Computer Vision Group
Prof. Daniel Cremers

Technische Universität München

Machine Learning for Applications in Computer Vision

Neural Networks and Deep Learning

Philip Häusser

Recap: What is Learning?

- problem (“classification”)
- data (images + labels)
- model (linear regression)
- cost function
- optimization algorithm

Recap: What is Learning?

- problem (“classification”)
- data (images + labels)
- model (linear regression)
- cost function
- optimization algorithm

What is “Deep”?

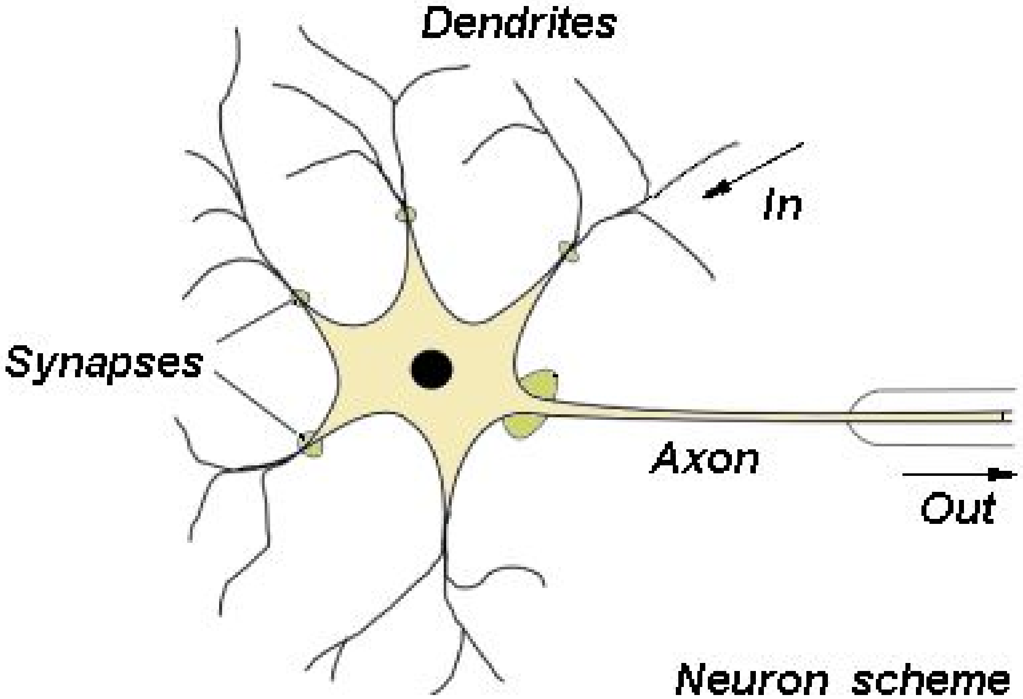
Recap: What is Learning?

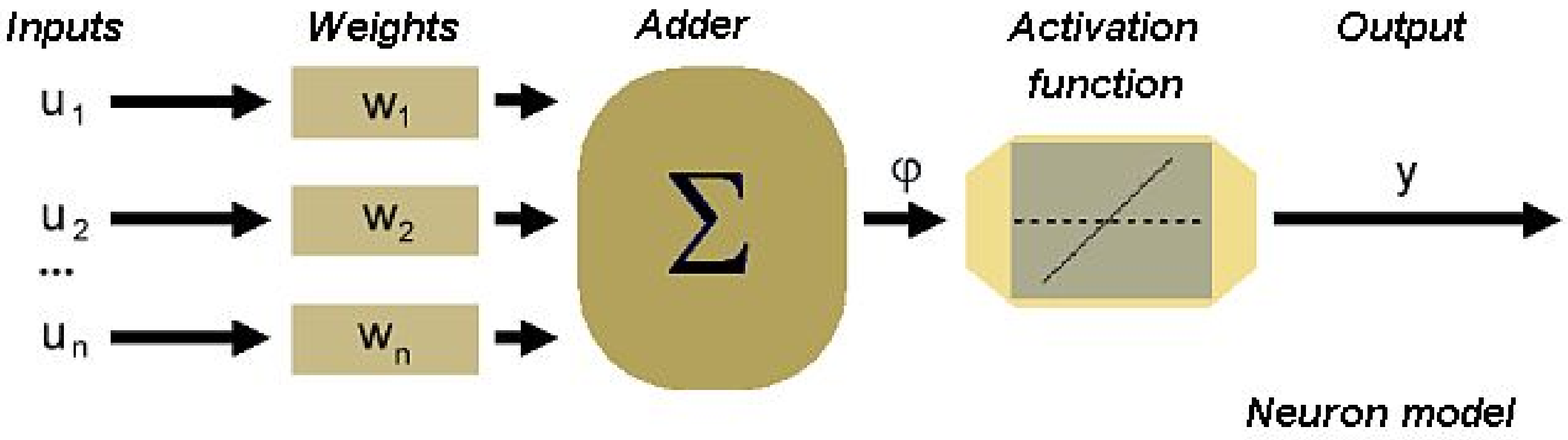
- problem ("classification")
- data (images + labels)
- **model (linear regression)**
- cost function
- optimization algorithm

What is "Deep"?

Perceptron

- The human brain (10^{10} cells) is the archetype of neural networks



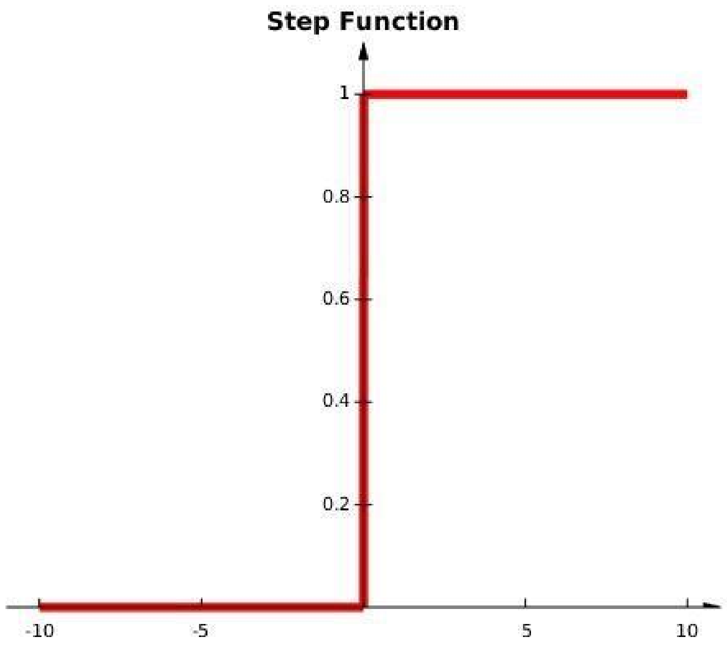


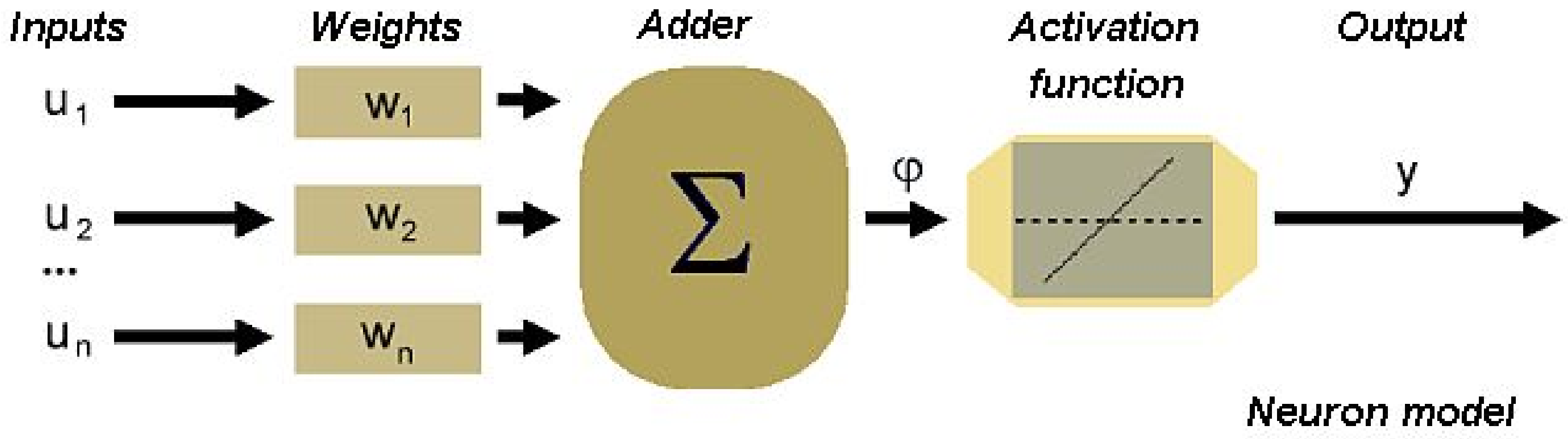
http://home.agh.edu.pl/~vlsi/AI/intro/

Neuron Activations

- Different type of activation functions

Threshold activation (binary classifier)



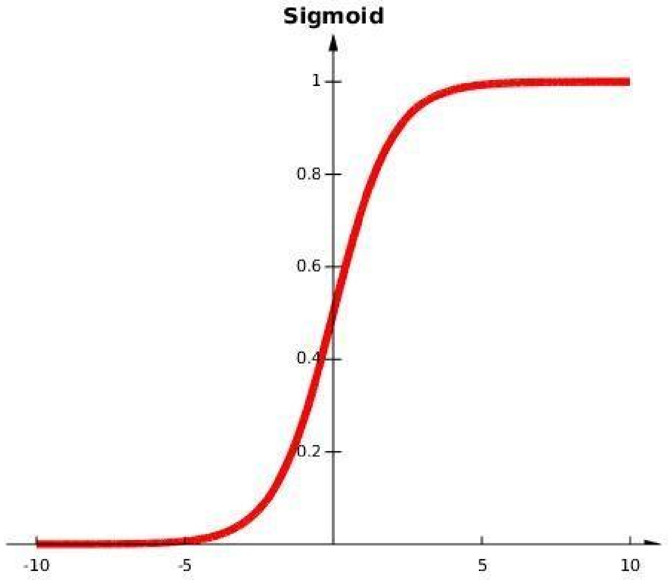


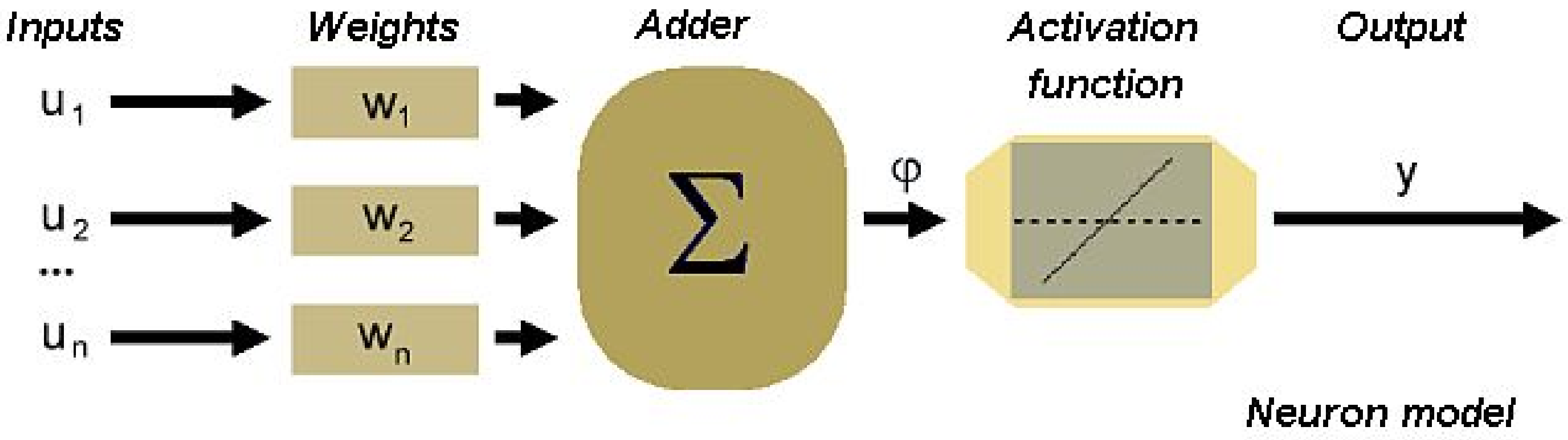
Neuron Activations

- Different type of activation functions

Sigmoid activation

$$y = \frac{1}{1 + \exp^{-\phi}}$$



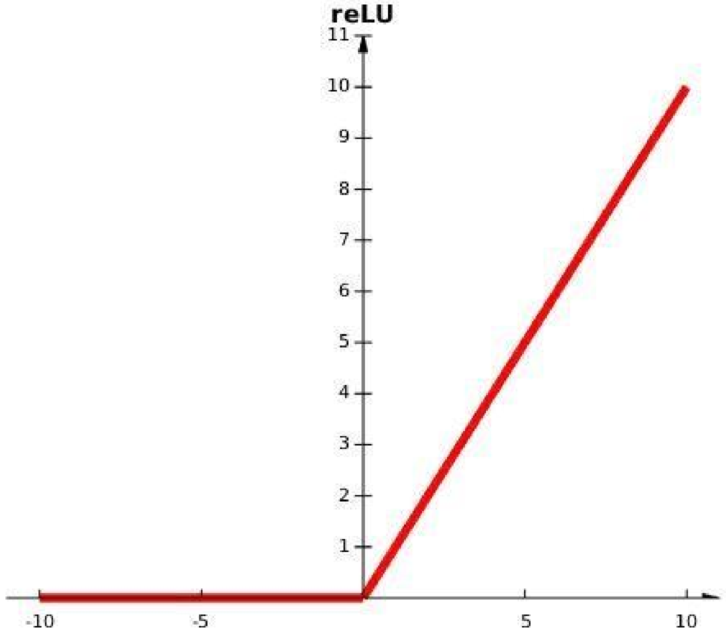


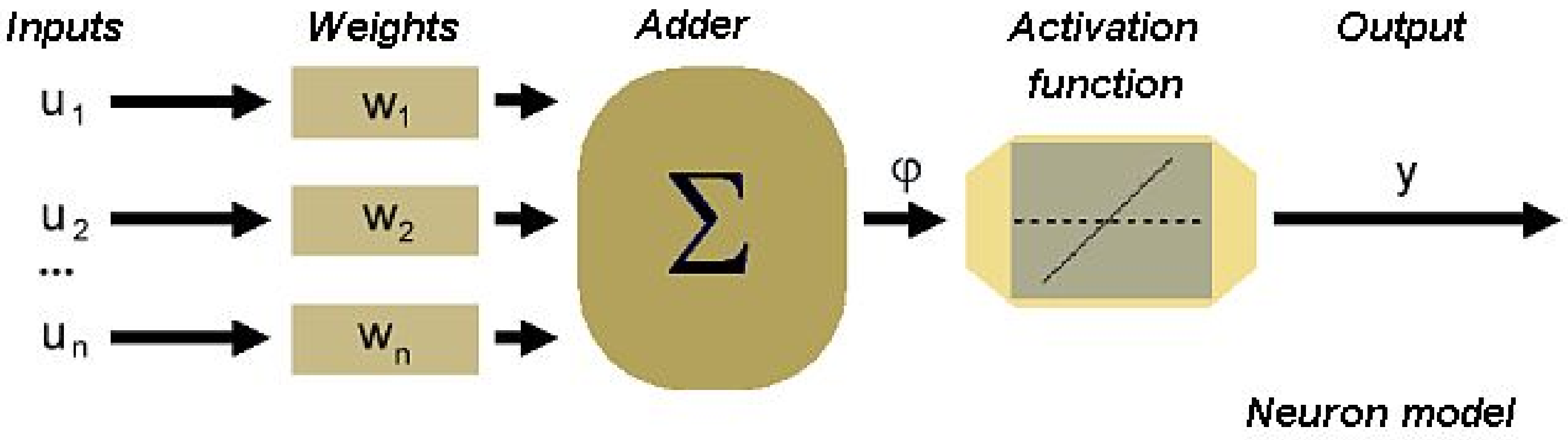
Neuron Activations

- Different type of activation functions

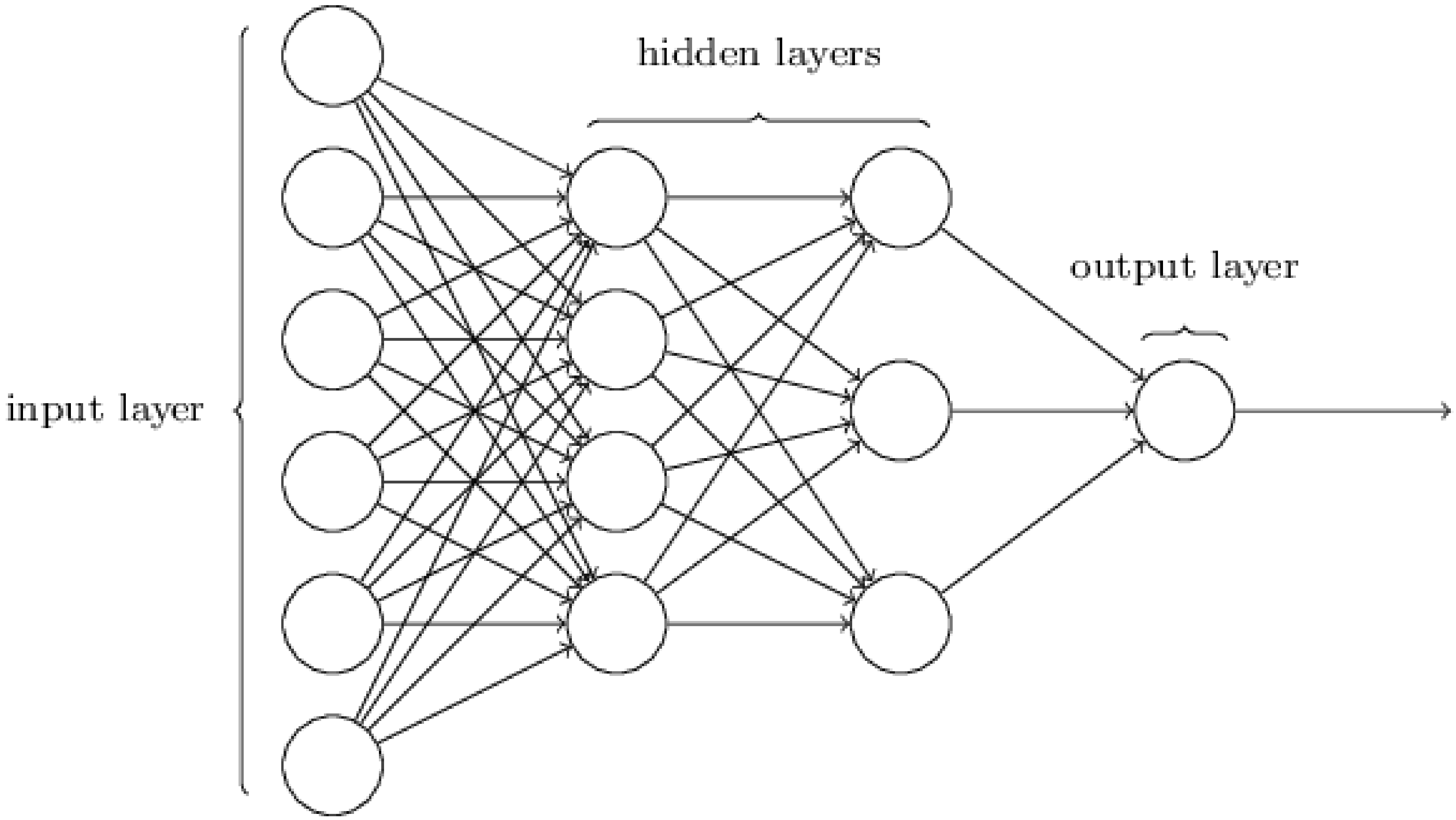
rectified Linear Unit activation

$$y = \max(0, \phi)$$



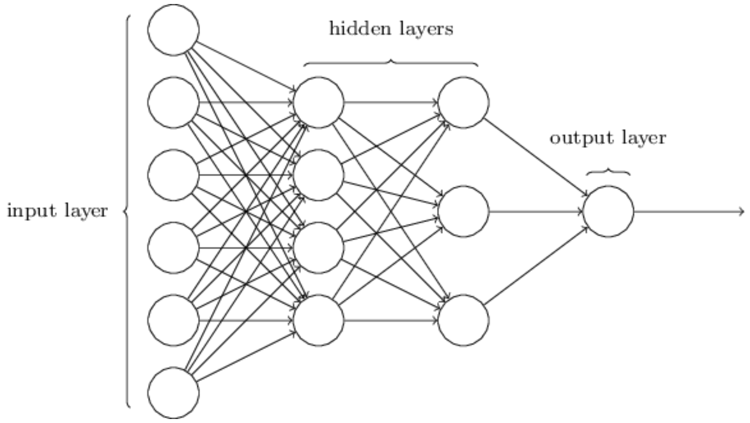


Simple Neural Network



- good reference: neuralnetworksanddeeplearning.com

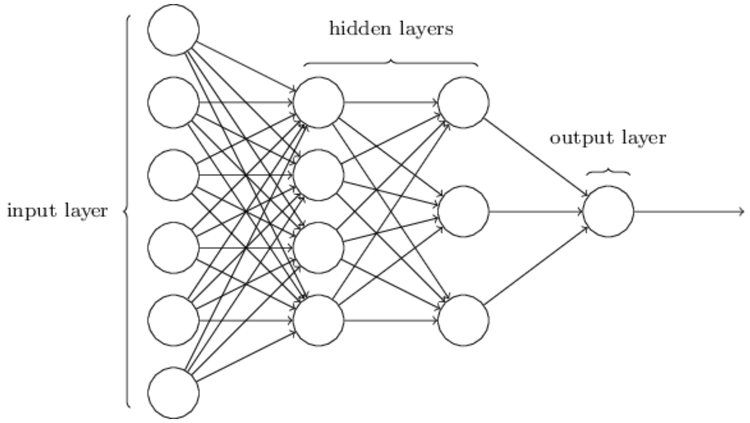
Simple Neural Network



Why is it good?

- inspired by brain
- highly non-linear → function approximator
- shown to solve complicated tasks successfully
- “self-organizing map” / parameters are learned
- data driven

Simple Neural Network



Why is it bad?

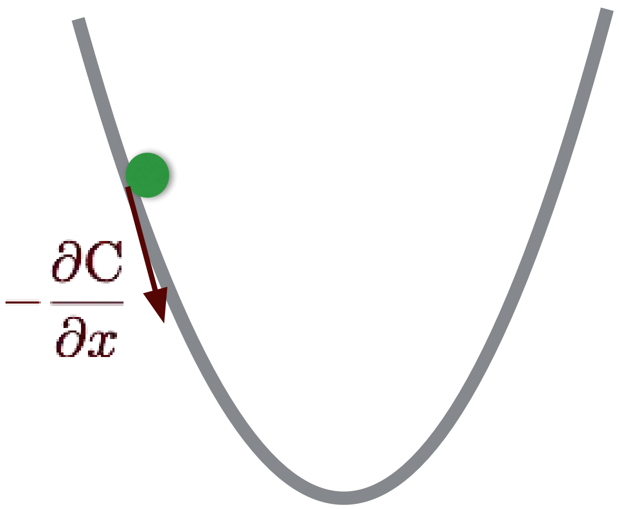
- inspired by brain
- highly non-linear
- high degree of freedom
- shown to solve complicated tasks successfully
- “self-organizing map”
  → parameters are learned
- data driven
- not much theory yet

Back Propagation

- Define a cost function

$$C(w,b) \equiv \frac{1}{2n}\sum_x \|\overset{\text{prediction}}{y(x)} - \overset{\text{ground truth}}{a}\|^2$$

- **Goal**: Find global minimum

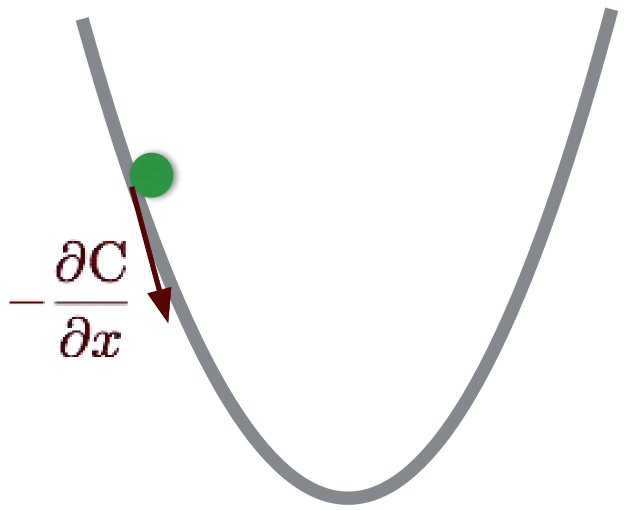


Back Propagation

- Define a cost function

$$C(w,b) \equiv \frac{1}{2n}\sum_x \|\overset{\text{prediction}}{\overset{\downarrow}{y(x)}} - \overset{\text{ground truth}}{\overset{\swarrow}{a}}\|^2$$

- **Goal**: Find global minimum



Gradient Descent

Back Propagation

- Define a cost function

$$C(w,b) \equiv \frac{1}{2n}\sum_x \|\overset{\text{prediction}}{y(x)} - \overset{\text{ground truth}}{a}\|^2$$

- Minimize the error (derivative of the cost) with gradient descent

Gradient descent update rule:

$$w_k \rightarrow w'_k = w_k - \eta\frac{\partial C}{\partial w_k}$$

$$b_l \rightarrow b'_l = b_l - \eta\frac{\partial C}{\partial b_l}$$

Learning rate

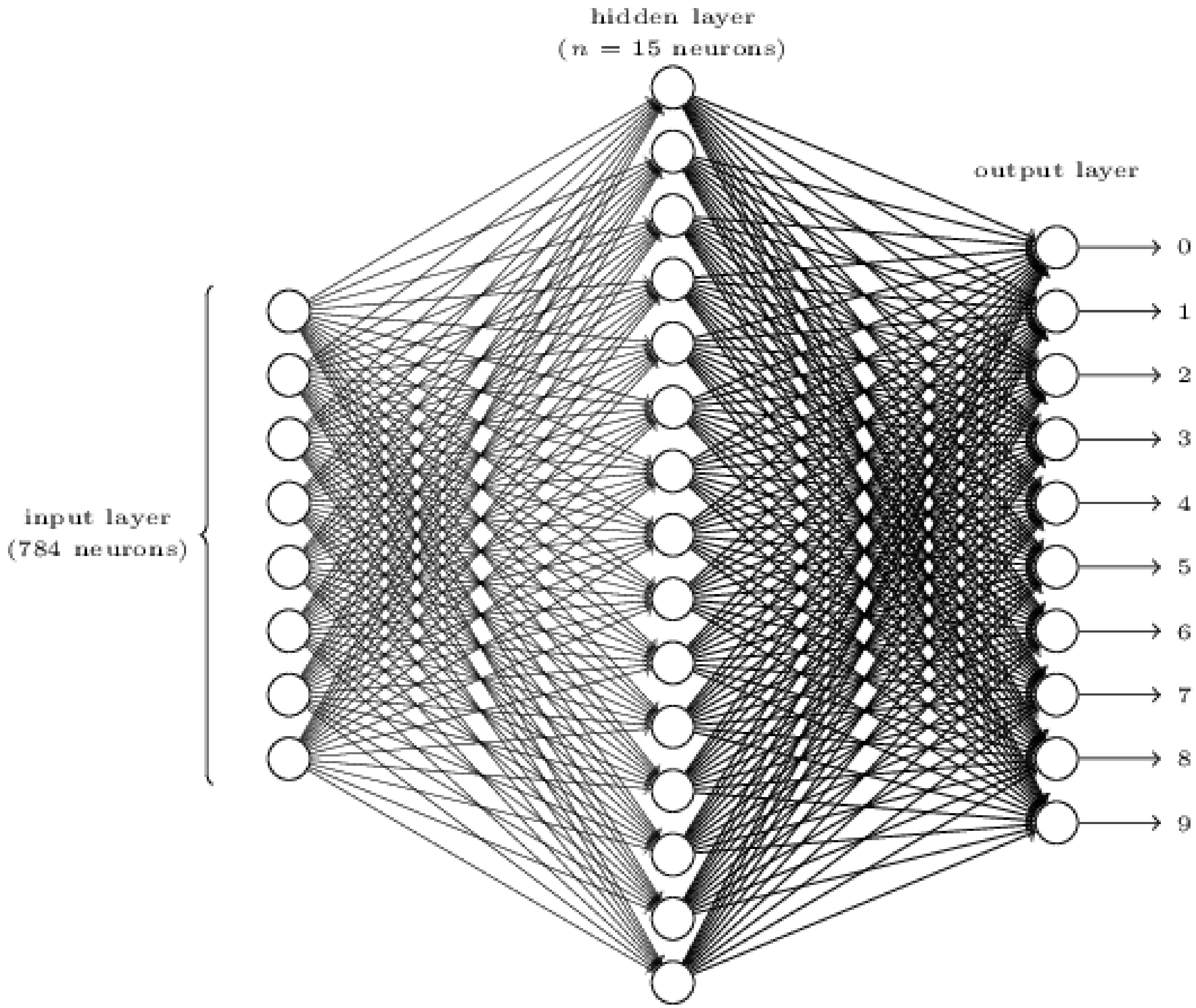
Back Propagation

- Define a cost function

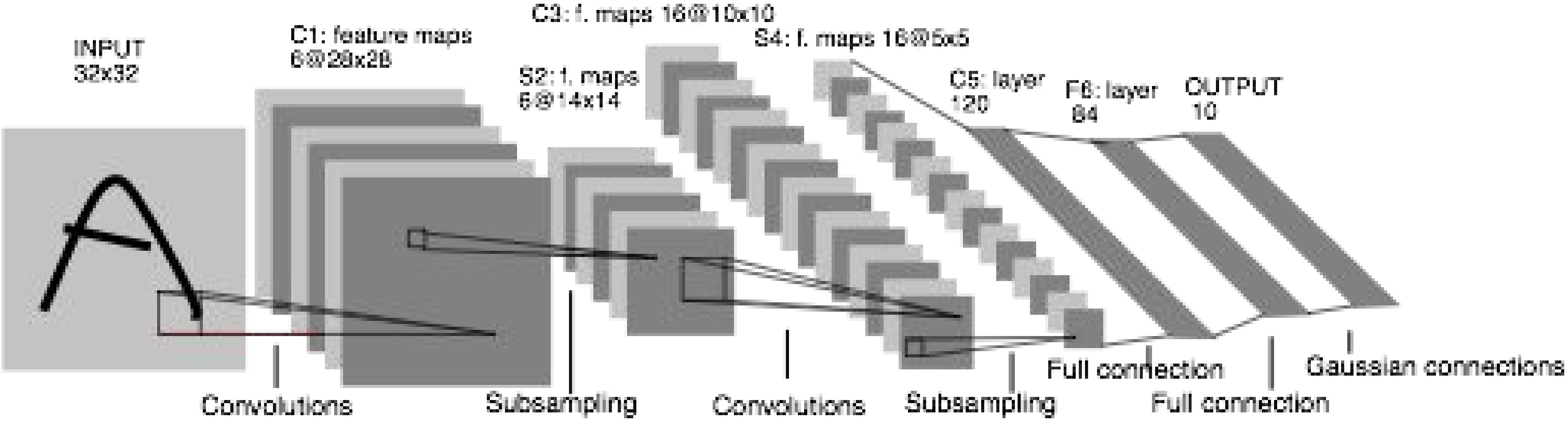
$$C(w,b) \equiv \frac{1}{2n}\sum_x \|\overset{\text{prediction}}{y(x)} - \overset{\text{ground truth}}{a}\|^2$$

- Propagate the error through layers
  - Take the derivative of the output of the layer w.r.t the input of the layer
  - Apply update rule for the parameters

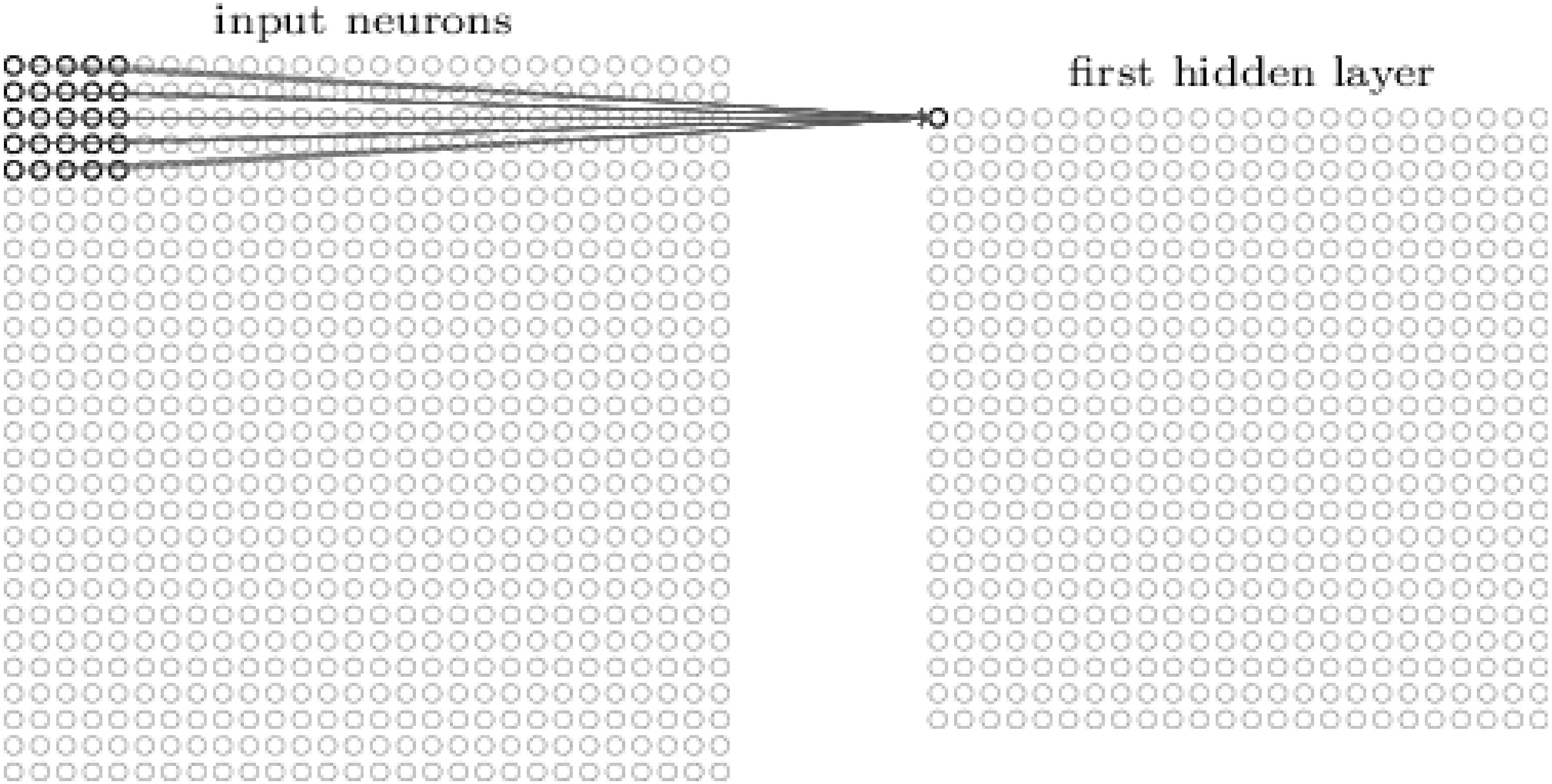
MNIST Digit Classification with NN



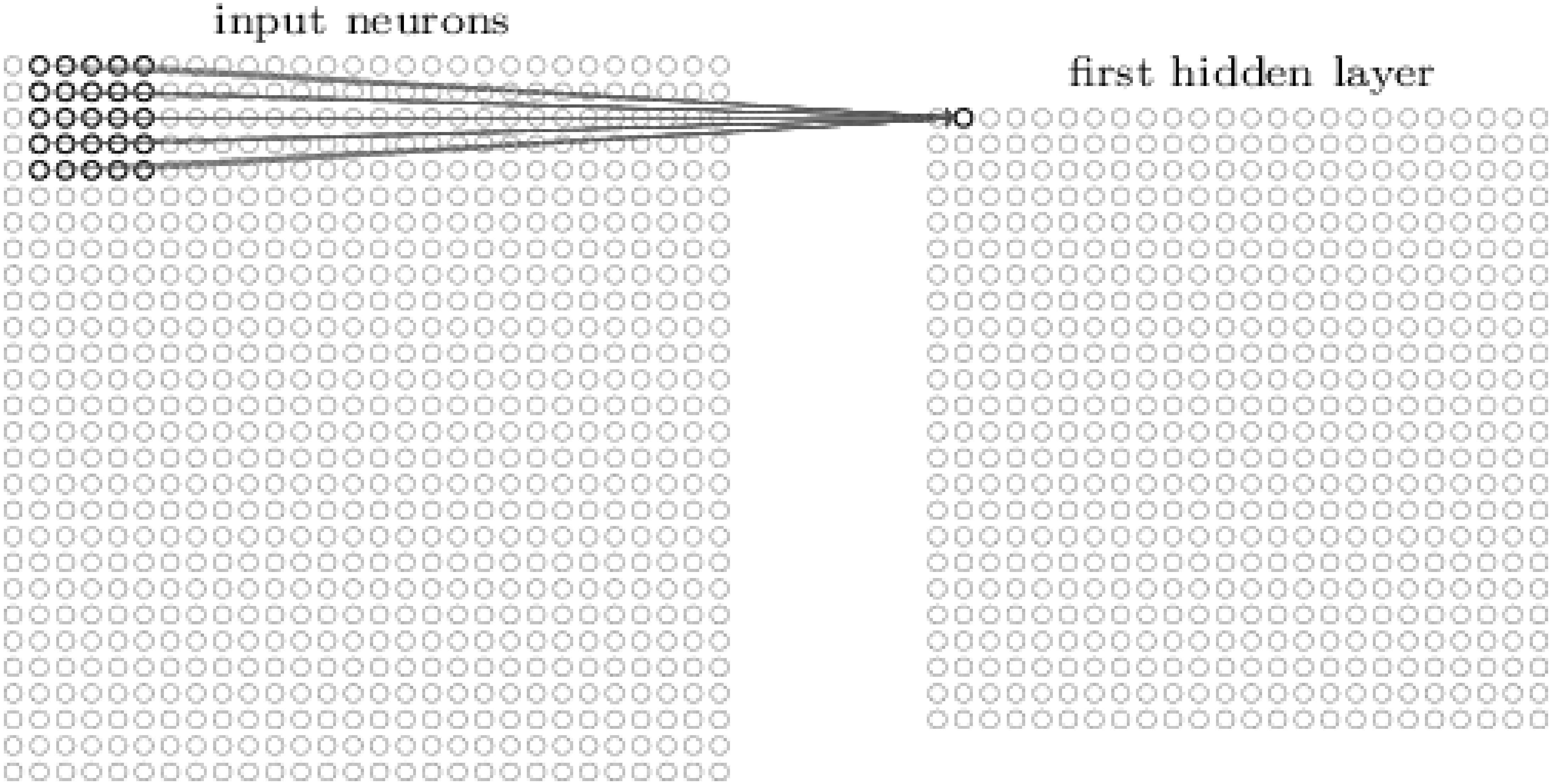
Convolutional Neural Networks



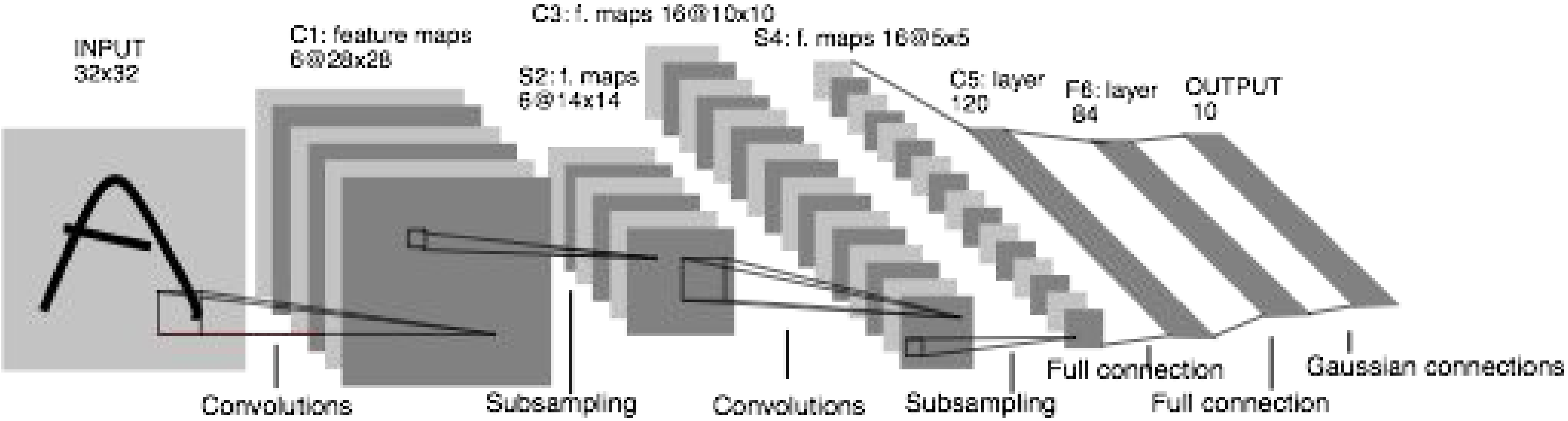
Convolutional Neural Networks



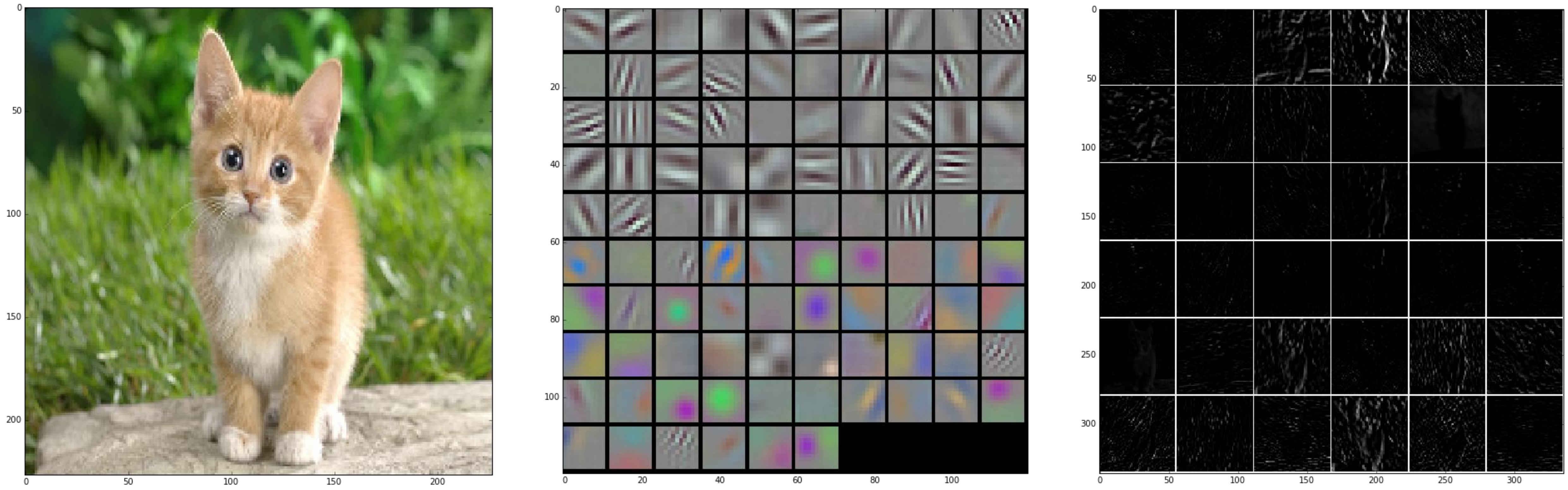
Convolutional Neural Networks



Convolutional Neural Networks



Convolutional Neural Networks

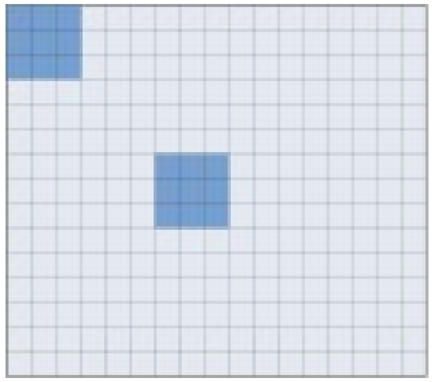


AlexNet
Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton.
"Imagenet classification with deep convolutional neural networks.”
Advances in neural information processing systems. 2012.

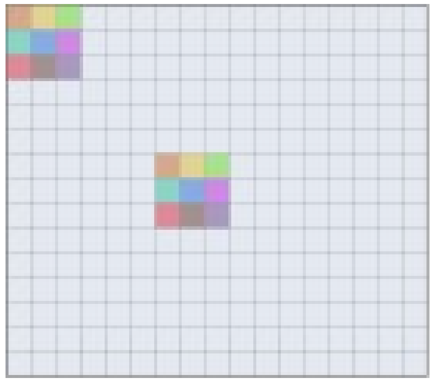
Convolutional Neural Networks

Convolutional networks take advantage of the properties of natural signals:

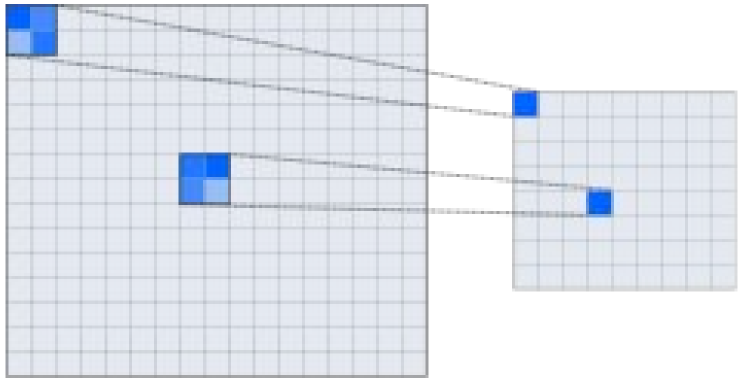
- local connections



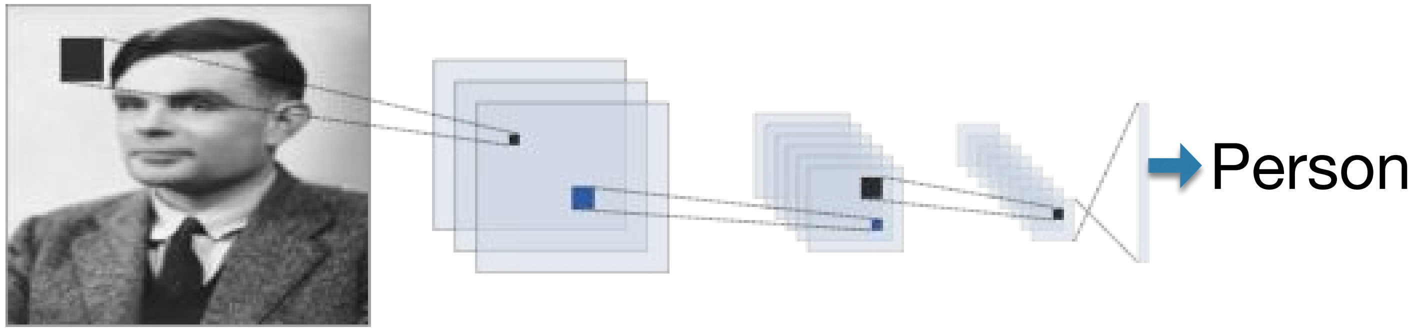
- shared weights



- pooling

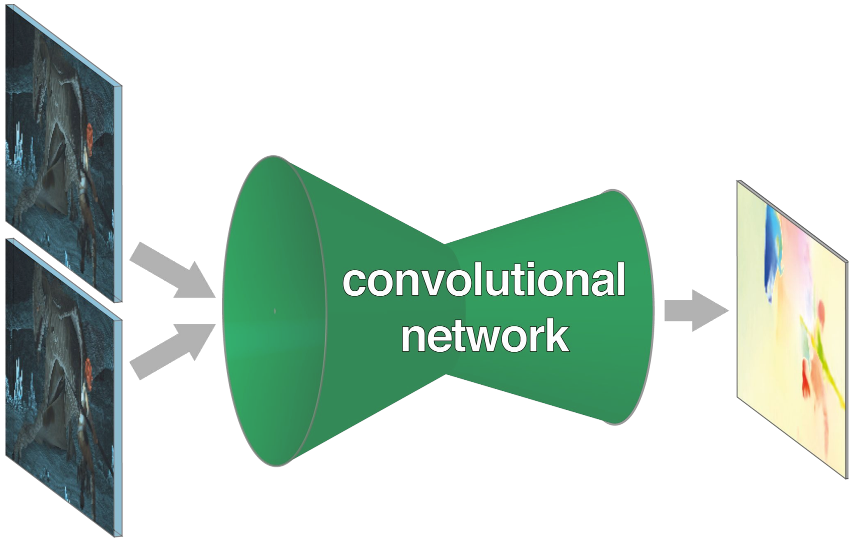


- the use of many layers

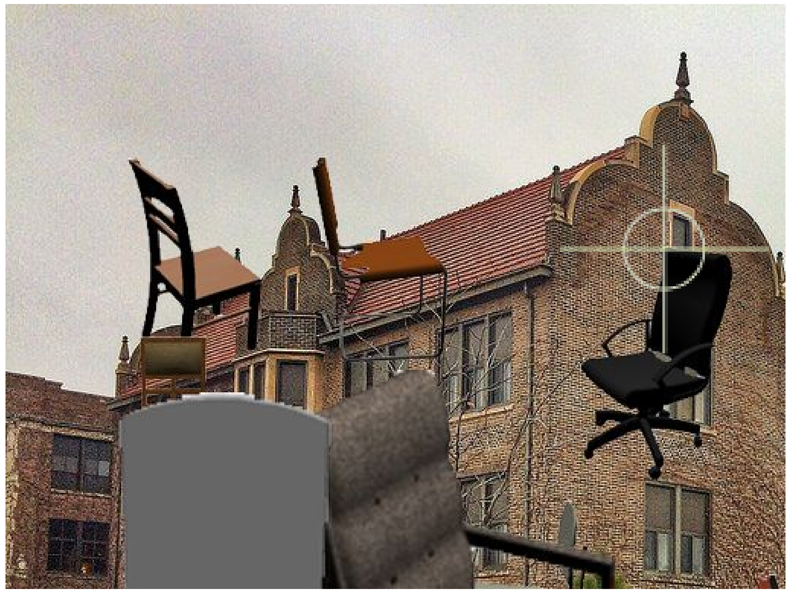


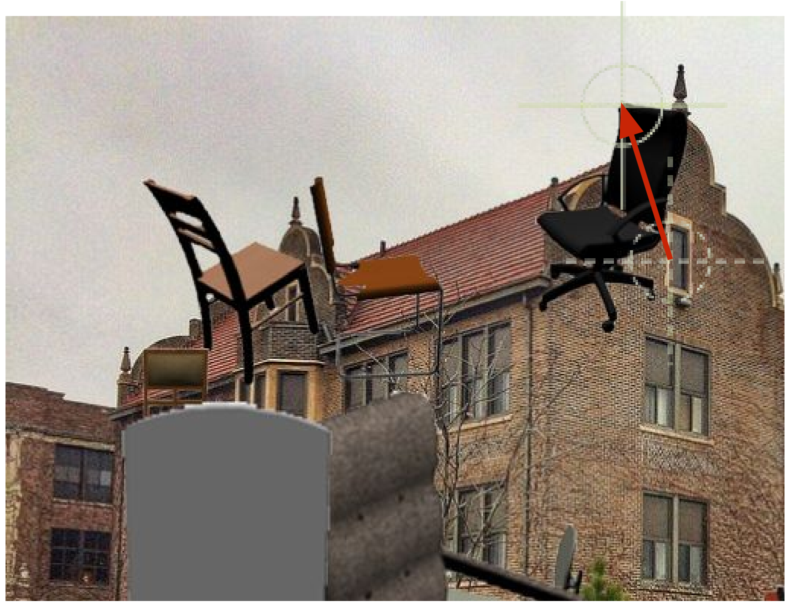
FlowNet: Learning Optical Flow with Convolutional Networks

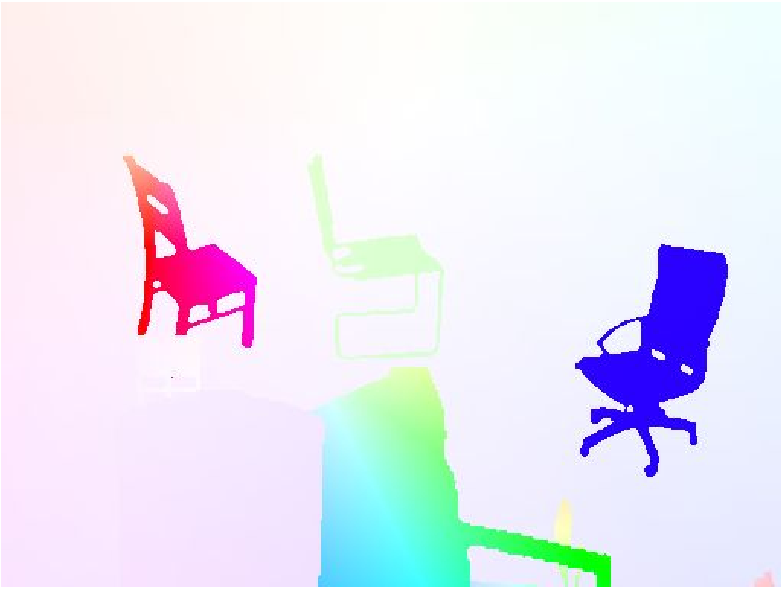
Philipp Fischer, Alexey Dosovitskiy, Eddy Ilg, Thomas Brox
Philip Häusser, **Caner Hazırbaş**, Vladimir Golkov, Daniel Cremers, Patrick van der Smagt

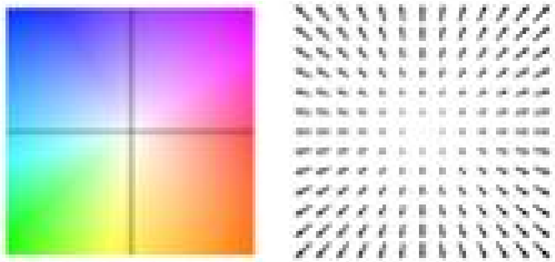


Flying Chairs

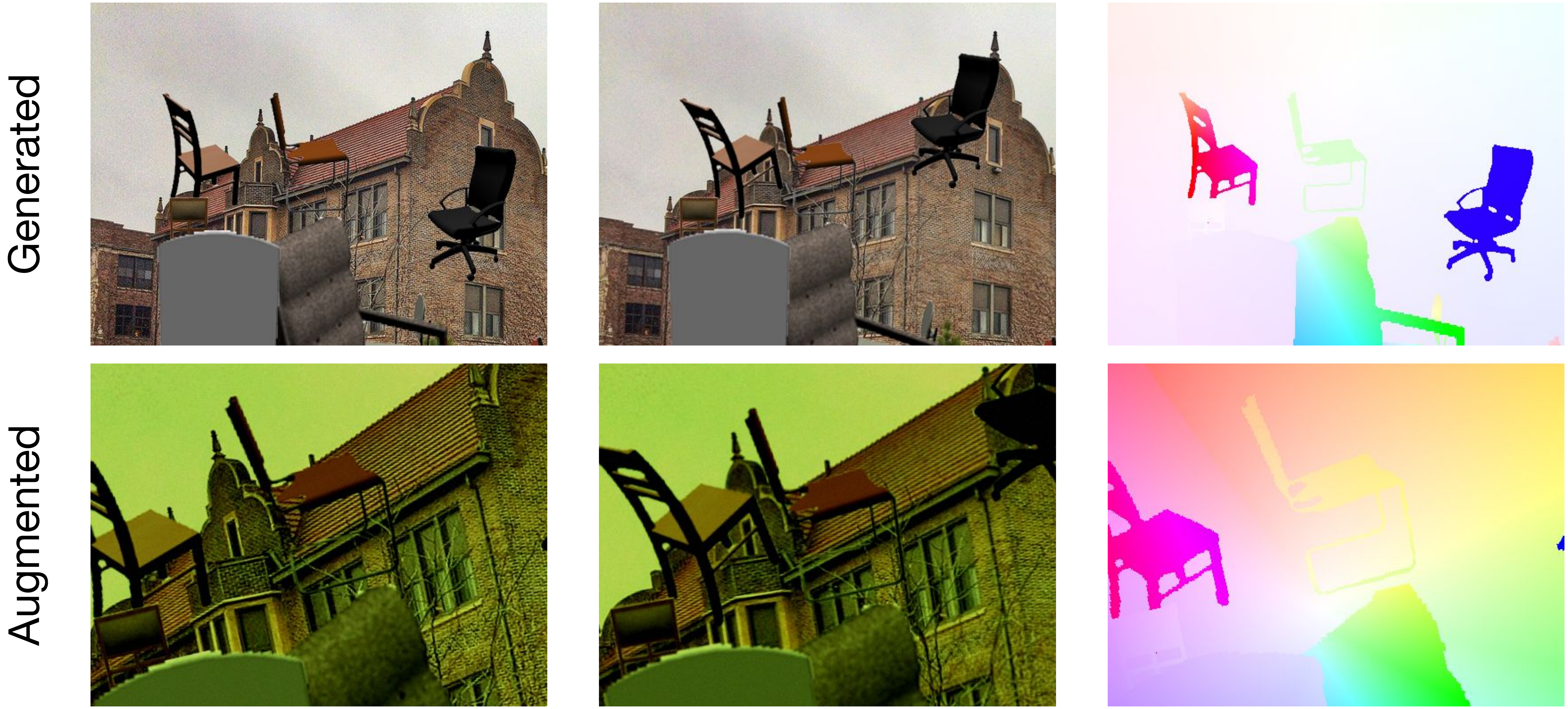






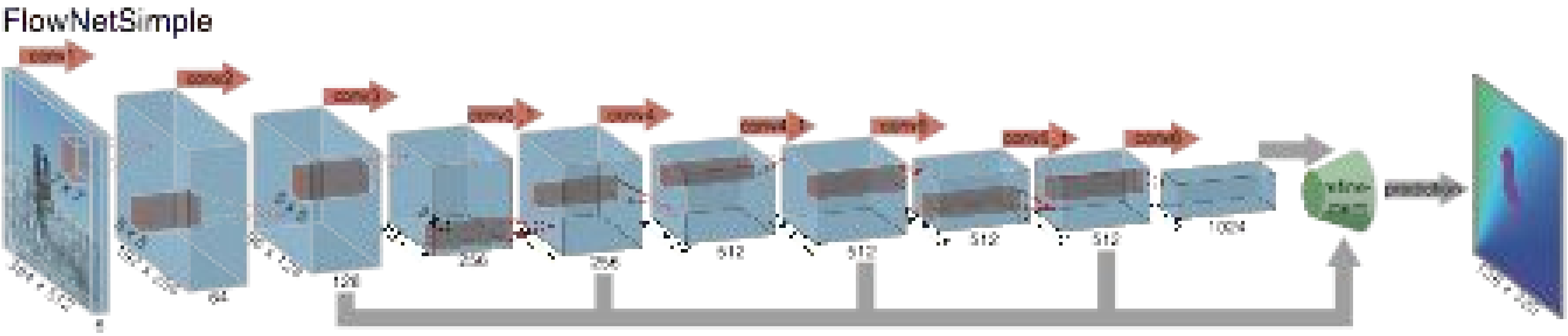


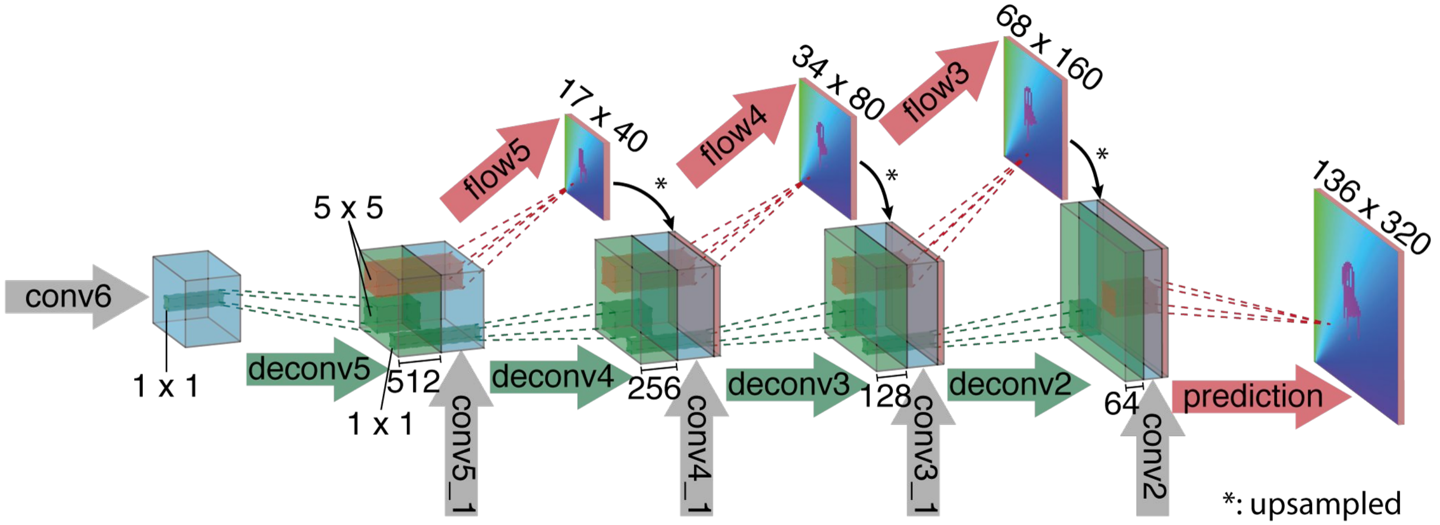
Data Augmentation



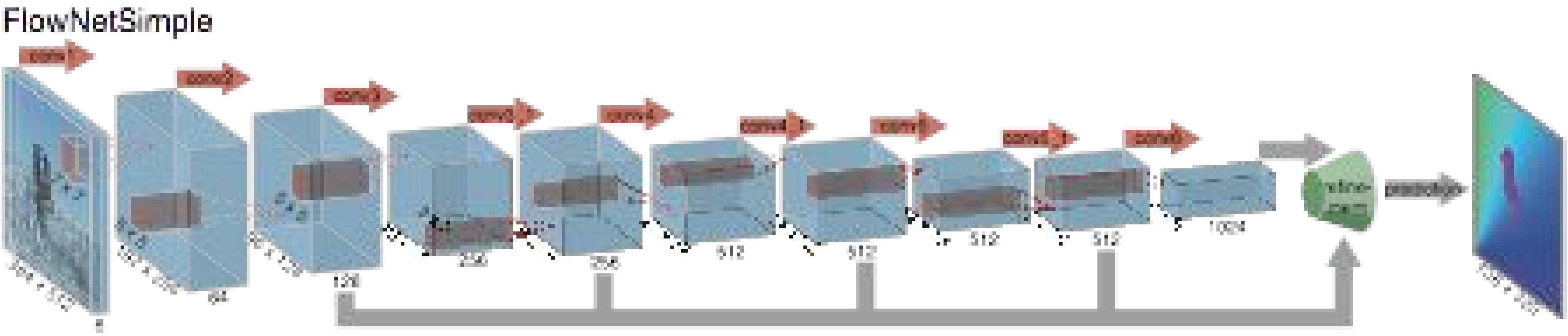
- *translation*, *rotation*, *scaling*, additive *Gaussian noise*
- changes in *brightness*, *contrast*, *gamma* and *colour*

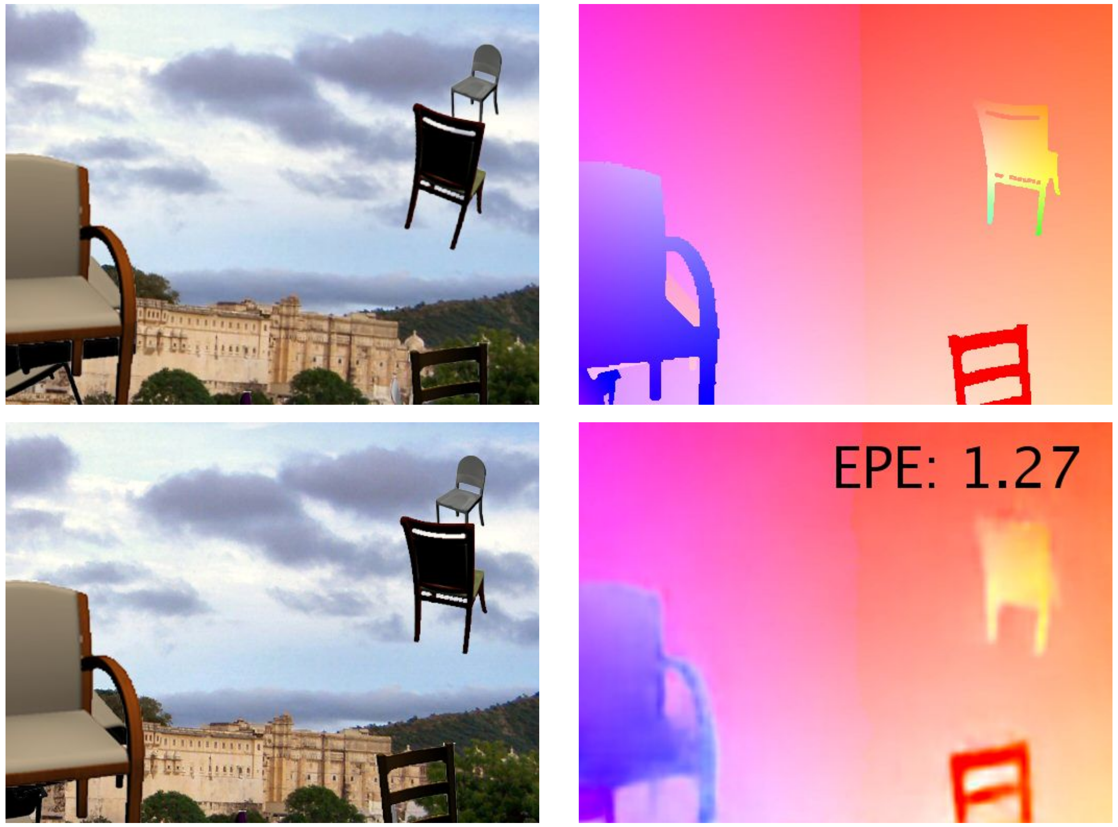
FlowNetSimple



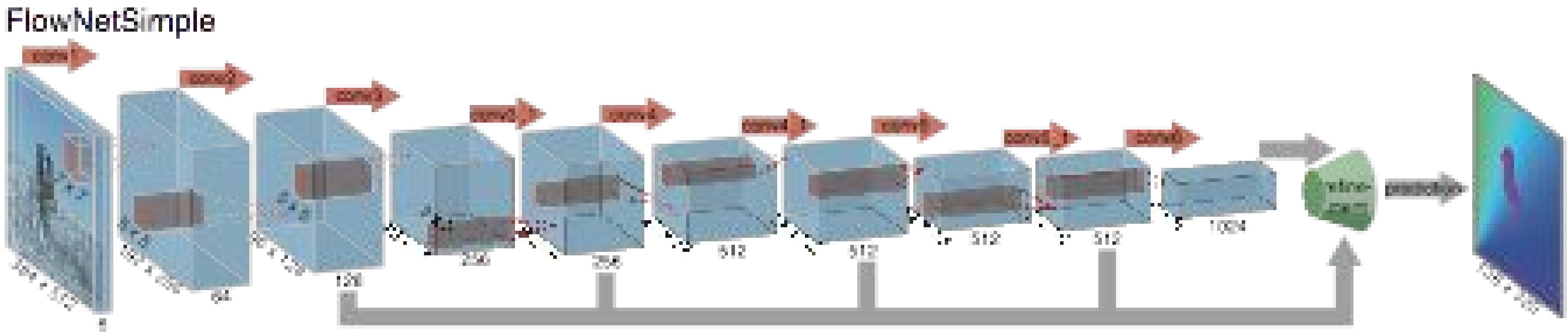


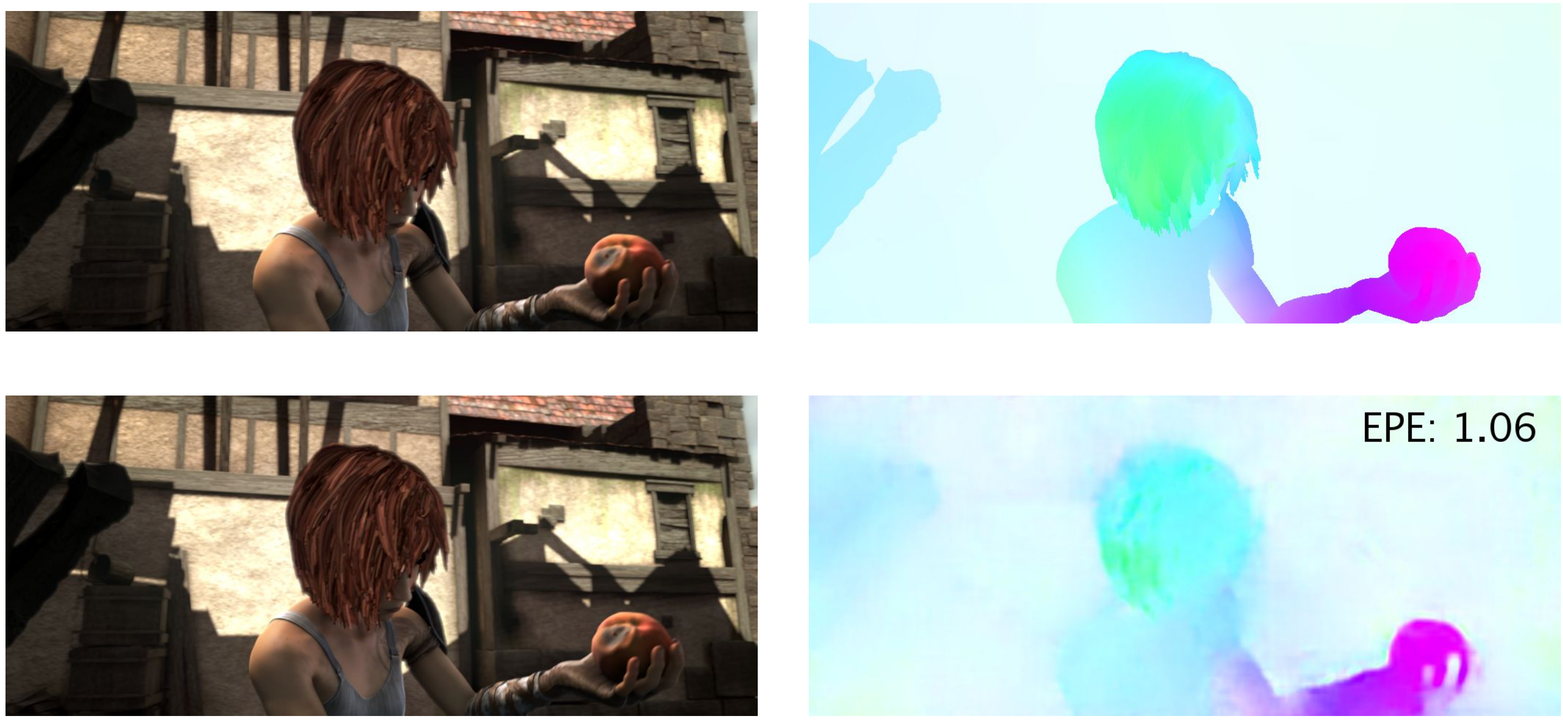
FlowNetSimple - Flying Chairs



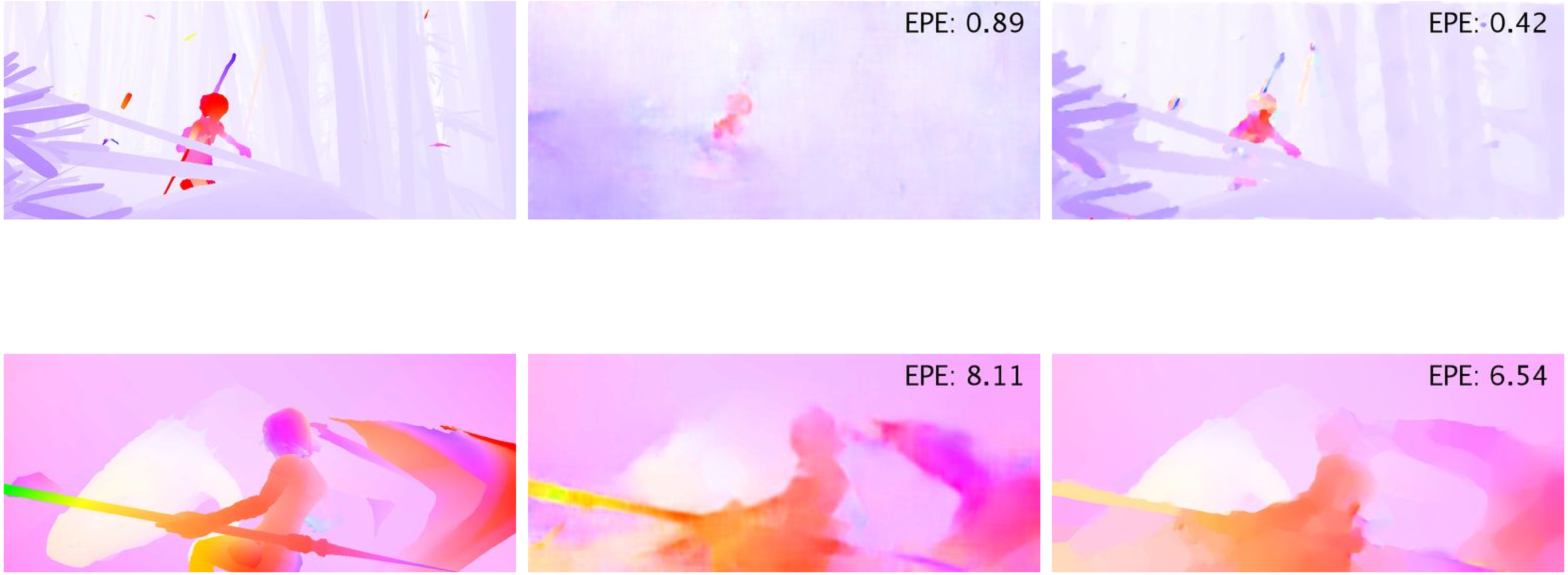


FlowNetSimple - Sintel

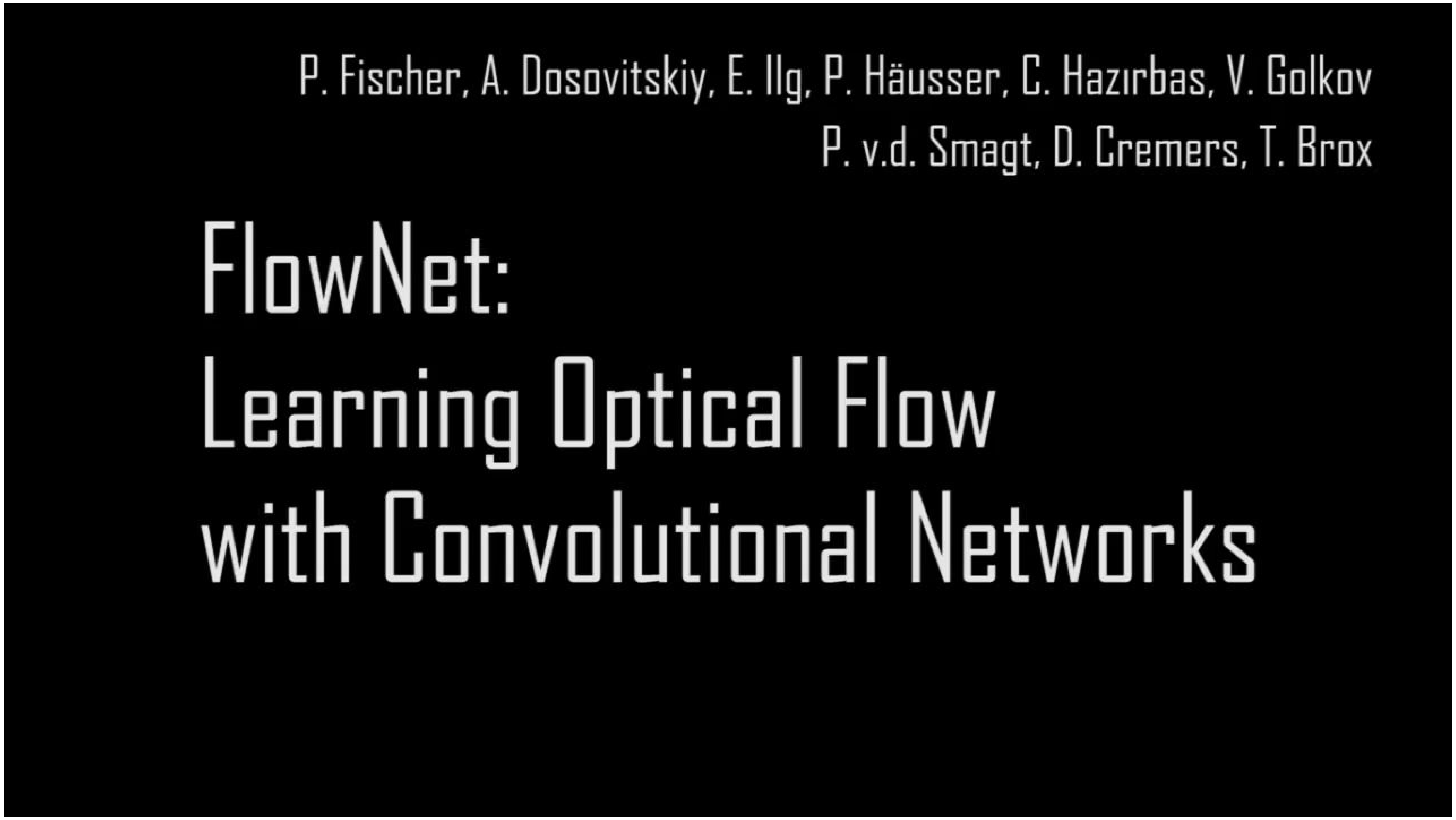




FlowNetSimple + Variational Smoothing



FlowNet: Learning Optical Flow with Convolutional Networks



P. Fischer, A. Dosovitskiy, E. Ilg, P. Häusser, C. Hazırbas, V. Golkov
P. v.d. Smagt, D. Cremers, T. Brox

FlowNet:
Learning Optical Flow
with Convolutional Networks