# The current landscape of Deep Learning

## Trends and Challenges

Soumith Chintala
Facebook AI Research
soumith@gmail.com

IIIT Hyderabad 18th February, 2015

# Deep Neural Networks

- Feed-forward Neural Networks
  - Convolutional neural networks
- Recurrent Neural Networks
- Boltzmann Machines

# Domains under discussion

- Vision (images, videos, RGB-D)
- Natural Language Processing
- Neural Machines

# Optimization and scaling

- Convex and non-convex optimizers
- Weight initialization and conditioning
- Distributed training

# ConvNets for Images: ImageNet Challenge

Top-5 Error on Imagenet Classification Challenge (1000 classes)

- AlexNet ~15%
- OverFeat ~ 13%
- ZeilerNet ~11%
- Oxford-VGG ~7%
- GoogLeNet ~6%, ~4.5%
- PReLU (MSR) ~4.9%
- Human Performance 3 to 5%

Deeper Networks
Smaller convolutions

## Trends in ConvNet Architectures

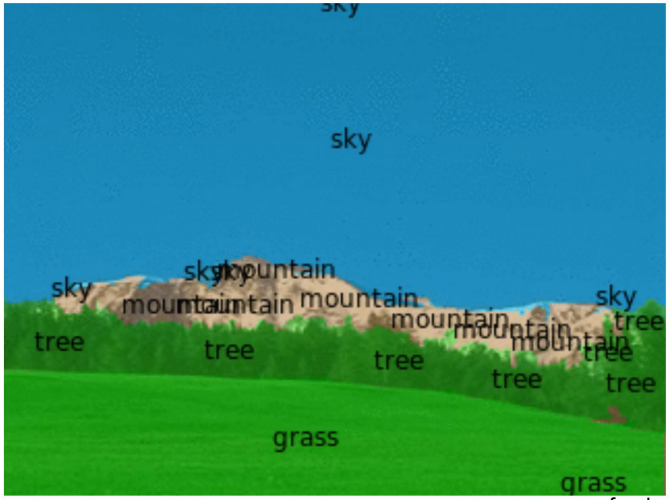
- Smaller kernels
  - increase in discriminative power
- Deeper Networks
- Regularized fully connected layers
  - Dropout, MaxOut, DropConnect

# ConvNets for Image Classification

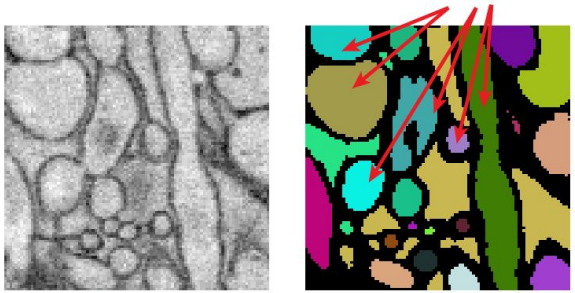
- How to win?
  - Pick a dataset
  - Train a ConvNet
  - Beat state-of-the-art
  - Publish
  - Win!

# ConvNets for Image Segmentation

- MALIS
- Watershed + ConvNets + Trees
- Plenty of scope for improvement
  - Clement Farabet, Camille Couprie, Srini Turaga

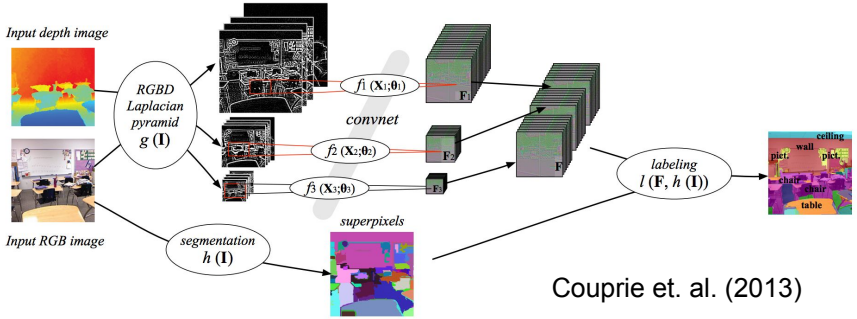



farabet (2012)

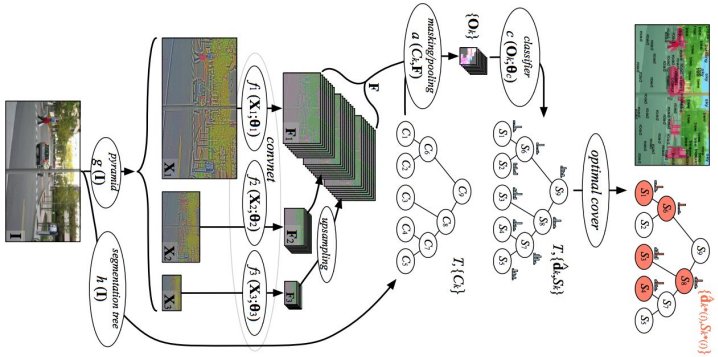



turaga (2009)

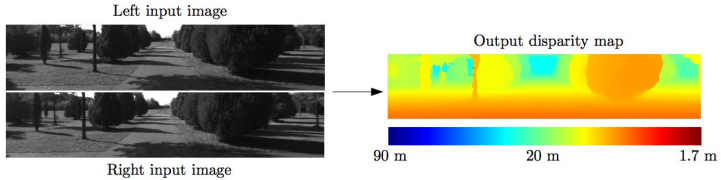
# Segmentation and Stereo

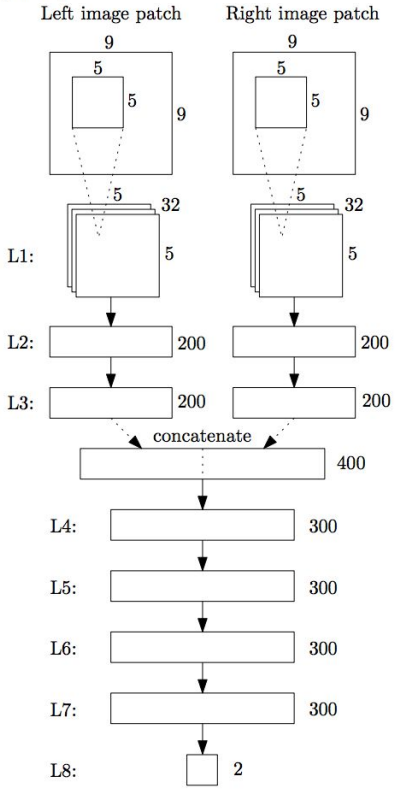



Couprie et. al. (2013)



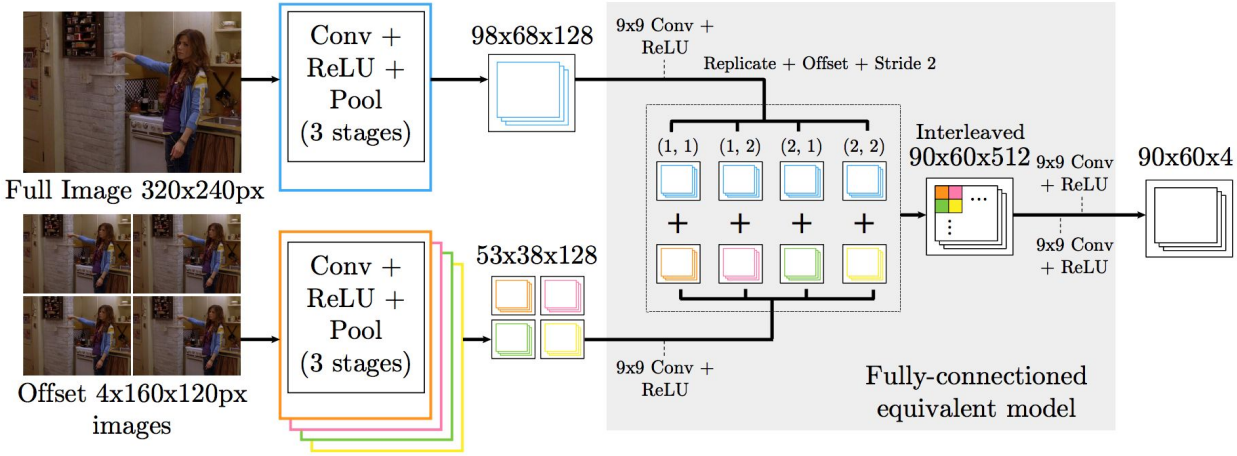

Farabet et. al. (2012)

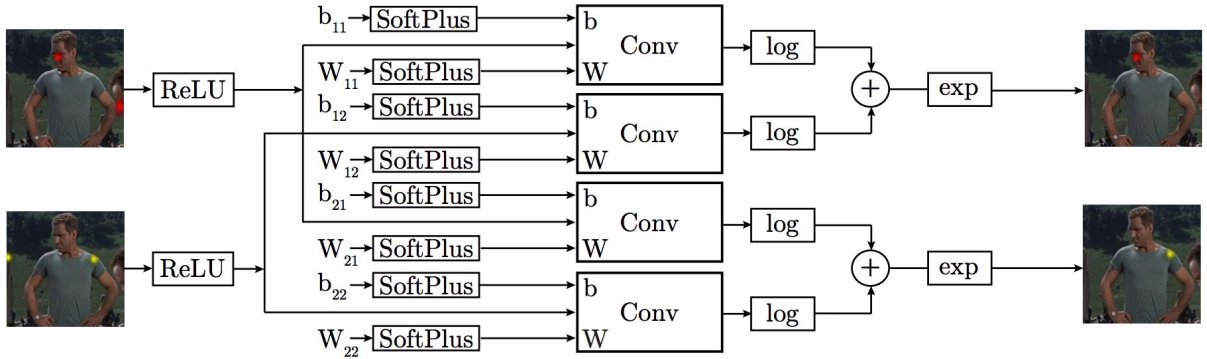


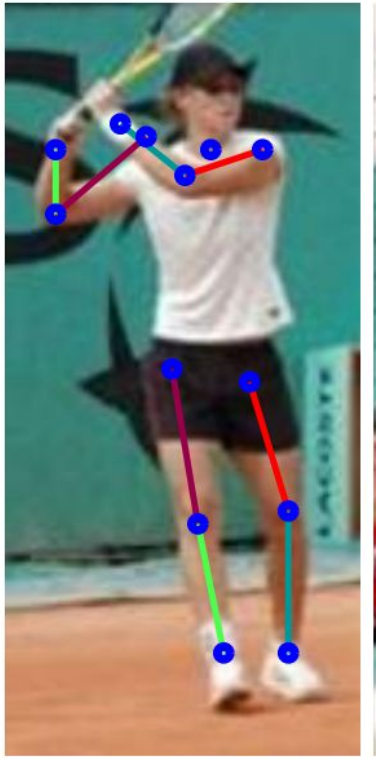




LeCun (2014)

# ConvNets + Graphical Model (Tompson et. al. 2014)



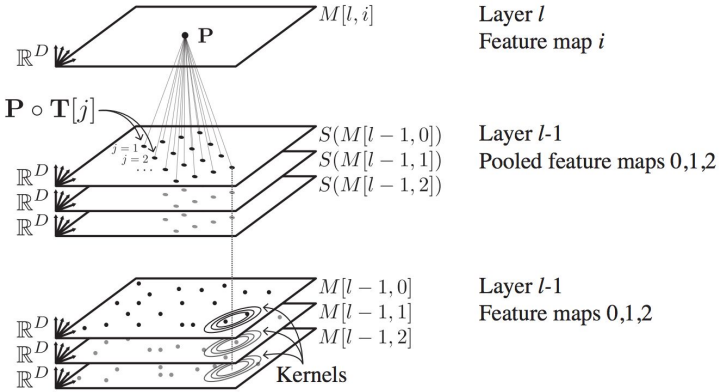



# Inferring groups and generalizing ConvNets

Chen et. al. (2014)

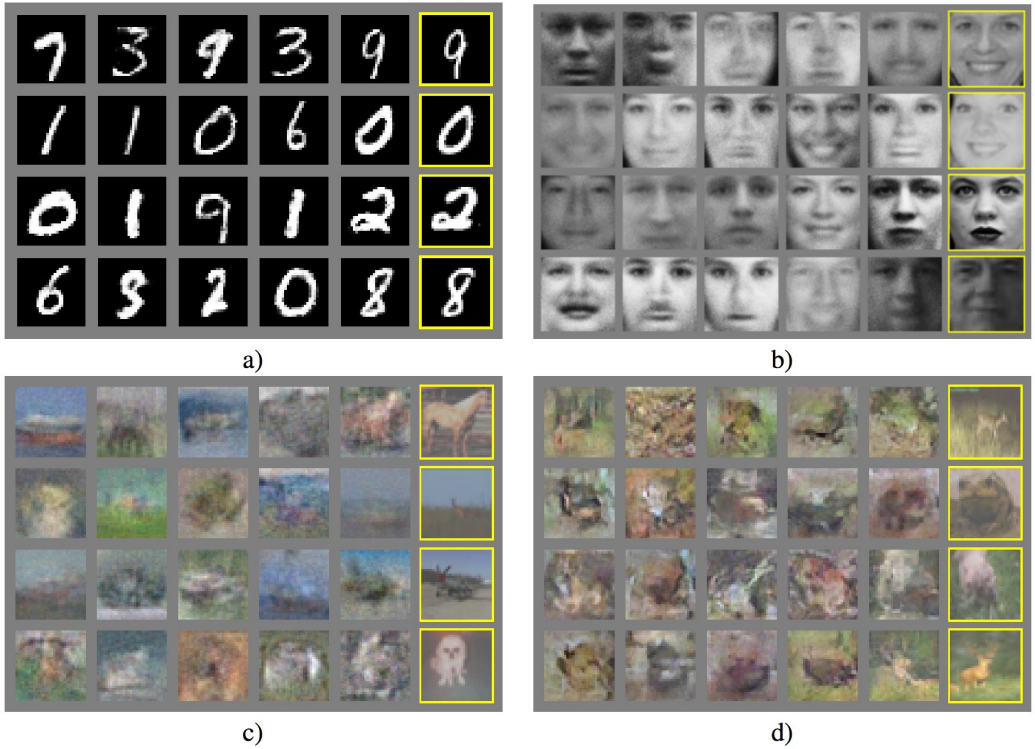
- Deep Haar Scattering on Graphs
- Deep Symmetry Networks

gens et. al. (2014)

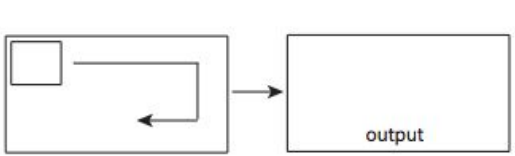


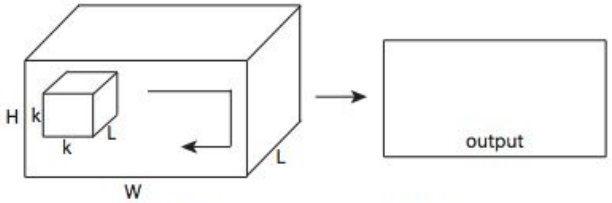
# Generative Adversarial Nets

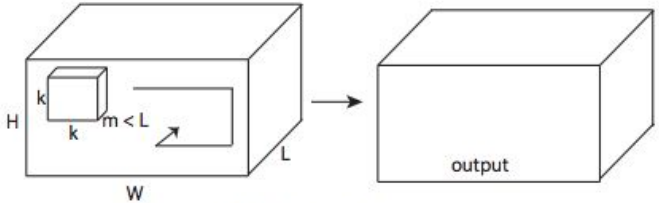
Goodfellow et. al. (2014)



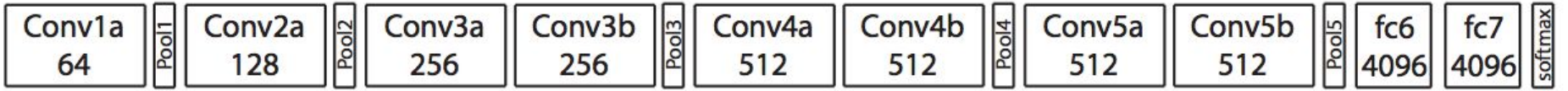
# ConvNets for Video



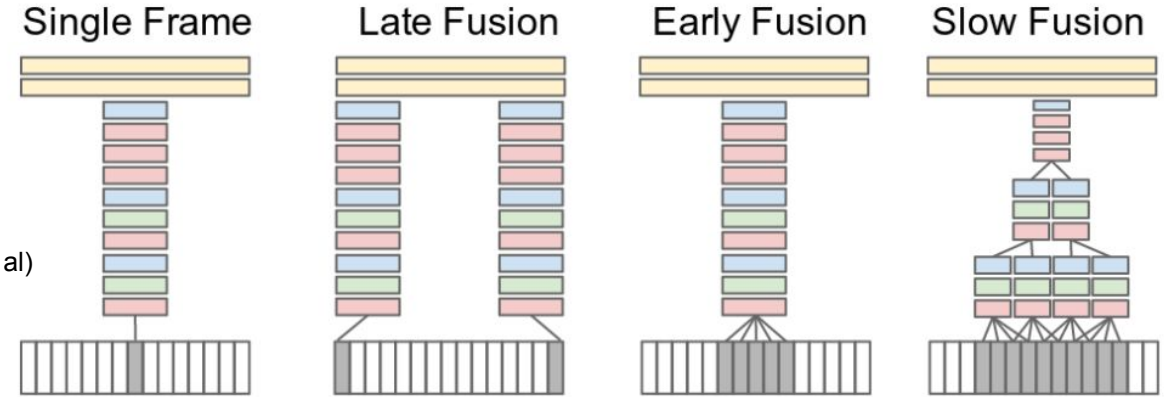



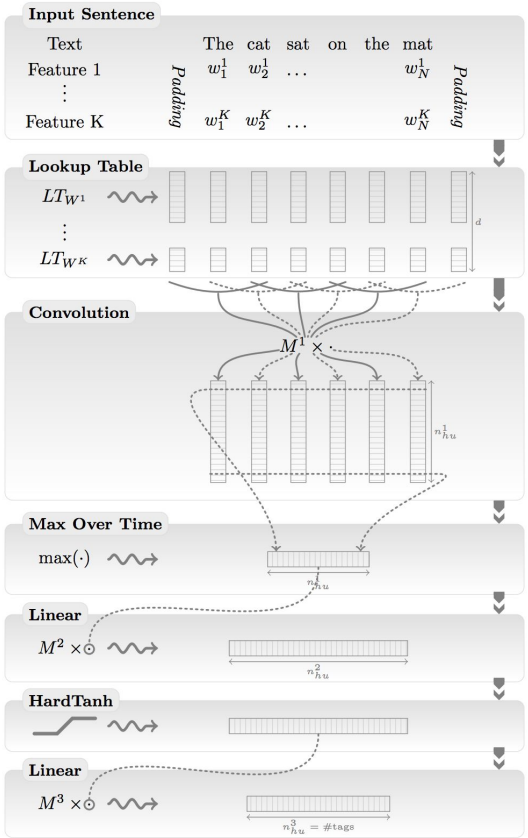

C3D (Tran et. al.) (a) 2D convolution (b) 2D convolution on multiple frames (c) 3D convolution




DeepVideo (Karpathy et. al)

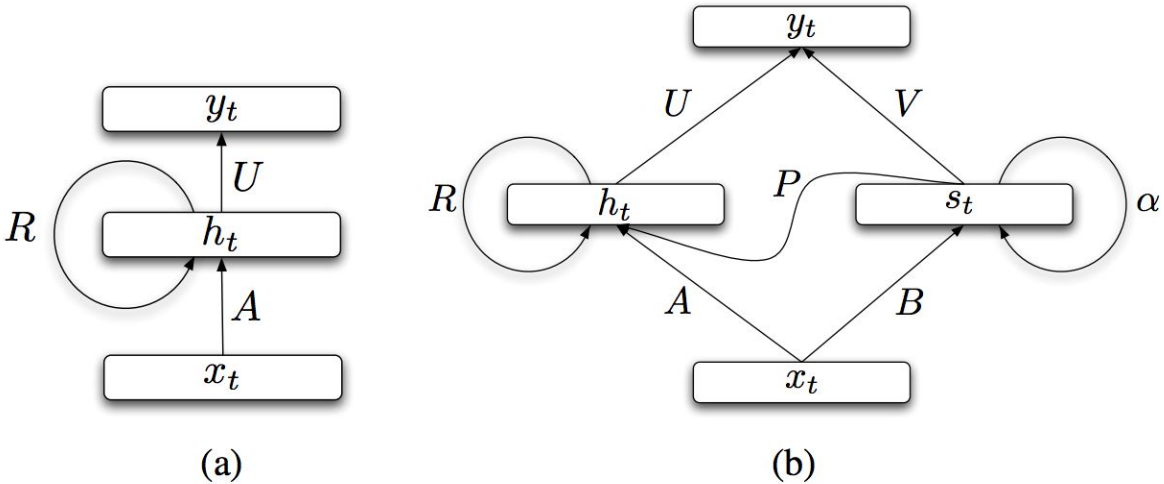


# ConvNets for NLP



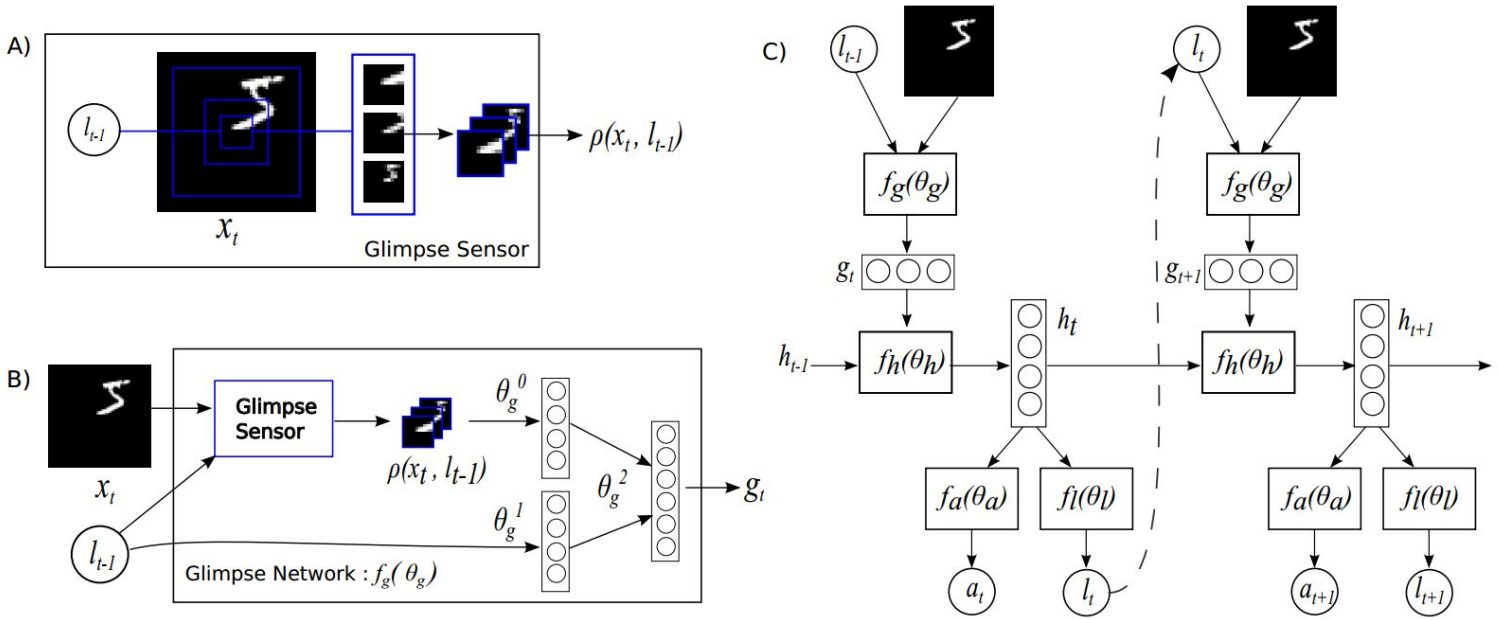

Collobert et. al. (2011)

# Recurrent Neural Networks



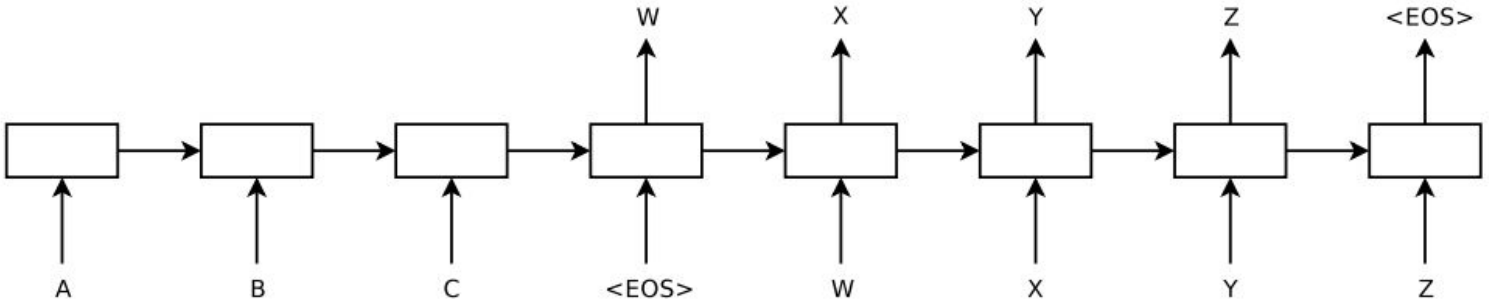

Mikolov 2014

# RNNs for visual attention



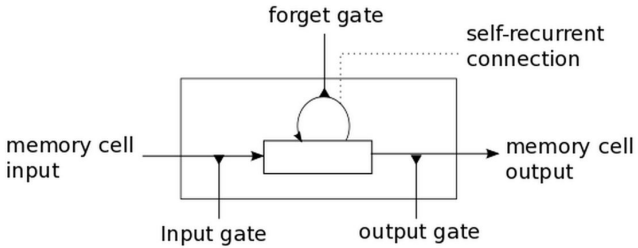

Mnih et. al. (DeepMind) 2014

# RNN-LSTMs




Sutskever et. al. (2014)

- Machine Translation
- Language Modeling
- Learning to execute (Python programs)




Basic LSTM unit (figure from deeplearning.net)

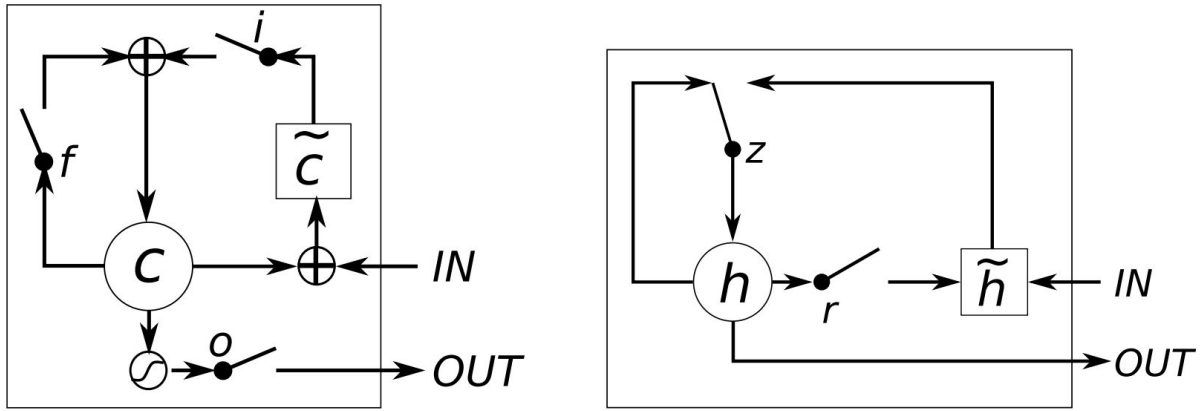
# RNN - Memory Networks for Q&A

Bilbo travelled to the cave. Gollum dropped the ring there. Bilbo took the ring.
Bilbo went back to the Shire. Bilbo left the ring there. Frodo got the ring.
Frodo journeyed to Mount-Doom. Frodo dropped the ring there. Sauron died.
Frodo went back to the Shire. Bilbo travelled to the Grey-havens. The End.
Where is the ring? A: Mount-Doom
Where is Bilbo now? A: Grey-havens
Where is Frodo now? A: Shire

Weston et. al. 2014

# Alternative memory

- Slow RNNs
  - Mikolov et. al. “Learning Longer MEMORY IN RECURRENT NEURAL NETWORKS”
- Gated Recurrent Units

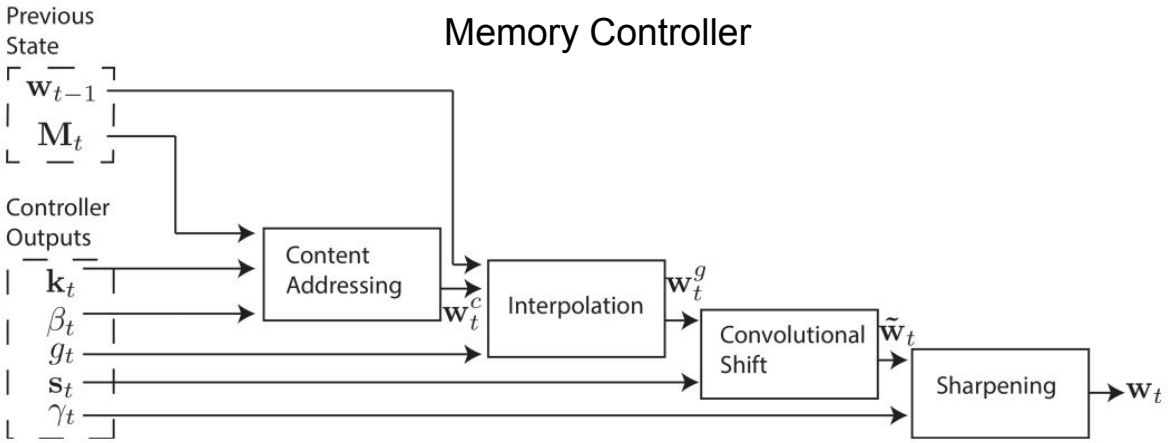



(a) Long Short-Term Memory (b) Gated Recurrent Unit

Chung et. al. (2015)

# Neural Machines: Turing Machine

Graves et. al. 2014 (DeepMind)




- Copying
- Sorting

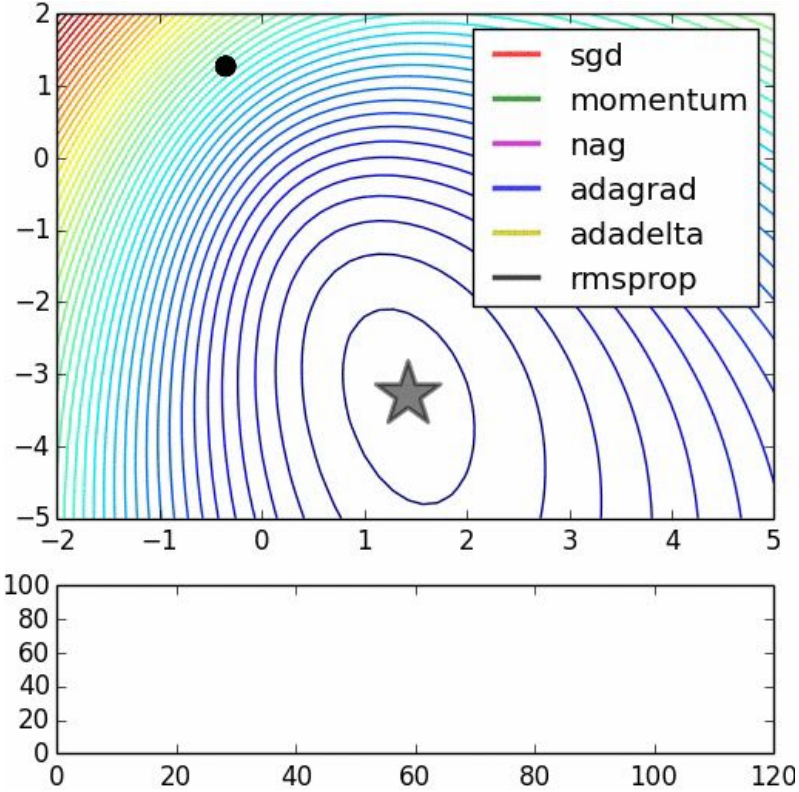
# Neural Machines

- Neural Turing Machine
  - Graves et al. 2014
- Memory Networks
  - Weston et. al. 2014
- Learning to Execute programs
  - Zaremba et. al. 2014
- Discovering Efficient Mathematical Identities
  - Zaremba et. al. 2014

# Optimization

- New algorithms
  - ADAM
- Distributed optimization
  - Hogwild, Dogwild
  - Downpour
  - Elastic Averaging SGD
- Understanding high-dimensional spaces
  - Loss Surfaces of Multilayer Networks - Chormanska et. al 2014
  - Exact solutions to the nonlinear dynamics of learning in deep linear neural networks - Saxe et. al. 2014
  - Identifying and attacking the saddle point problem in high-dimensional non-convex optimization - Dauphin et. al. 2014

Visualization by Alec Radford (Indico)



# Trends

- Deeper nets
- Smaller convolutions
- RNN + LSTM
- Multiple GPUs + Multiple Machines
- Neural Machines
- Other kinds of memory units
- Better weight initialization
- Meta-problems

# Challenges

- Scaling up for big data (videos, social networks etc.)
- Discrete optimization
- Memory that works