Vienna Deep Learning Meetup

June 20, 2017 @ FH Technikum Wien

Thomas Lidy

Jan Schlüter

Alex Schindler

Organizing Host:
Thomas Faast

Vienna 12th Deep Learning Meetup

Agenda:

- Welcome by the Organizers (Tom Lidy, Jan Schlüter)
- Welcome by the Host (Prof. Alexander Hoffmann, Head of MSc Program Game Engineering & Simulation, FH Technikum Wien)
- **Microsoft Cognitive Toolkit and Applications in Image Object Recognition** (Philipp Kranen, Microsoft)
- **Generative Adversarial Networks** (Michal Šustr, TU Prague - Blindspot Solutions)
- AI Summit Vienna (Michael Platzer)
- Hot Topics & Latest News (Jan Schlüter, Alex Schindler, Gregor M.)
- Discussion

Vienna Deep Learning Meetup

Announcements

ToolSense

What do we do?

We're looking for people with Embedded, Electrical Eng. & ML background → stefan.oettl@toolsense.io

1 → 2 → 3

Connect

Analyze

- **gather** sensor data
- **analyze** sensor data

Monetize

!

Classification of Sensordata on Microcontroller!

Sepp Hochreiter, **JKU Linz**
Inventor of LSTM

Dave Elliott, **Google**
Global Product Lead

Ulla Kruhse-Lehtonen
DAIN Studios

Calvin Seward, **Zalando**
Deep Learning Research

Our Promise: a non-profit, content-rich AI summit tailored to a forward-thinking and tech-savvy audience

→ mostly.ai/summit

Sepp Hochreiter, **JKU Linz**
Inventor of LSTM

Deep Learning is Evolving into the Key Technology of Artificial Intelligence

Deep Learning has emerged as one of the most successful fields of machine learning and artificial intelligence with overwhelming success in industrial speech, language and vision benchmarks. Consequently it became the central field of research for IT giants like Google, facebook, Microsoft, Baidu, and Amazon. Deep Learning is founded on novel neural network techniques, the recent availability of very fast computers, and massive data sets. In its core, Deep Learning discovers multiple levels of abstract representations of the input.

The main obstacle to learning deep neural networks is the vanishing gradient problem. The **vanishing gradient** impedes credit assignment to the first layers of a deep network or early elements of a sequence, therefore limits model selection. I will show that the success of Deep Learning concepts like unsupervised stacking, ReLUs, residual networks, highway networks, and LSTM can be traced back to avoiding the vanishing gradient. I will present our current research on **self-normalizing neural networks** which do not suffer for exploding or vanishing gradients and converge to optimal learning states. Currently, we advance generative adversarial networks (GANs) by stochastic approximation theory which allows proofing the convergence of GAN learning.

In the **AUDI Deep Learning Center**, which I am heading, we apply Deep Learning to advance autonomous driving using LSTMs for sequence prediction and for attention. We use it or **natural language processing** (NLP) in collaboration with companies like Zalando, e.g. to analyze fashion blogs. Using Deep Learning, we won the NIH Tox21 challenge and identified unknown **side effects of drug candidates** based on data from bioassays and high content imaging.

Dave Elliott, **Google**
Global Product Lead

Machine Learning with Google Cloud

Artificial intelligence and machine learning has great potential to improve our lives. Google has made machine learning a central part of how they design services, and Google Cloud has developed several ways for both data scientist and software developers to realize that potential.

In this session we will walk through examples of **how Google uses machine learning internally**, and then introduce different ways people are working with Google and Google Cloud, including:

- **TensorFlow**: an open-source machine learning platform that is fast, flexible, and production-ready, and scales from research to production.
- **Google Cloud Machine Learning Engine**: a managed service that enables you to build machine learning models that work on any type of data, of any size.
- **APIs**: pre-trained models available via application programming interfaces (APIs) that allow software developers to realize ML benefits in domains such as image analysis, video analysis, speech recognition, text analysis, and dynamic translation.

Ulla Kruhse-Lehtonen
DAIN Studios

Seizing the Machine Learning Opportunity

Gartner predicts 2017 to be the year, when Machine Learning reaches its peak in terms of inflated expectations. This is the right time to not only talk about the sea of opportunity, but also provide a no-nonsense view of the practical obstacles which organizations will encounter on their path to seizing them.

In this talk I'll cover these topics from a **managerial perspective**:

- Machine Learning **use cases** from the industry
- **Factors of success** and non-success
- Ideation process for **data products**
- Efficient **organizational setup**
- **Recruiting** & retaining of data scientists

Vienna
Deep Learning
Meetup

Calvin Seward, **Zalando**
Deep Learning Research

Deep Learning: More Than Classification

Many of Deep learning's first big breakthroughs were in the field of classification, for example recognizing handwritten digits or imagenet images. While such results were impressive, they are only a drop in the sea of possible deep learning applications, a sea whose extent we are only just now beginning to discover. In this talk, I'll demonstrate a sampling of and exciting **deep learning applications** we're working on **at Zalando** in cooperation with Prof. Dr. Sepp Hochreiter's team at JKU Linz. These include image generation with **Generative Adversarial Networks**, Semi-Supervised Semantic Segmentation, **Warehouse Optimization** and **Recommender Systems**. I'll then wrap up with a few technical tips and tricks on how you too can start learning your own deep neural networks.

Vienna Deep Learning Meetup

Sepp Hochreiter, **JKU Linz**
Inventor of LSTM

Dave Elliott, **Google**
Global Product Lead

Ulla Kruhse-Lehtonen
DAIN Studios

Calvin Seward, **Zalando**
Deep Learning Research

Our Promise: a non-profit, content-rich AI summit tailored to a forward-thinking and tech-savvy audience

⟶ mostly.ai/summit

Modern Data Science Tools Meetup
Tue, Jun 27th

Building Streaming Applications with Kafka

https://www.meetup.com/Vienna-Modern-Data-Science/events/240730623/

Vienna
Deep Learning
Meetup

You're brilliant? We're hiring.

- Junior Data Scientist
  https://recruiting.novomatic.com/Vacancies/776/Description/2

- Senior Data Scientist
  https://recruiting.novomatic.com/Vacancies/777/Description/2

NOVOMATIC

Vienna Deep Learning Meetup

PhD Career Opportunity at WU Wien

- Institute for Service Marketing at WU Wien
- Strong emphasis on quantitative research methods and empirical validation
- Research focus on applying Deep Learning within the marketing domain, in particular CRM
- Master’s degree of equivalent in statistics, mathematics, computer science or econometrics
- Working knowledge in quantitative marketing or economics are an advantage but not required
- Interdisciplinary and international team

If interested please reach out to thomas.reutterer@wu.ac.at

Vienna Deep Learning Meetup

Digital
Heat

Kaspar Möhring
Co-Founder

www.digitalheat.at
kaspar.moehring@digitalheat.at
+43 660 123 72 09

Latest News
Hot Topics

Jan Schlüter - Alex Schindler -
Gregor Mitscha-Baude

a 5-10 min block at every meetup
to briefly present “trending topics”

Send us contributions (tom.lidy@gmail.com)
or come with slides to do a 5-10 min block yourself!

Vienna Deep Learning Meetup

Deep Voice 2: Multi-Speaker Neural Text-to-Speech

- Improved architecture of Deep Voice 1
- Improved Tacotron with a WaveNet instead of Griffin-Lim
- Extended both to multi-speaker systems, with learnable, ~50-dimensional speaker embeddings fed at various points into the networks

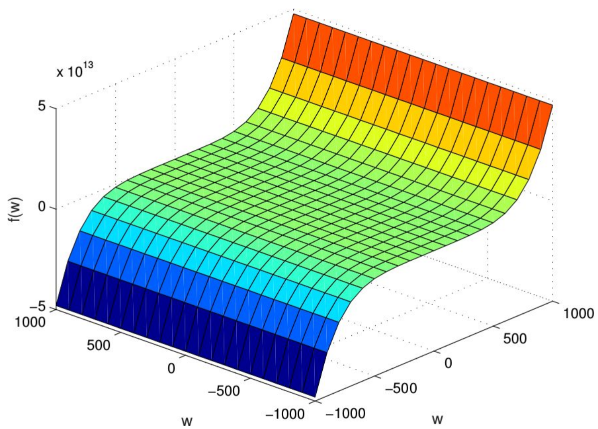
- Experimental results:
  - Better quality than Deep Voice 1 and Tacotron according to human judgements
  - Faithful reproduction of voices according to a speech identification system
- Can learn hundreds of speakers from ~30 min audio per speaker

http://research.baidu.com/deep-voice-2-multi-speaker-neural-text-speech/

http://arxiv.org/abs/1705.08947 (May 24th)

Are Saddles Good Enough for Deep Learning?

- Common idea: optimization finds a local minimum, needs to avoid getting stuck at saddle points
- This paper: optimization actually finds saddle points, even more so for methods that try to avoid saddles
- Catch: It finds *degenerate saddles* with zero Eigenvalues of the Hessian. Found by training a network on CIFAR-10 small enough to allow computing the full Hessian.
- May have implications on the design of optimization methods.



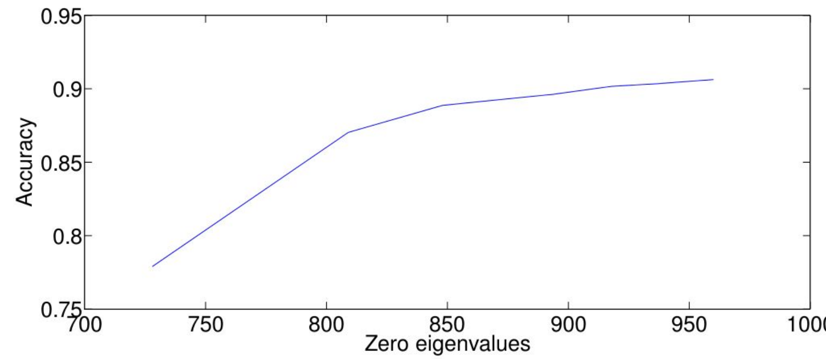


Fig. 3. Steady increase of zero eigenvalues while training

http://arxiv.org/abs/1706.02052 (June 7th)

Vienna Deep Learning Meetup

Self-Normalizing Neural Networks

- Problem: Vanishing/Exploding gradient
- Solution: Scale inputs to zero mean / unit variance, initialize weights such that mean/variance is kept through the layers

- Problem: When weights change, mean/variance changes
- Solutions: Batch normalization, weight normalization

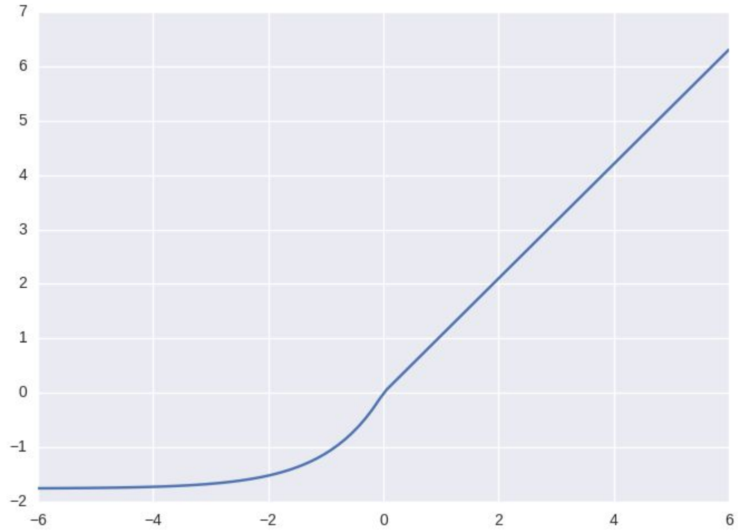
- Problem: Introduces noise in training, works for CNNs/RNNs due to weight sharing, but not so well for deep MLPs
- Solution: Design nonlinearity such that layer output not only *keeps* zero mean / unit variance, but is *drawn towards it*. This way, more layers mean better standardization.

$$\mathrm{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leqslant 0 \end{cases}$$

http://arxiv.org/abs/1706.02515 (June 8th)

Self-Normalizing Neural Networks

alpha = 1.6732632423543772848170429916717
lambda = 1.0507009873554804934193349852946

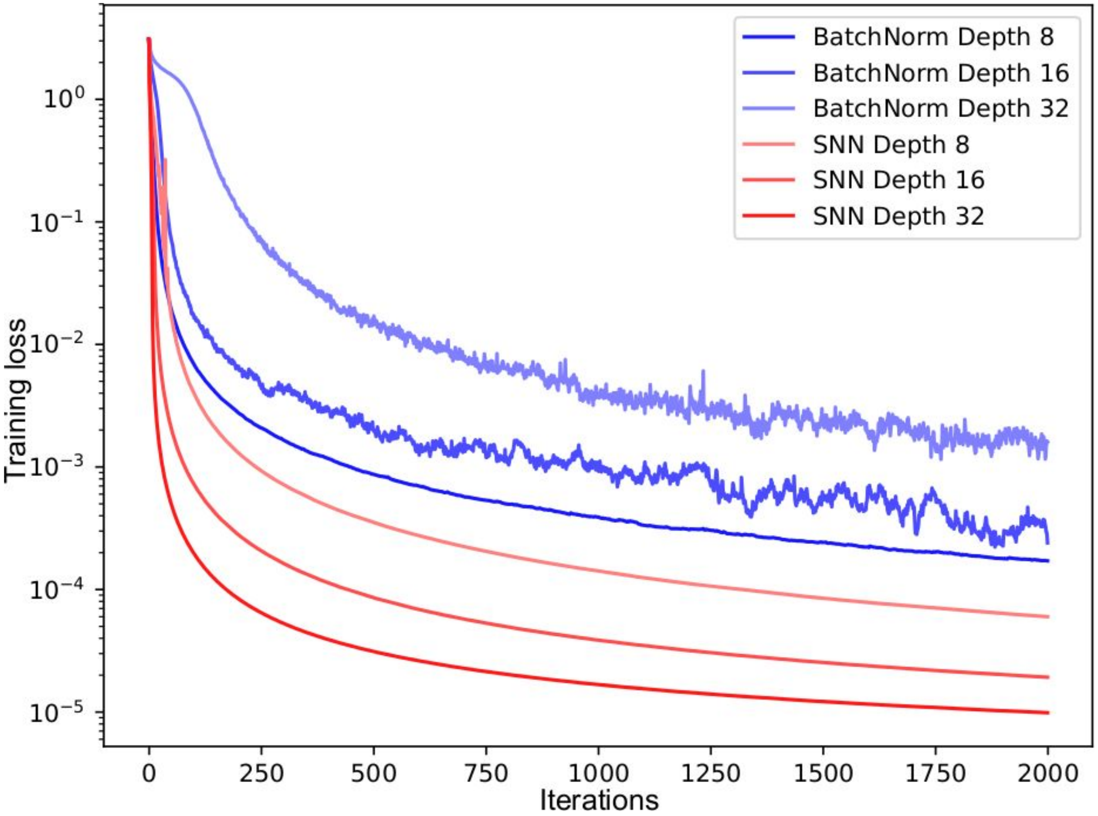


$$\mathrm{selu}(x) = \lambda \begin{cases} x & \text{if } x > 0 \\ \alpha e^x - \alpha & \text{if } x \leqslant 0 \end{cases}$$

http://arxiv.org/abs/1706.02515 (June 8th)

Vienna Deep Learning Meetup

Self-Normalizing Neural Networks



http://arxiv.org/abs/1706.02515 (June 8th)

Vienna Deep Learning Meetup

Depthwise Separable Convolutions for Neural Machine Translation

Google Brain, 2017 http://arxiv.org/abs/1706.03059v1

- Trend: Convolutional architectures for sequence problems
  - Often better/faster than RNN state of the art
  - WaveNet for audio generation (http://arxiv.org/abs/1609.03499v2)
  - ByteNet for machine translation (http://arxiv.org/abs/1610.10099v2)
  - ConvS2S by Facebook (https://arxiv.org/abs/1705.03122v2)

Depthwise Separable Convolutions for Neural Machine Translation

Google Brain, 2017 http://arxiv.org/abs/1706.03059v1

Normal (2D) convolution:

$$(\mathrm{W} * \mathrm{x})_{ij}^{n} = \sum_{k,l,m} W_{kl}^{nm} x_{i+k,j+l}^{m}$$

“Depthwise separable” convolution:

$$(\mathrm{W} *^{\mathrm{ds}} x)_{ij}^{n} = \sum_{k,l,m} W_{kl}^{n} W^{nm} x_{i+k,j+l}^{m}$$

- Kernel/Channel weights separated
- Parameters: $O(K^2N^2)$ vs. $O(K^2N + N^2)$

- Has also been applied to images (Xception)
- For sequence models, no dilated convolutions

Vienna Deep Learning Meetup

Depthwise Separable Convolutions for Neural Machine Translation

Google Brain, 2017 http://arxiv.org/abs/1706.03059v1

Model	BLEU (newstest14)
SliceNet (Full, 2048)	25.5
SliceNet (Super 2/3, 3072)	**26.1**
ByteNet [11]	23.8
GNMT [23]	24.6
ConvS2S [7]	25.1
GNMT+Mixture of Experts [17]	26.0

Table 3: Performance of our larger models compared to best published results.

Vienna Deep Learning Meetup

Attention Is All You Need

Google Brain, 2017 http://arxiv.org/abs/1706.03762v1

- Another novel architecture for MT
- No convolutions or RNN units, only attention
- Massive improvement on SOTA

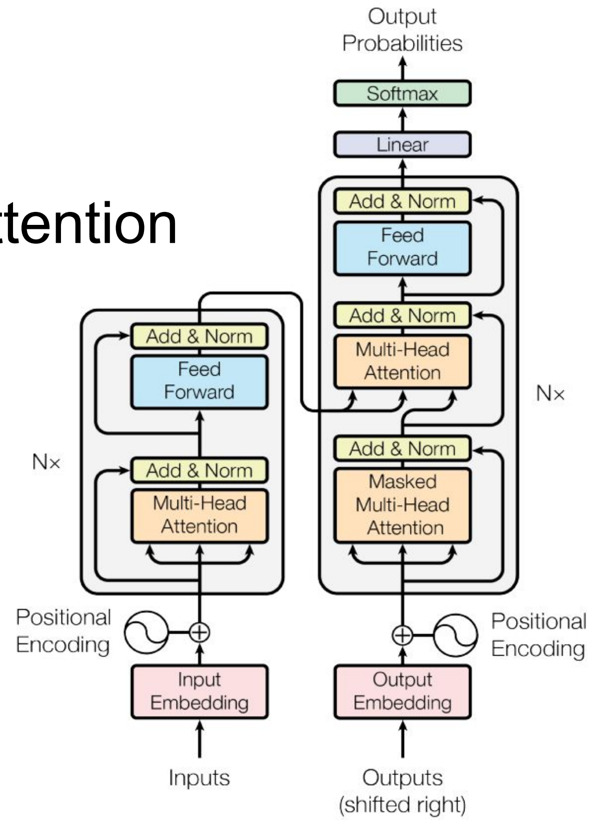


Figure 1: The Transformer - model architecture.

Vienna Deep Learning Meetup

Attention Is All You Need

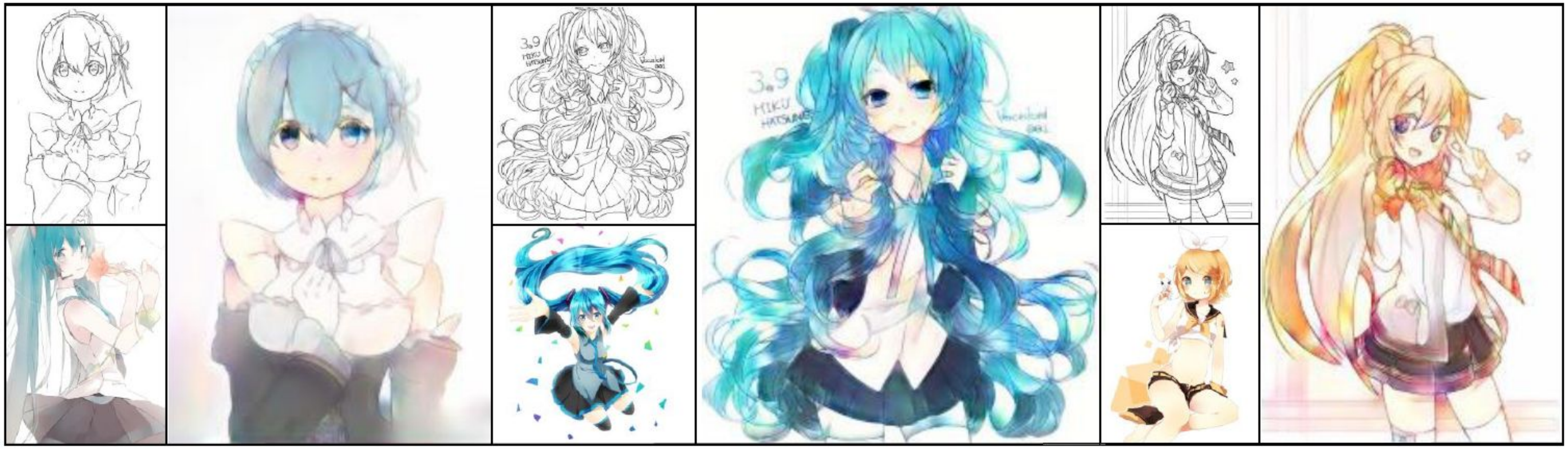
Google Brain, 2017 http://arxiv.org/abs/1706.03762v1

Table 2: The Transformer achieves better BLEU scores than previous state-of-the-art models on the English-to-German and English-to-French newstest2014 tests at a fraction of the training cost.

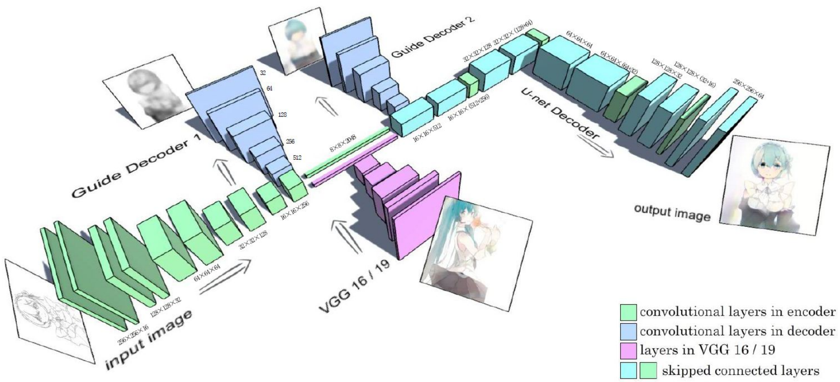
Model	BLEU		Training Cost (FLOPs)	
	EN-DE	EN-FR	EN-DE	EN-FR
ByteNet [16]	23.75			
Deep-Att + PosUnk [35]		39.2		$1.0 \cdot 10^{20}$
GNMT + RL [34]	24.6	39.92	$2.3 \cdot 10^{19}$	$1.4 \cdot 10^{20}$
ConvS2S [9]	25.16	40.46	$9.6 \cdot 10^{18}$	$1.5 \cdot 10^{20}$
MoE [29]	26.03	40.56	$2.0 \cdot 10^{19}$	$1.2 \cdot 10^{20}$
Deep-Att + PosUnk Ensemble [35]		40.4		$8.0 \cdot 10^{20}$
GNMT + RL Ensemble [34]	26.30	41.16	$1.8 \cdot 10^{20}$	$1.1 \cdot 10^{21}$
ConvS2S Ensemble [9]	26.36	**41.29**	$7.7 \cdot 10^{19}$	$1.2 \cdot 10^{21}$
Transformer (base model)	27.3	38.1	$\mathbf{3.3 \cdot 10^{18}}$	
Transformer (big)	**28.4**	**41.0**	$2.3 \cdot 10^{19}$	

Vienna Deep Learning Meetup

Style Transfer for Anime Sketches with Enhanced Residual U-net and Auxiliary Classifier GAN



- Apply the style of a painting to a grayscale sketch
- Residual U-Net
  - Image segmentation
  - Large hints from VGG 19 outputs
- with Auxiliary Classifier Generative Adversarial Network (AC-GAN)

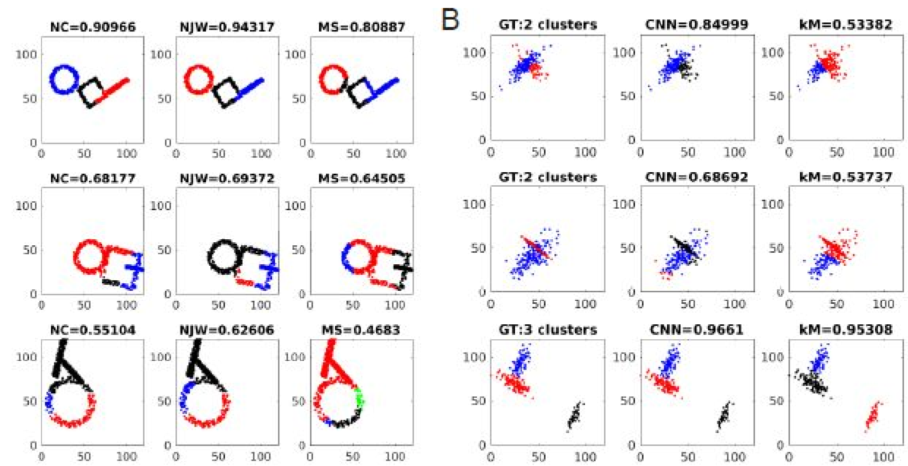


https://arxiv.org/abs/1706.03319

Vienna Deep Learning Meetup

CNN based Clustering

- Human vision inspired clustering (CNN-based)
- U-Net encoder-decoder Architecture
- Promising results, even with noisy data



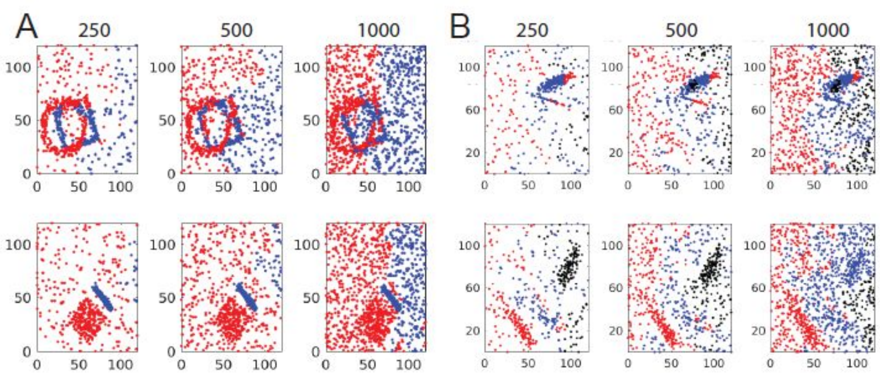


Table 1 shows the results of the first five experiments. CNN wins over all models in all experiments with a large margin (e.g., about 7% improvement over the second best in experiment 1). This large margin hints that maybe even the best optimization of the compared algorithms will not be able to compete with the proposed CNN. All other models perform about the same.

https://arxiv.org/abs/1706.05048

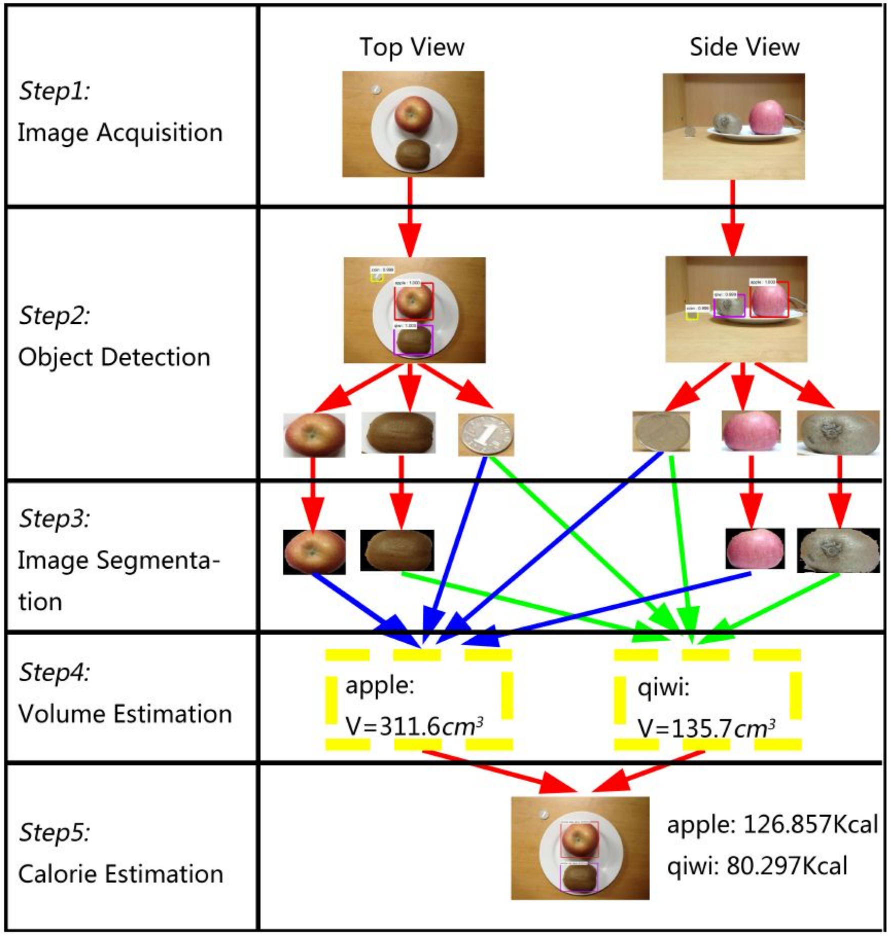
Vienna Deep Learning Meetup

Honorable Mention
Deep Learning-Based Food Calorie Estimation Method

- Estimate calories from food images
- 1. Estimate volume
    - Fast R-CNN for object localization
- 2. Estimate mass
- 3. Calculate calories

- [Dataset provided on GitHub]

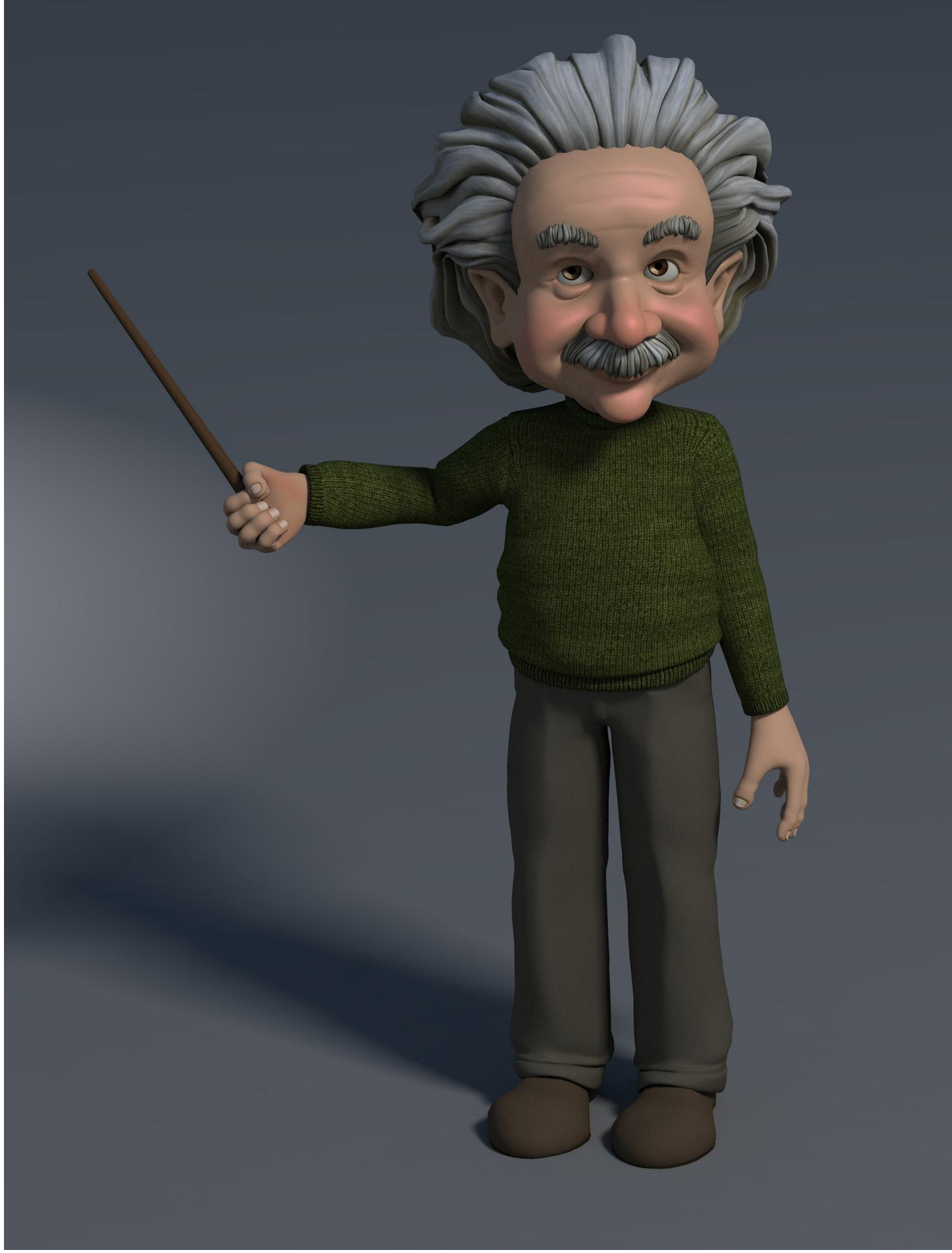
Food Type	estimation image number	mean volume	mean estimation volume	volume error (%)	mean mass	mean estimation mass	mass error (%)
apple	158	332.78	320.65	-3.65	263.60	250.02	-5.15
banana	92	162.17	128.13	-20.99	146.61	116.92	-20.25
bread	20	155.00	102.03	-34.17	29.04	18.71	-35.57
bun	56	247.50	237.02	-4.23	78.19	89.91	15.00
doughnut	10	166.00	197.58	19.03	63.44	59.44	-6.31
egg	42	52.38	56.40	7.67	61.20	65.80	7.51
fired dough twist	38	64.74	70.73	9.26	40.60	41.72	2.76
grape	26	240.00	203.98	-15.01	219.50	203.07	-7.48
lemon	112	96.79	100.54	3.88	94.24	94.71	0.49
litchi	46	43.48	45.35	4.30	44.01	44.38	0.84



https://arxiv.org/abs/1706.04062

Vienna Deep Learning Meetup

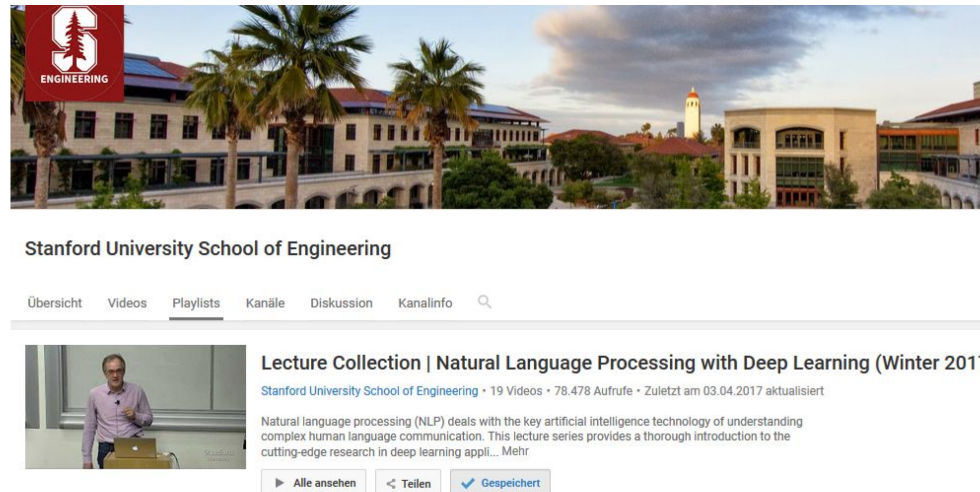
Tutorials
Courses
Books



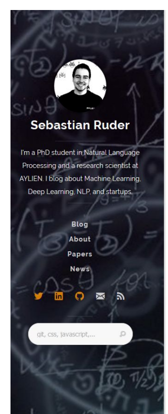
Vienna
Deep Learning
Meetup

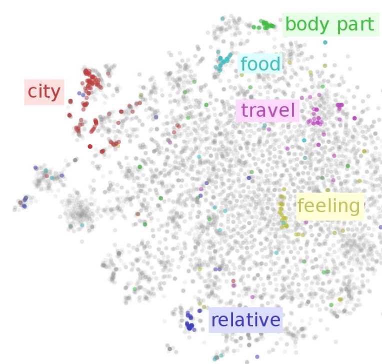
Natural Language Processing With Deep Learning

- [Complete Stanford Lecture on Youtube]
- 18 Lectures
- Bleeding Edge research topics
  - Word Embeddings
  - Machine Translation
  - Topic modelling
  - Speech Processing
  - Question Answering



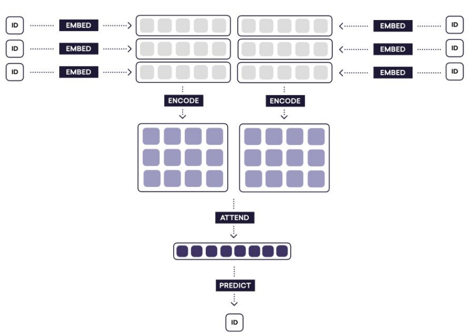
- [Blog Post by Sebastian Ruder]
  - 4-part blog post on Word Embeddings and DNN





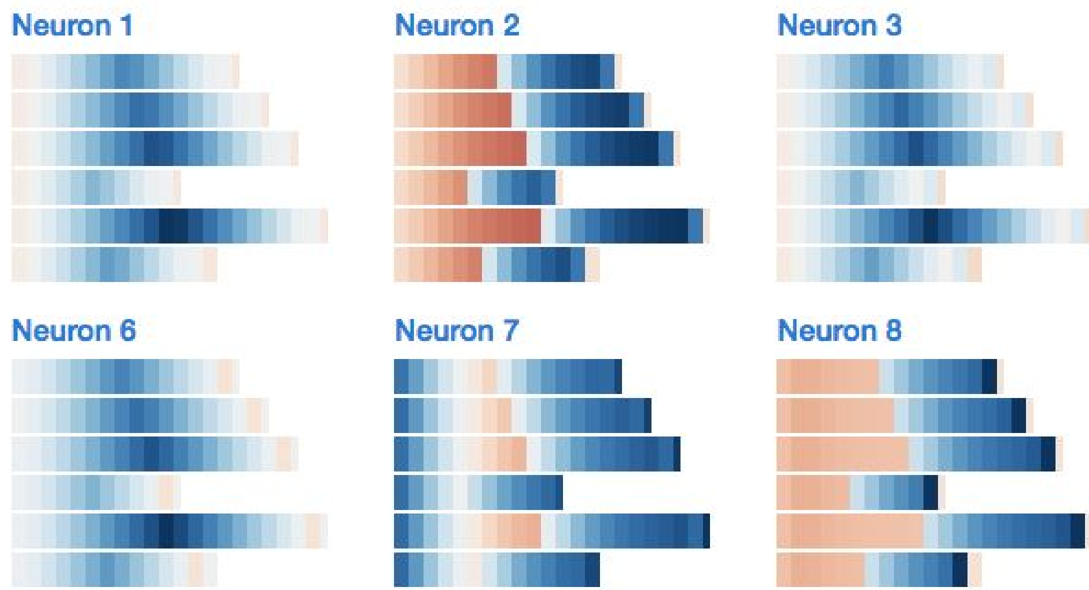
On word embeddings - Part 1

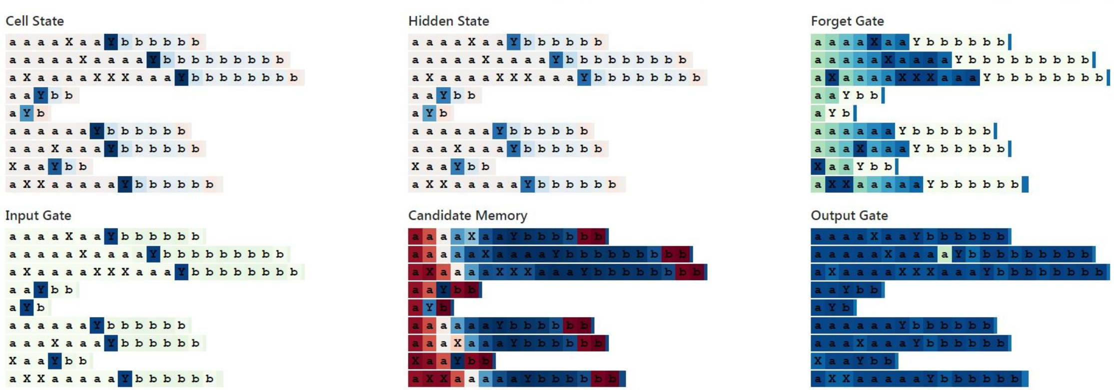
- [Blogpost by Matthew Honnibal]
  - Step by Step intro
  - Decomposable Attention Model
  - spaCy and Keras code on GitHub



Long-Short Term Memory (LSTM)

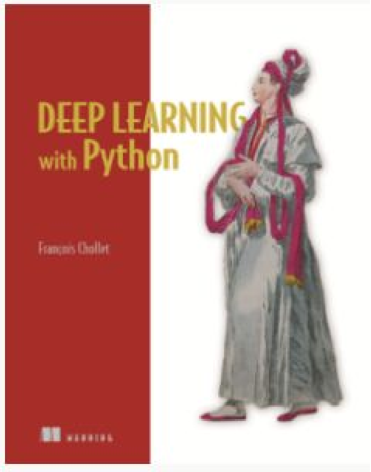
- Blog Post by Edwin Chen
  - Extensive introduction to LSTMs
  - Lots of visualizations



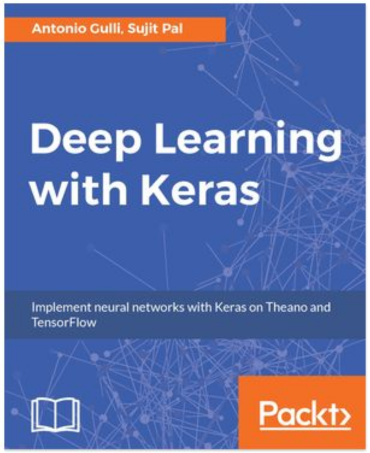


New Books on Deep Learning

- **Deep Learning with Python**
  - By Francois Chollet (Keras creator)



- **Deep Learning with Kears**
  - By Antonio Gulli, Sujit Pal



Vienna Deep Learning Meetup

Thank you for coming!

Next Event:

AI Summit Vienna:

4 Sep 2017 @ WU Wien Learning Center

mostly.ai/summit

Michael Platzer

Thomas Lidy

Jan Schlüter

Alex Schindler

Vienna Deep Learning Meetup