

LSTM and Recurrent Neural Nets

Part 1

Sepp Hochreiter

LIT AI Lab, Institute for Machine Learning
Johannes Kepler University

Course Overview

Basic Course of Master Artificial Intelligence

Class:	3 ECTS	2 SWS VO
Exercise:	1.5 ECTS	1 SWS UE

Exams

Class: written **exam at the end of the semester**

Exercise: 6 **bi-weekly homeworks** (evaluated) and **exam** at the end

Content

Recurrent Neural Networks

1 Simple Recurrent Networks

1.1 Jordan Network

1.2 Elman Network

1.3 Fully Recurrent Network

1.4 Backpropagation through Time

1.5 Other RNN Learning Algorithms

1.5.1 Truncated Backpropagation Through Time

1.5.2 Real-Time Recurrent Learning

1.6 Vanishing and Exploding Gradients

1.7 NARX Recurrent Neural Networks

1.8 Time-Delay Neural Networks

Content

Recurrent Neural Networks

2 Long Short-Term Memory

2.1 Backpropagation for LSTM

2.2 Forget Gate

2.3 Tricks of the Trade

2.4 Example: Generating Discrete Sequences with an LSTM

3 Gated Recurrent Unit

3.1 Backpropagation for GRU

3.2 Application

4 Attractor Networks

4.1 Backpropagation for Attractor Networks with continuous time

4.2 Hopfield Networks

4.3 Boltzmann Machine

4.4 Sequence Processing with RNNs

Topics Covered

Architectures

- Elman nets
- Jordan nets
- NARX nets
- time-delay neural networks
- LSTM nets
- GRU nets
- attractor networks
- Hopfield networks
- Boltzmann machines
- echo state networks
- bidirectional RNNs
- recursive networks

LSTM-based Architectures

- ConvLSTM
- Grid-LSTM
- Pyramid-LSTM
- Neural Turing Machine
- Sequence-to-Sequence Learning
- Tree-LSTM

Topics Covered

Learning methods

- Real-Time Recurrent Learning (RTRL)
- Backpropagation through time (BPTT)
- truncated BPTT
- focused backpropagation

Regularization methods

- weight decay
- early stopping
- dropout
- keeping state changes small

Topics Covered

Vanishing and exploding gradient problem for RNNs

Attention mechanisms

- Attention
- Transformer Networks and BERT
- Tinker

Meta-Learning with RNNs

- Meta-Learning with LSTM

Topics Covered

Reinforcement Learning with RNNs

- RUDDER
- LSTM agents
- recurrent actor-critic
- Bram Bakker's architecture

RNNs for credit assignment

- uniform credit assignment
- attention mechanisms for credit assignment

Topics Covered

Applications

- RNN in speech
- RNN for text analysis
- RNN for translation
- RNN for time series prediction
- RNN in finance
- RNN in hydrology

Highlights

- OpenAI uses LSTM Agents for Dota2
- DeepMind uses LSTM Agents for StarCraftII

RNNs win challenges

- speech
- language
- text
- translation

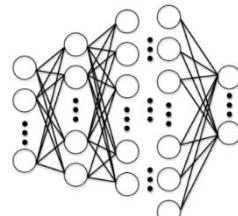
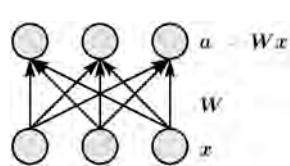
Introduction and Motivation

Why recurrent neural networks?

Why LSTM?

Deep Learning

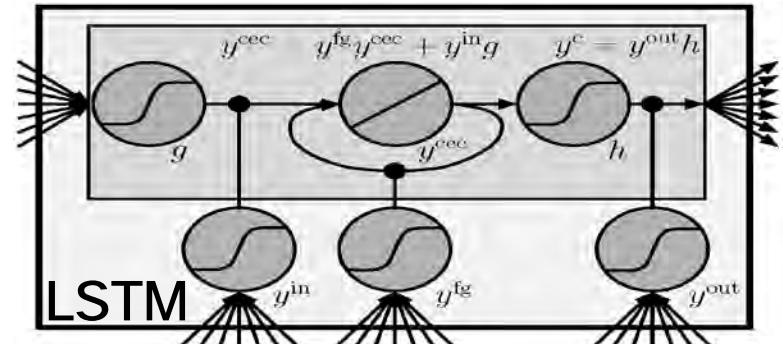
- Neural networks & very fast computers & massive data sets



- Multiple levels or sparse representations → higher levels code abstract concepts



- Started already in 1991 with **LSTM** by Hochreiter



KEY MOMENTS IN DEEP-LEARNING HISTORY 1989-1997

1989

French researcher Yann LeCun, then at Bell Labs, begins foundational work on a type of neural net that becomes crucial for image recognition.

1991

German researchers Sepp Hochreiter and Jürgen Schmidhuber pioneer a neural net with memory features, which eventually proves superior for natural-language processing.

1997

IBM's Deep Blue beats **world champion Garry Kasparov** [right] in chess using traditional AI techniques.



Long Short-Term Memory

- 1991: invented by Hochreiter
- 1997: pub. Hochreiter&Schmidhuber
- 2009: wins French & Arabic handwriting
- 2011: wins offline Chinese handwriting
- 2012: **Google's Android speech recogn.**
- 2015: **Google's Voice transcription**
- 2016: **Apple's iOS 10 → Quicktype**
- 2016: **Google's Translate**
- 2016: **Amazon's Alexa**
- 2017: **Facebook's Translation**

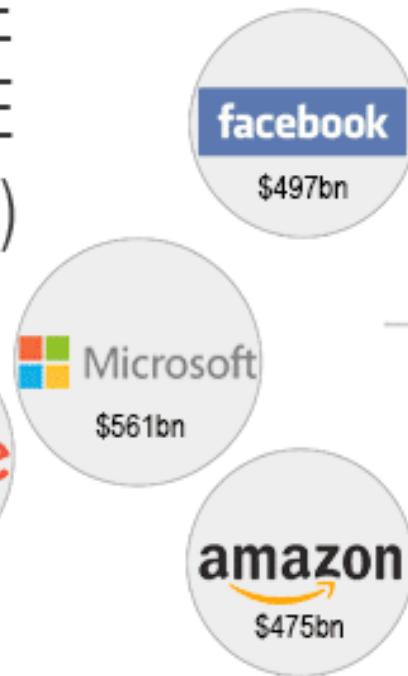


Long Short-Term Memory

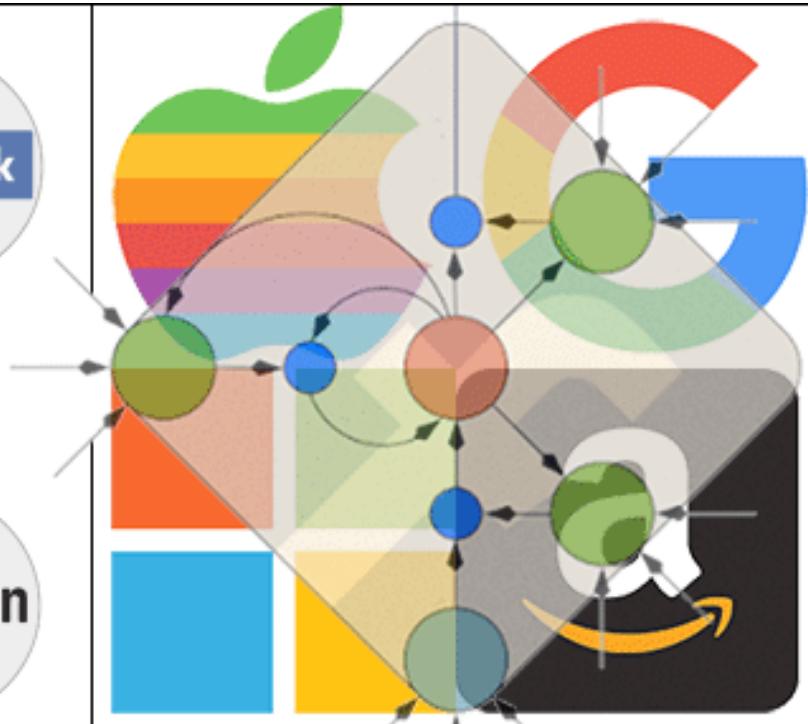
- Google's speech recognition for Android phones etc. based on LSTM: Google Research Blog, Aug/Sep 2015; Dramatic improvement of Google's speech recognition by LSTM: Alphr Technology or 9to5google, Jul 2015
- Apple's iPhone uses LSTM, e.g., TechCrunch, Jul 2016, or noJitter, Jun 2016
- Apple's Siri uses LSTM for various tasks, e.g., BGR.com, Jun 2016
- Microsoft's speech recognition also uses LSTM, e.g., TheRegister, Oct 2016 or Business Insider, Oct 2016
- Microsoft 2017 Conversational Speech Recognition System with LSTM (Xiong et al., Aug 20, 2017)
- Baidu's speech recognition also uses LSTM, e.g., VentureBeat, Jan 2016
- Amazon uses our LSTM for Alexa & Echo, e.g., Vogels' Blog, Nov 2016
- Google's image caption generation with LSTM: arXiv PDF, Nov 2014
- Google's automatic email answering with LSTM: WIRED, Mar 2015
- Google's smart assistant Allo with LSTM: Google Research Blog, May 2016
- Google's dramatically improved Google Translate based on LSTM, e.g., arXiv report, Sep 2016, or HotHardWare, Sep 2016, or WIRED, Sep 2016, or siliconAngle, Sep 2016
- IBM uses LSTM to analyze emotions (2014)
- Microsoft uses LSTM for photo-real talking heads (2014)
- Microsoft uses LSTM for learning to write programs (2017)
- Facebook is using LSTM to handle over 4 billion automatic translations per day (The Verge, August 4, 2017); see also Facebook blog by J.M. Pino, A. Sidorov, N.F. Ayan (August 3, 2017)
- Samsung uses LSTM, too (2016). Samsung is the world's most profitable public company.

Long Short-Term Memory

OUR A.I.'S IMPACT ON THE
WORLD'S 5 MOST VALUABLE
PUBLIC COMPANIES (2017)



JS, August 9, 2017



From Jürgen Schmidhuber: <http://people.idsia.ch/~juergen/impact-on-most-valuable-companies.html>

facebook



Artificial Intelligence



Facebook DeepText



Artificial Intelligence

“Ok Google, Shazam this song.”

Discover music without lifting a finger.

Shazam + Google app. Now you're talking.



Amazon AI

Bringing powerful artificial intelligence to all developers



“Alexa....please tell
Baxter the Robot to
machine this part.”



Siri



JYU



“Hey Siri,
what's the best
sushi place in
town?”

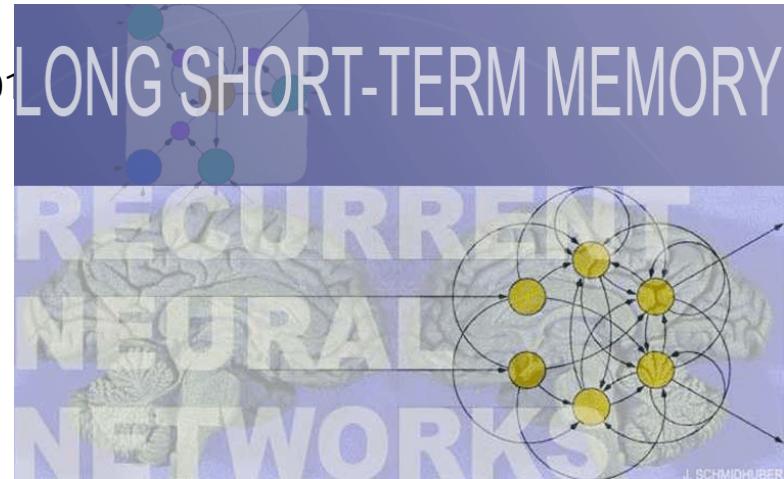
LSTM: Translation, Speech, Lang.

Some benchmark records of 2014 achieved by LSTM:

1. Text-to-speech synthesis (Fan et al., Microsoft, Interspeech 2014)
2. Language identification (Gonzalez-Dominguez et al., Google, Interspeech 2014)
3. Large vocabulary speech recognition (Sak et al., Google, Interspeech 2014)
4. Prosody contour prediction (Fernandez et al., IBM, Interspeech 2014)
5. Medium vocabulary speech recognition (Geiger et al., Interspeech 2014)
6. English to French translation (Sutskever et al., Google, NIPS 2014)
7. Audio onset detection (Marchi et al., ICASSP 2014)
8. Social signal classification (Brueckner & Schulter, ICASSP 2014)
9. Arabic handwriting recognition (Bluche et al., DAS 2014)
10. Image caption generation (Vinyals et al., Google, 2014)
11. Video to textual description (Donahue et al., 2014)

Records of 2013:

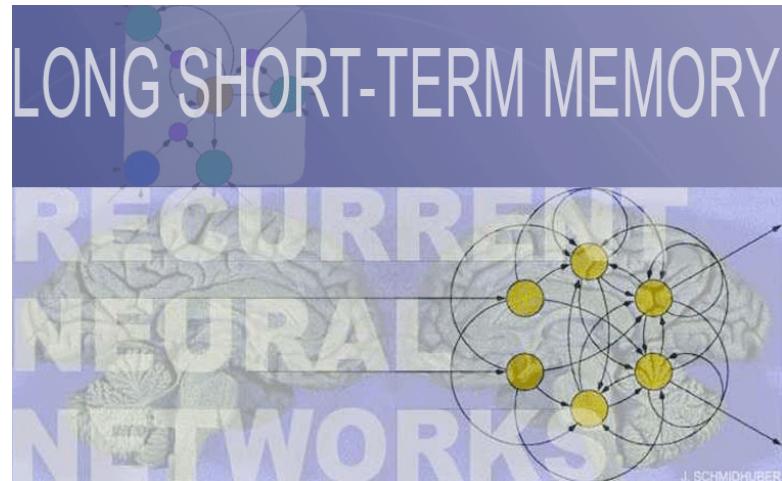
1. TIMIT phoneme recognition (Graves et al., ICASSP 2013)
2. Optical character recognition (Breuel et al., ICDAR 2013)



LSTM: Translation, Speech, Lang.

LSTM at ICASSP 2015 (only papers with LSTM in title):

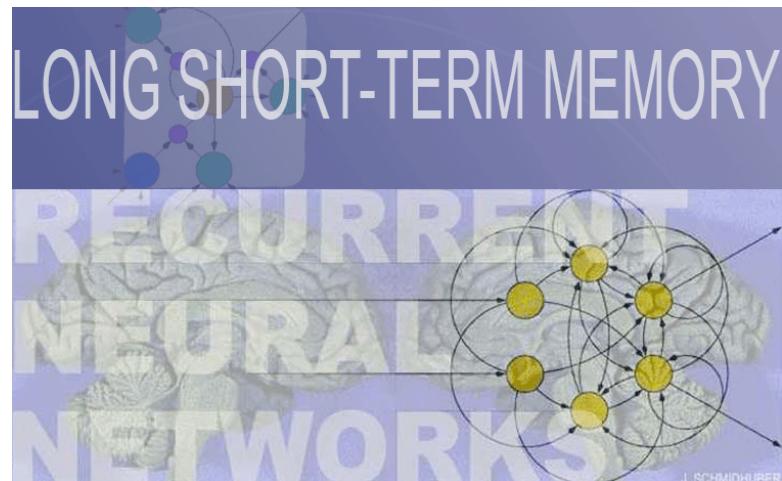
1. Photo-Real Talking Microsoft (Fan et al., [Microsoft](#))
2. Large Vocabulary Speech Recognition (Li and Wu)
3. Convolutional LSTM (Sainath et al., [Google](#))
4. Grapheme-To-Phoneme Conversion (Rao et al., [Google](#))
5. Morph./ Speech Recognition (Renshaw and Hall, [Google](#))
6. Maxout Units (Li and Wu)
7. keyword spotting (Chen et al., [Google](#))
8. Speech Synthesis (Zen and Sak, [Google](#))
9. Voice Conversion (Sun et al.)
10. Acoustic Novelty Detection (Marchi et al.)
11. Phone Models (Senior et al., [Google](#))
12. Parallelization / GPU (Hwang and Sung)
13. Keyboard Gesture Decoding (Alsharif et al., [Google](#))



LSTM: Translation, Speech, Lang.

1. Speaker Adaptation in Speech Recognition
2. Speaker-Aware Training ([Microsoft](#), [MIT](#))
3. Large Vocabulary Speech Recognition ([Microsoft](#))
4. Speech with Deep Networks ([Microsoft](#))
5. CD-CTC-SMBR LSTM Models ([Google](#))
6. Speech Synthesis Cross-Lingual Information
7. Talking Avatar
8. Speech Recognition and Adaptation ([Microsoft](#))
9. Emotion Recognition in Video
10. Speech Recognition and Word Error ([Mitsubishi](#))
11. Simplifying LSTMs for Fast Training ([Microsoft](#))
12. Lipreading with LSTMs
13. Distant Speech Recognition ([Microsoft](#), [MIT](#))
14. Music Dynamic Motion Recognition

LSTM at ICASSP 2016 (only papers with LSTM in title)



LSTM @ facebook



<https://www.facebook.com/FBAIResearch/posts/362517620591864>

“Basically, the neural net acts as a ‘reasoning’ engine that stores and retrieves data to be operated on from a separate memory. There have been ideas along those lines going back to the early 1990's with Hochreiter and Schmidhuber's ‘Long Short-Term Memory’ (LSTM). But new ideas are being explored by teams at Facebook, Google and elsewhere.”

LSTM @ Google

- Frame-based Automatic Language Identification in Short Utterances using **LSTM**
- Large Vocabulary Speech Recognition by **LSTM**
- Automatic Language Identification using **LSTM**
- Sequence Discriminative Dist. Training of **LSTM**
- Translation English to French: Sequence to Sequence Learning with multilayered **LSTM**
- Biologically Plausible Speech Recognition with **LSTM** Neural Nets
- Large Scale Deep Learning with Sequential **LSTM**
- Statistical Parametric Speech Synthesis using **LSTM**
- A Probabilistic Model for Melodies with **LSTM** recurrent networks
- One Billion Word Benchmark using **LSTM** Neural Networks for Language Modeling

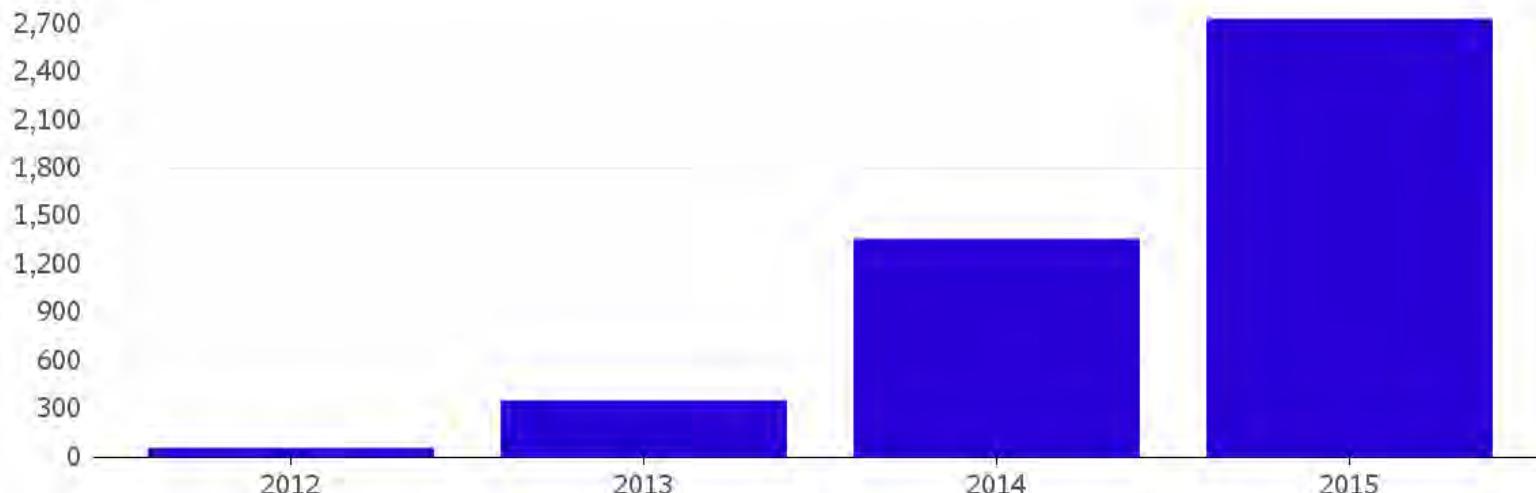


Research
at Google

LSTM @ Google

Artificial Intelligence Takes Off at Google

Number of software projects within Google that uses a key AI technology, called Deep Learning.



Source: Google

Note: 2015 data does not incorporate data from Q4

LSTM @ Microsoft



Microsoft®
Research

- Photo-Real Talking Head with Deep Bidirectional LSTM
- LSTM for audio/visual modeling in photo-real talking head system
- Spoken Language Understanding using LSTM: Record-setting performances in natural language
- TTS Synthesis with Bidirectional LSTM based Recurrent Neural Networks

LSTM Applications

Activity Recognition

Input:
Sequence
of Frames

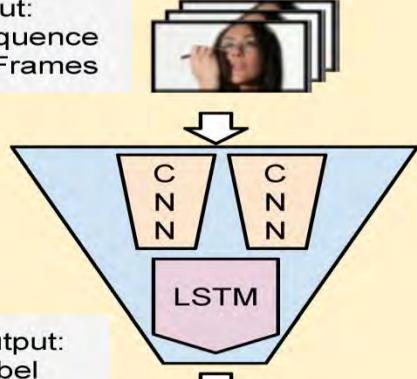
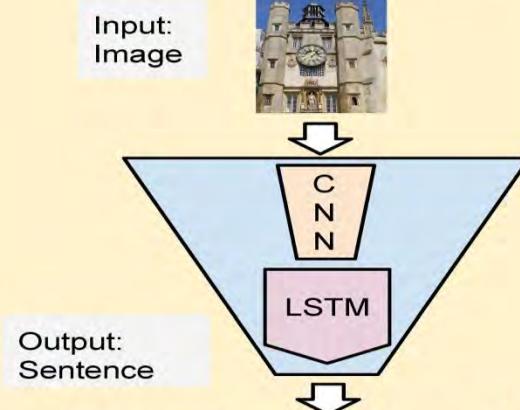


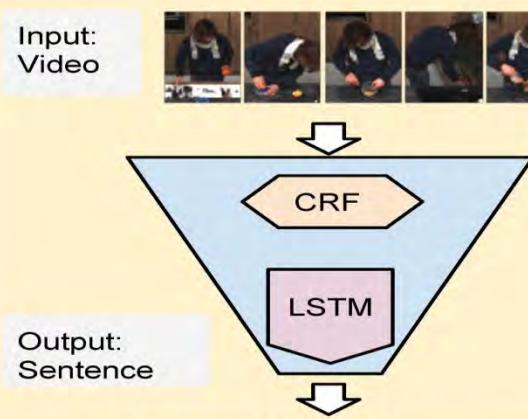
Image Description

Input:
Image

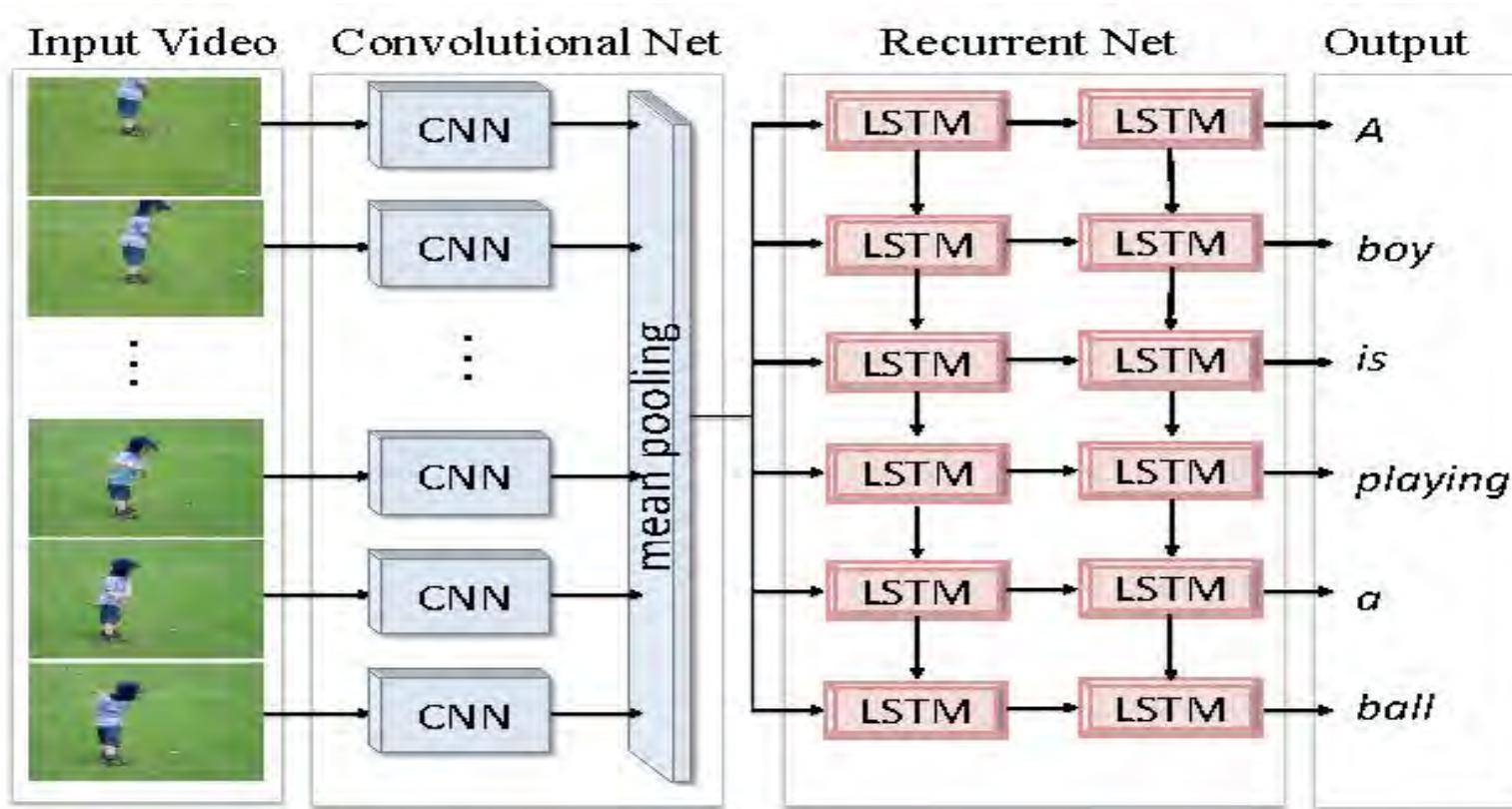


Video Description

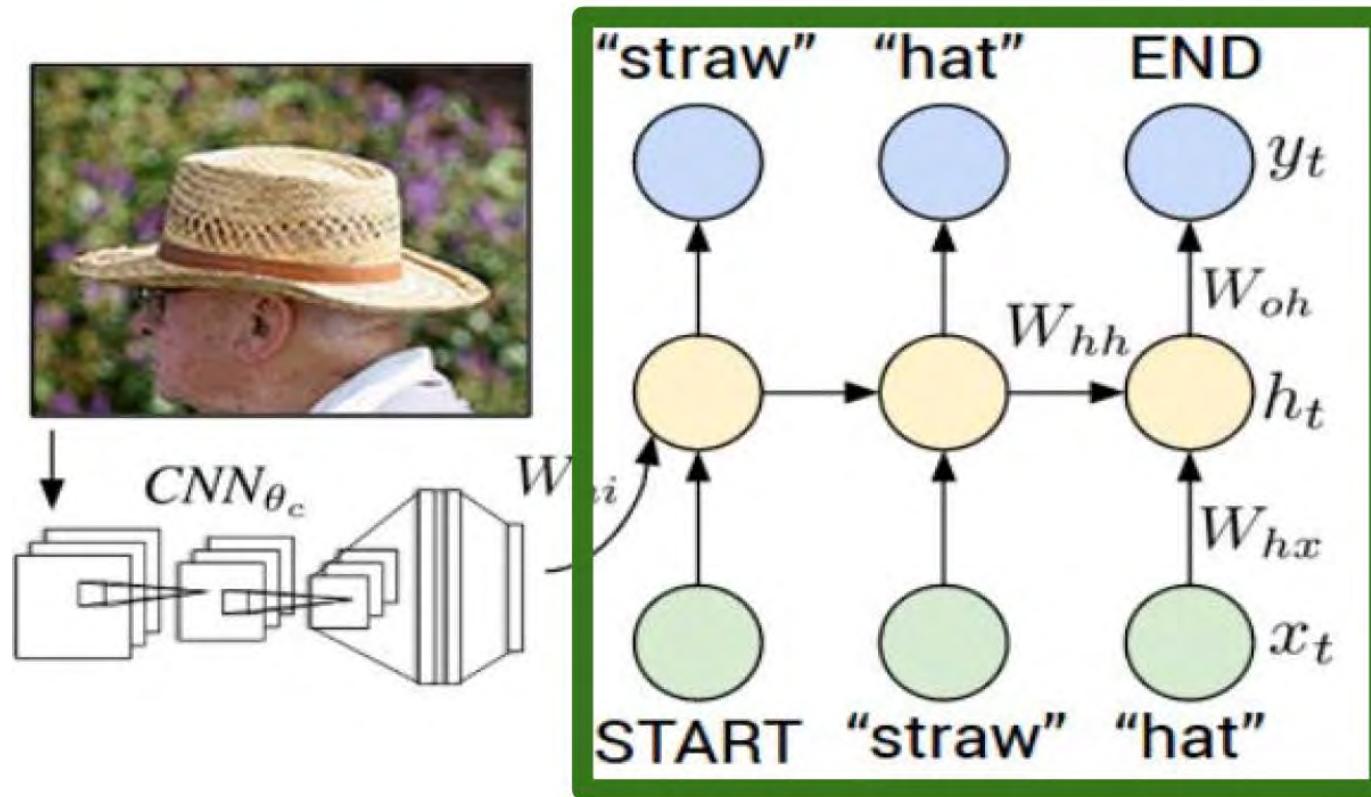
Input:
Video



LSTM Image Captions



LSTM Image Captions



LSTM Image Captions



"man in black shirt is playing guitar."



"construction worker in orange safety vest is working on road."



"two young girls are playing with lego toy."



"boy is doing backflip on wakeboard."



"girl in pink dress is jumping in air."



"black and white dog jumps over bar."



"young girl in pink shirt is swinging on swing."



"man in blue wetsuit is surfing on wave."

LSTM Image Captions



a man standing next to a clock on a wall
logprob: -10.08



a young boy is holding a
baseball bat
logprob: -7.65



a cat is sitting on a couch with a remote control
logprob: -12.45

LSTM Text Generation

Good afternoon. God bless you.

The United States will step up to the cost of a new challenges of the American people that will share the fact that we created the problem. They were attacked and so that they have to say that all the task of the final days of war that I will not be able to get this done. The promise of the men and women who were still going to take out the fact that the American people have fought to make sure that they have to be able to protect our part. It was a chance to stand together to completely look for the commitment to borrow from the American people. And the fact is the men and women in uniform and the millions of our country with the law system that we should be a strong stretches of the forces that we can afford to increase our spirit of the American people and the leadership of our country who are on the Internet of American lives.

Thank you very much. God bless you, and God bless the United States of America.

LSTM Question Answering



Q: What is there in yellow?

A: Bananas



Q: What is there on the grass, except the person?

A: Sheep



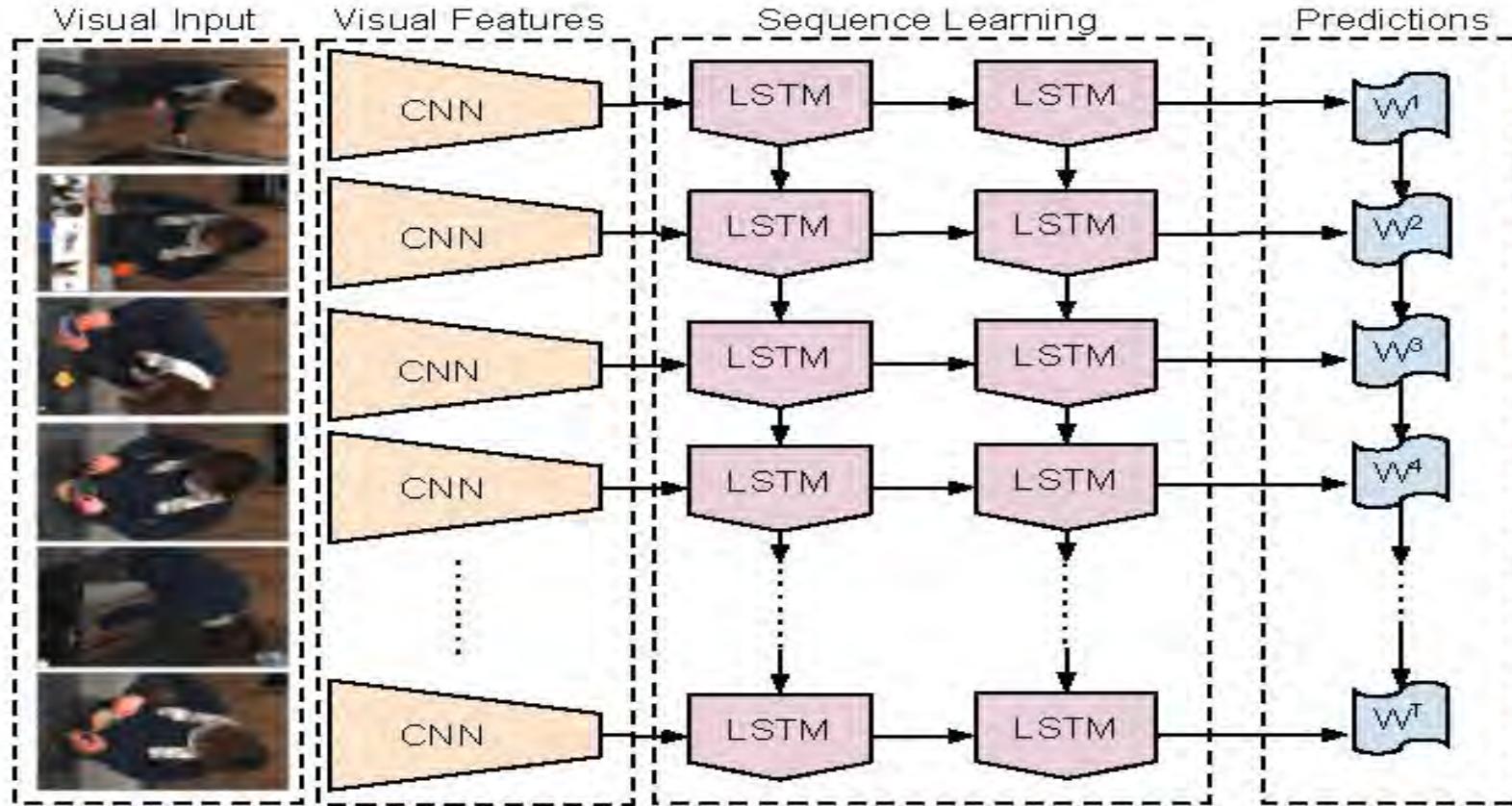
Q: Where is the kitty?

A: On the chair

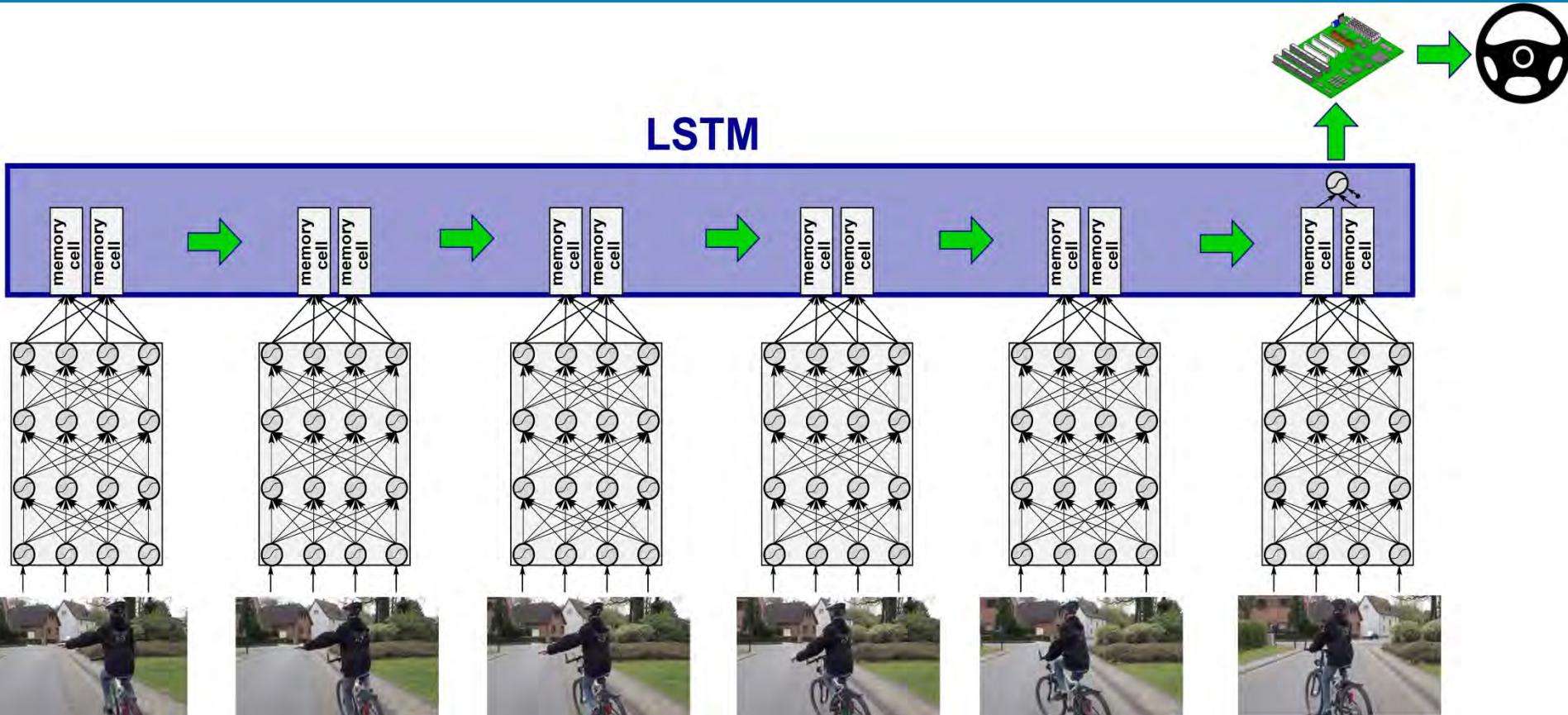
LSTM Projects

- Zalando: Fashion Blogs including images
- Pharma: Patent analysis
- Stock market: analyzing news articles
- Automotive: Self-driving cars
- LSTM attention: drug cocktails, environment samples, self-driving cars, high-content imaging

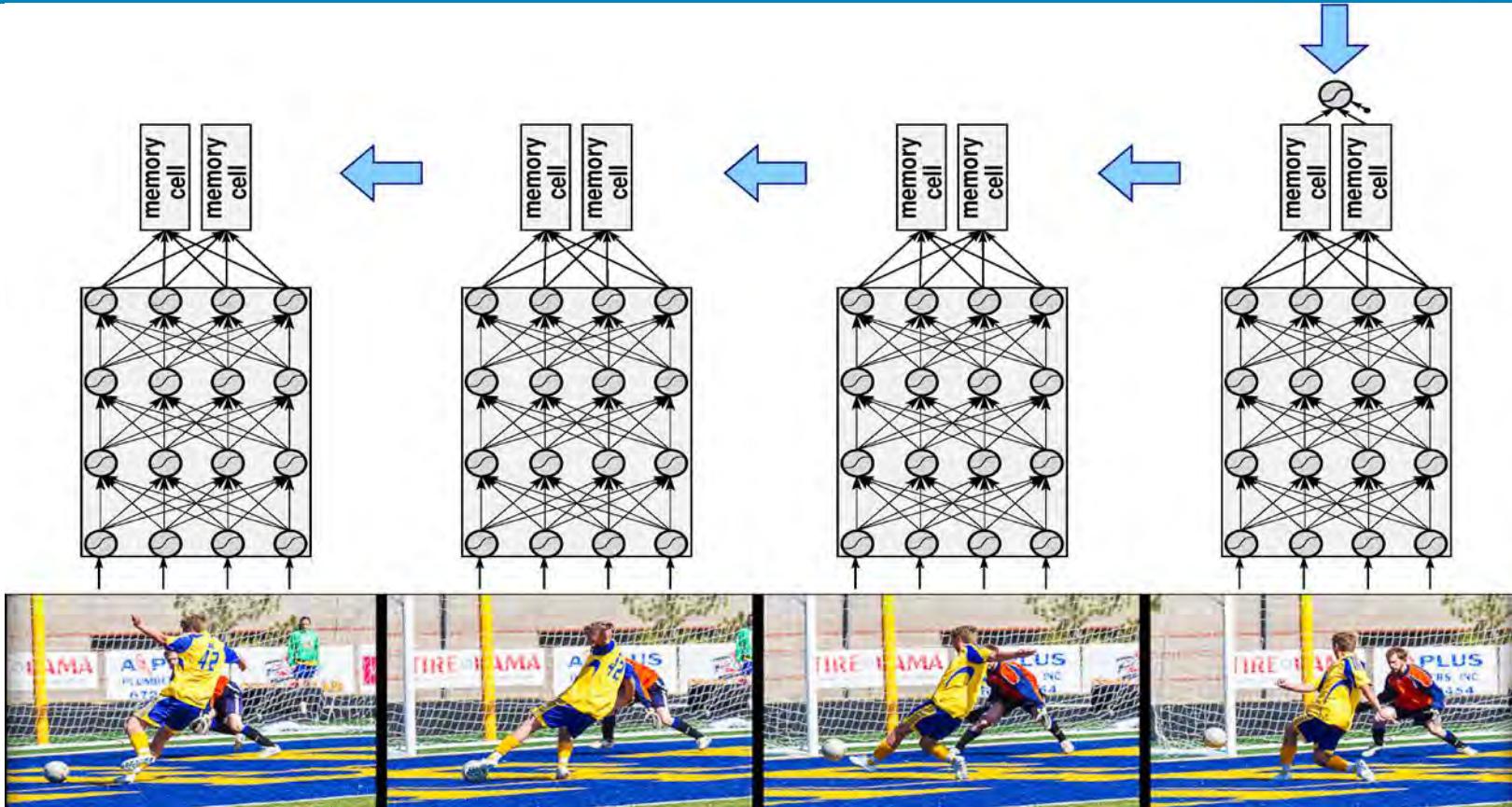
LSTM Action Prediction



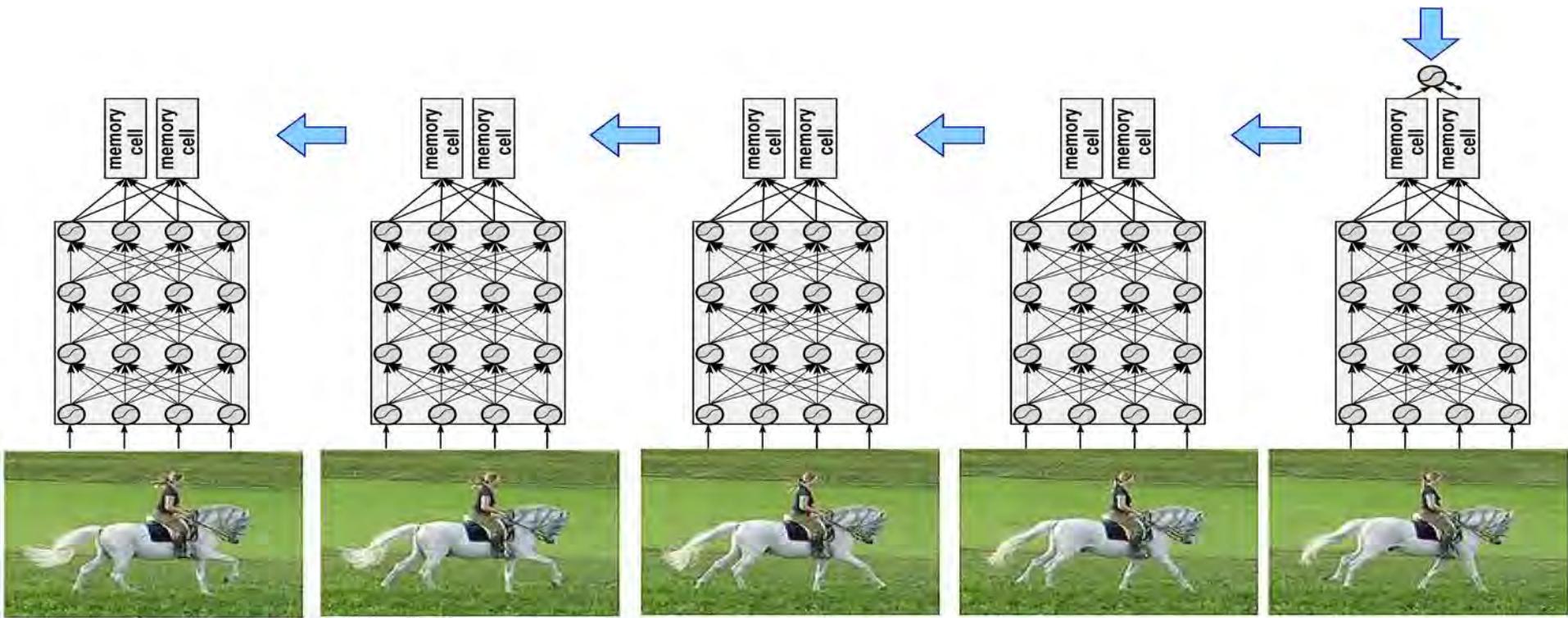
Long Short-Term Memory



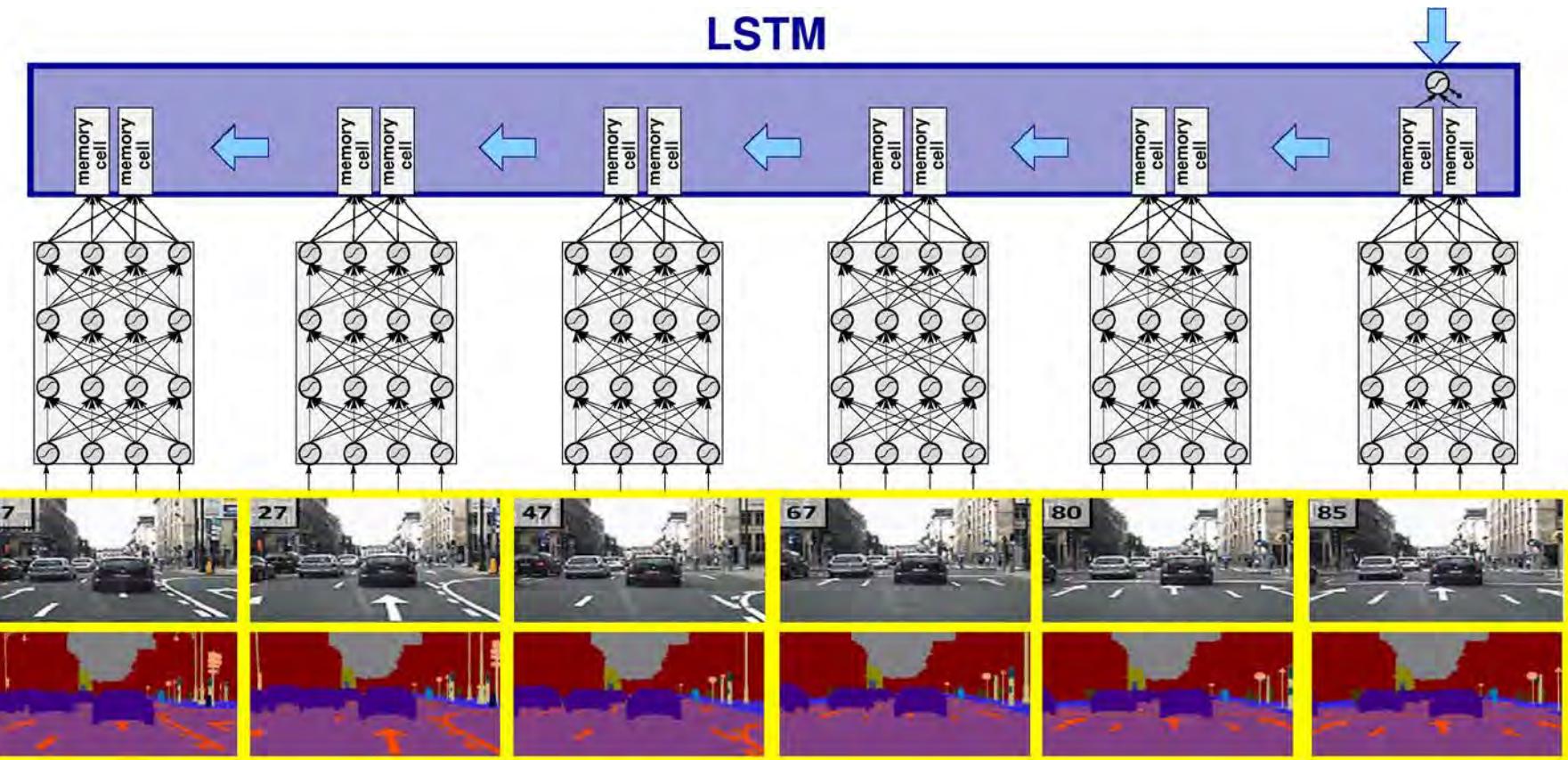
LSTM Action Prediction



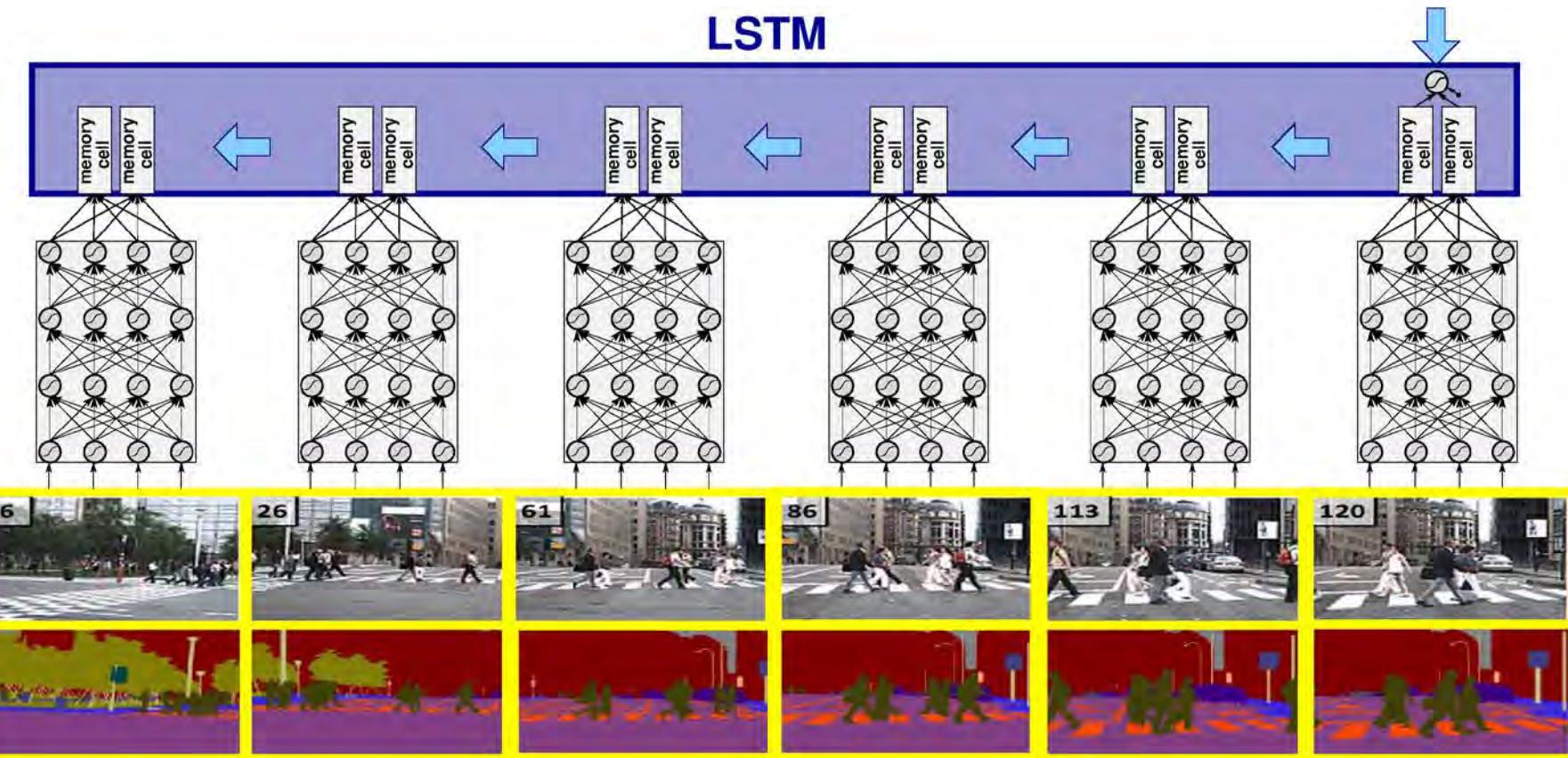
LSTM Action Prediction



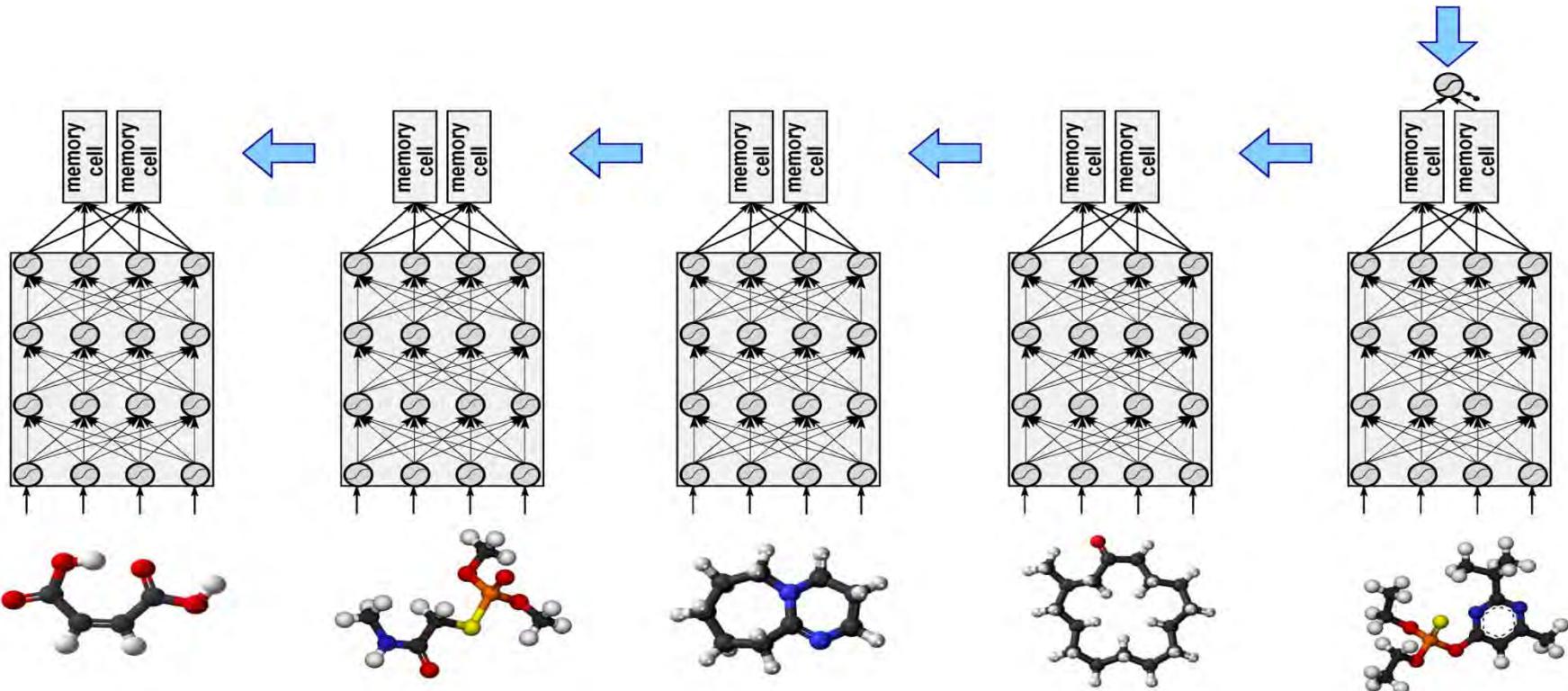
LSTM Action Prediction



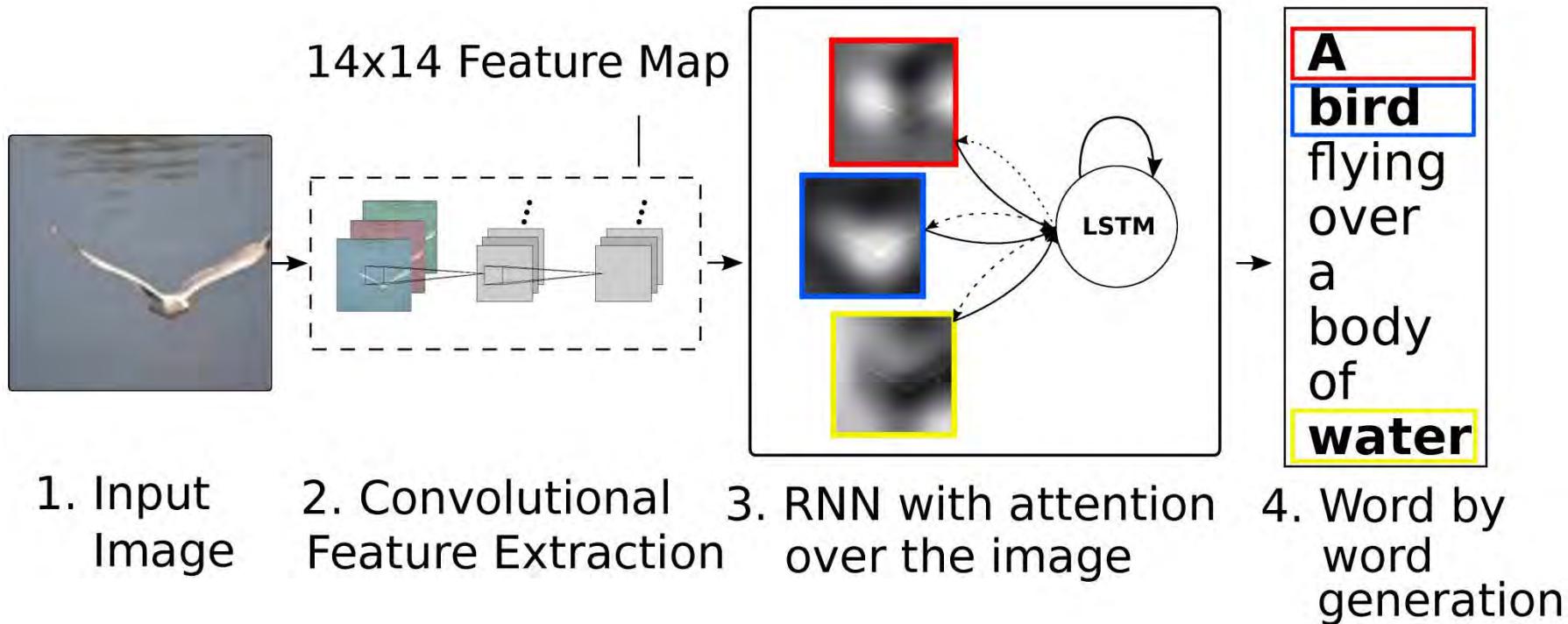
LSTM Action Prediction



LSTM Drug Mixture Effects



LSTM for Attention



LSTM for Attention



A woman is throwing a frisbee in a park.

A dog is standing on a hardwood floor.

A stop sign is on a road with a mountain in the background.



A little girl sitting on a bed with a teddy bear.



A group of people sitting on a boat in the water.



A giraffe standing in a forest with trees in the background.

LSTM for Attention

Detected green traffic light!

Received reward:



LSTM for Attention



Long Short-Term Memory

DOTA 2



OpenAI

Long Short-Term Memory StarCraft II



DeepMind

LSTM in Projects

JKU
Partners



LSTM in Projects

JKU



JKU
Partners

LSTM in Projects



LSTM in Projects

Deep learning for fashion
images and fashion blogs

- text and language analysis
- image analysis

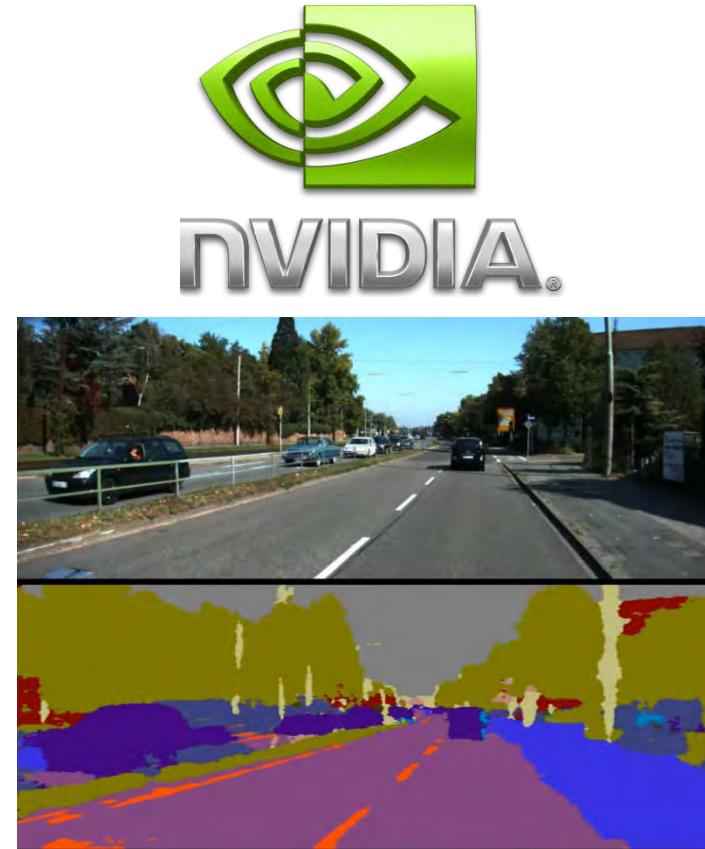


LSTM in Projects



Deep learning for autonomous driving

- Prediction of situations (LSTM)
- Attention (LSTM)
- Sensor (GPS, Radar, Lidar) fusion
- Rare events (dangerous situations)



LSTM in Projects

Johnson & Johnson



janssen



LSTM in Projects



LSTM in Projects



LSTM in Projects



Deep Learning for Target Prediction

- Target prediction
- Toxicity prediction
- Image analysis



LSTM in Projects

ExCape: ExaScale machine learning and HPC in drug design:

- Toxicity prediction
- Drug target prediction
- Focus on deep learning

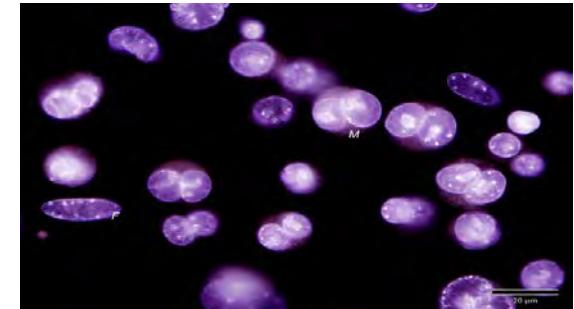
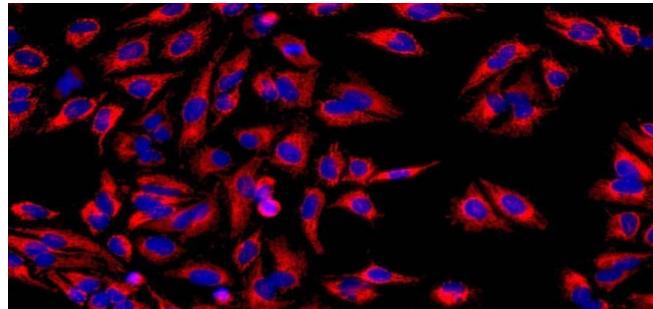
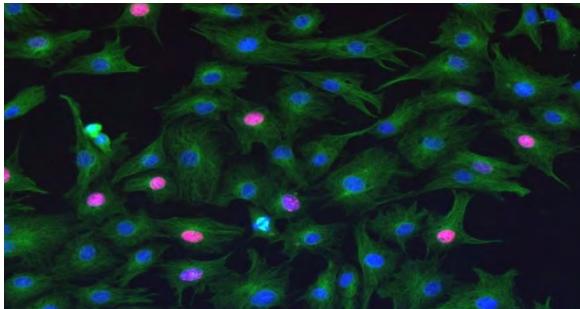


LSTM in Projects



Exaptation: Imaging and deep learning

- International consortium
- 650,000 high content imaging data



Chapter 1

Recurrent Neural Networks