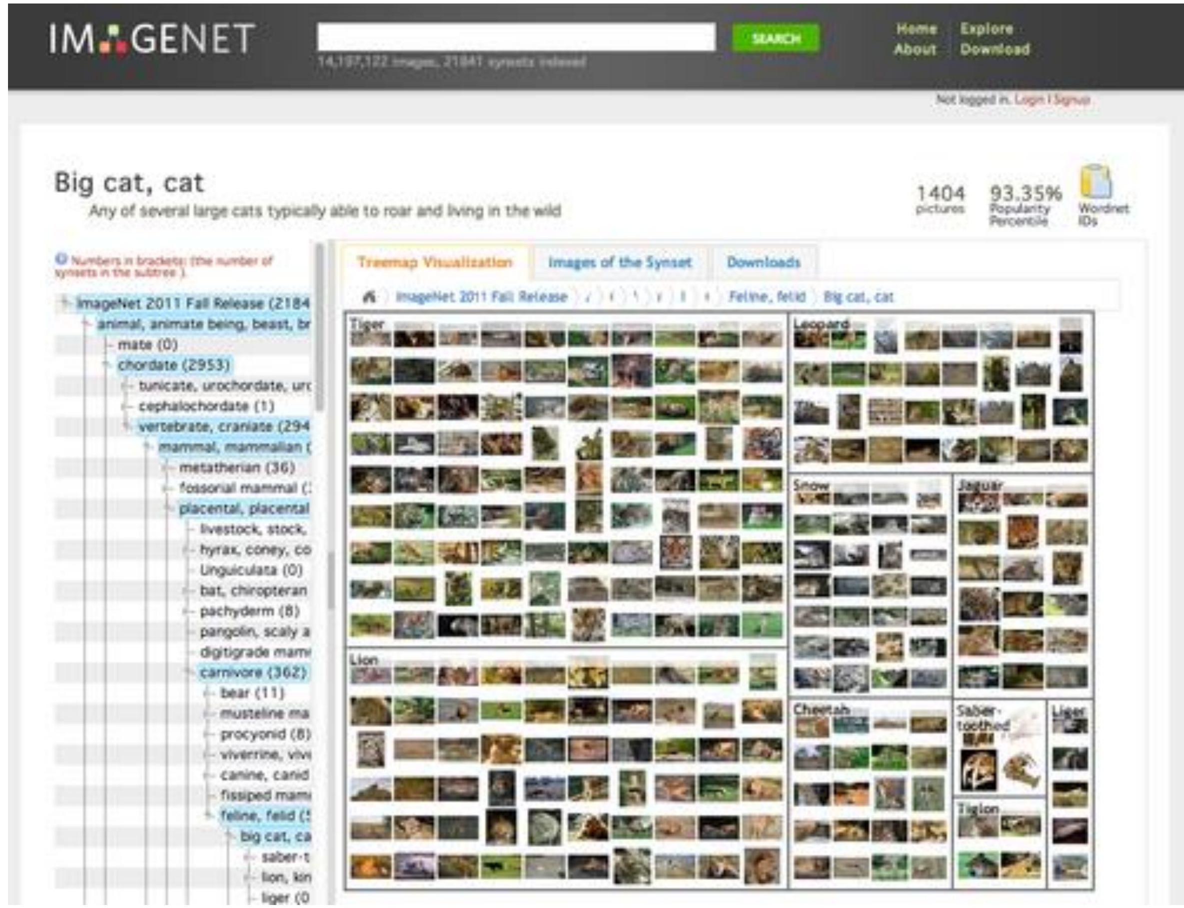
Hindex		Publisher	Conference Details
3	98	Springer	ECCV : European Conference on Computer Vision Sep 8, 2018 - Sep 14, 2018 - Munich , Germany https://eccv2018.org/
55	43	Springer	BMVC : British Machine Vision Conference Sep 5, 2018 - Sep 8, 2018 - Northumbria University , United Kingdom https://www.northumbria.ac.uk/about-us/news-events/events/2018/09/british-machine-vision-conference/
88	35	Springer	ACCV : Asian Conference on Computer Vision Dec 2, 2018 - Dec 6, 2018 - Perth Western , Australia http://accv2018.net/
98	34	Springer	MICCAI : Medical Image Computing and Computer Assisted Intervention Sep 16, 2018 - Sep 20, 2018 - Granada , Spain http://www.miccai2018.org/en/
319	15	Springer	ISVC : International Symposium on Visual Computing Dec 12, 2016 - Dec 14, 2016 - Las Vegas , United States http://www.isvc.net
365	13	Springer	ICIAP : International Conference on Image Analysis and Processing Sep 11, 2017 - Sep 15, 2017 - Catania , Italy http://www.iciap2017.com/

Machine Learning Problems

	Supervised Learning	Unsupervised Learning
Discrete	classification or categorization	clustering
Continuous	regression	dimensionality reduction

ImageNet

- Images for each category of WordNet
- 1000 classes
- 1.2mil images
- 100k test



The machine learning framework

- Apply a prediction function to a feature representation of the image to get the desired output:

f() = “apple”

f() = “tomato”

f() = “cow”

Slide credit: L. Lazebnik

The machine learning framework

$$\mathbf{f}(\mathbf{x}) = \mathbf{y}$$

Prediction function or *classifier*

Image feature

Output (label)

Training: Given a *training set* of labeled examples:

$$\{(\mathbf{x}_1,y_1), \ldots, (\mathbf{x}_N,y_N)\}$$

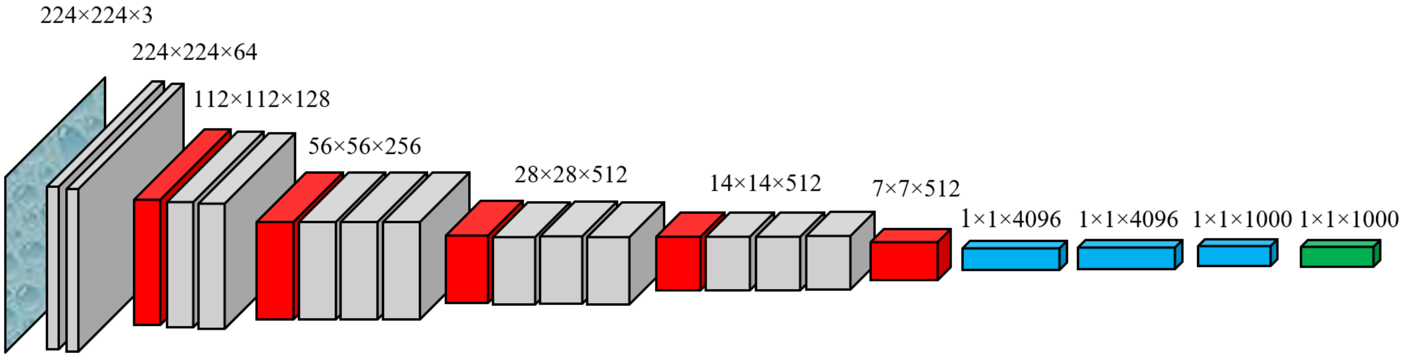
Estimate the prediction function f by minimizing the prediction error on the training set.

Testing: Apply f to a unseen *test example* $\mathbf{x}_u$ and output the predicted value $y_u = f(\mathbf{x}_u)$ to *classify* $\mathbf{x}_u$.

Slide credit: L. Lazebnik

State of the Art

- With enough training data, computer vision nearly matches human vision at most recognition tasks
- Deep learning has been an enormous disruption to the field. Many techniques are being “deepified”.
- The world of computer vision changed when deep learning arrived. For most of the computer vision tasks, deep learning models (Convolutional Neural Network) were built and trained which started outpacing the counter-part old machine learning methods.
- Applications are:Image classification, Object detection, Object tracking, Pose estimation, Text detection & recognition, Visual saliency detection, Action recognition, Scene labelling

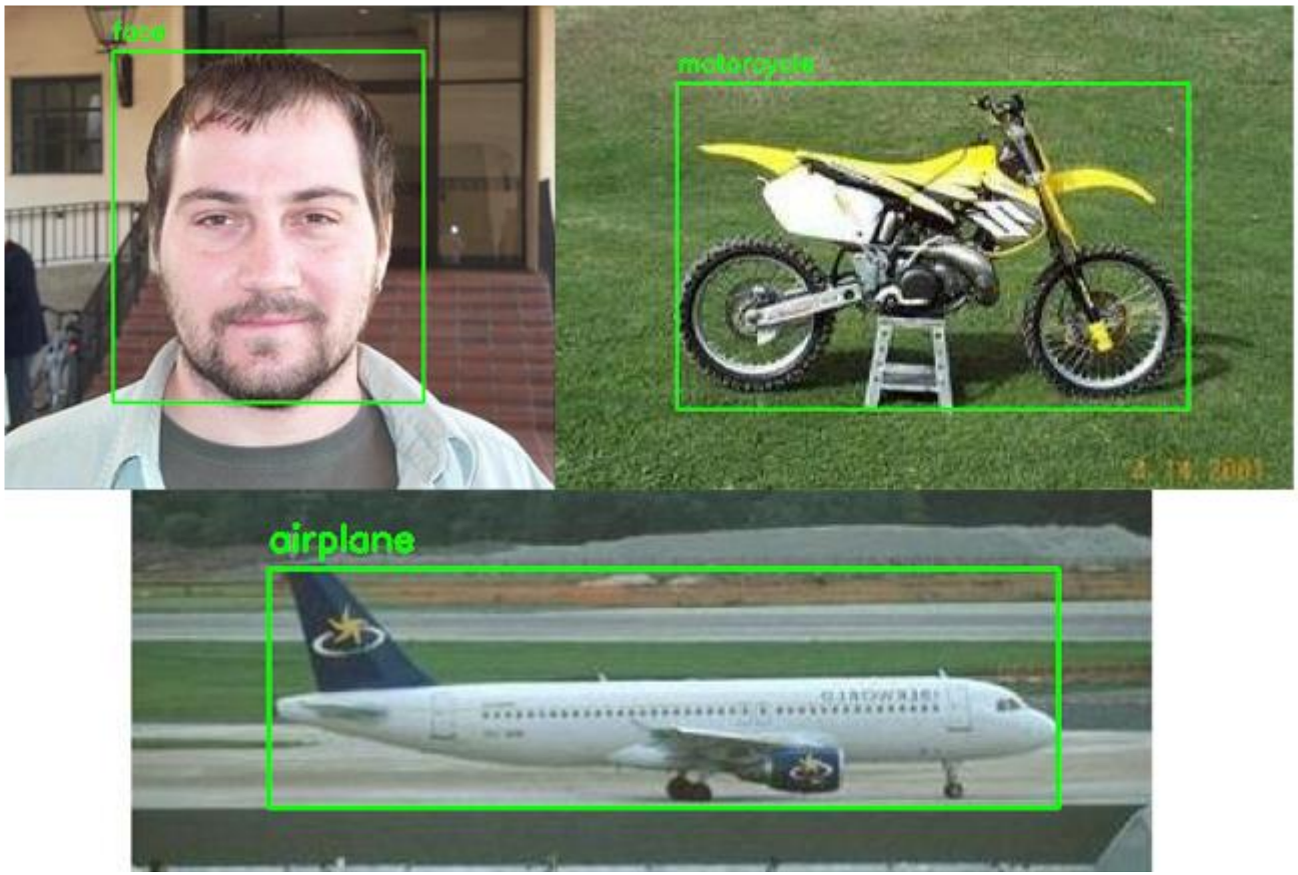


VGG19

Microsoft COCO: Common Objects in Context

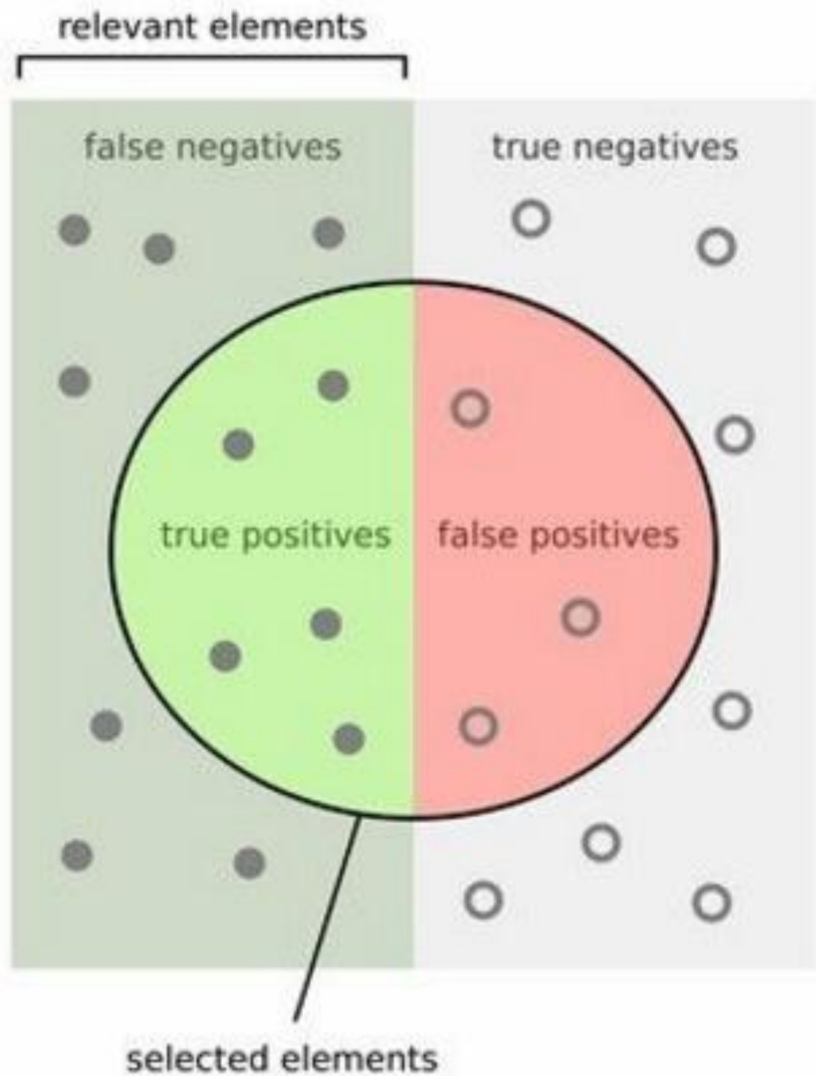
Object Detection

Regression/Classification and Object Detection

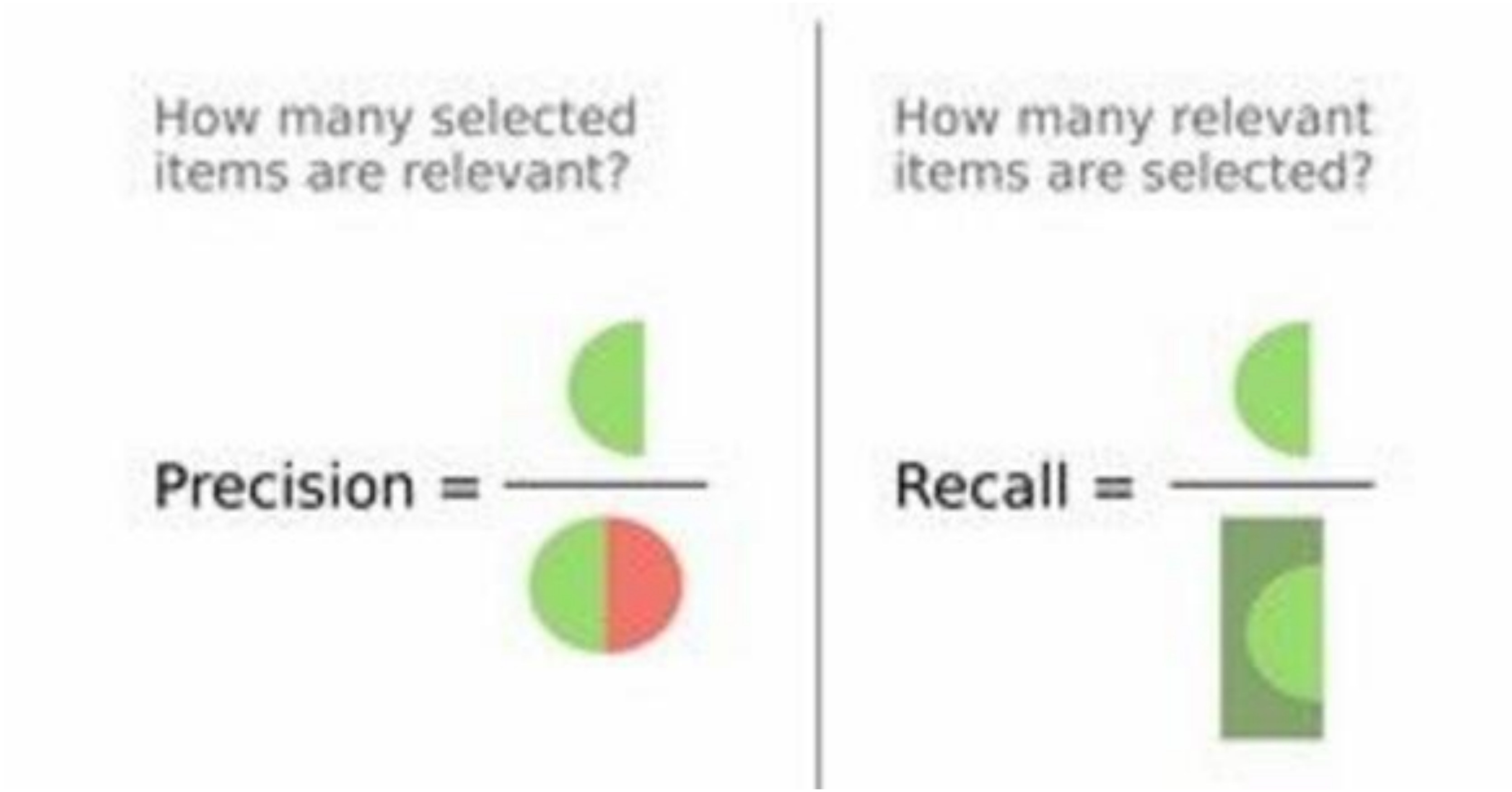


It aims to find the box parameters which are considered analog.

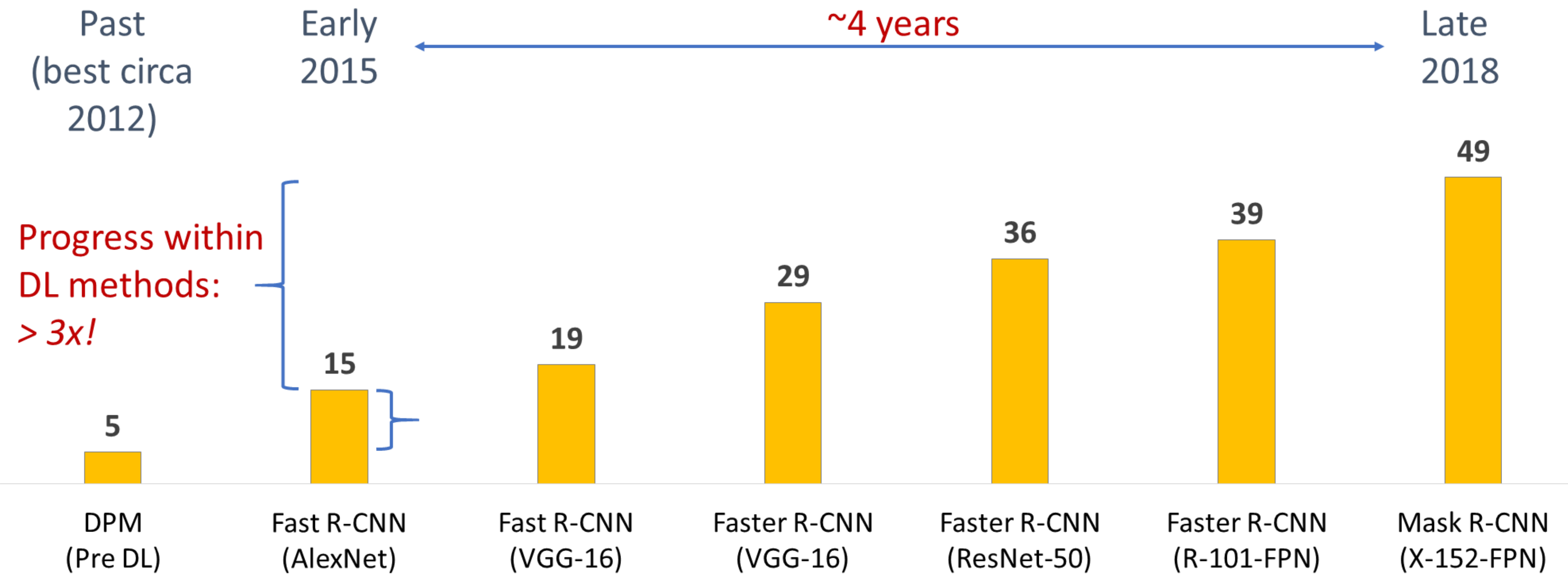
TP/FP


relevant elements
false negatives
true negatives
true positives
false positives
selected elements

Precision


How many selected items are relevant?
Precision =
How many relevant items are selected?
Recall =

COCO Object Detection Average Precision (%)



Microsoft COCO: Common Objects in Context

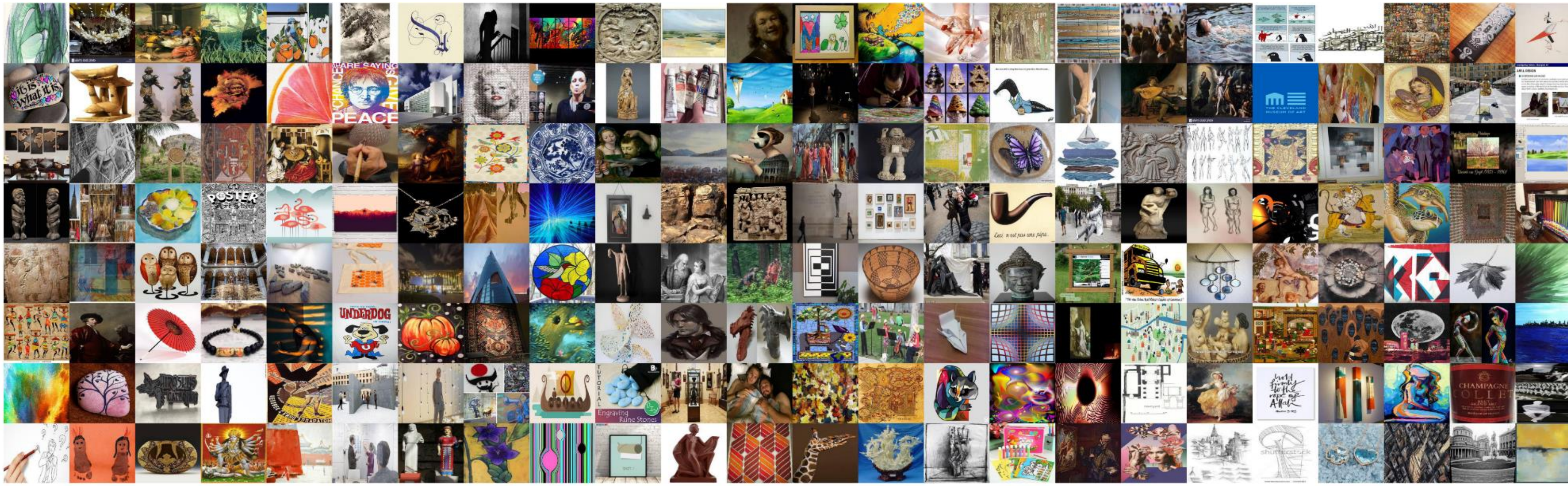
Why These Improvements in Performance?

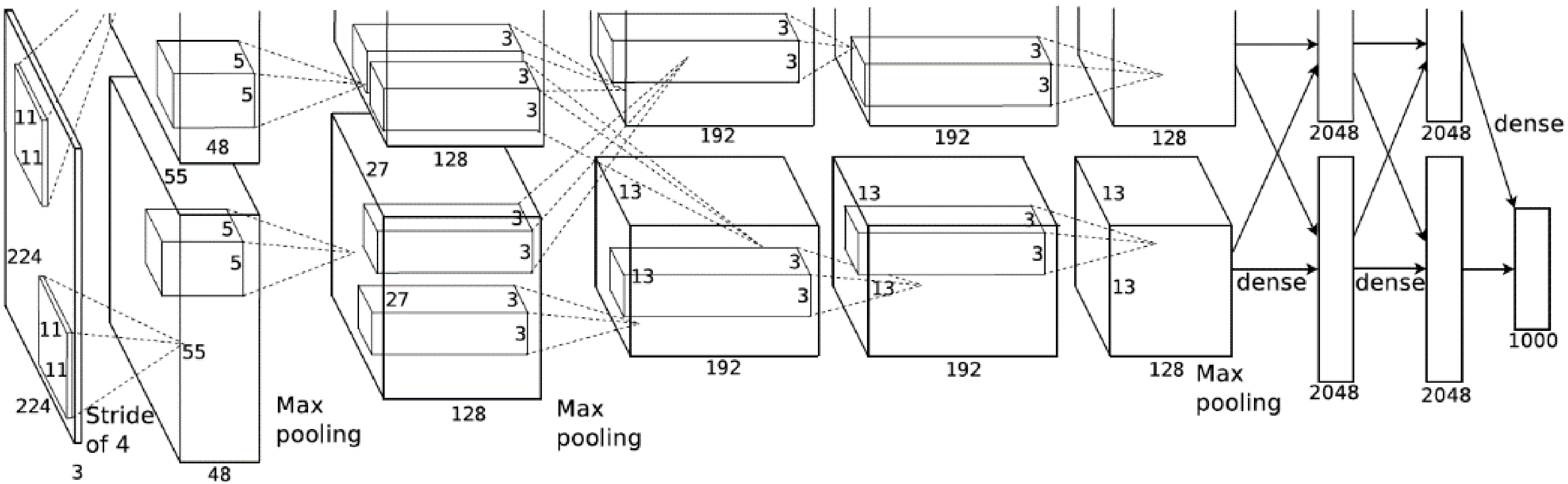
- Features are learned rather than hand-crafted.
- More layers capture more invariances[1].
- More data to train deeper networks.
- More computing (GPUs).
- Better regularization: Dropout.
- New nonlinearities (Max pooling, Rectified linear units (ReLU)).
- Better optimization techniques[2].
- Several learning libraries have emerged as winners which provide a lot of support and convenience to train deep learning models for visual recognition and other visual tasks

[1] Razavian, Azizpour, Sullivan, Carlsson, CNN Features off-the-shelf: an Astounding Baseline for Recognition. CVPRW’14.
[2] Diederik P. Kingma and Jimmy Lei Ba. Adam: a Method for Stochastic Optimization. International Conference on Learning Representations, pages 1–13, 2015.

The way you see the world will be changed by the following CNN-based vision applications.

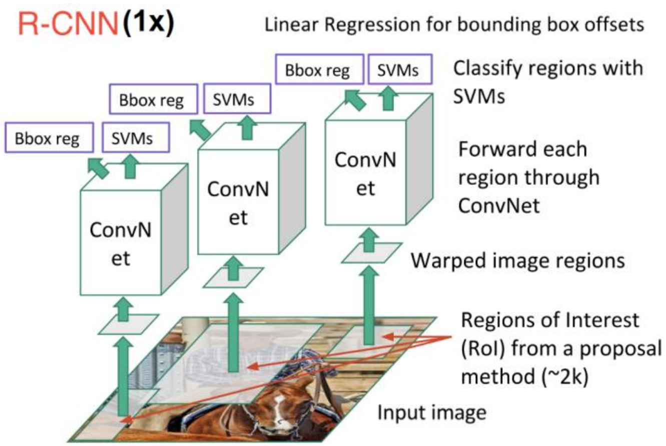
1- Image Classification

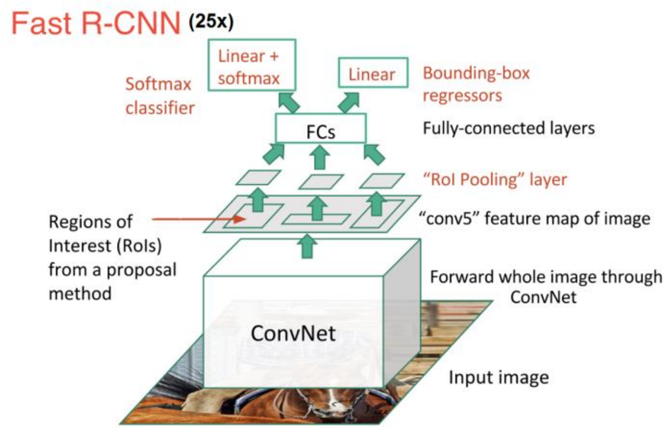


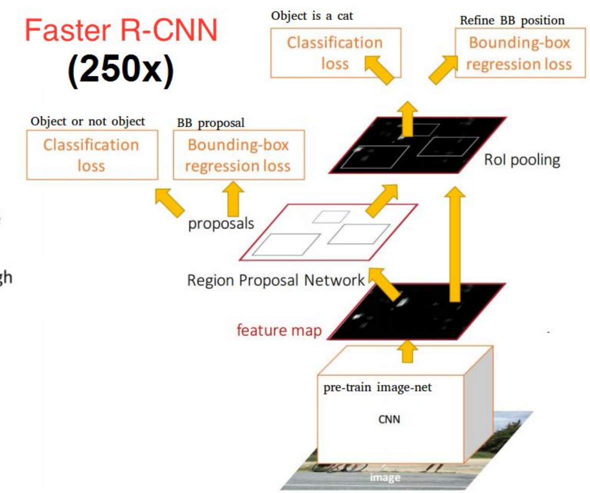


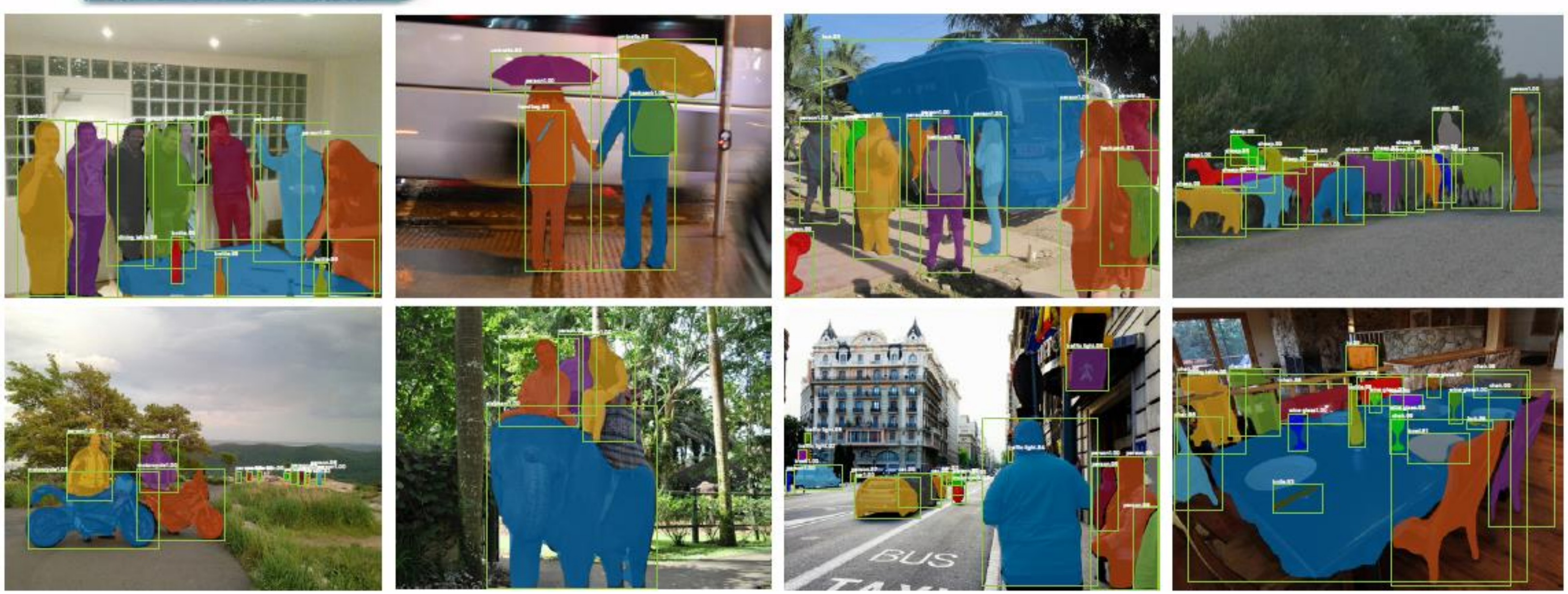
Since AlexNet (The winner of the 1st ImageNet competition, Alex Krizhevsky (Neural Information Processing - NIPS 2012)), there have been multiple new models using CNN as their backbone architecture and achieving excellent results in ImageNet: ZFNet (2013), GoogLeNet (2014), VGGNet (2014), Residual Net ResNet (2015), DenseNet (2016) etc.

2- Object Detection (CVPR-2014, ICCV-2015, PAMI-2017)



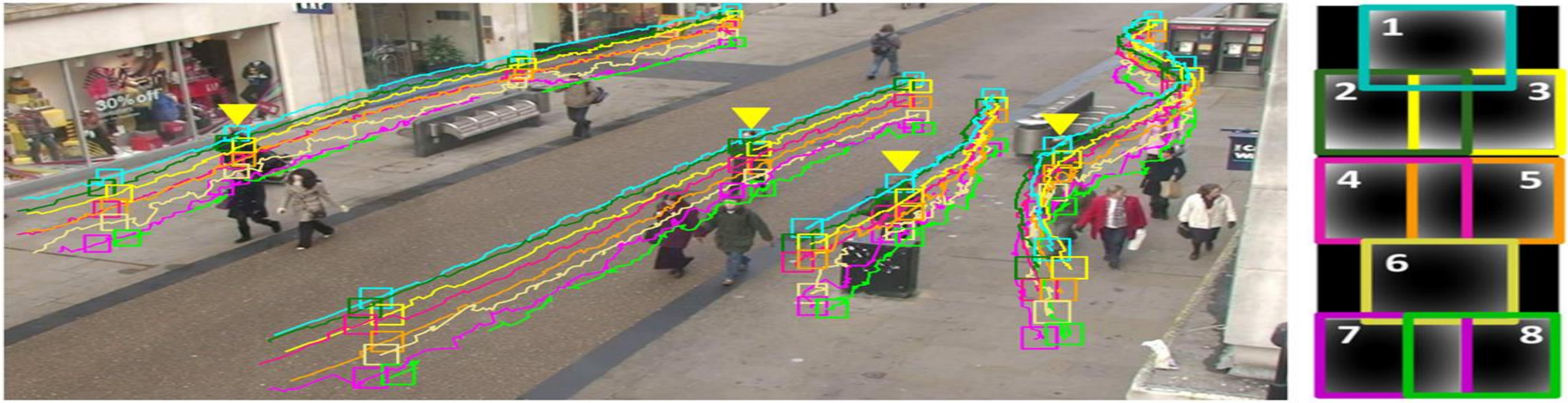




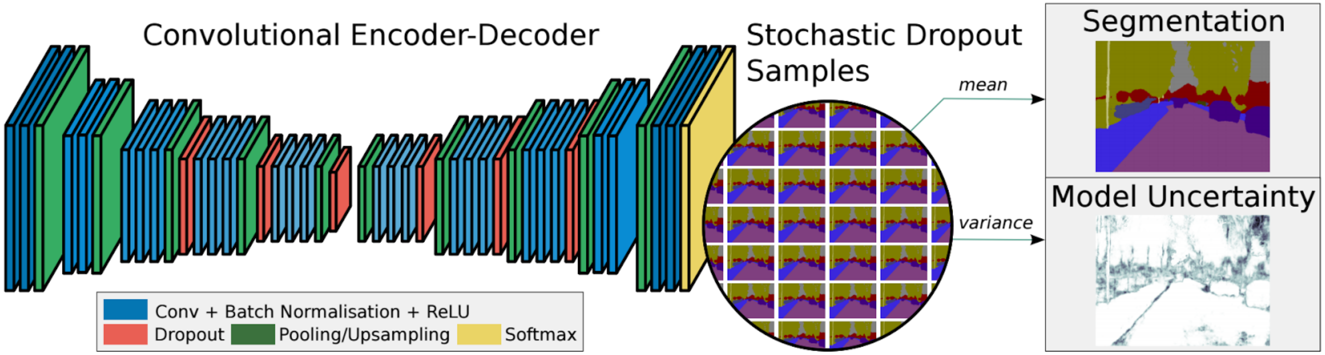


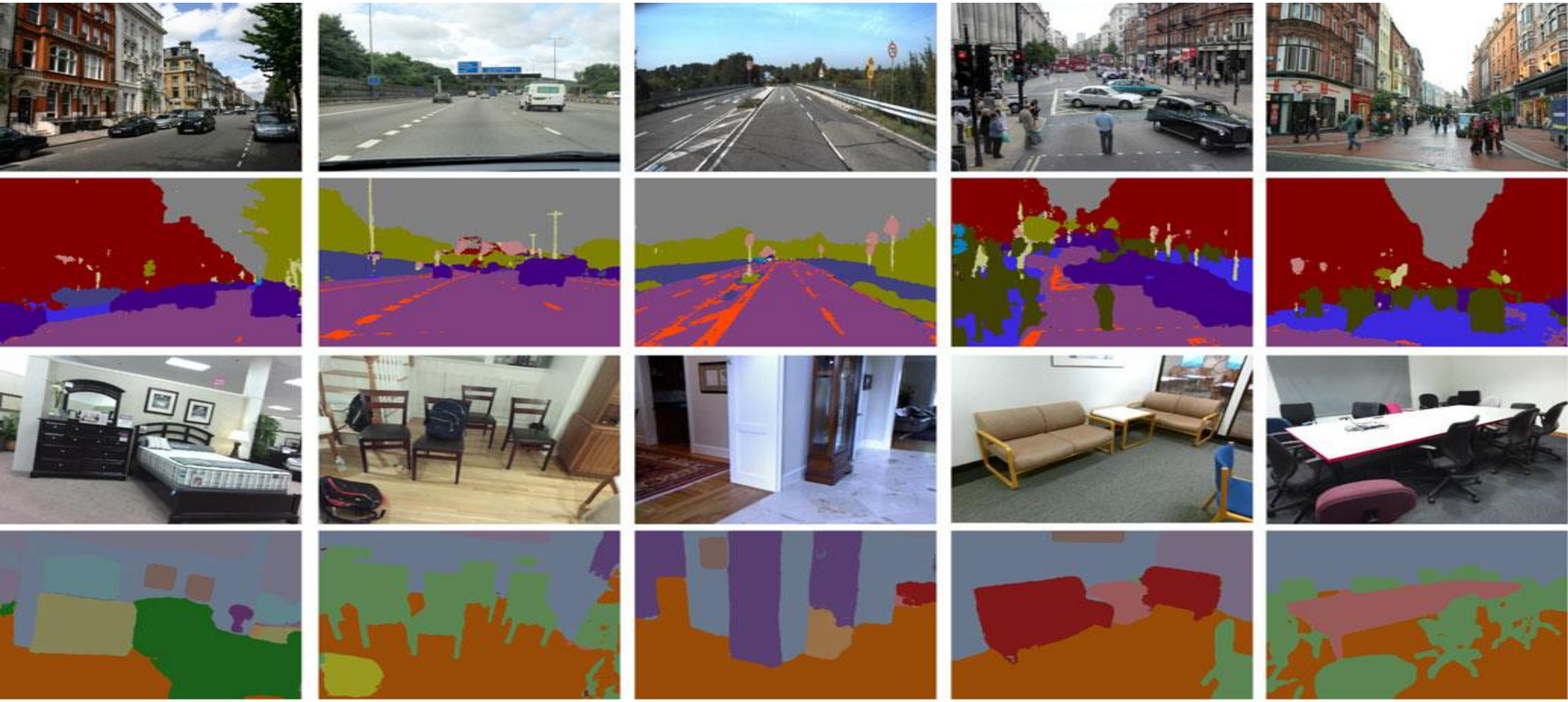
3- Object Tracking

Learning Multi-Domain Convolutional Neural Networks for Visual Tracking (CVPR 2016)



4- Semantic Segmentation: SegNet - IEEE PAMI 2017

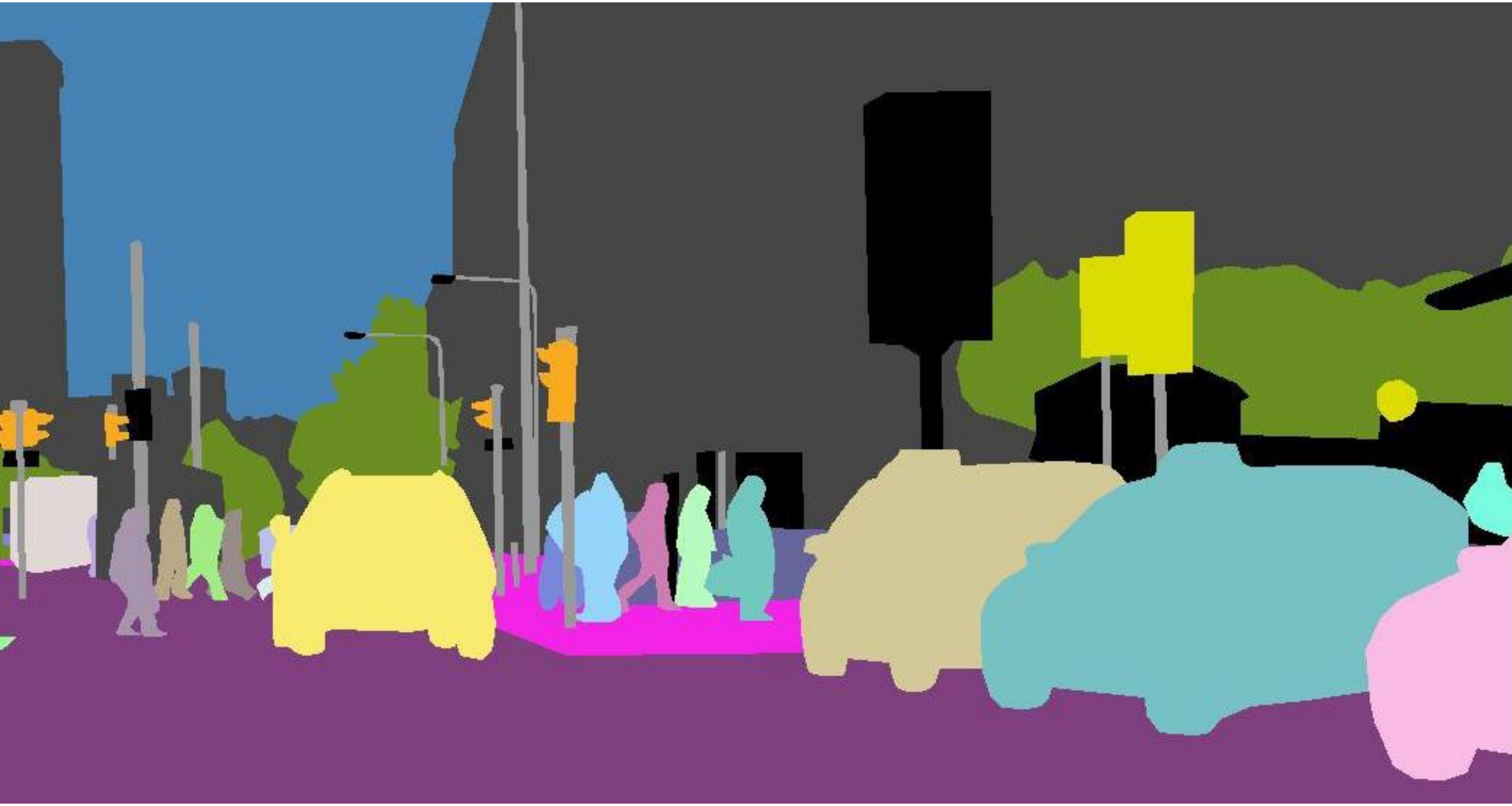




Recent research in Semantic Segmentation all relies heavily on fully convolutional networks, such as Dilated Convolutions, DeepLab (IEEE PAMI 2018), and RefineNet.

5- Instance Segmentation:

Mask R-CNN (arXiv:1703.06870v3 [cs.CV] 24 Jan 2018)



Some Non-CNN Based Techniques Still Evolve

- IEEE PAMI 2019, Image Projective Invariants
- IEEE PAMI 2018, Towards Reaching Human Performance in Pedestrian Detection
- IEEE PAMI 2018, SIFT Meets CNN: A Decade Survey of Instance Retrieval
- ICCV 2017, Reconfiguring the Imaging Pipeline for Computer Vision
- CVPR 2018: Five-point Fundamental Matrix Estimation for Uncalibrated Cameras
- CVPR 2018: Single View Stereo Matching