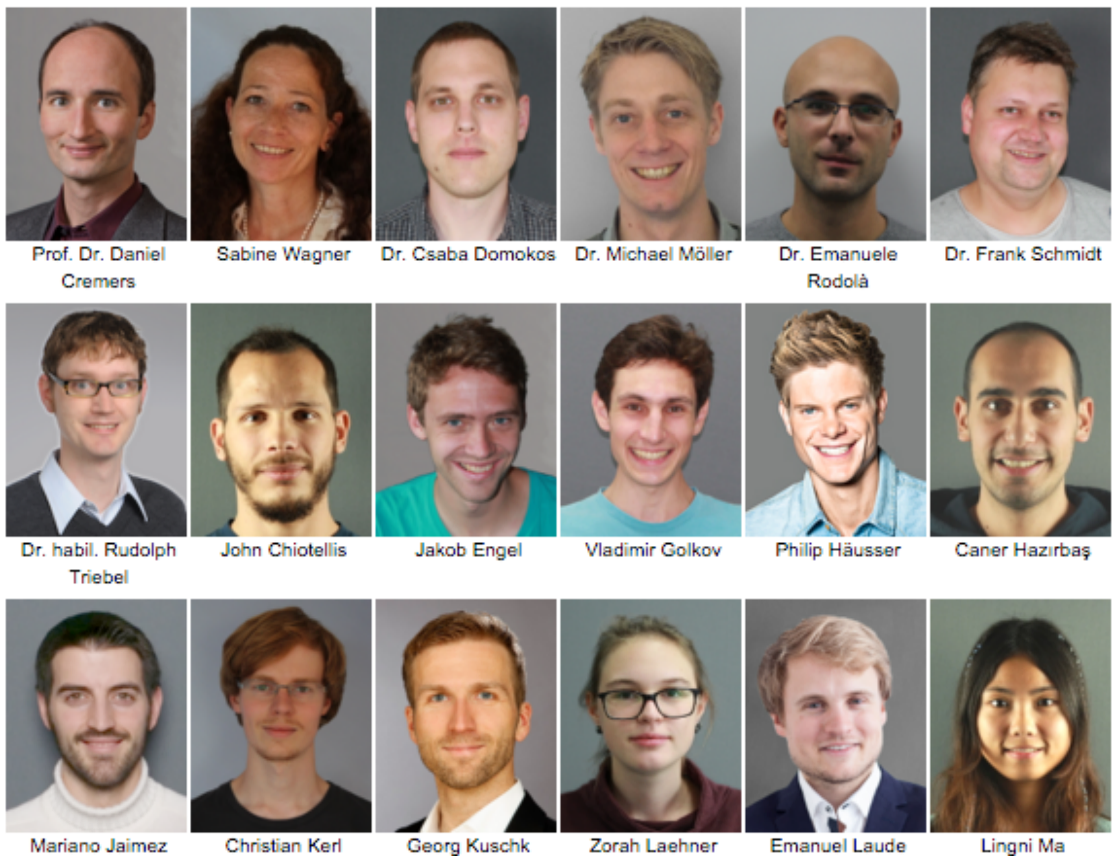
Convolutional Neural Networks for Computer Vision

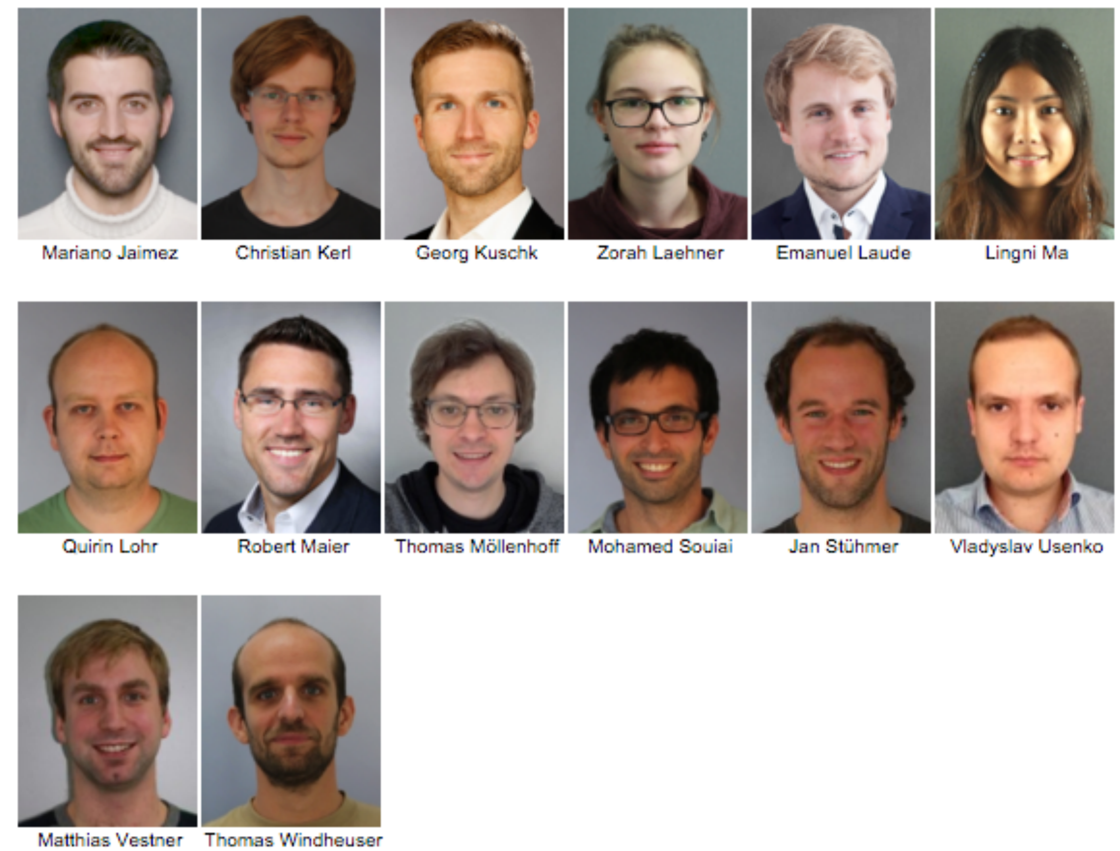
Caner Hazırbaş

Centrum für Informations- und Sprachverarbeitung
24. November ’15

LUDWIG-MAXIMILIANS-UNIVERSITÄT MÜNCHEN

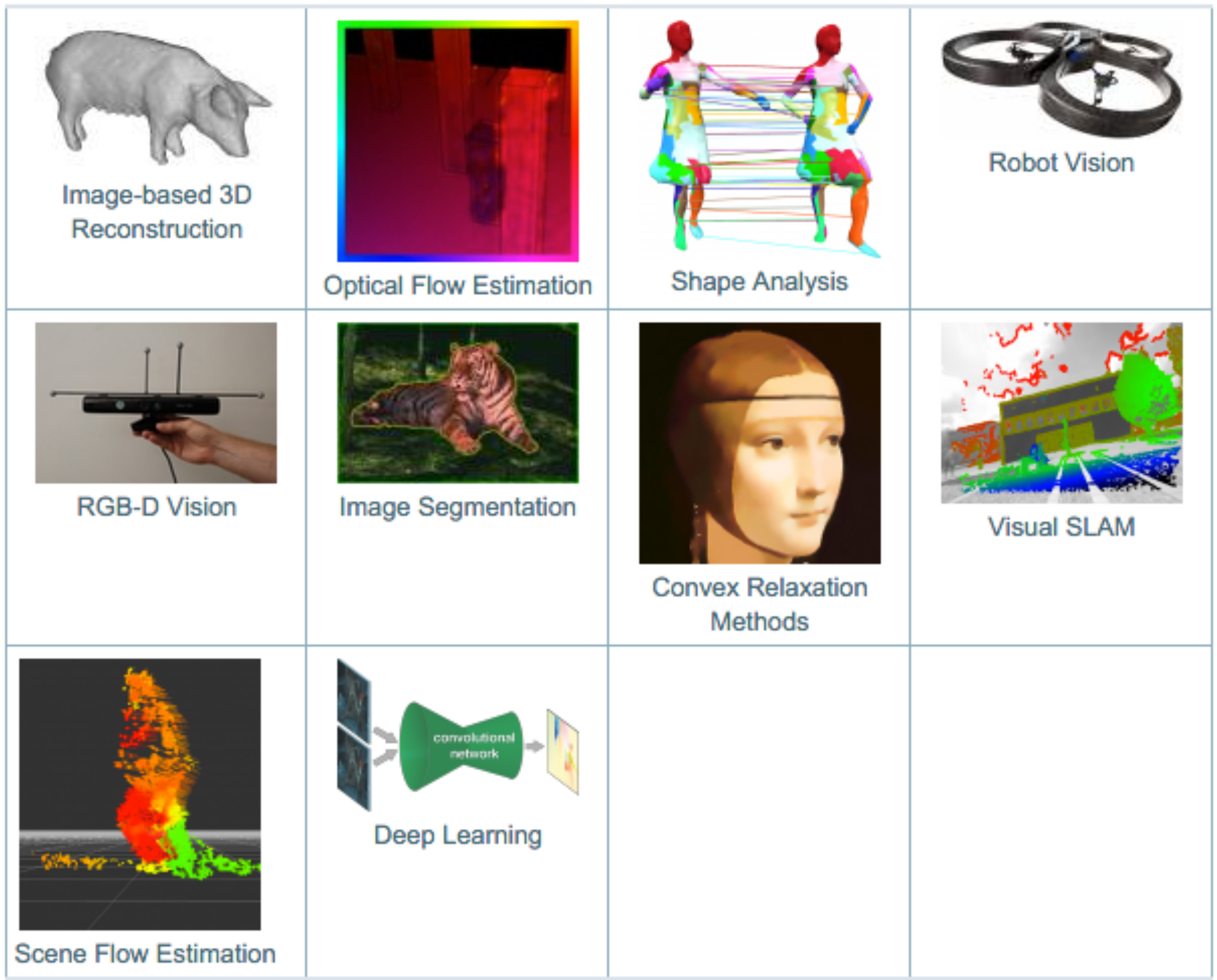
Computer Vision Group





5 Postdocs, 24 PhD students

Research in Computer Vision



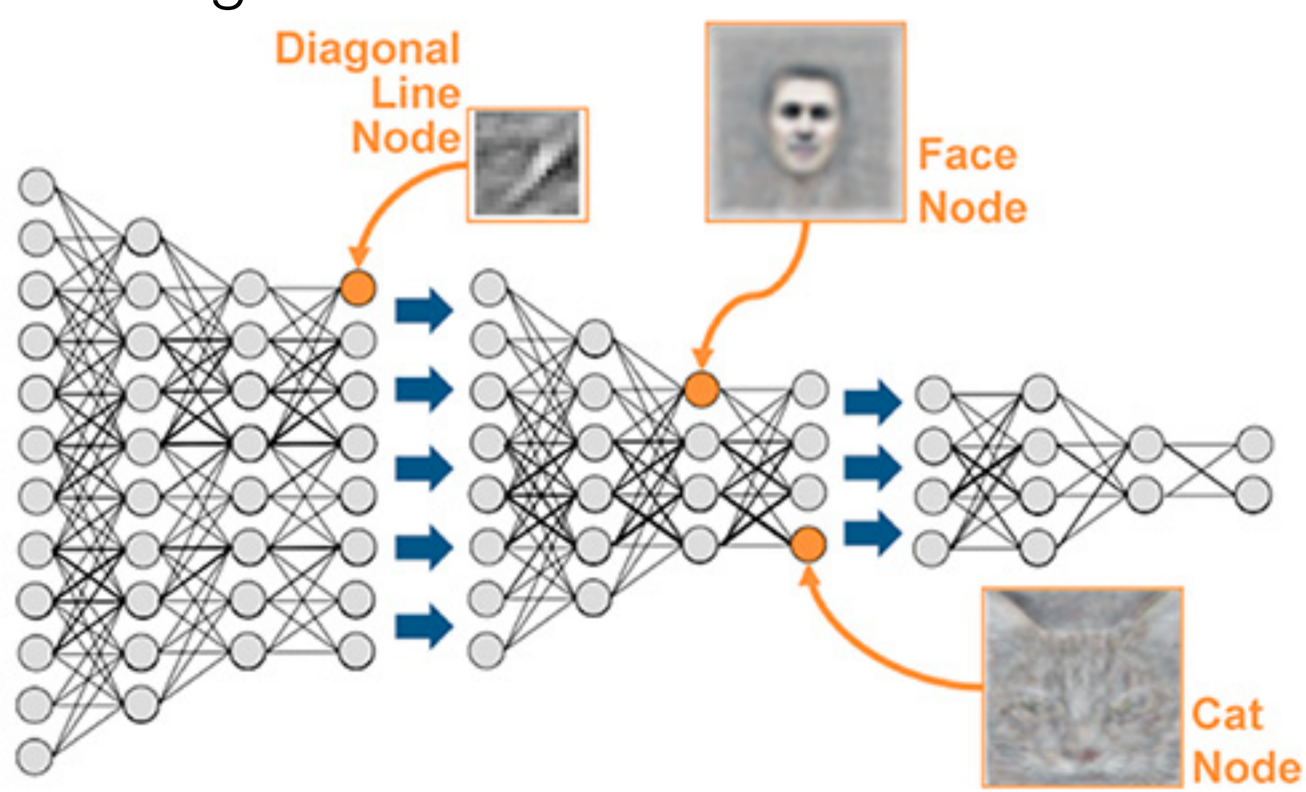
Convolutional Neural Networks for Computer Vision

What is deep learning ?

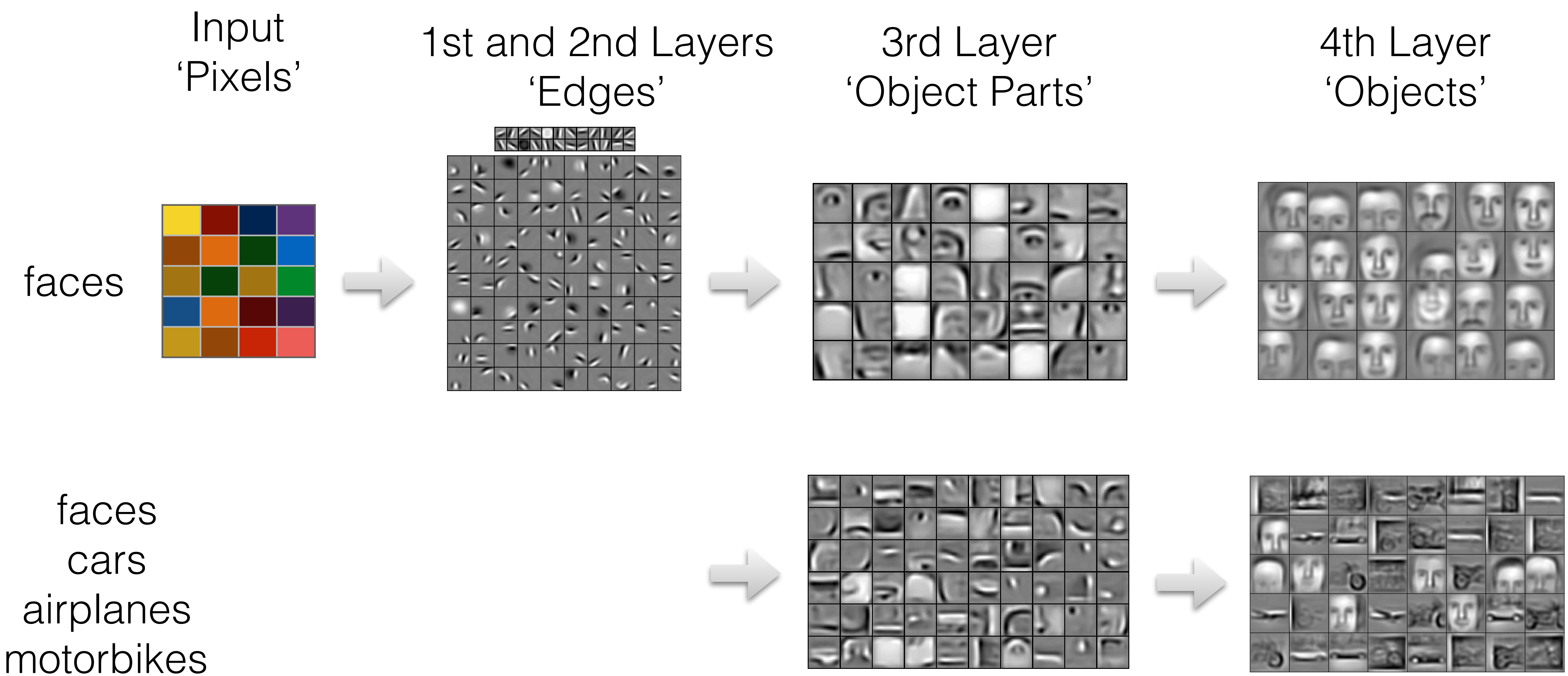
- Representation learning method
  Learning good features automatically from raw data

- Learning representations of data with multiple levels of abstraction

Google's cat detection neural network



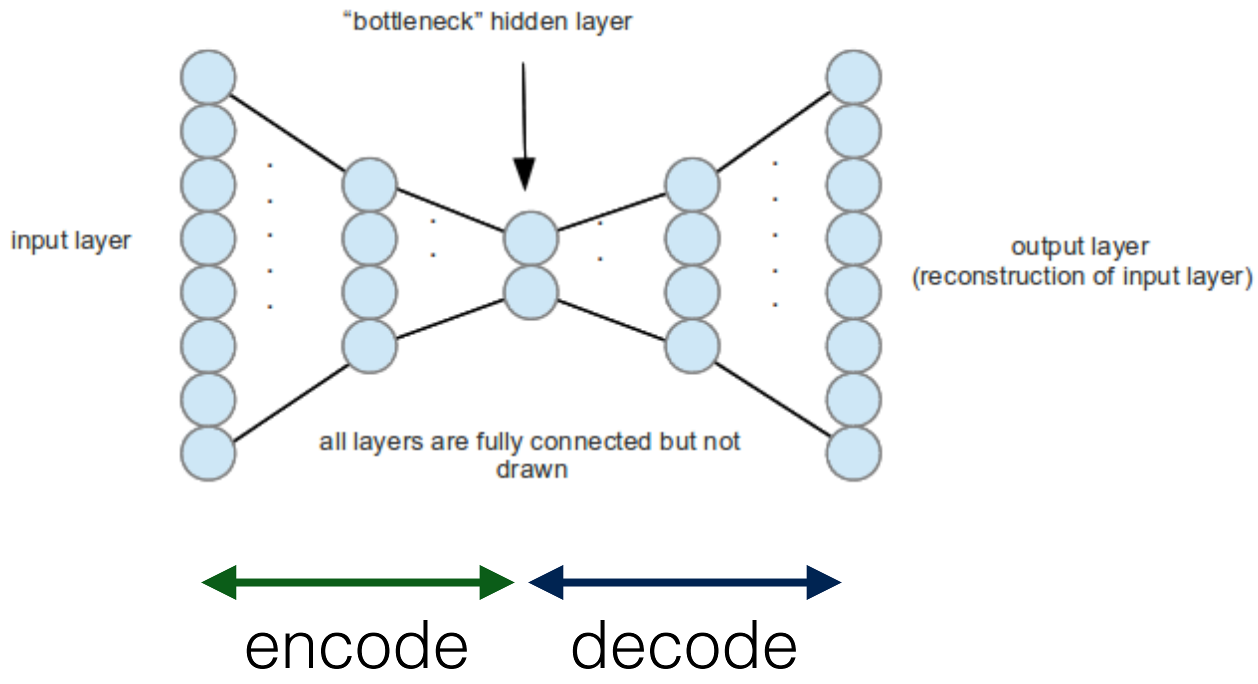
Going deeper in the network

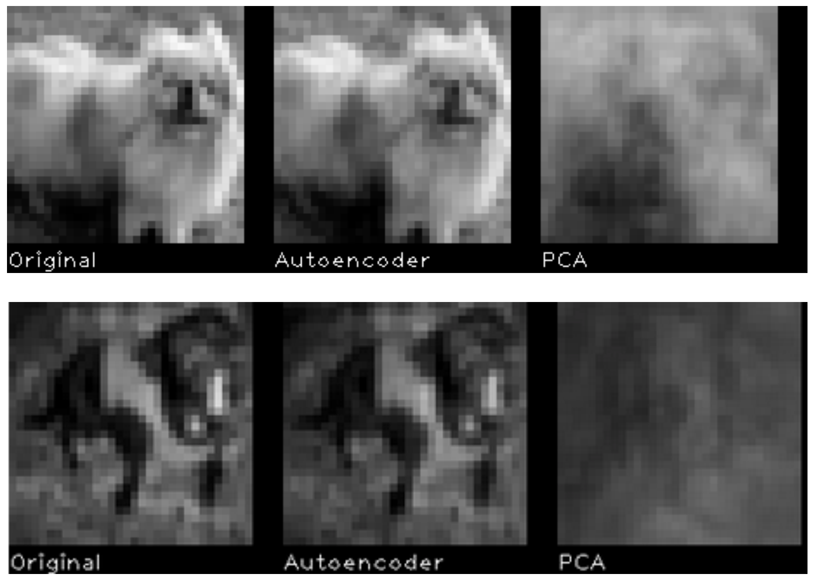


Deep Learning Methods

Unsupervised Methods

- Restricted Boltzmann Machines
- Deep Belief Networks
- Auto encoders: unsupervised feature extraction/learning

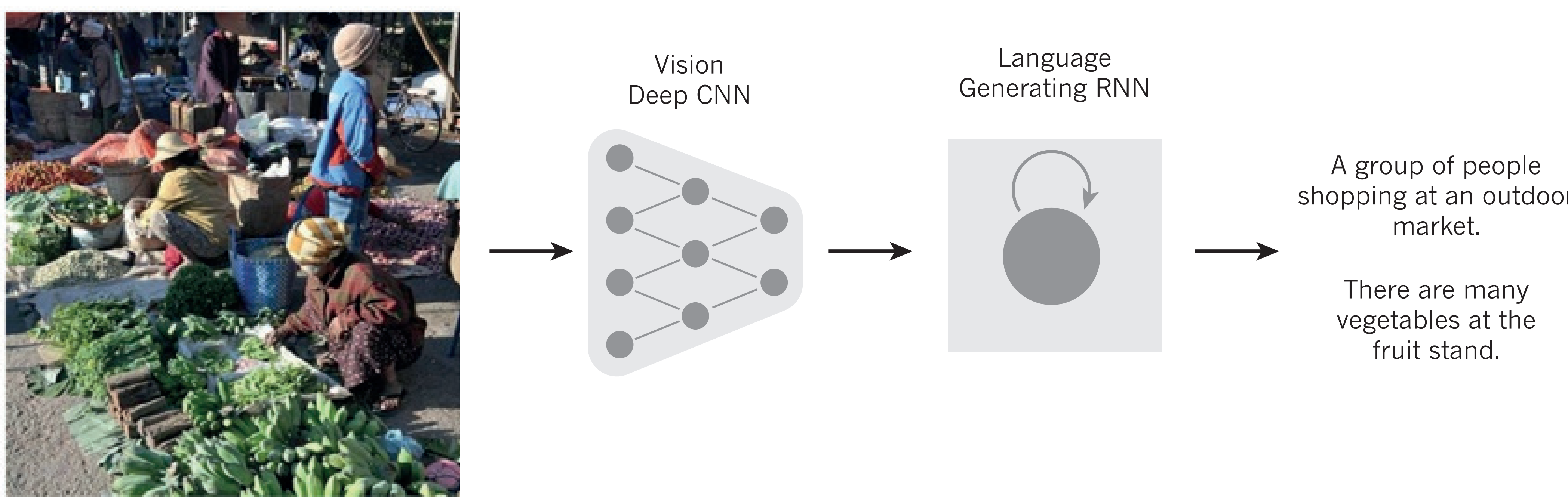




Deep Learning Methods

Supervised Methods

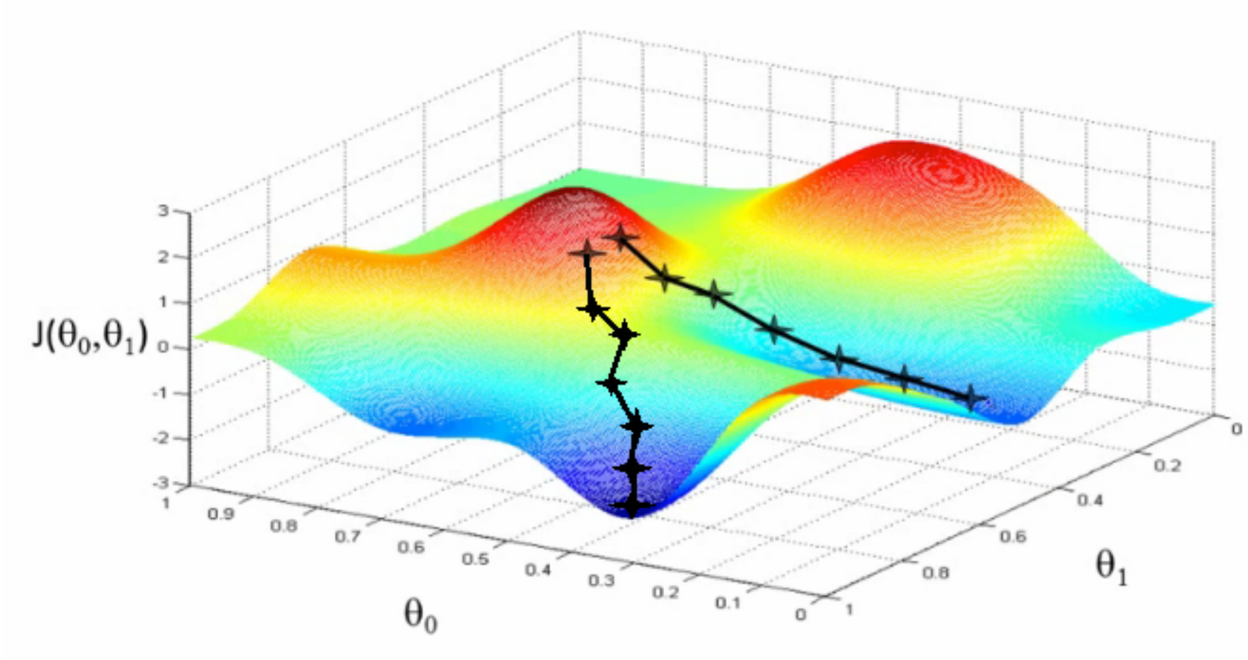
- Deep Neural Networks
- Recurrent Neural Networks
- Convolutional Neural Networks



How to train a deep network ?

Stochastic Gradient Descent — *supervised learning*

- show input vector of few examples
- compute the output and the errors
- compute average gradient
- update the weights accordingly



How to train a deep network ?

Alternatives:

- AdaGrad, AdaDelta, NAG (Nesterov's Accelerated Gradient)…
- **ADAM** (now in Caffe - `http://caffe.berkeleyvision.org/tutorial/solver.html`)
  The Adam is a gradient-based optimization method (like SGD). This includes an "adaptive moment estimation" (m_t,v_t) and can be regarded as a generalization of AdaGrad. The update formulas are:

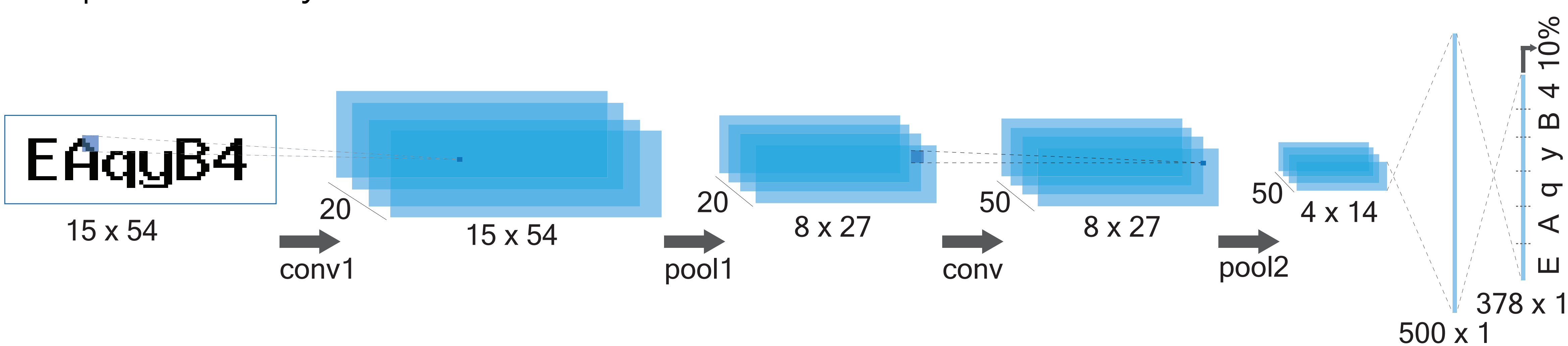
$$(m_t)_i = \beta_1 (m_{t-1})_i + (1 - \beta_1)(\nabla L(W_t))_i,$$
$$(v_t)_i = \beta_2 (v_{t-1})_i + (1 - \beta_2)(\nabla L(W_t))_i^2$$

$$(W_{t+1})_i = (W_t)_i - \alpha \frac{\sqrt{1 - (\beta_2)_i^t}}{1 - (\beta_1)_i^t} \frac{(m_t)_i}{\sqrt{(v_t)_i} + \varepsilon}.$$

D. Kingma, J. Ba. Adam: A Method for Stochastic Optimization. International Conference for Learning Representations, 2015

Convolutional Neural Networks

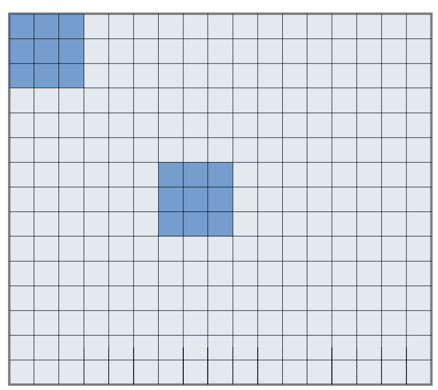
- CNNs are designed to process the data in the form of multiple arrays (e.g. 2D images, 3D video/volumetric images)
- Typical architecture is composed of series of stages: ***convolutional*** *layers* and ***pooling*** *layers*
- Each unit is connected to local patches in the feature maps of the previous layer



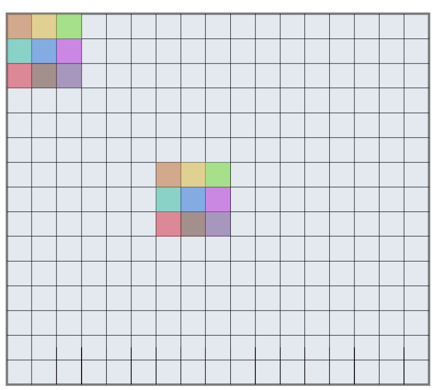
Key Idea behind Convolutional Networks

Convolutional networks take advantage of the properties of natural signals:

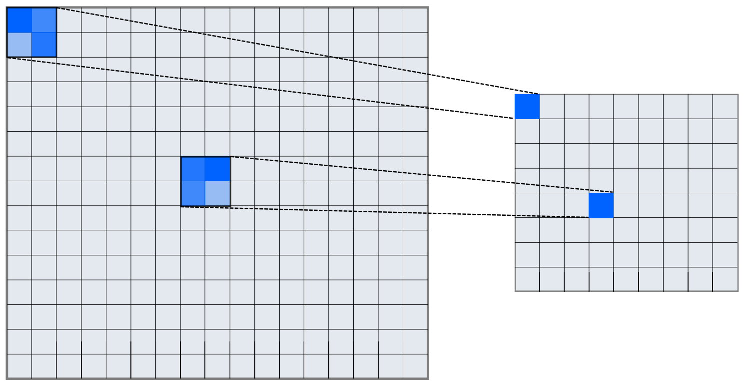
- local connections



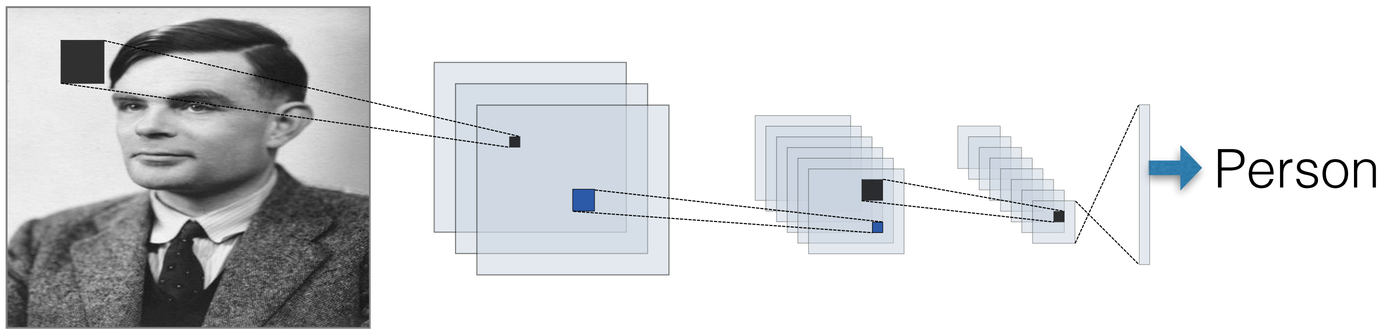
- shared weights



- pooling

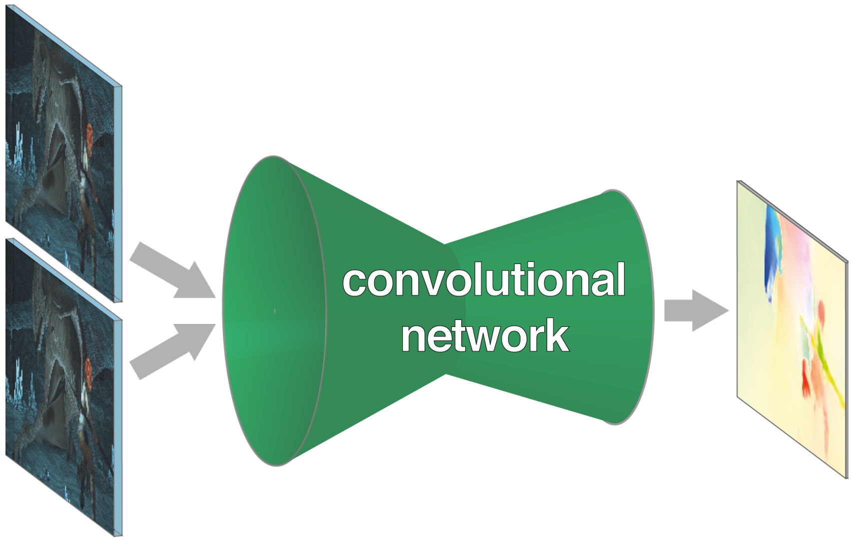


- the use of many layers

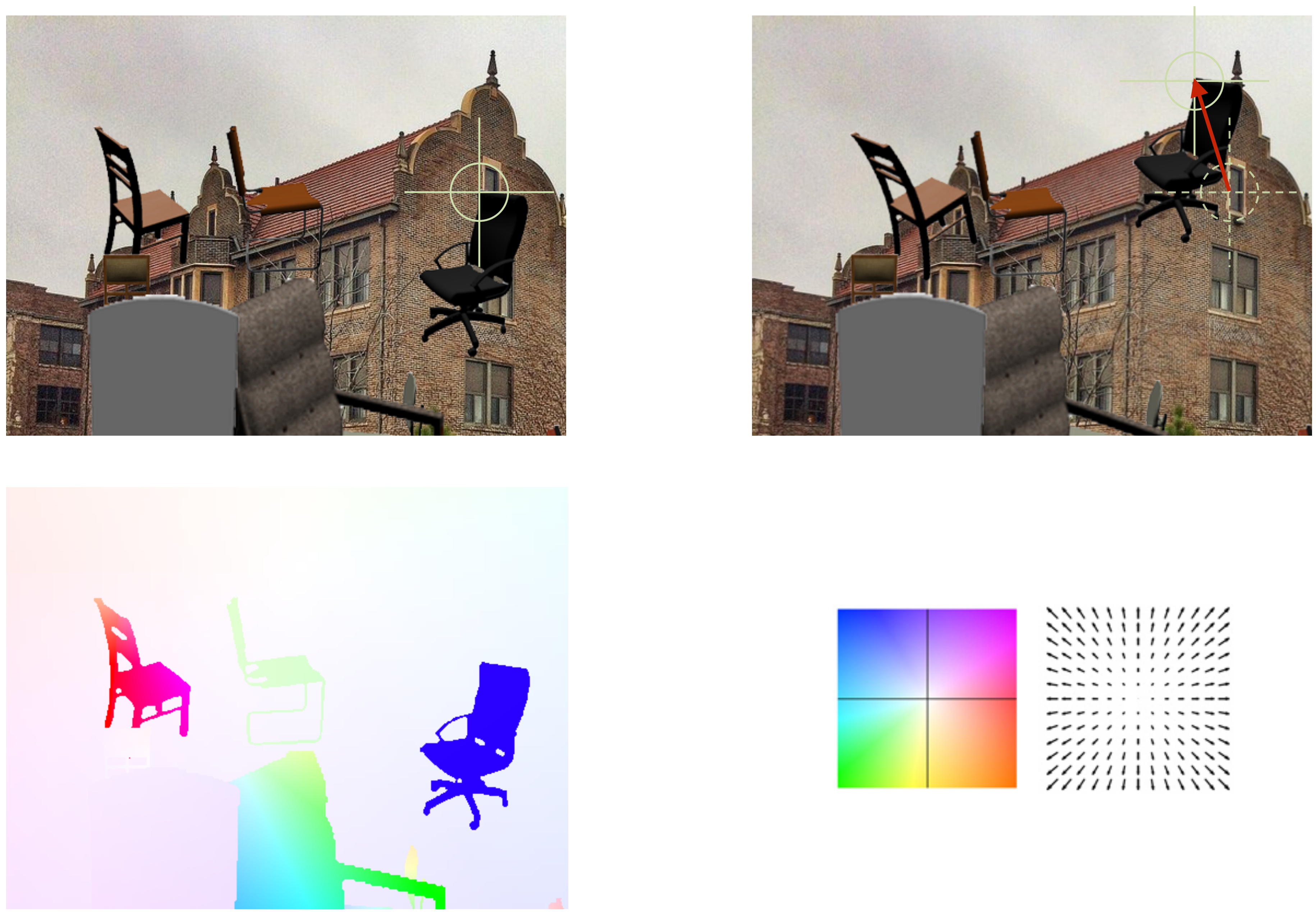


FlowNet: Learning Optical Flow with Convolutional Networks

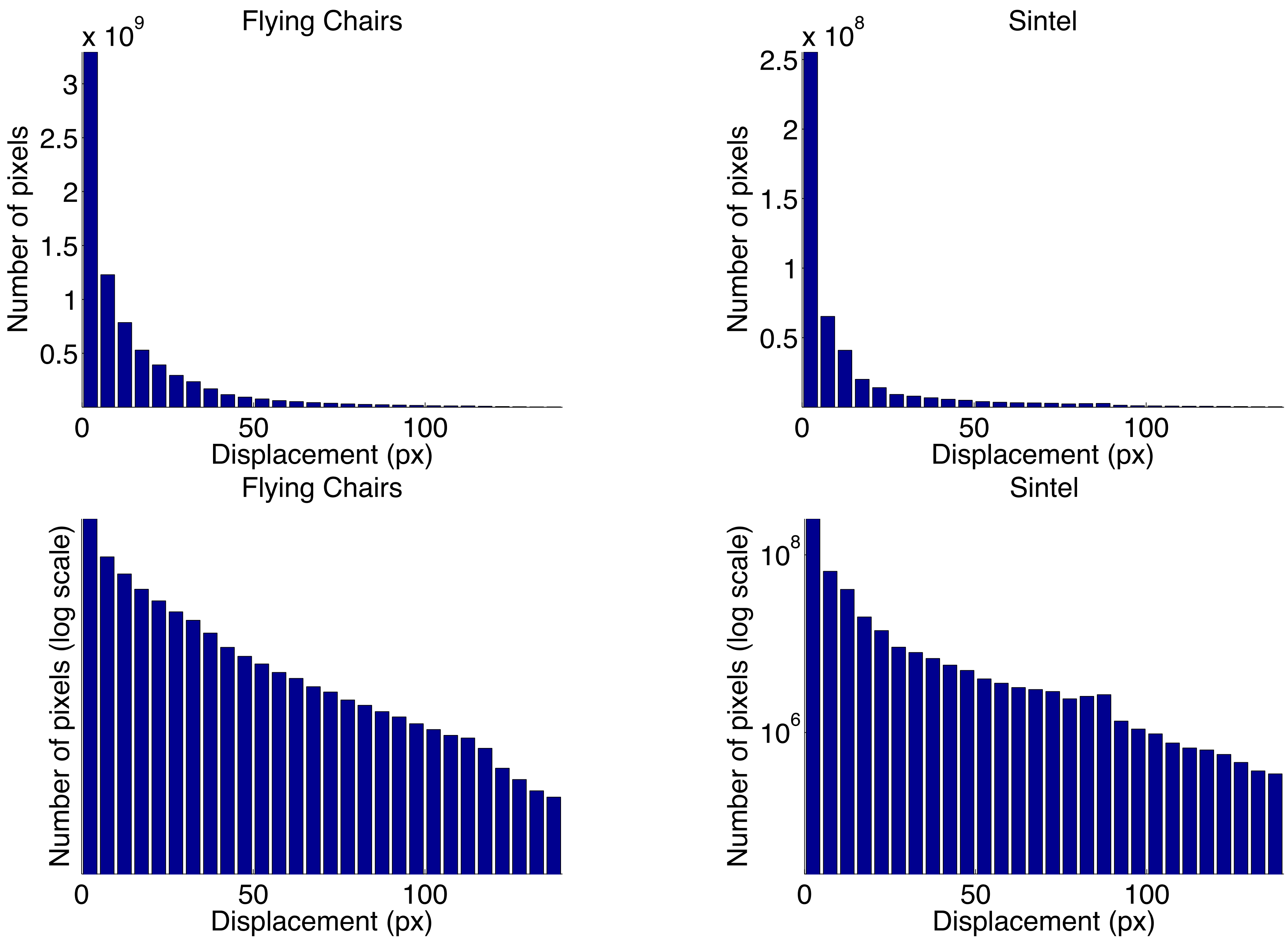
Philipp Fischer, Alexey Dosovitskiy, Eddy Ilg, Thomas Brox
Philip Häusser, Caner Hazırbaş, Vladimir Golkov, Daniel Cremers, Patrick van der Smagt



Flying Chairs



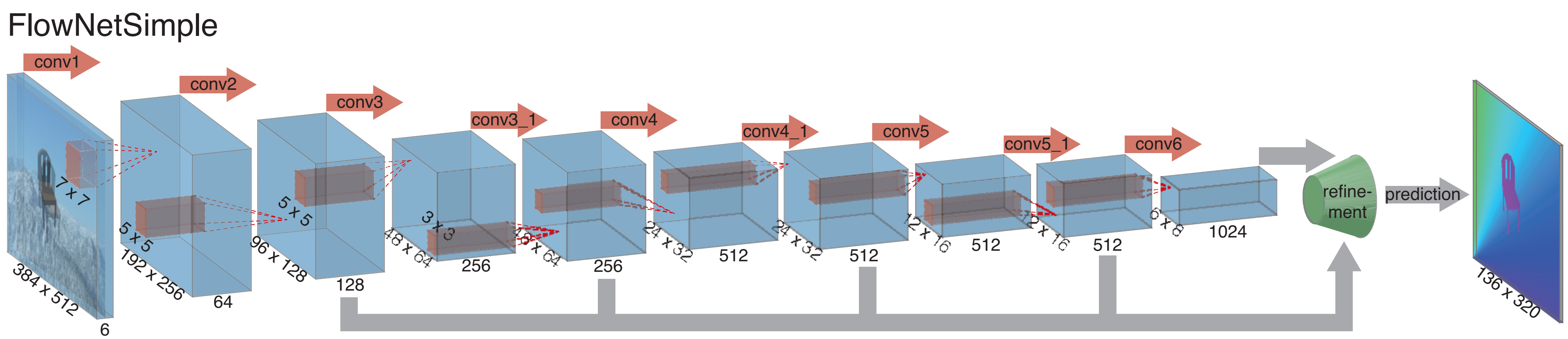
Flying Chairs

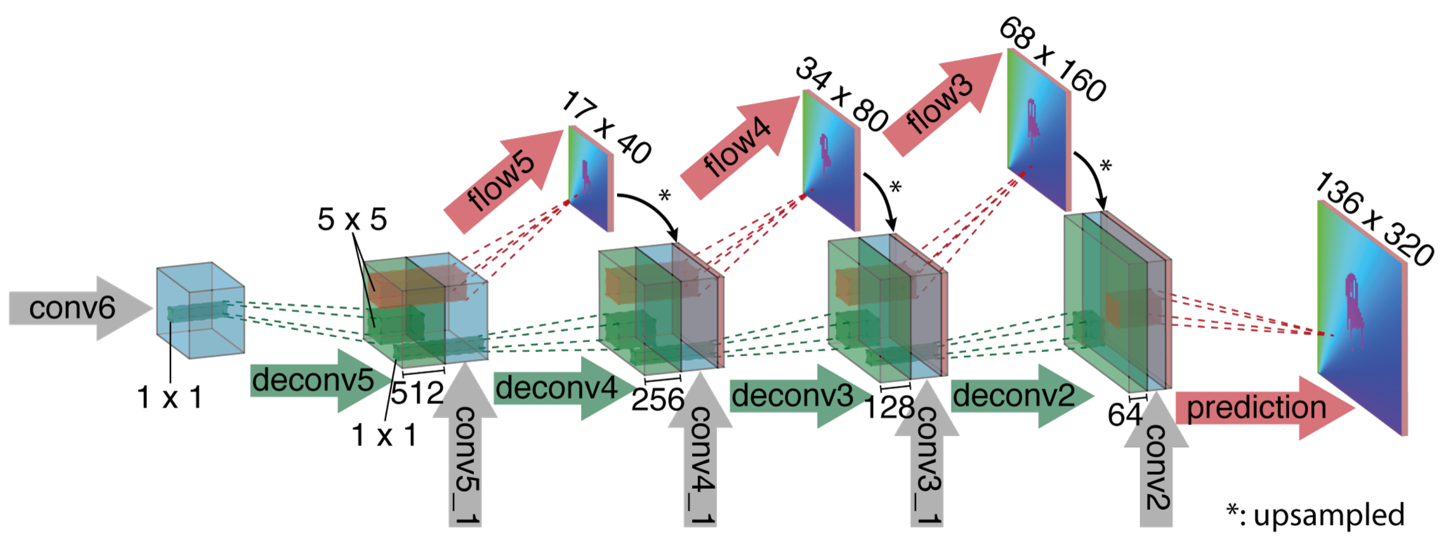


Data Augmentation

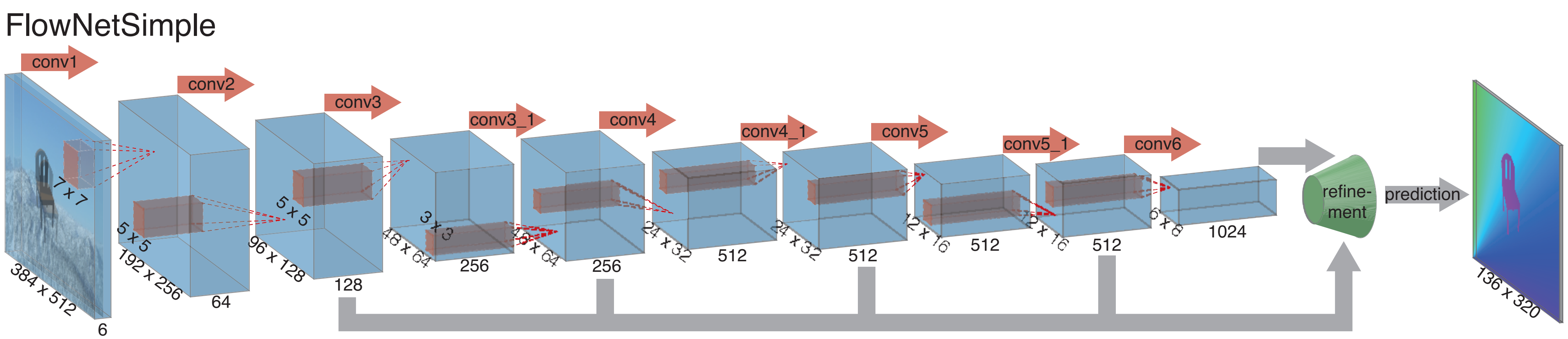
- *translation*, *rotation*, *scaling*, additive *Gaussian noise*
- changes in *brightness*, *contrast*, *gamma* and *colour*

FlowNetSimple

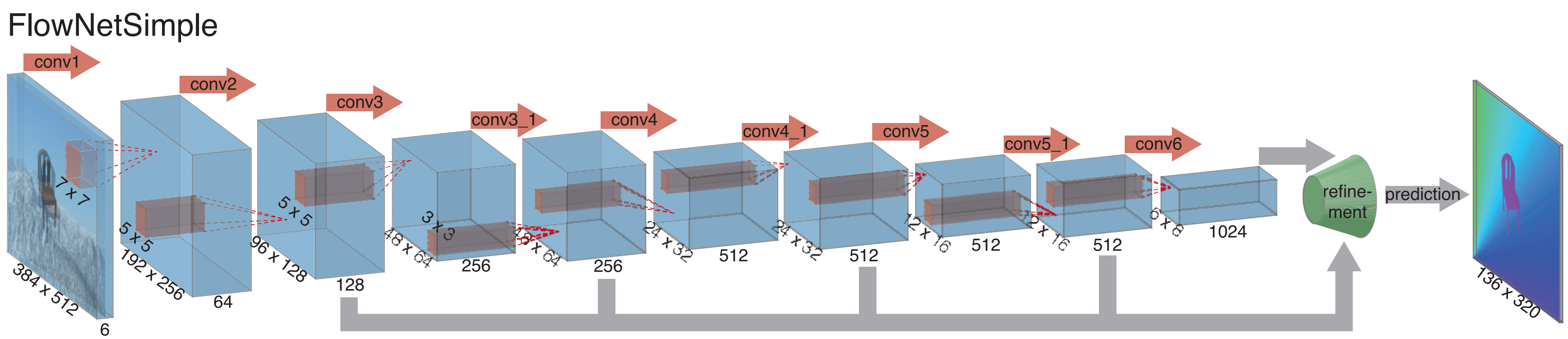




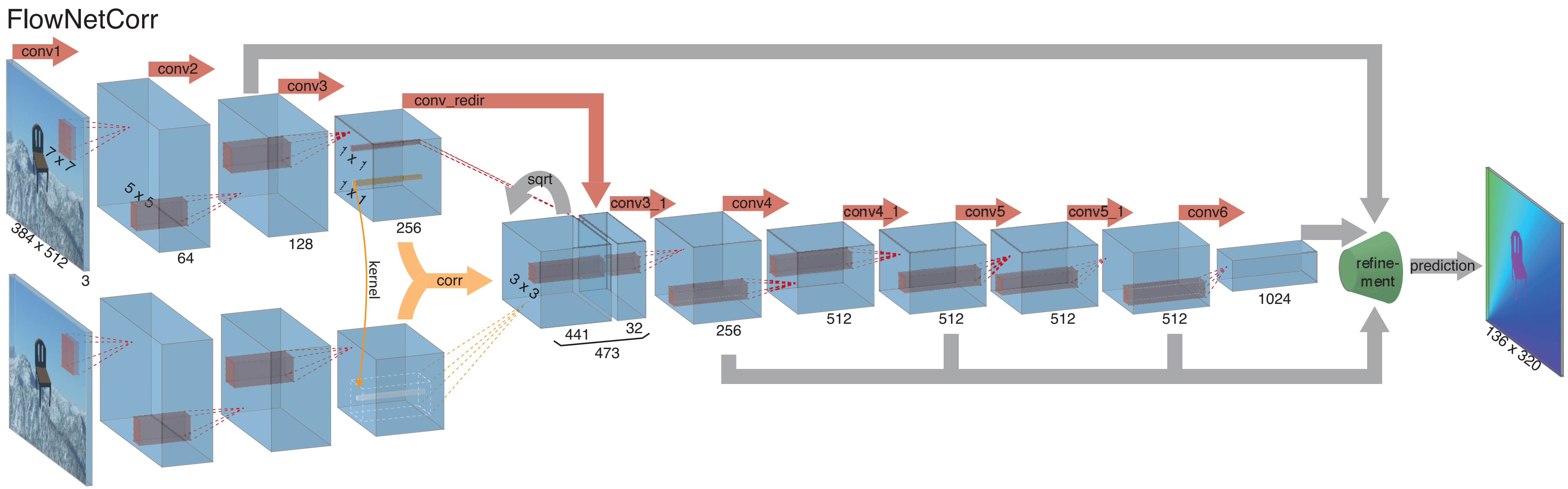
FlowNetSimple - Flying Chairs



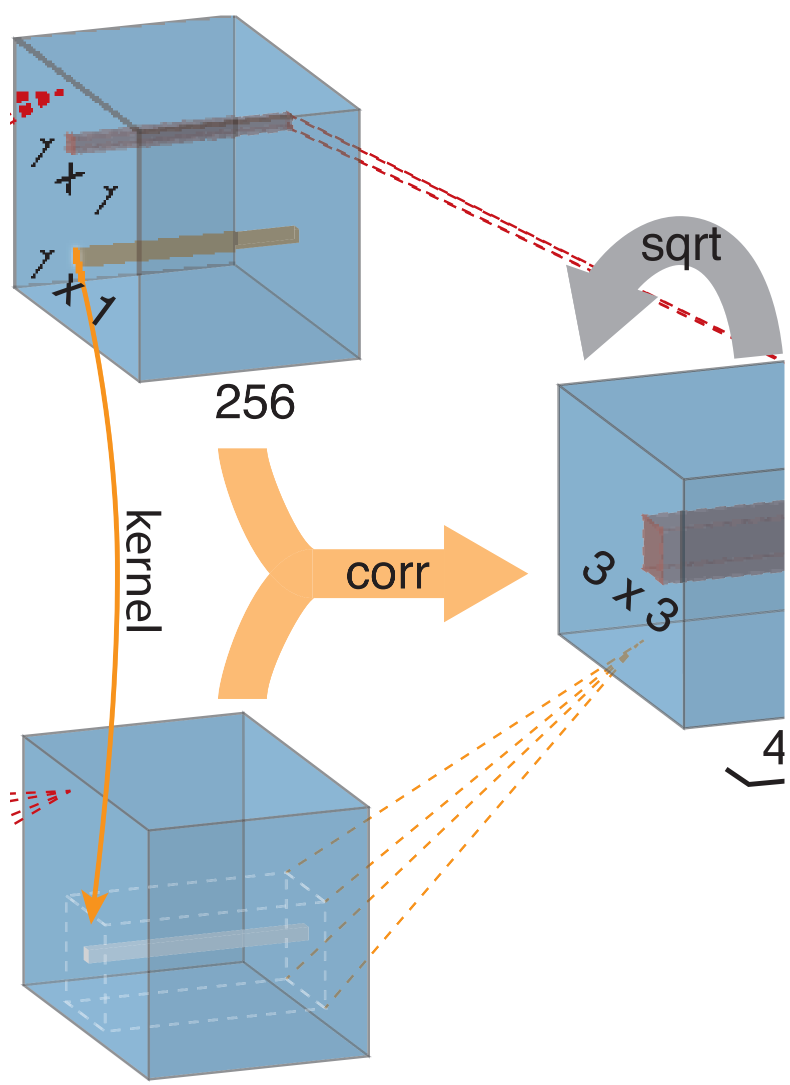
FlowNetSimple - Sintel

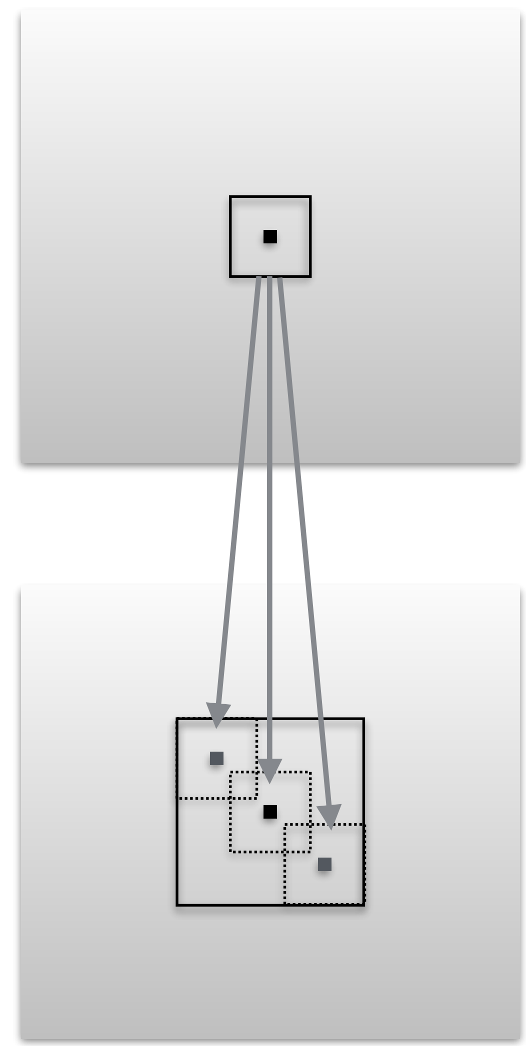


FlowNetCorr

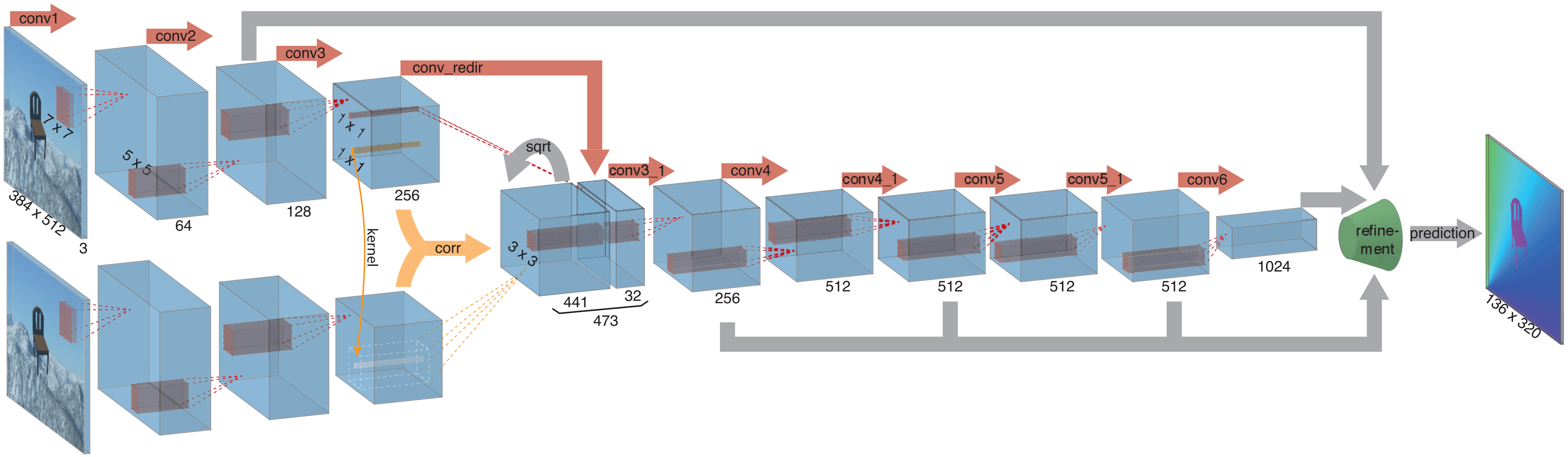


Correlation Layer



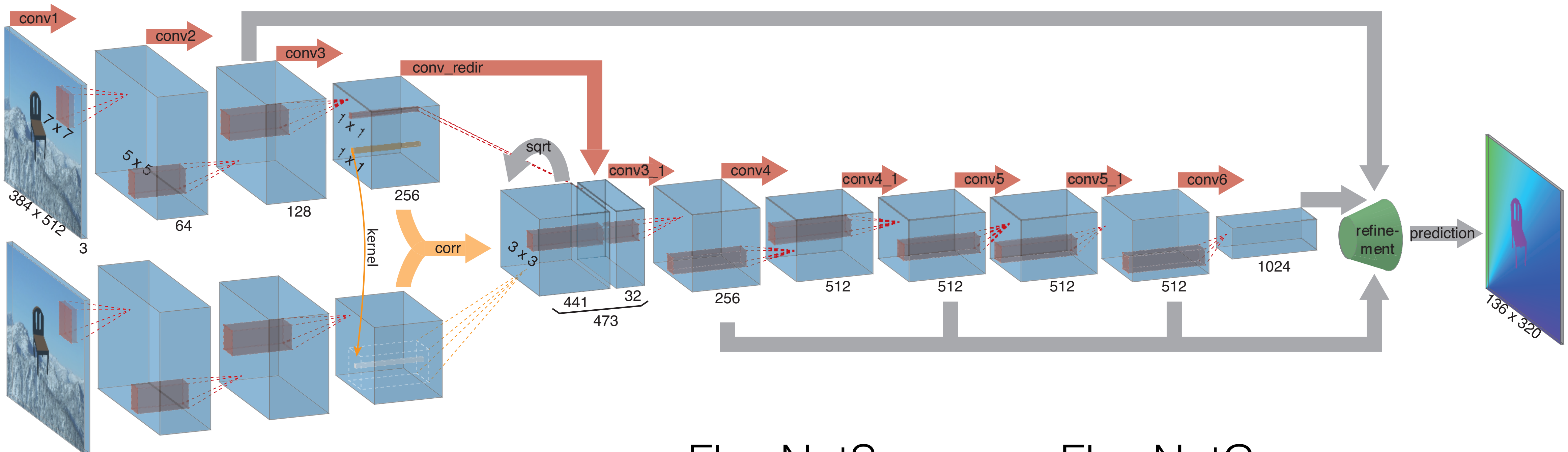


FlowNetCorr - Flying Chairs



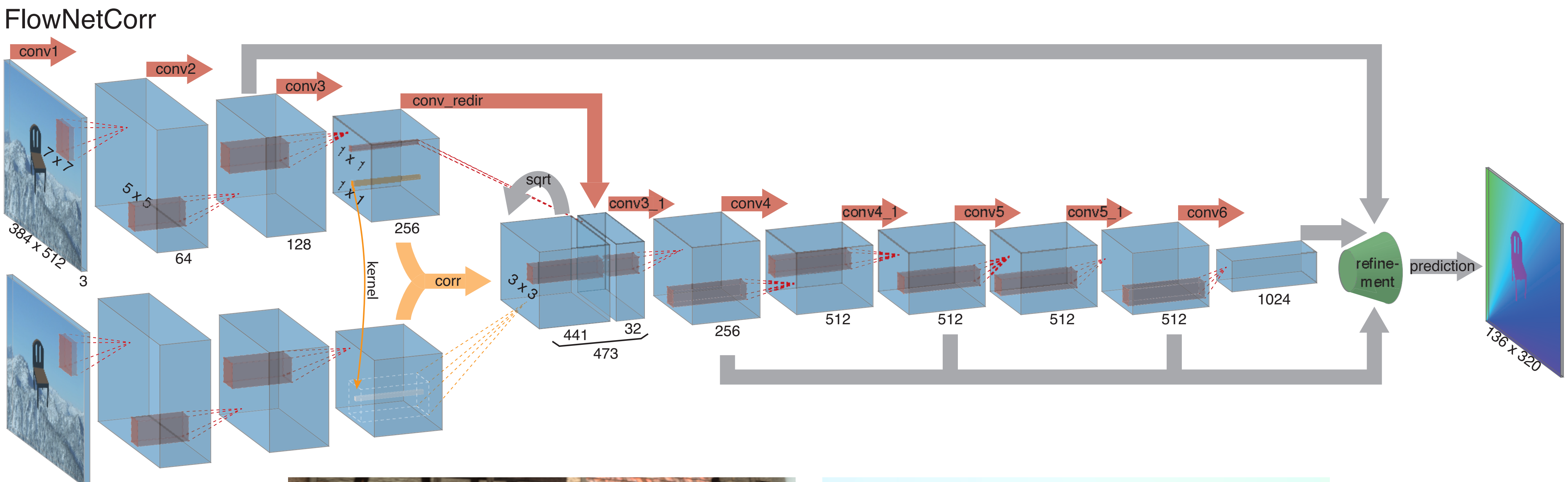
Simple vs. Corr - Flying Chairs

FlowNetCorr

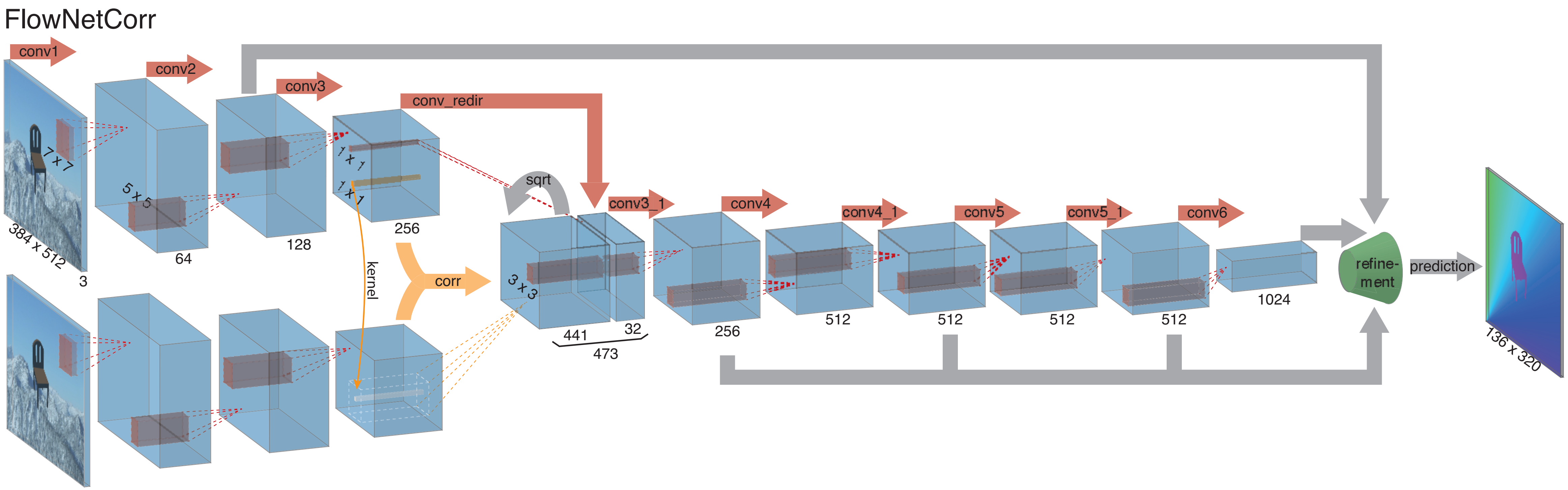


FlowNetS FlowNetCorr

FlowNetCorr - Sintel

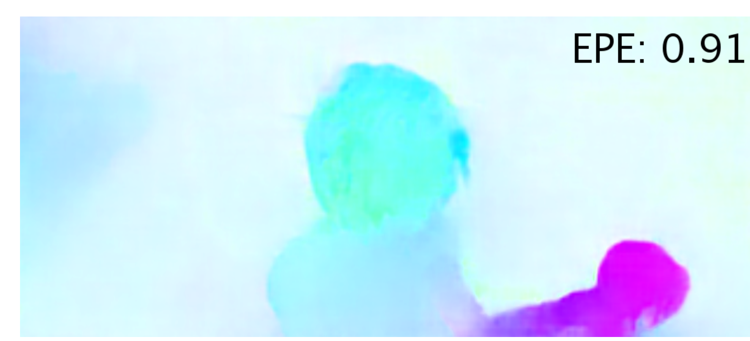


Simple vs. Corr - Sintel



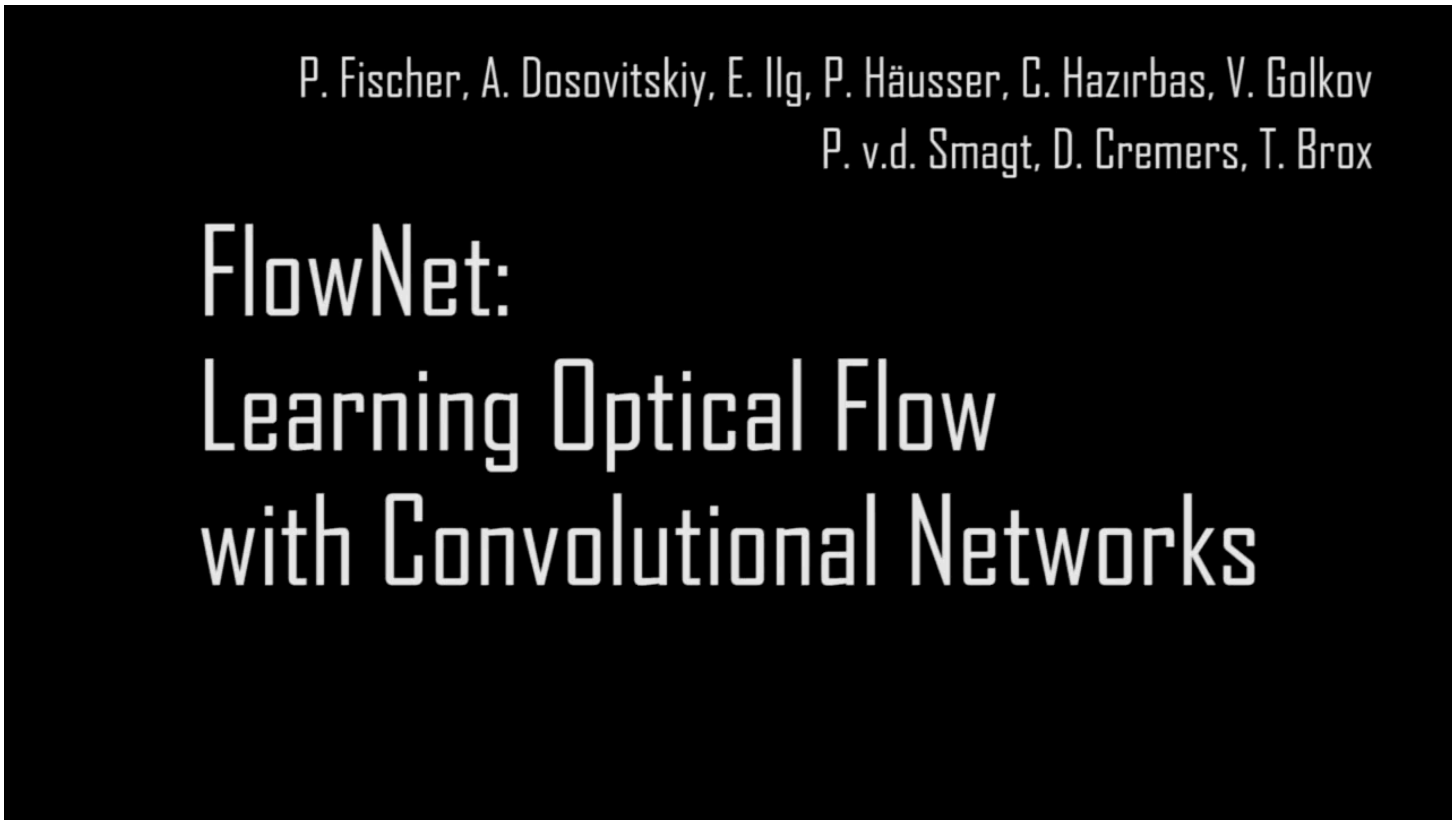
FlowNetS

FlowNetCorr



FlowNetSimple + Variational Smoothing

FlowNet: Learning Optical Flow with Convolutional Networks



References

- Building High-level Features Using Large Scale Unsupervised Learning
  Quoc V. Le , Rajat Monga , Matthieu Devin , Kai Chen , Greg S. Corrado , Jeff Dean , Andrew Y. Ng
  ICML'12

- Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations
  Honglak Lee Roger Grosse Rajesh Ranganath Andrew Y. Ng ICML'09

- ImageNet Classification with Deep Convolutional Neural Networks
  Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton NIPS'12

- Gradient-based learning applied to document recognition.
  Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner Proceedings of the IEEE'98

- FlowNet: Learning Optical Flow with Convolutional Networks
  Philipp Fischer, Alexey Dosovitskiy, Eddy Ilg, Philip Häusser, Caner Hazırbaş, Vladimir Golkov, Patrick van der Smagt, Daniel Cremers, Thomas Brox

References

- Google's cat detection neural network http://www.resnap.com/image-selection-technology/deep-learning-image-classification/
- Example auto-encoder : http://nghiaho.com/?p=1765
- SGD : http://blog.datumbox.com/tuning-the-learning-rate-in-gradient-descent/