JOHANNES KEPLER UNIVERSITY LINZ

LSTM and Recurrent Neural Nets

Part 1

Sepp Hochreiter

LIT AI Lab, Institute for Machine Learning
Johannes Kepler University

Course Overview

JOHANNES KEPLER UNIVERSITY LINZ

Basic Course of Master Artificial Intelligence

Class:	3 ECTS	2 SWS VO	Mo 15:30-17:00 (HS 7)
Exercise:	1.5 ECTS	1 SWS UE	Mo 13:45-15:15 (HS 7)

Exams

Class: **2** written intermediate **exams**: at 25.11.2019 and at 27.01.2020
(at 17.02.2020 re-examination)

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

Master AI 1. Semester

JOHANNES KEPLER UNIVERSITY LINZ

Monday	Tuesday	Wednesday	Thursday	Friday
8:30 *alternatingly blocked:* KV Artificial Intelligence in Society VL Control Systems (Track 1) UE Control Systems (Track 1) 11:45		8:30 VL Pervasive Computing: Systems & Environments (Track 2) 10:00	8:30 *simultaneously:* VL Optimum and Adaptive Signal Processing Sys. (Track 2) KV Model Checking (Track 3) 10:00	
		8:30 VL Pervasive Computing: Design and Developm. (Track 2, 3rd semester) 10:00	10:15 VL Production Automation Systems (Track 1) 11:45	10:15 VL Deep Learning and Neural Nets I 11:45
	12:00 KV Computer Vision [Bimber] 13:30	12:00 UE Pervasive Computing: Systems & Environments (Track 2) 12:45	12:00 *simultaneously:* UE Optimum and Adaptive Signal Processing Sys. (blocked, Track 2, Gr. A) 13:30	12:00 VL AI and Law I 13:30
		12:45 UE Pervasive Computing: Design and Developm. (Track 2, 3rd semester) 13:30	UE Machine Learning: Supervised Techniques 13:45 UE Optimum and Adaptive Signal Processing Sys. (blocked, Track 2, Gr. B) 15:15	
13:45 *alternatingly blocked:* UE LSTM and Recurrent Neural Nets UE Deep Learning and Neural Nets I 15:15	14:00 VL Lecture Series Artificial Intelligence (recommended) 15:00			
15:30 VL LSTM and Recurrent Neural Nets 17:00	15:30 UE Computer Algebra (Track 3, Groups A, B) 16:15 16:15 VL Computer Algebra (Track 3) 18:00	15:30 VL Machine Learning: Supervised Techniques (Admission course) 17:00	15:30 KV Basic Methods of Data Analysis (Admission course) 17:00	
17:15 KV Computer Vision (blocked, Group A) 18:45		17:15 KV Computer Vision (blocked, Group B) 18:45		

Bachelor AI 1. Semester

JOHANNES KEPLER UNIVERSITY LINZ

Monday	Tuesday	Wednesday	Thursday	Friday
8:30 KV Hands-on AI I (blocked) 11:45	8:30 VL Logic 10:00			
	10:15 UE Logic 11:45	10:15 VL Introduction to AI 11:45	10:15 VL Mathematics for AI I 11:45	
12:00 KV Responsible AI 13:30		12:00 KV Programming in Python I 15:15	12:00 UE Mathematics for AI I (Groups A1, A2, ...) 13:30	
	14:00 VL Lecture Series Artificial Intelligence (blocked) 15:00			13:45 UE Mathematics for AI I (Group C) 15:15
	15:30 VL Mathematics for AI I 17:00			
	18:00 Tutorium: Mathematics for AI I (Group A) 19:45			
		19:00 Tutorium: Mathematics for AI I (Group B) 20:30	19:00 UE Mathematics for AI I (Group B) 20:30	

Content

JOHANNES KEPLER UNIVERSITY LINZ

Recurrent Neural Networks

Content

Topics Covered

JOHANNES KEPLER UNIVERSITY LINZ

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

JOHANNES KEPLER UNIVERSITY LINZ

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

JOHANNES KEPLER UNIVERSITY LINZ

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

JOHANNES KEPLER UNIVERSITY LINZ

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

JOHANNES KEPLER UNIVERSITY LINZ

Applications

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

RNNs win challenges

- speech
- language
- text
- translation

Highlights

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

Introduction

JOHANNES KEPLER UNIVERSITY LINZ

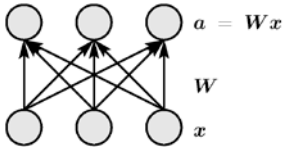
Introduction and Motivation

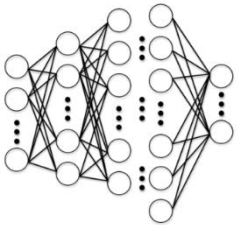
Why recurrent neural networks?

Why LSTM?

Deep Learning

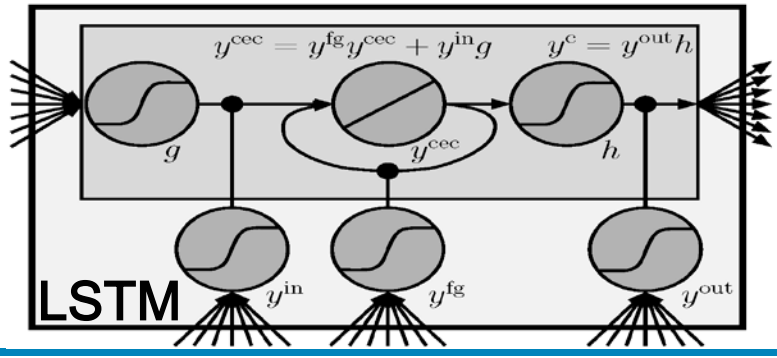
- Neural networks & very fast computers & massive data sets





- Multiple levels of 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.

Fortune 28.09.2016, "The AI Revolution: Why Deep Learning is suddenly changing your life"

STAN HONDA—AFP/GETTY IMAGES

Long Short-Term Memory

JOHANNES KEPLER UNIVERSITY LINZ

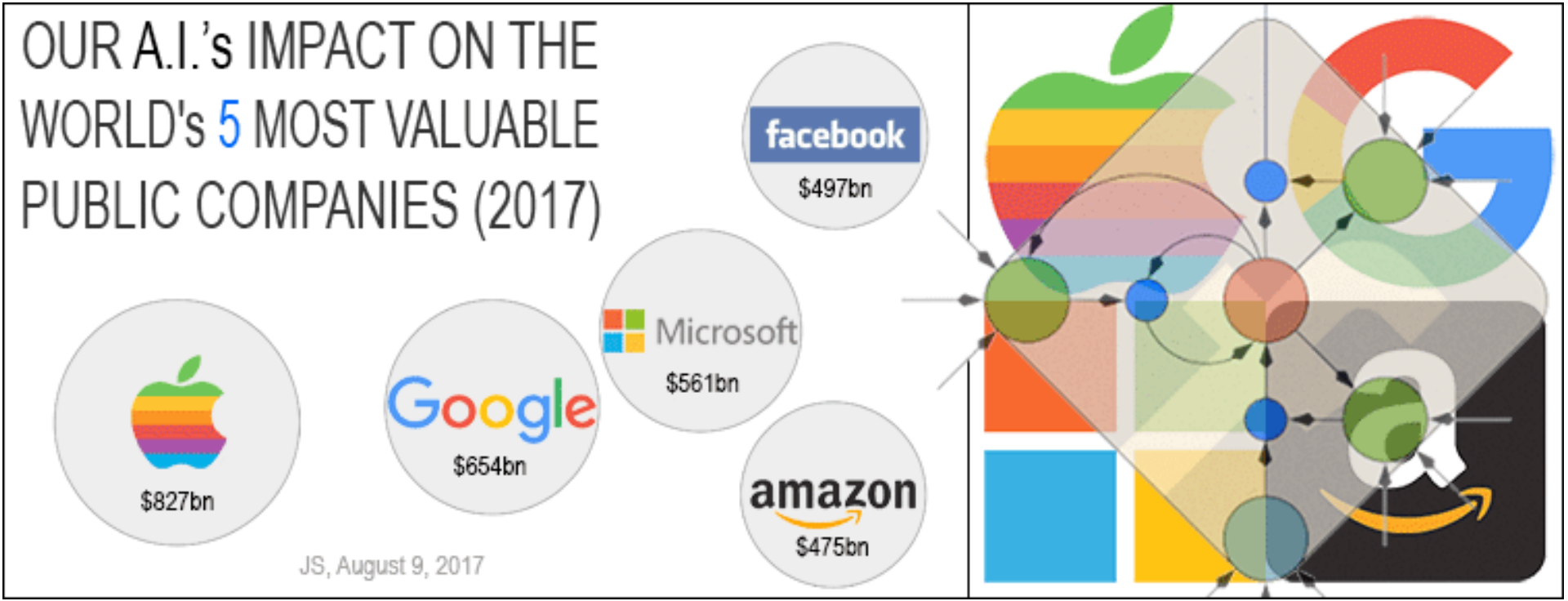
- 1991: invented by Hochreiter
- 1997: pub. Hochreiter&Schmidhuber
- 2009: wins French & Arabic handwriting
- 2011: wins offline Chinese handwriting
- 2012: Google's Android speech recogn.
- 2015: Google's Voice transcription
- 2016: Apple's iOS 10 → Quicktype
- 2016: Google's Translate
- 2016: Amazon's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

Long Short-Term Memory

JOHANNES KEPLER UNIVERSITY LINZ

- 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

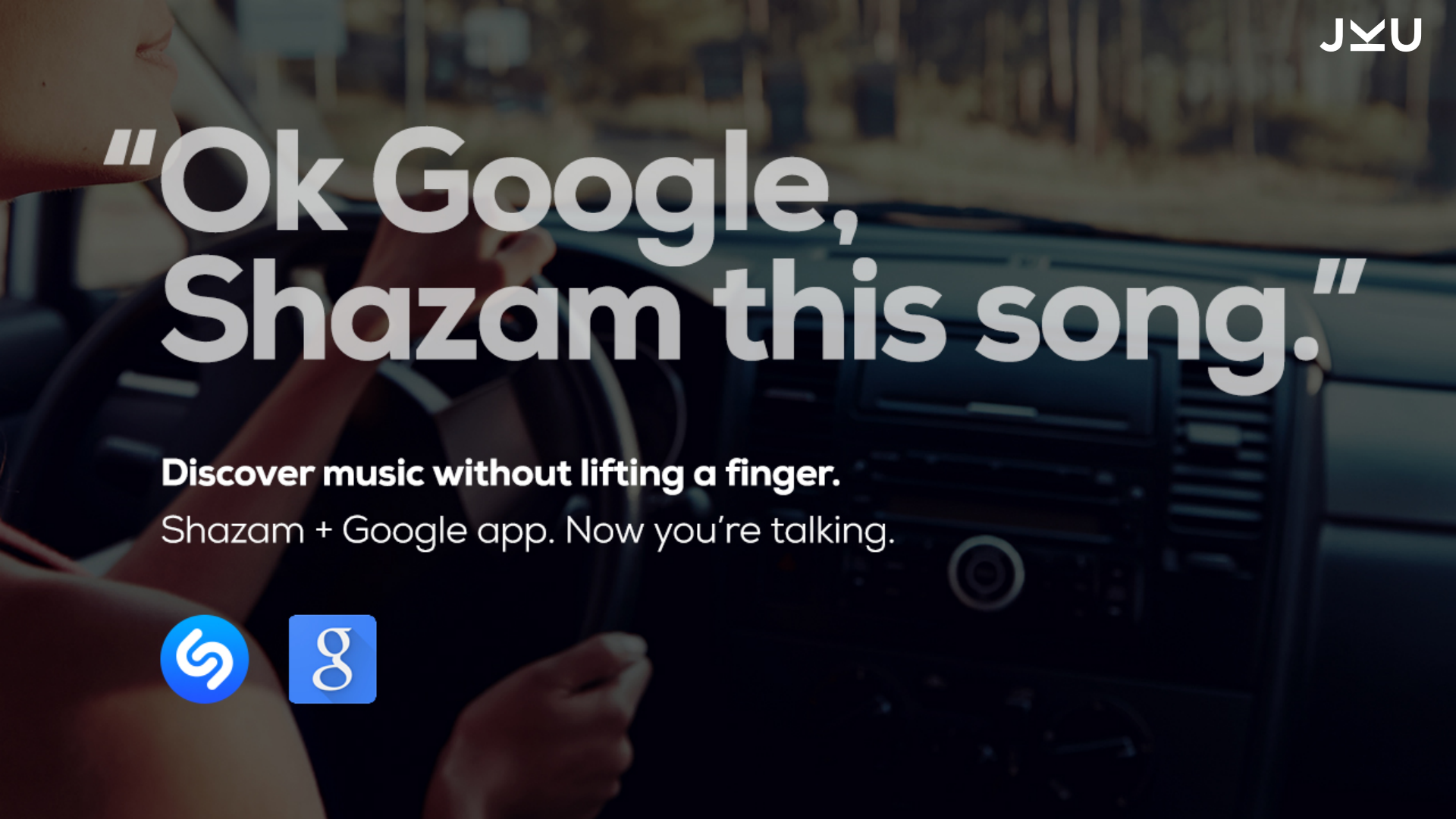


From Jürgen Schmidhuber: http://people.idsia.ch/~juergen/impact-on-most-valuable-companies.html


facebook
Artificial Intelligence

Facebook Deep Text


JKU
Google
Artificial Intelligence


"Ok Google,
Shazam this song."
Discover music without lifting a finger.
Shazam + Google app. Now you're talking.


JKU
Amazon AI
Bringing powerful artificial intelligence to all developers


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


Siri

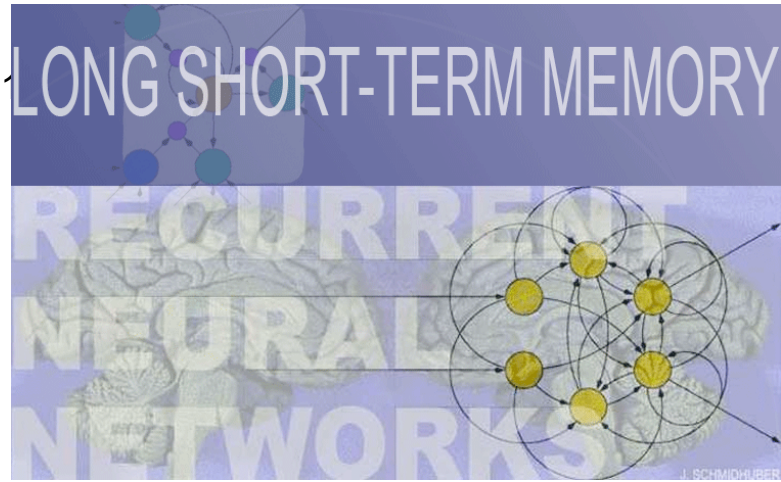

“Hey Siri, what’s the best sushi place in town?”

LSTM: Translation, Speech, Lang.

JOHANNES KEPLER UNIVERSITY LINZ

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 201
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:

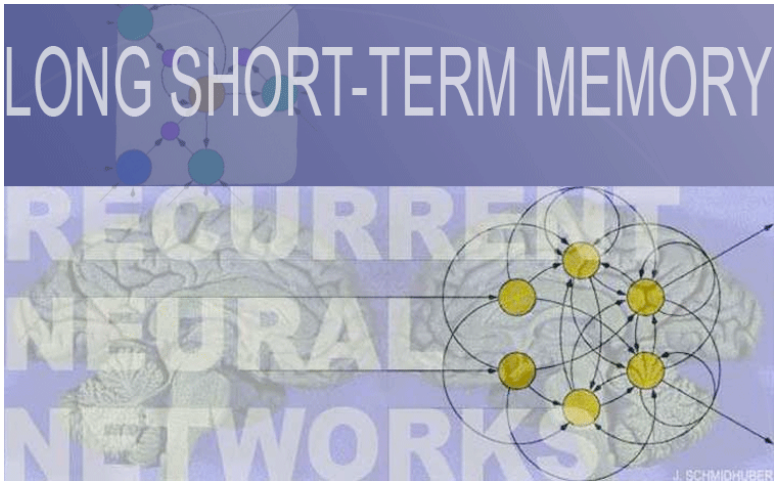
1. TIMIT phoneme recognition (Graves et al., ICASSP 2013)
2. Optical character recognition (Breuel et al., ICDAR 2013)

LSTM: Translation, Speech, Lang.

JOHANNES KEPLER UNIVERSITY LINZ

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)

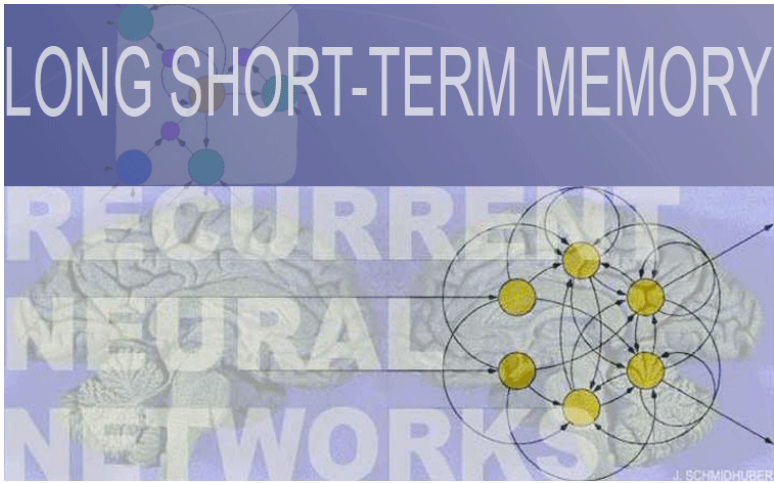


LSTM: Translation, Speech, Lang.

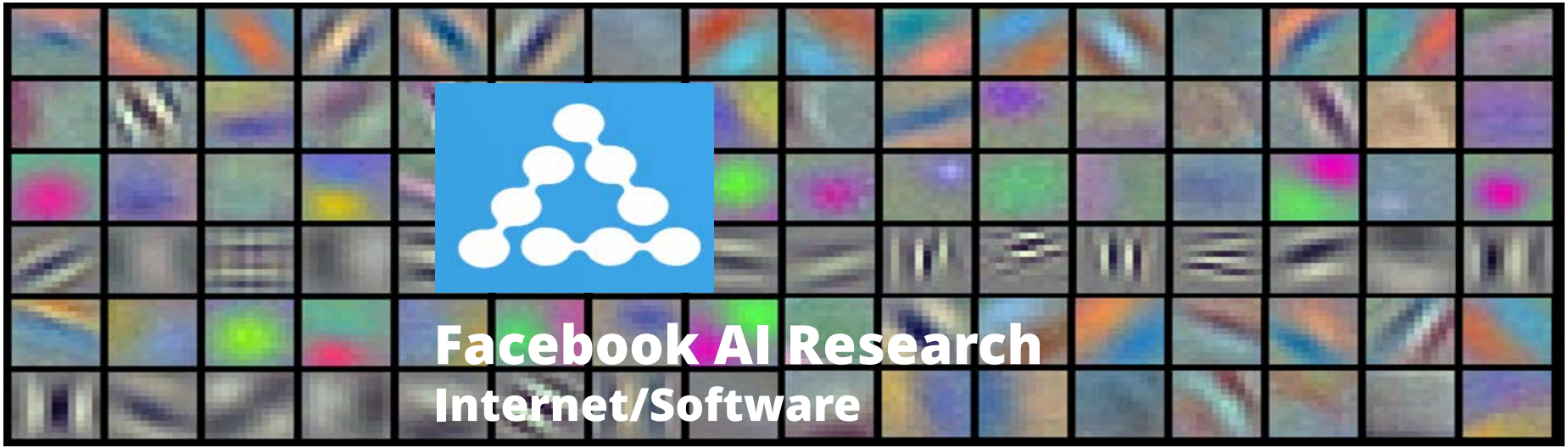
JOHANNES KEPLER UNIVERSITY LINZ

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

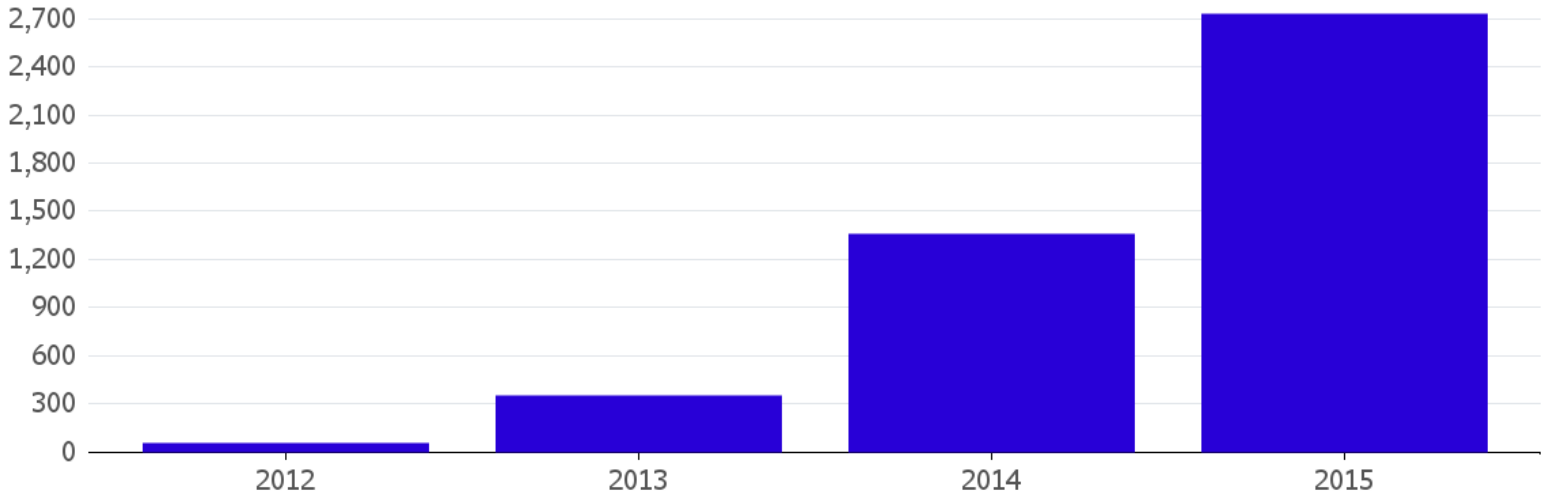
- Frame-based Automatic Language Identification in Short Utterances using LSTM
- Large Vocabulary Speech Recognition by LSTM
- Automatic Language Identification using LSTM
- Sequence Discriminative Dist. Training of LSTM
- Translation English to French: Sequence to Sequence Learning with multilayered 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

Bloomberg

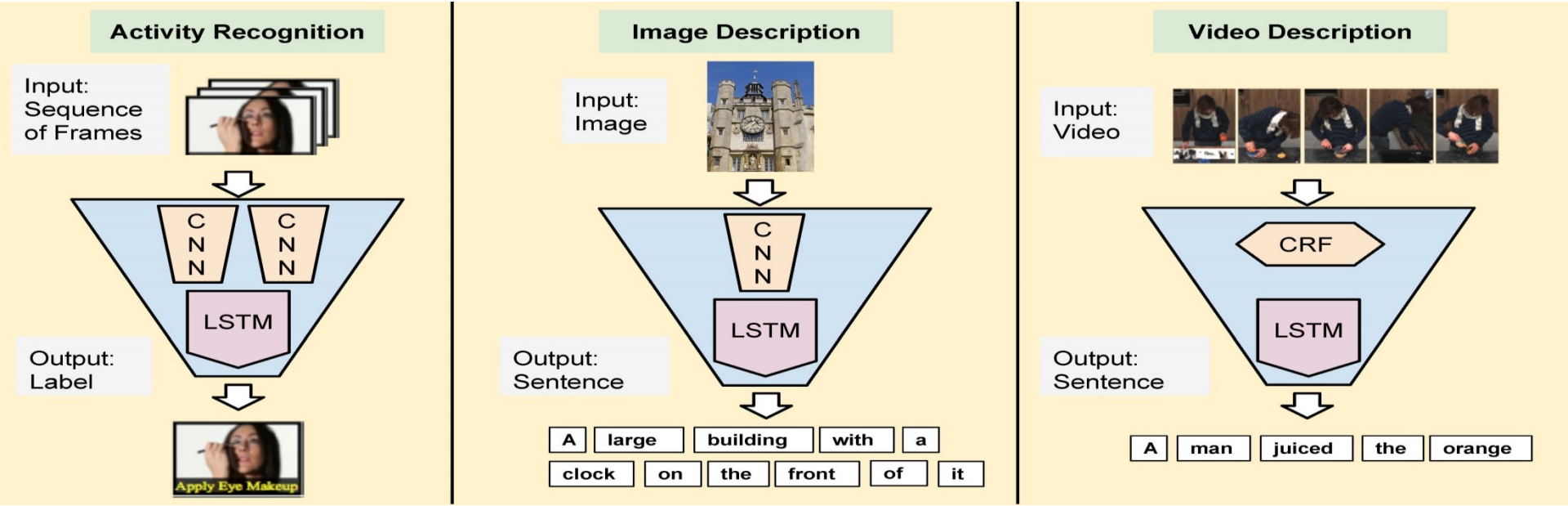
LSTM @ Microsoft

Microsoft Research

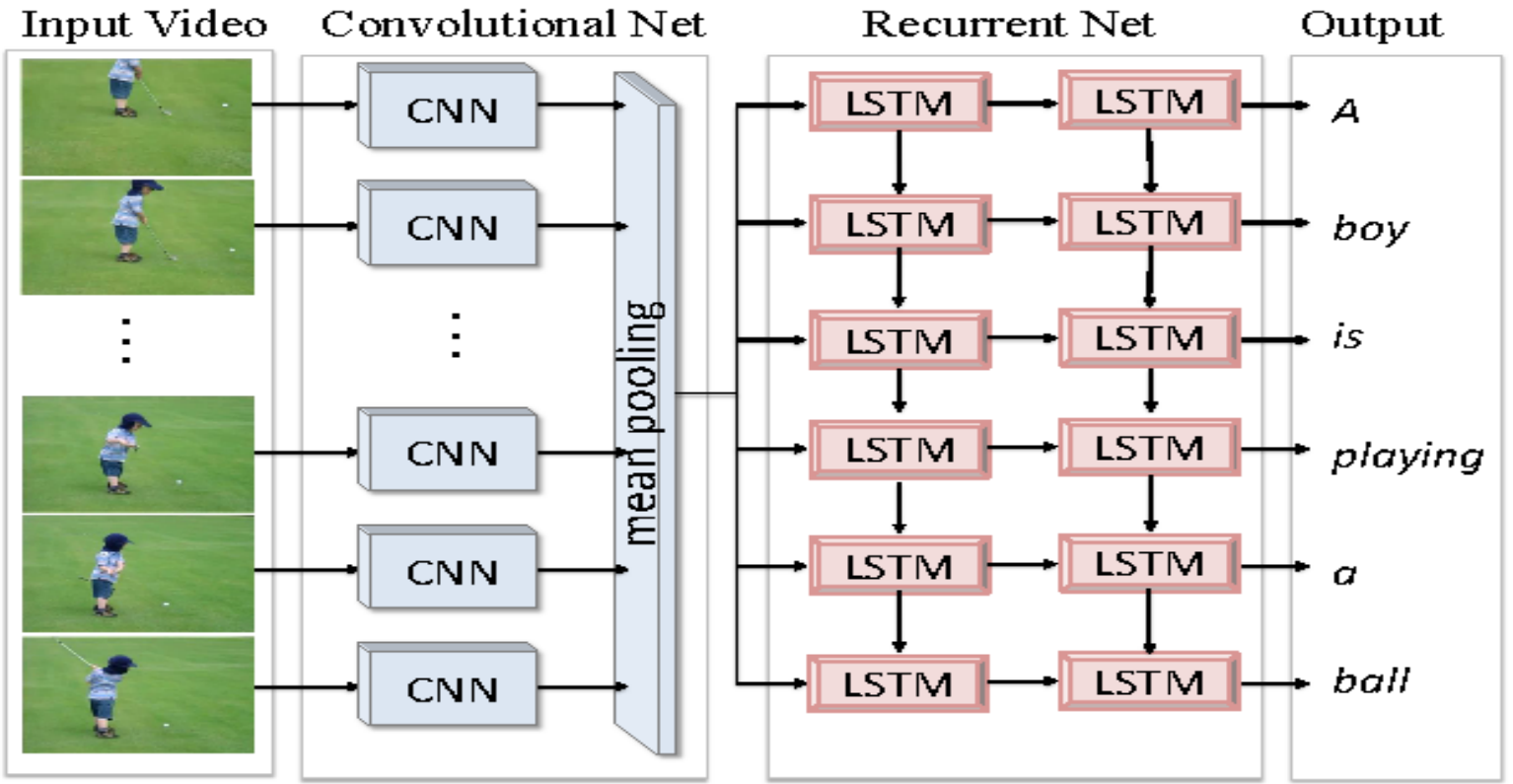
- Photo-Real Talking Head with Deep Bidirectional LSTM
- LSTM for audio/visual modeling in photo-real talking head system
- Spoken Language Understanding using LSTM: Record-setting performances in natural language
- TTS Synthesis with Bidirectional LSTM based Recurrent Neural Networks

LSTM Applications

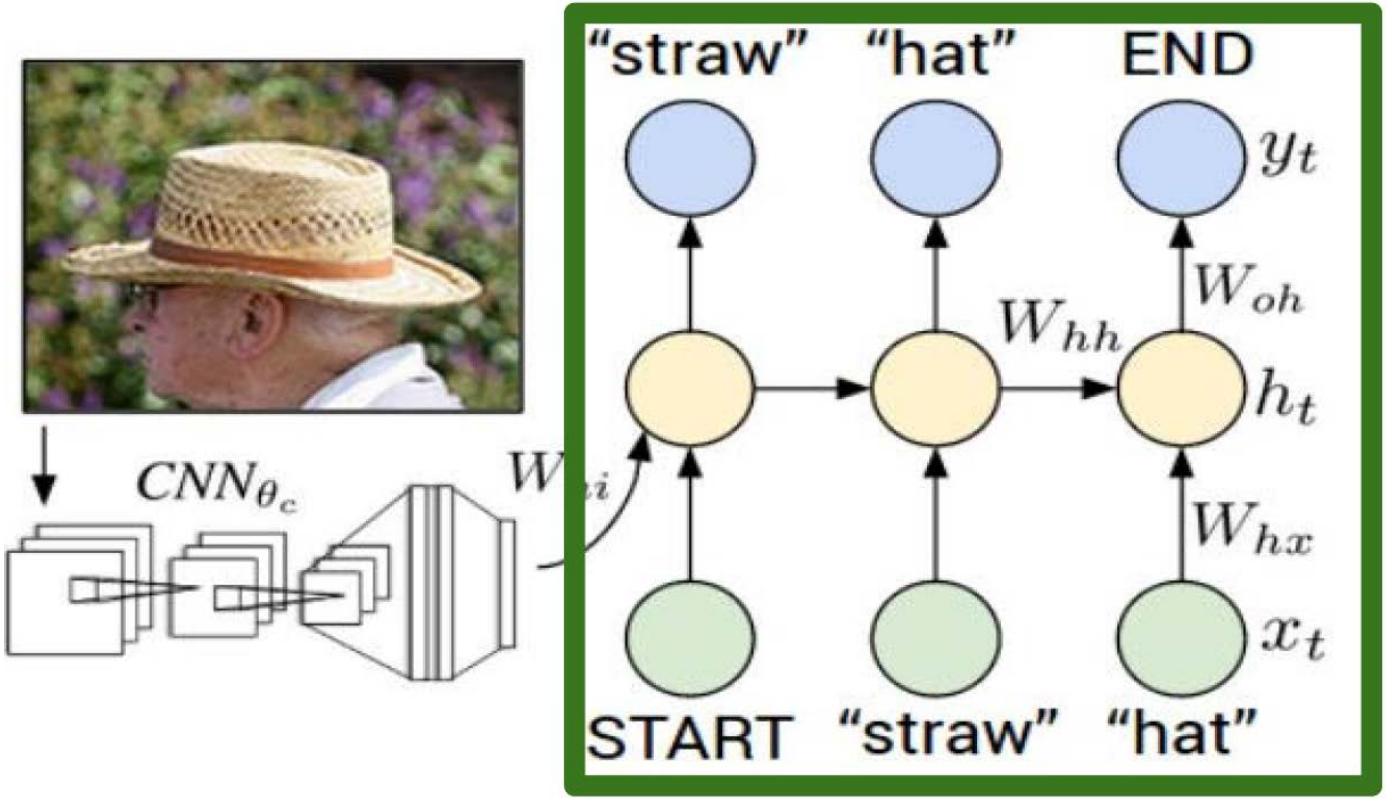
JOHANNES KEPLER UNIVERSITY LINZ



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

JOHANNES KEPLER UNIVERSITY LINZ

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 stretcks 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?

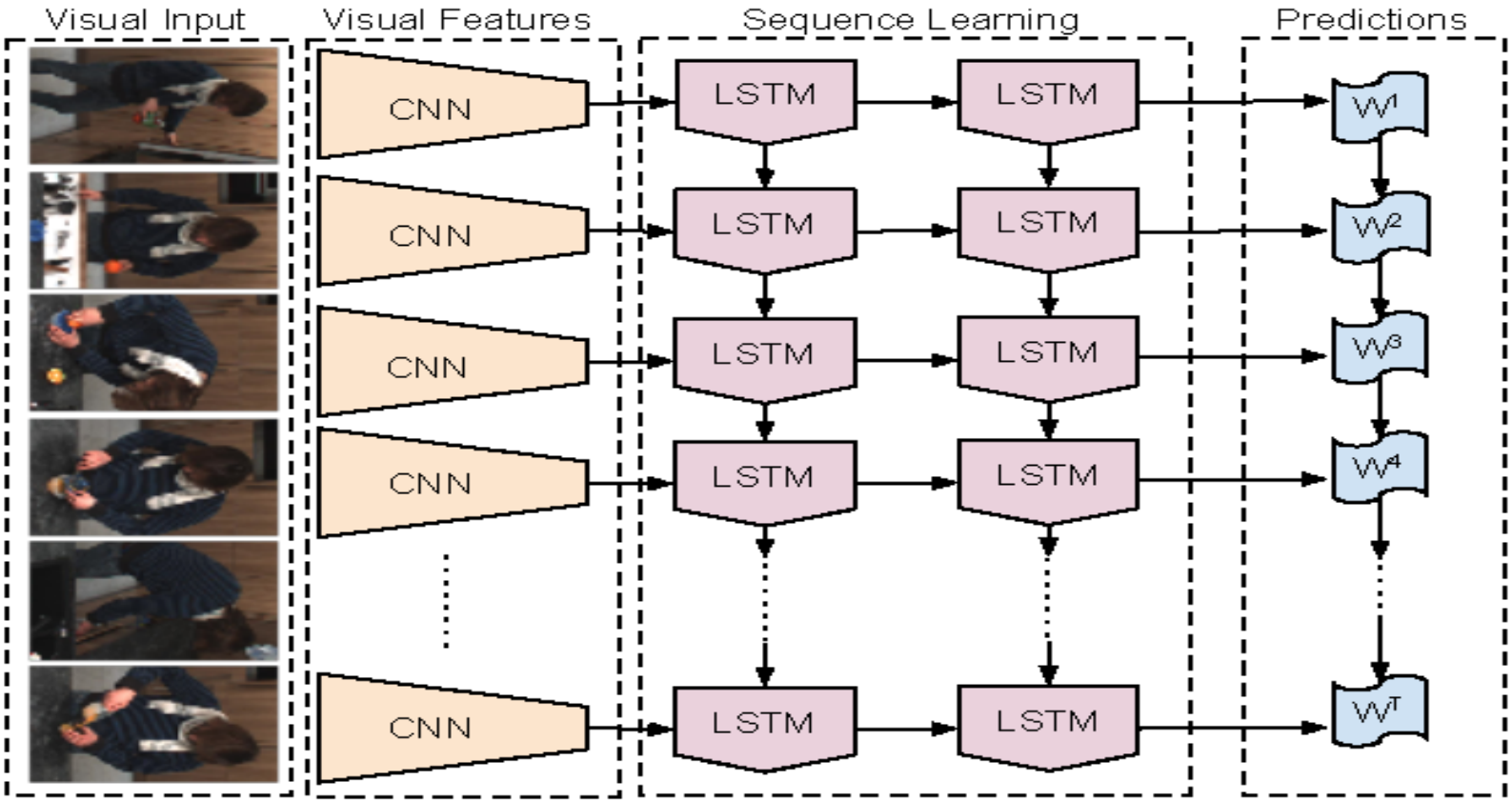
A: On the chair

LSTM Projects

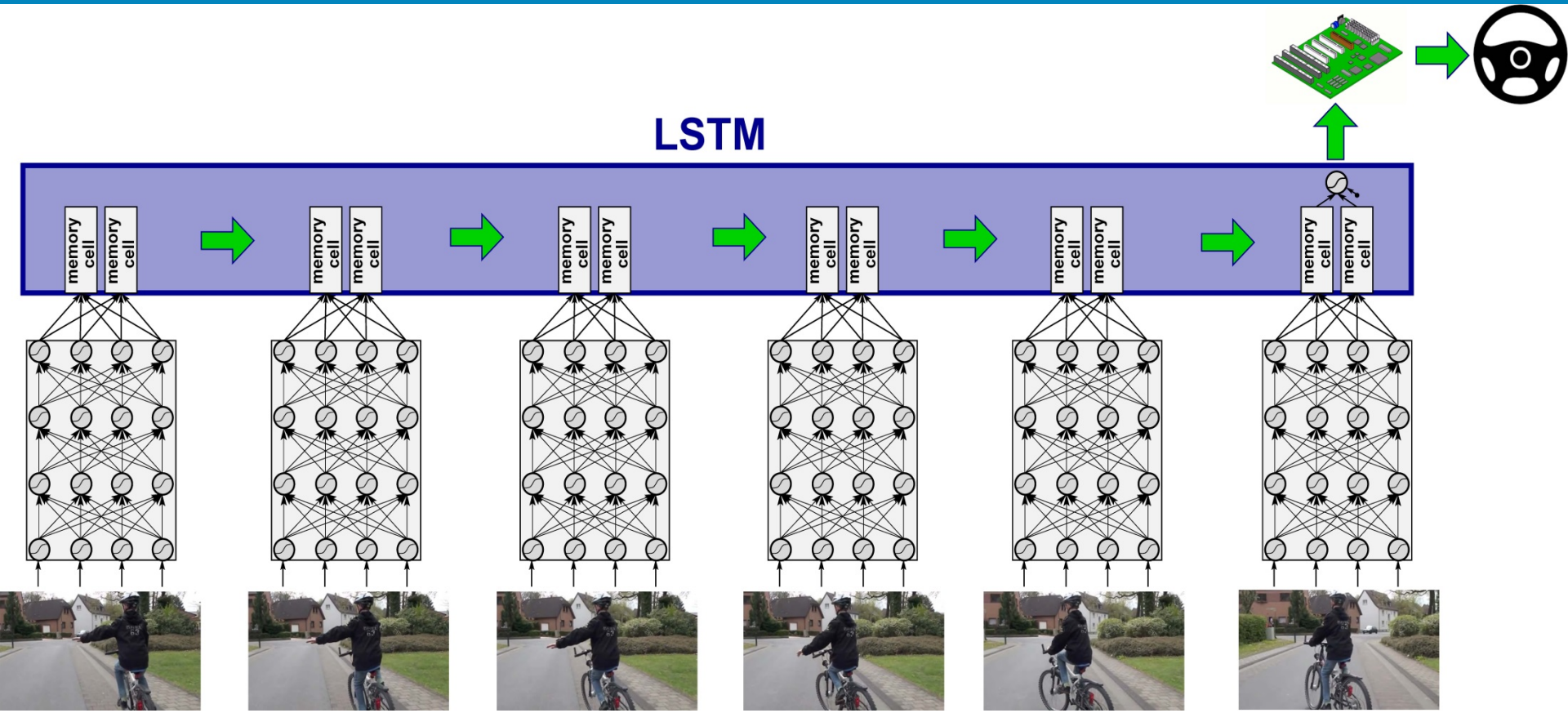
JOHANNES KEPLER UNIVERSITY LINZ

- 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

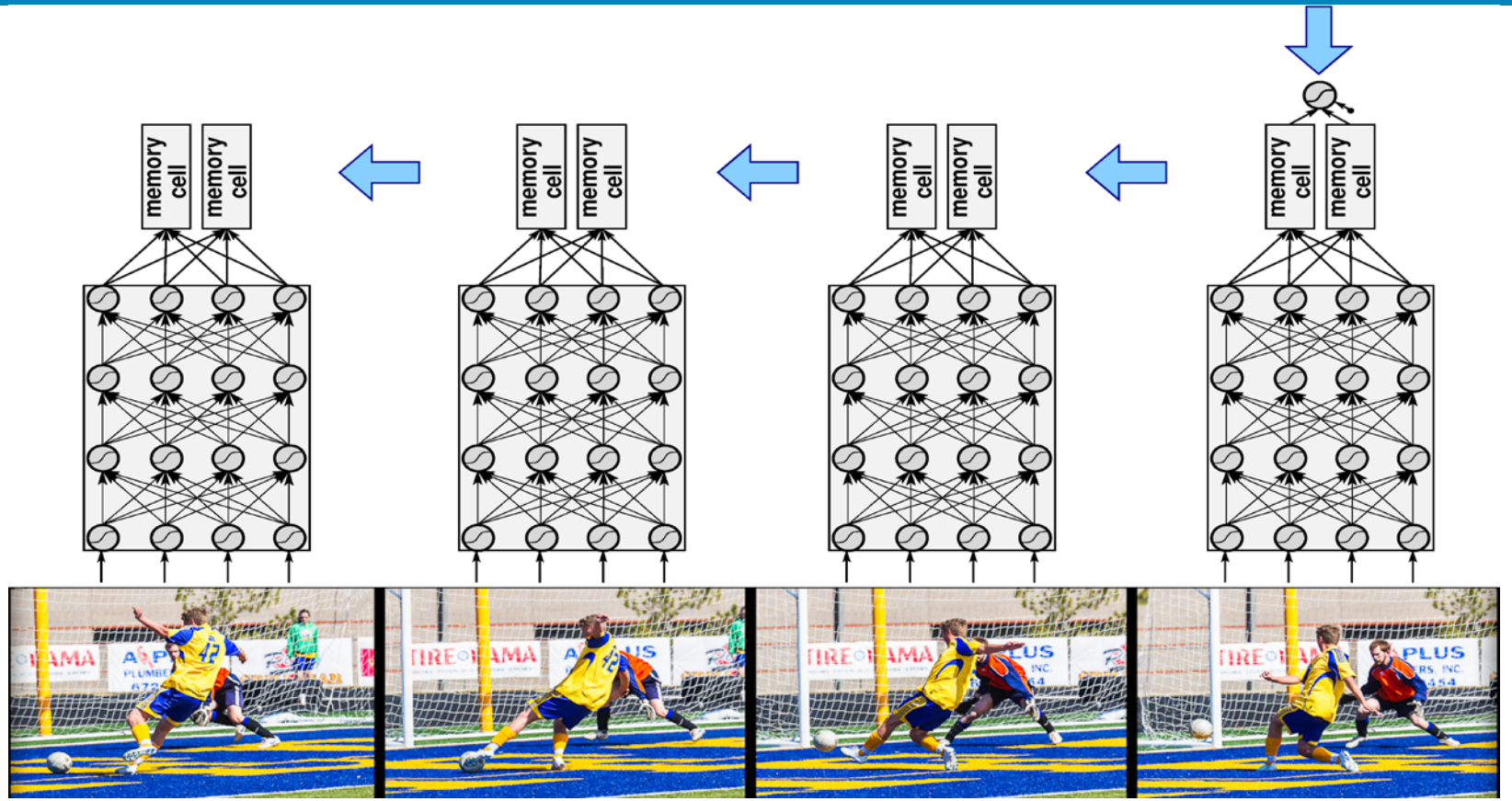
LSTM Action Prediction



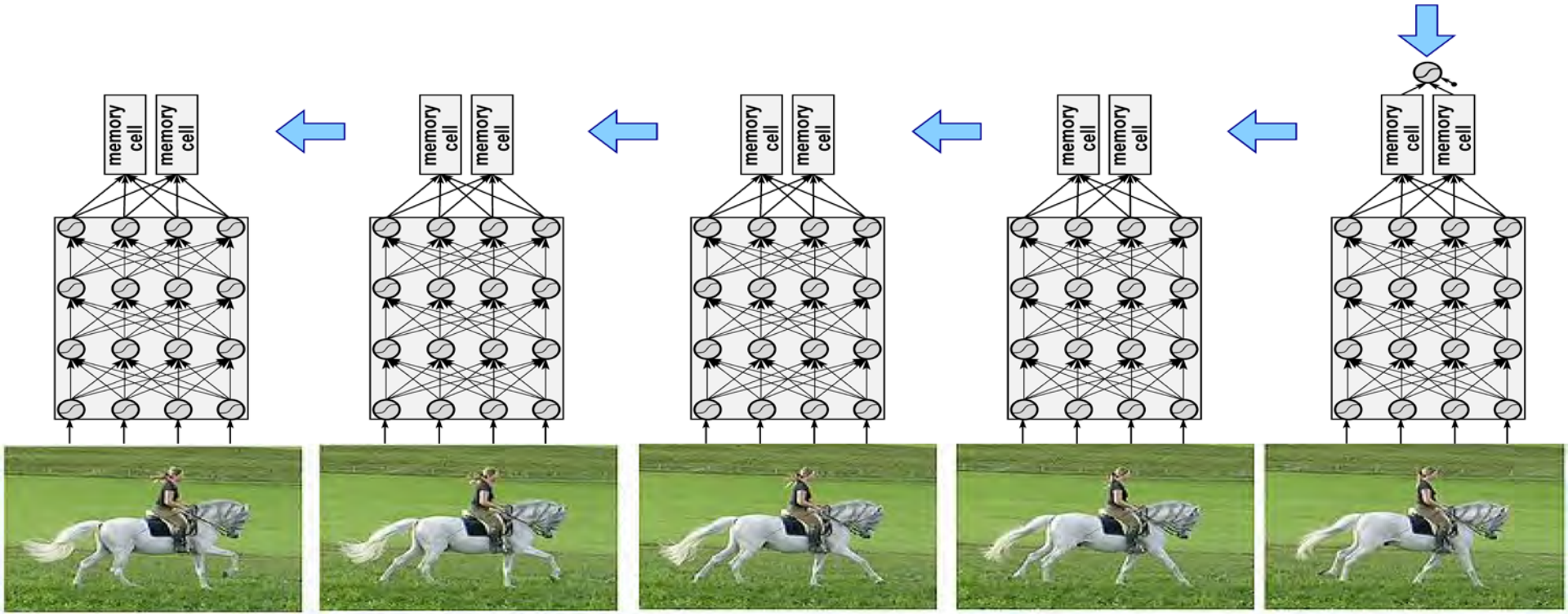
Long Short-Term Memory



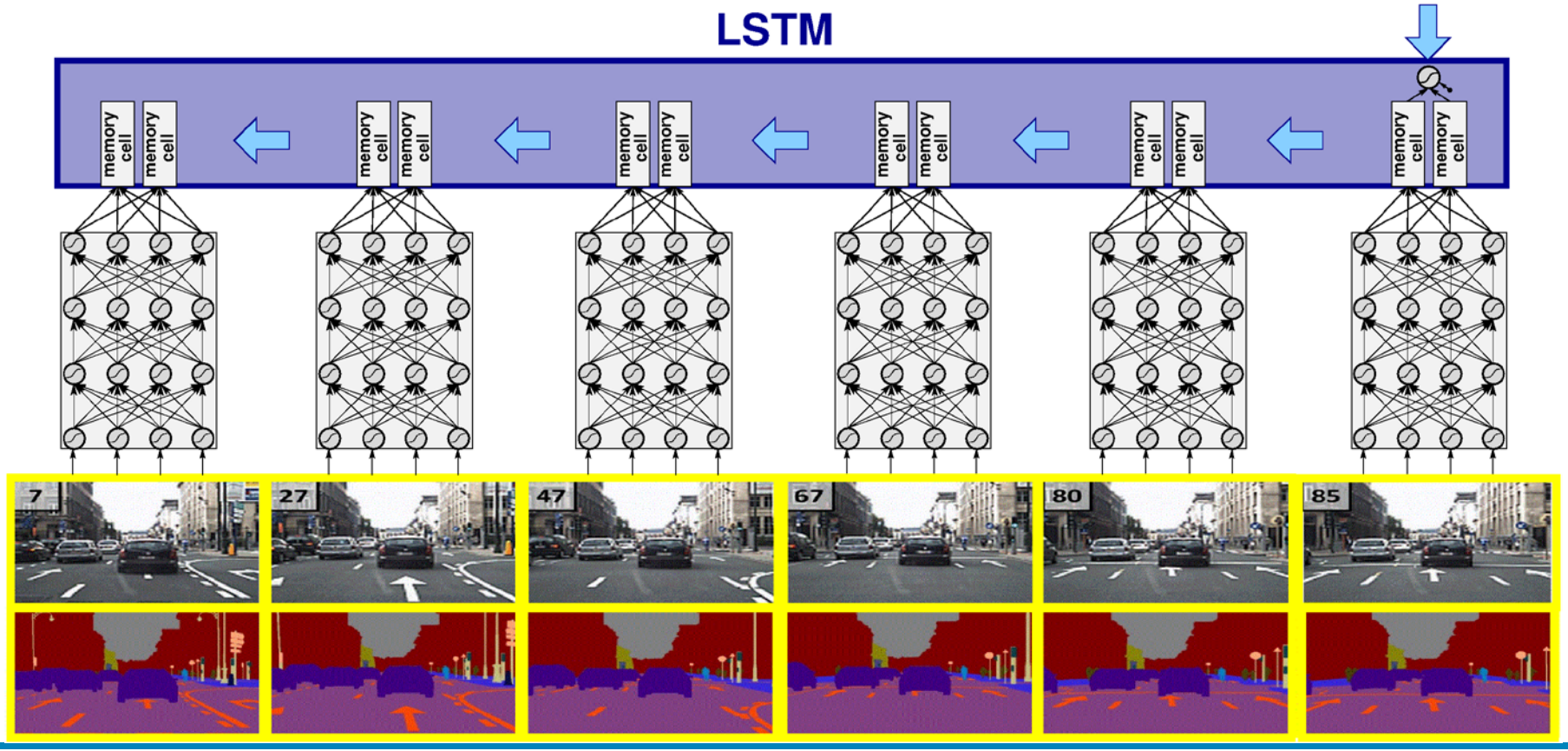
LSTM Action Prediction



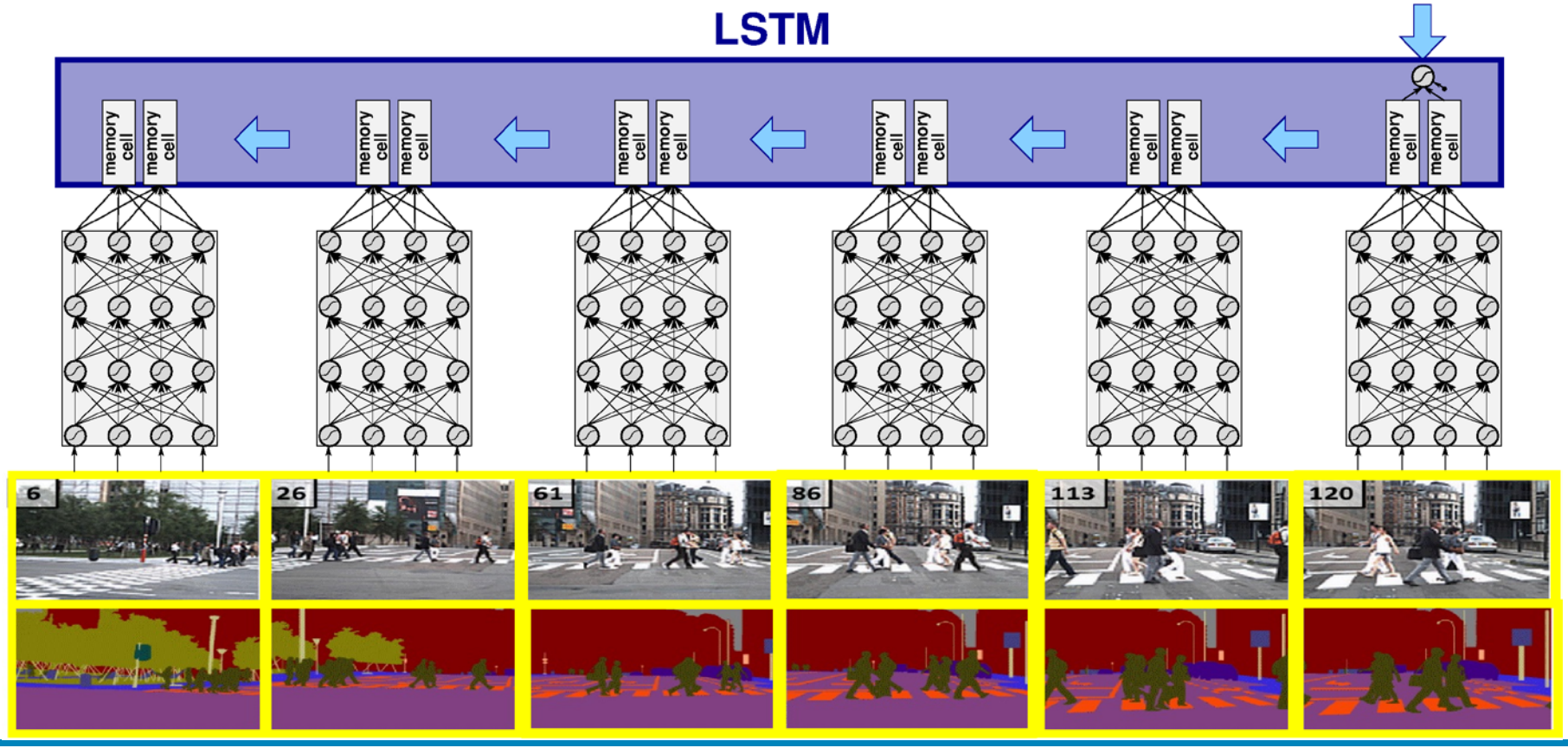
LSTM Action Prediction



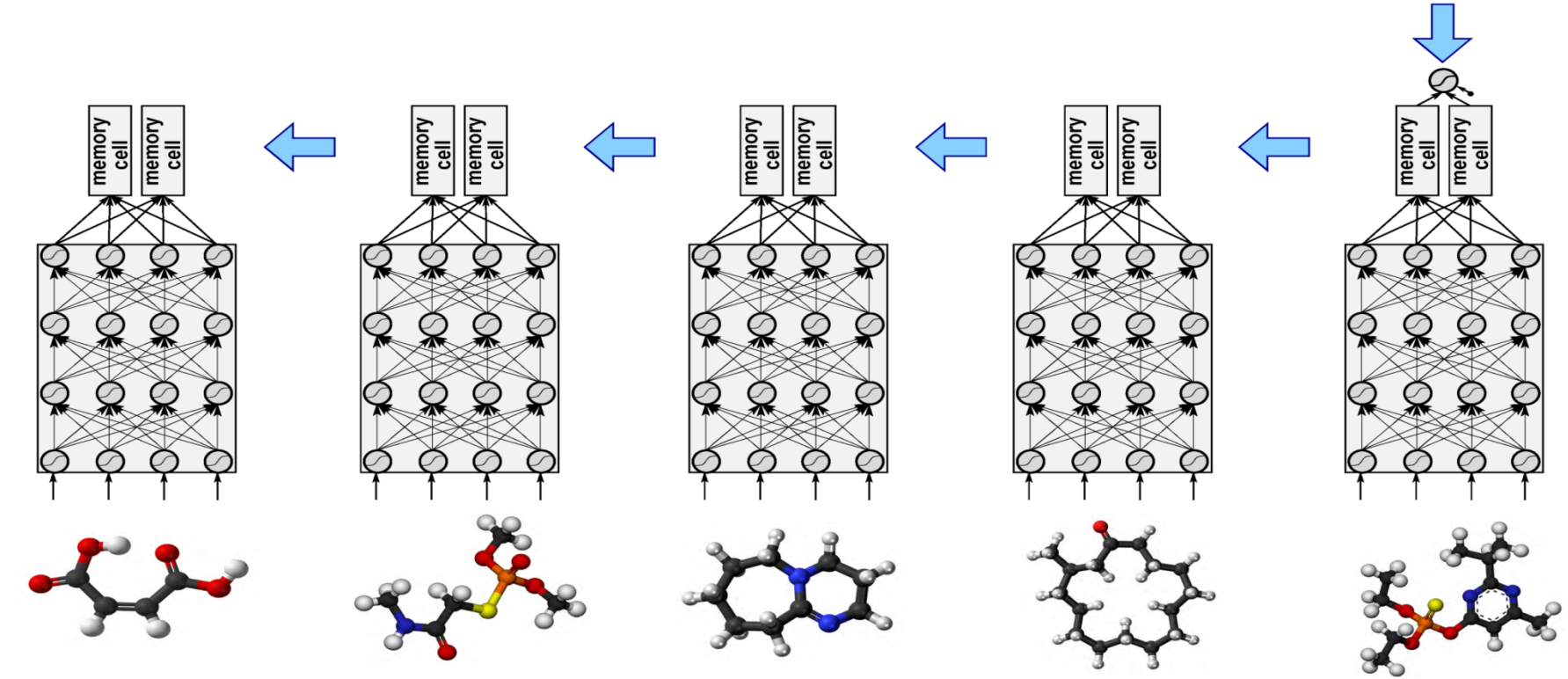
LSTM Action Prediction



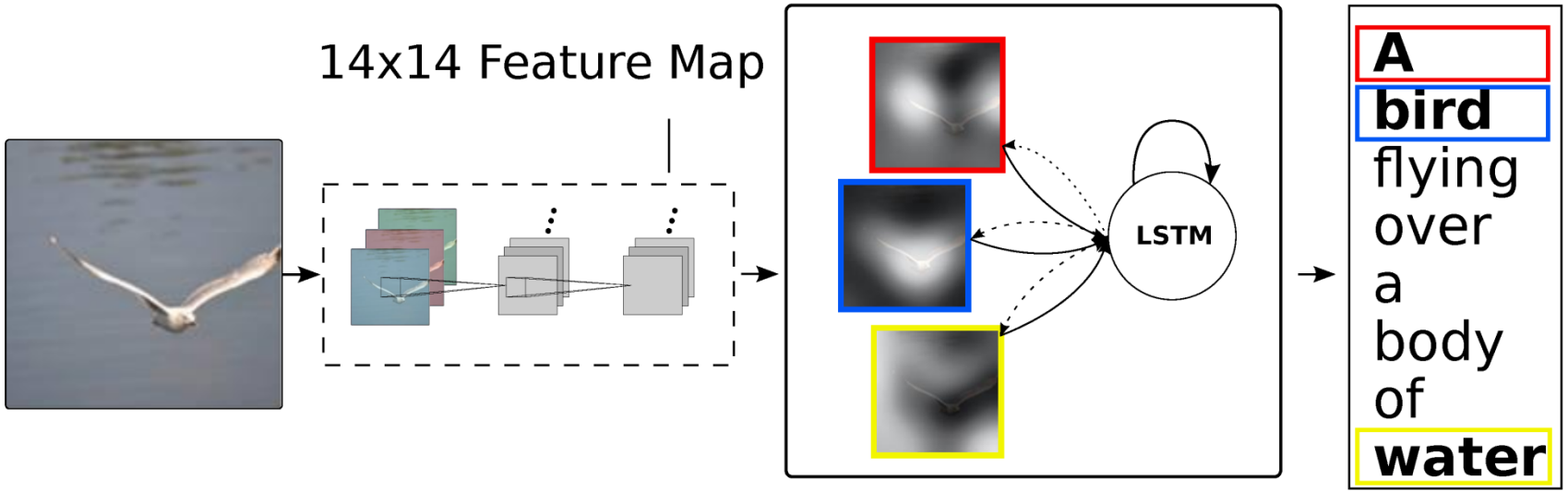
LSTM Action Prediction



LSTM Drug Mixture Effects



LSTM for Attention



1. Input Image
2. Convolutional Feature Extraction
3. RNN with attention over the image
4. Word by word generation

LSTM for Attention

JOHANNES KEPLER UNIVERSITY LINZ

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

LSTM for Attention

Long Short-Term Memory

Long Short-Term Memory

StarCraft II

LSTM in Projects

LSTM in Projects

JKU

Partners

LSTM in Projects

JOHANNES KEPLER UNIVERSITY LINZ

LSTM in Projects

JOHANNES KEPLER UNIVERSITY LINZ

Deep learning for fashion images and fashion blogs

- text and language analysis
- image analysis

zalando

LSTM in Projects

JOHANNES KEPLER UNIVERSITY LINZ

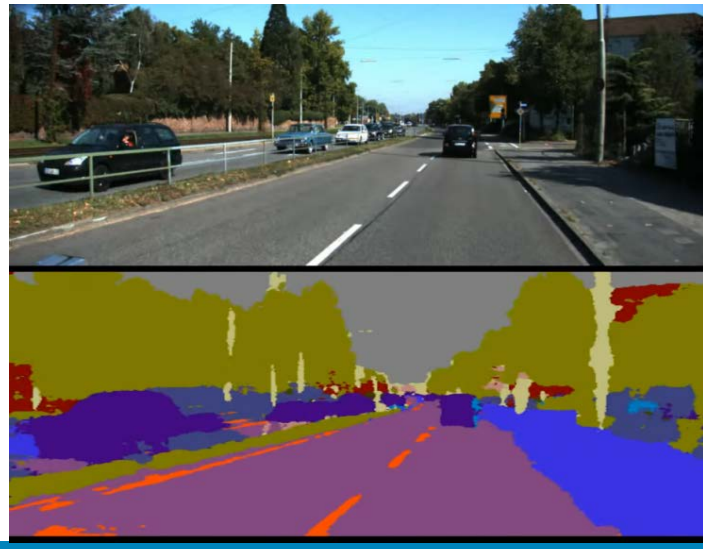
Audi | JKU

deep learning center

NVIDIA

Deep learning for autonomous driving

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



LSTM in Projects

JOHANNES KEPLER UNIVERSITY LINZ

Johnson & Johnson

janssen

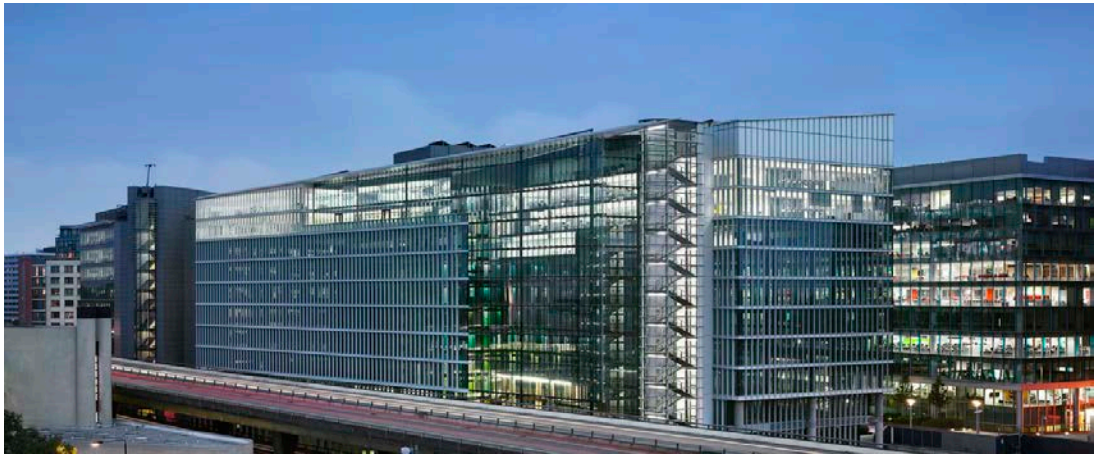
LSTM in Projects

JOHANNES KEPLER UNIVERSITY LINZ

LSTM in Projects

JOHANNES KEPLER UNIVERSITY LINZ

AstraZeneca



LSTM in Projects

JOHANNES KEPLER UNIVERSITY LINZ

MERCK

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**

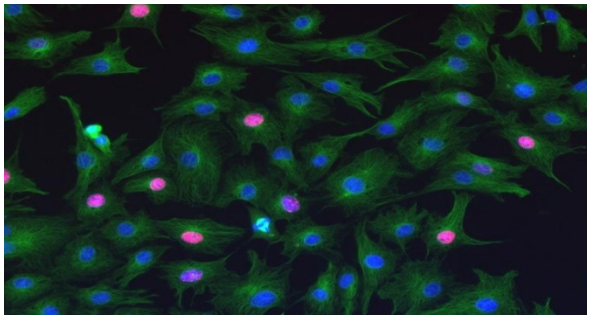
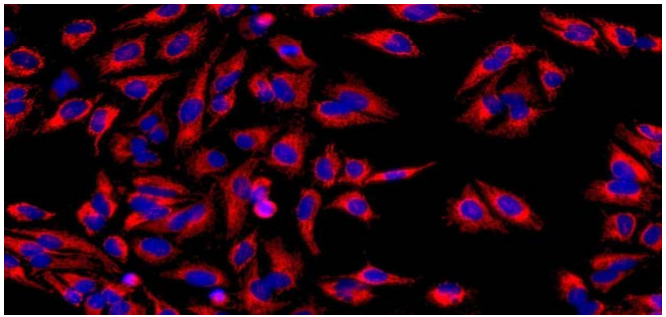
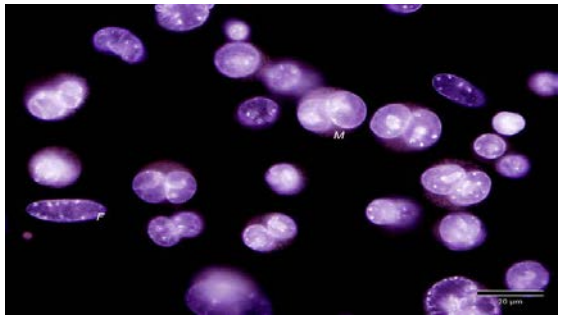
LSTM in Projects

JOHANNES KEPLER UNIVERSITY LINZ

janssen intel imec KU LEUVEN Universiteit Antwerpen UNIVERSITEIT GENT

Exaptation: Imaging and **deep learning**

- International consortium
- 650,000 high content imaging data

Chapter 1
Recurrent Neural Networks