Zürcher Hochschule
für Angewandte Wissenschaften

Overview of what AI is & how DL works

Distinguished lecture, University of Engineering & Management, Kolkata
September 18, 2020

Thilo Stadelmann

Source: http://www.gymbsb.de/wp-content/uploads/2017/10/fall-1072821_1920.jpg

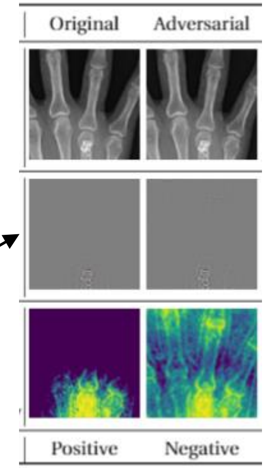
About us & our work

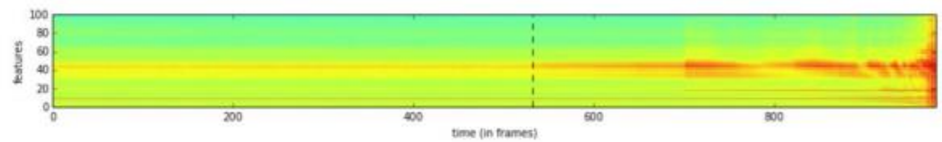
ZHAW School of Engineering, Winterthur, Switzerland

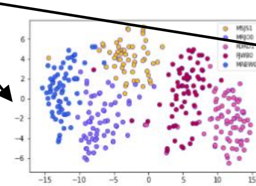
Zürcher Hochschule
für Angewandte Wissenschaften

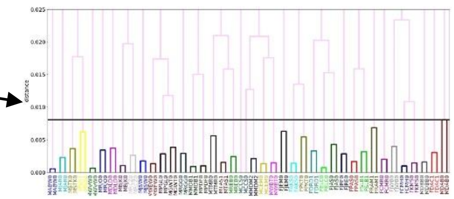
Machine learning-based Pattern Recognition

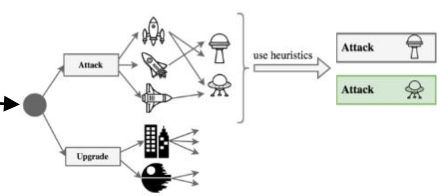
- Robust Deep Learning
- Voice Recognition
- Document Analysis
- Learning to Learn & Control

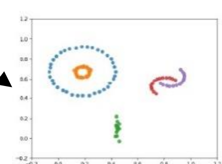












Zürcher Hochschule für Angewandte Wissenschaften

1

What is AI?

What is AI?

thinking

"The exciting new effort to **make computers think**... machines with minds, in the full and literal sense."
(Haugeland, 1985)

"[The automation of] **activities that we associate with human thinking**, activities such as decision-making, problem solving, learning..." (Bellman, 1978)

"The study of mental faculties through the use of **computational models**."
(Charniak and McDermott, 1985)

"The study of the **computations that** make it possible to **perceive, reason, and act**." (Winston, 1992)

humanly ← Standard measured by → rationally

↑ concerned with ↓

"The art **of creating machines that perform functions** that require **intelligence** when performed by people."
(Kurzweil, 1990)

"The study of **how to make computers do things** at which, at the moment, **people are better**."
(Rich and Knight, 1991)

"Computational Intelligence is **the study of the design of intelligent agents**."
(Poole et al., 1998)

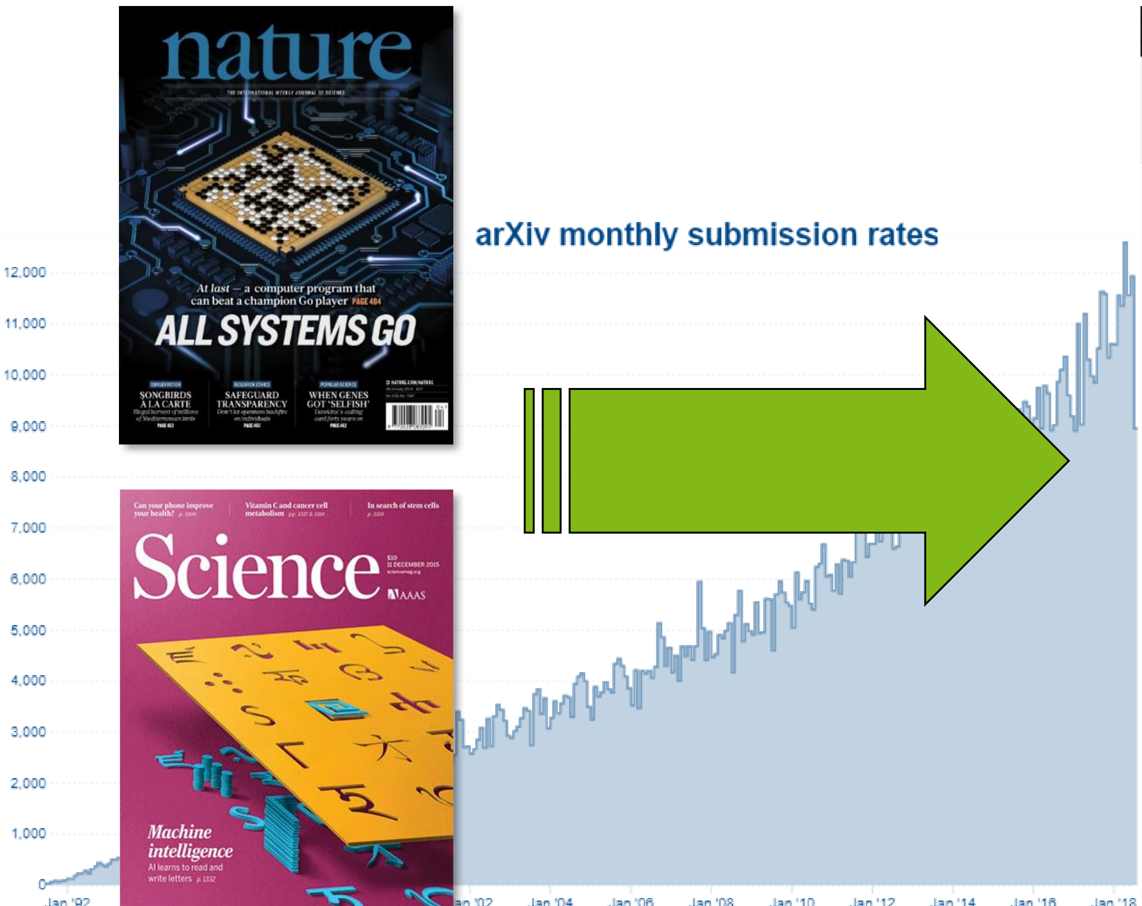
"AI... is concerned with **intelligent behaviour in artefacts**." (Nilsson, 1998)

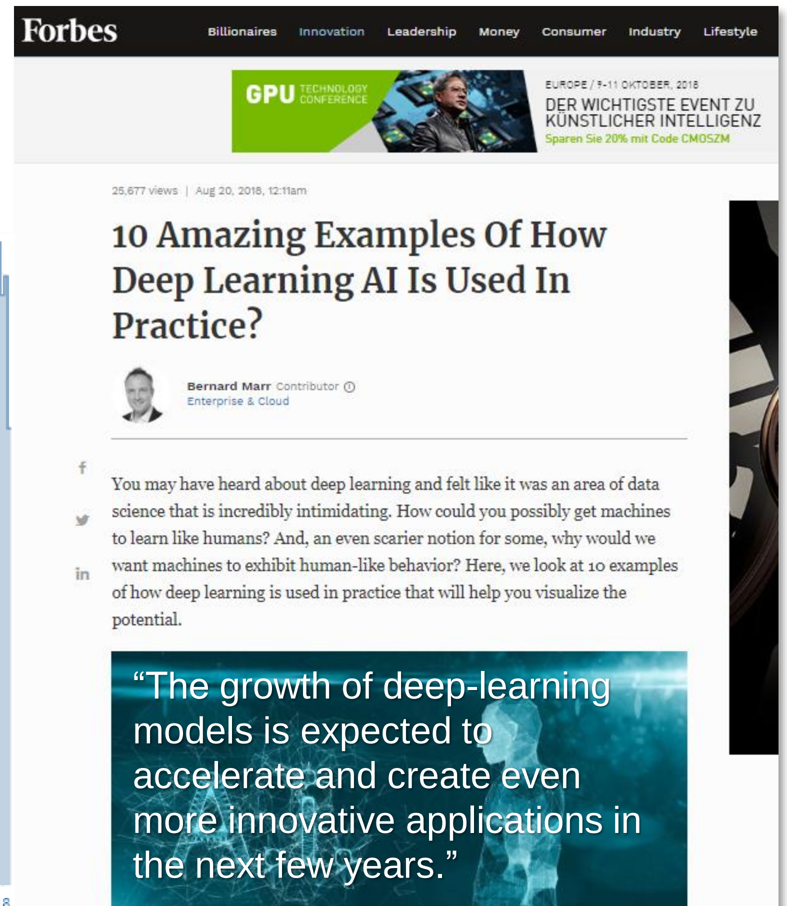
acting

The scientific field concerned with complex computer applications

Zürcher Hochschule
für Angewandte Wissenschaften

Why?




Forbes

Billionaires Innovation Leadership Money Consumer Industry Lifestyle

GPU TECHNOLOGY CONFERENCE

EUROPE / 9-11 OKTOBER, 2018

DER WICHTIGSTE EVENT ZU KÜNSTLICHER INTELLIGENZ

Sparen Sie 20% mit Code CMOSZM

25,677 views | Aug 20, 2018, 12:11am

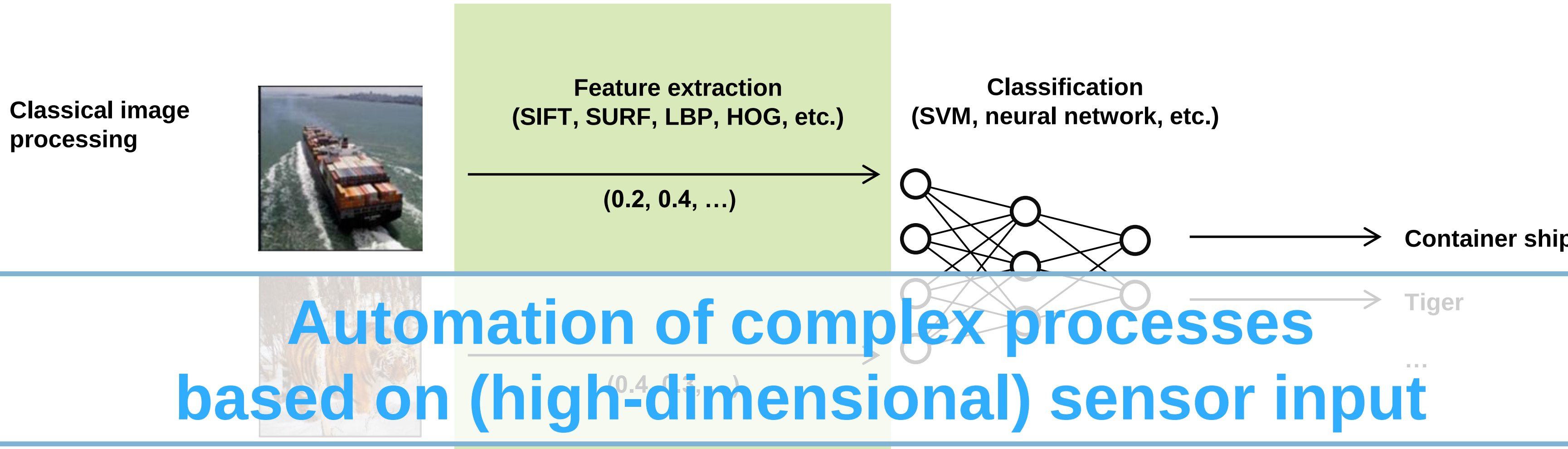
10 Amazing Examples Of How Deep Learning AI Is Used In Practice?

Bernard Marr Contributor
Enterprise & Cloud

You may have heard about deep learning and felt like it was an area of data science that is incredibly intimidating. How could you possibly get machines to learn like humans? And, an even scarier notion for some, why would we want machines to exhibit human-like behavior? Here, we look at 10 examples of how deep learning is used in practice that will help you visualize the potential.

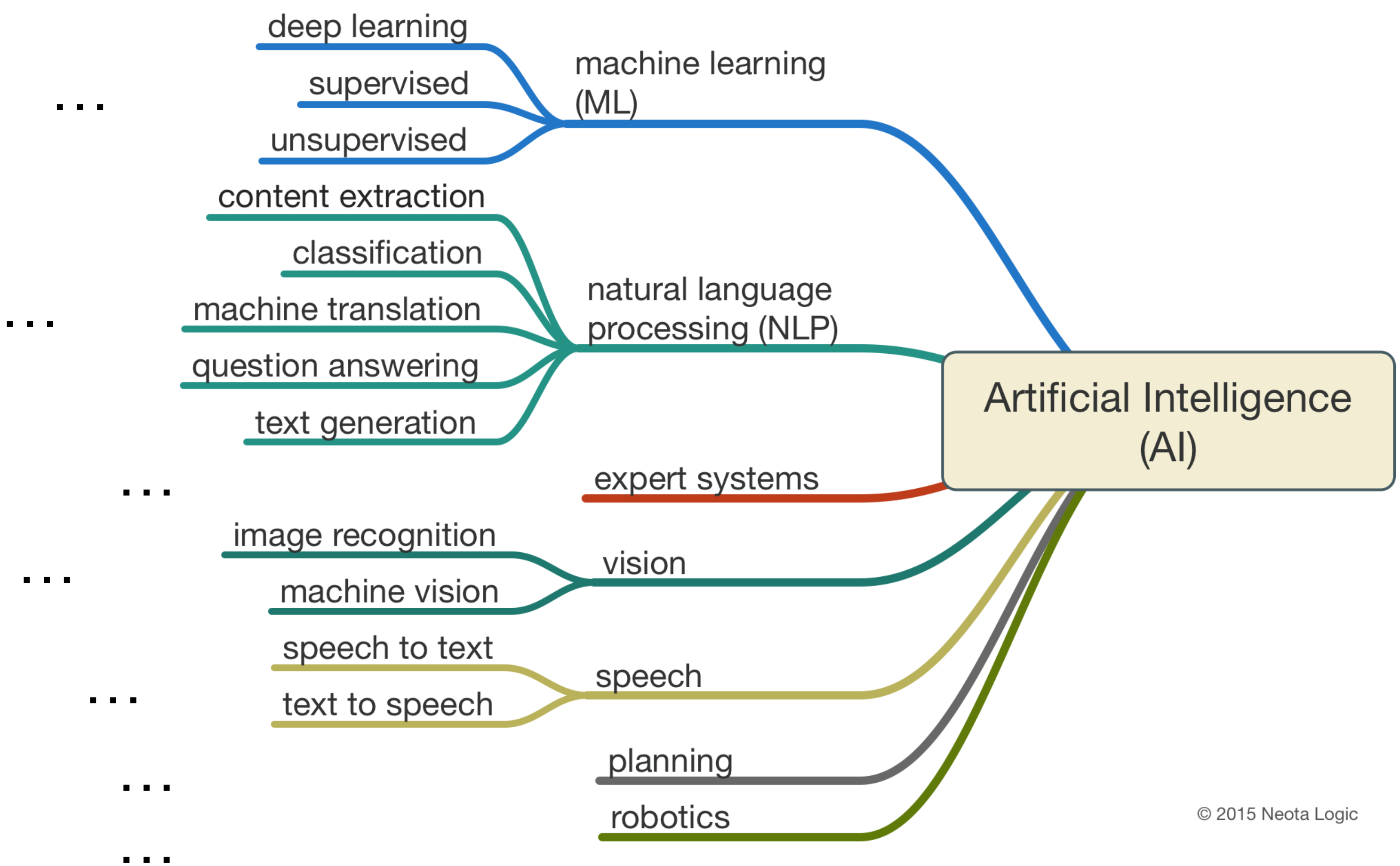
"The growth of deep-learning models is expected to accelerate and create even more innovative applications in the next few years."

Idea: Add depth to learn features automatically



Automation of complex processes based on (high-dimensional) sensor input

What belongs to AI?



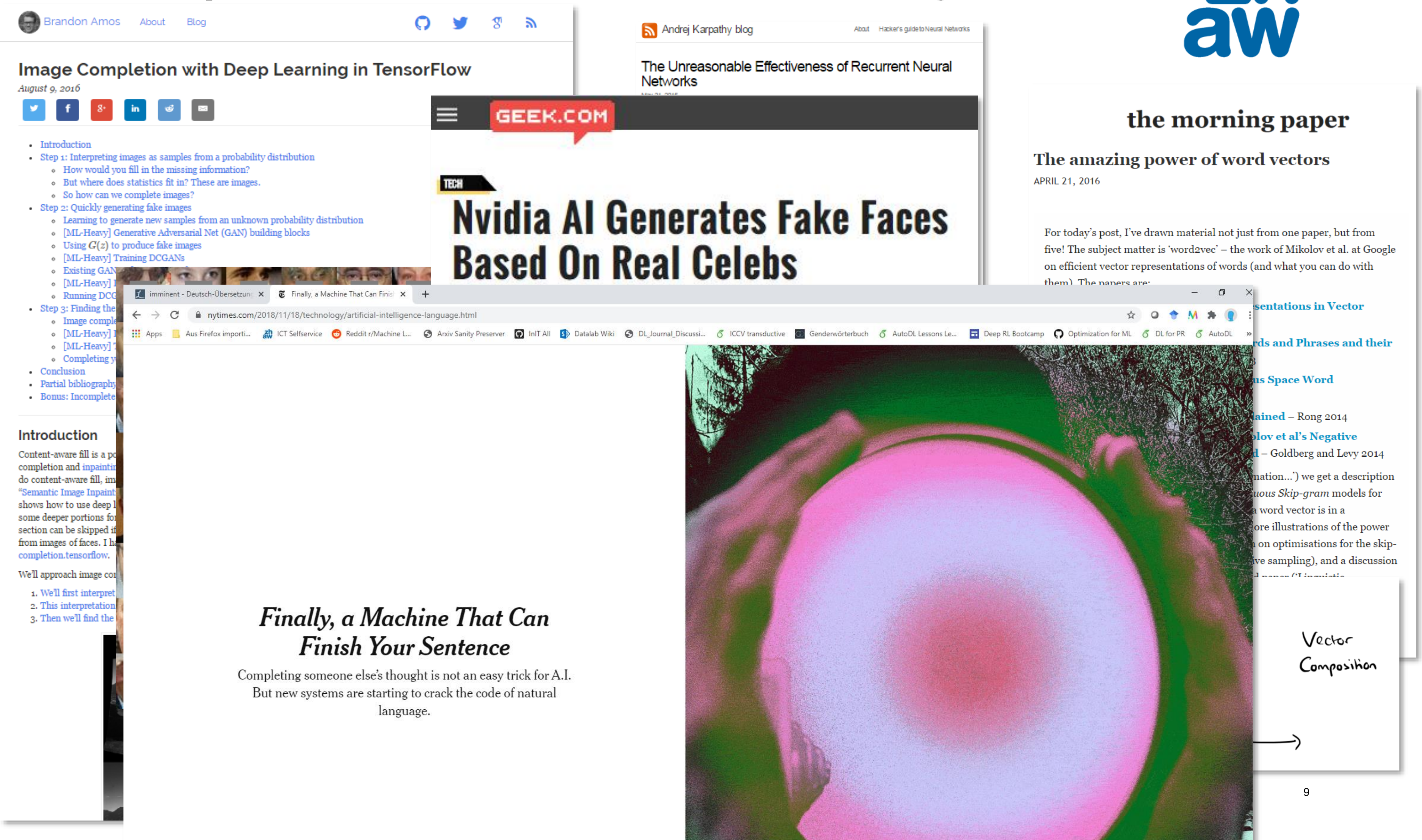
Zürcher Hochschule
für Angewandte Wissenschaften

2

How does Deep Learning Work?

Zürcher Hochschule
für Angewandte Wissenschaften

Examples of «AI» in the media in recent years



Brandon Amos About Blog

Image Completion with Deep Learning in TensorFlow

August 9, 2016

- Introduction
- Step 1: Interpreting images as samples from a probability distribution
  - How would you fill in the missing information?
  - But where does statistics fit in? These are images.
  - So how can we complete images?
- Step 2: Quickly generating fake images
  - Learning to generate new samples from an unknown probability distribution
  - [ML-Heavy] Generative Adversarial Net (GAN) building blocks
  - Using $G(z)$ to produce fake images
  - [ML-Heavy] Training DCGANs
  - Existing GAN
  - [ML-Heavy]
  - Running DCG
- Step 3: Finding the
  - Image compl
  - [ML-Heavy]
  - [ML-Heavy]
  - Completing y
- Conclusion
- Partial bibliography
- Bonus: Incomplete

Introduction

Content-aware fill is a p completion and inpaintin do content-aware fill, im "Semantic Image Inpaint shows how to use deep l some deeper portions fo section can be skipped i from images of faces. I h completion.tensorflow.

We'll approach image co

1. We'll first interpret
2. This interpretation
3. Then we'll find the

Andrej Karpathy blog About Hacker's guide to Neural Networks

The Unreasonable Effectiveness of Recurrent Neural Networks

GEEK.COM

TECH

Nvidia AI Generates Fake Faces Based On Real Celebs

the morning paper

The amazing power of word vectors

APRIL 21, 2016

For today's post, I've drawn material not just from one paper, but from five! The subject matter is 'word2vec' – the work of Mikolov et al. at Google on efficient vector representations of words (and what you can do with

sentations in Vector

rds and Phrases and their

us Space Word

ained – Rong 2014

lov et al's Negative

d – Goldberg and Levy 2014

nation...') we get a description

uous Skip-gram models for

a word vector is in a

ore illustrations of the power

on optimisations for the skip-

ve sampling), and a discussion

Vector Composition

imminent - Deutsch-Übersetzung | Finally, a Machine That Can Finis

nytimes.com/2018/11/18/technology/artificial-intelligence-language.html

Finally, a Machine That Can Finish Your Sentence

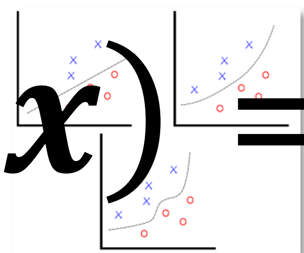
Completing someone else's thought is not an easy trick for A.I. But new systems are starting to crack the code of natural language.

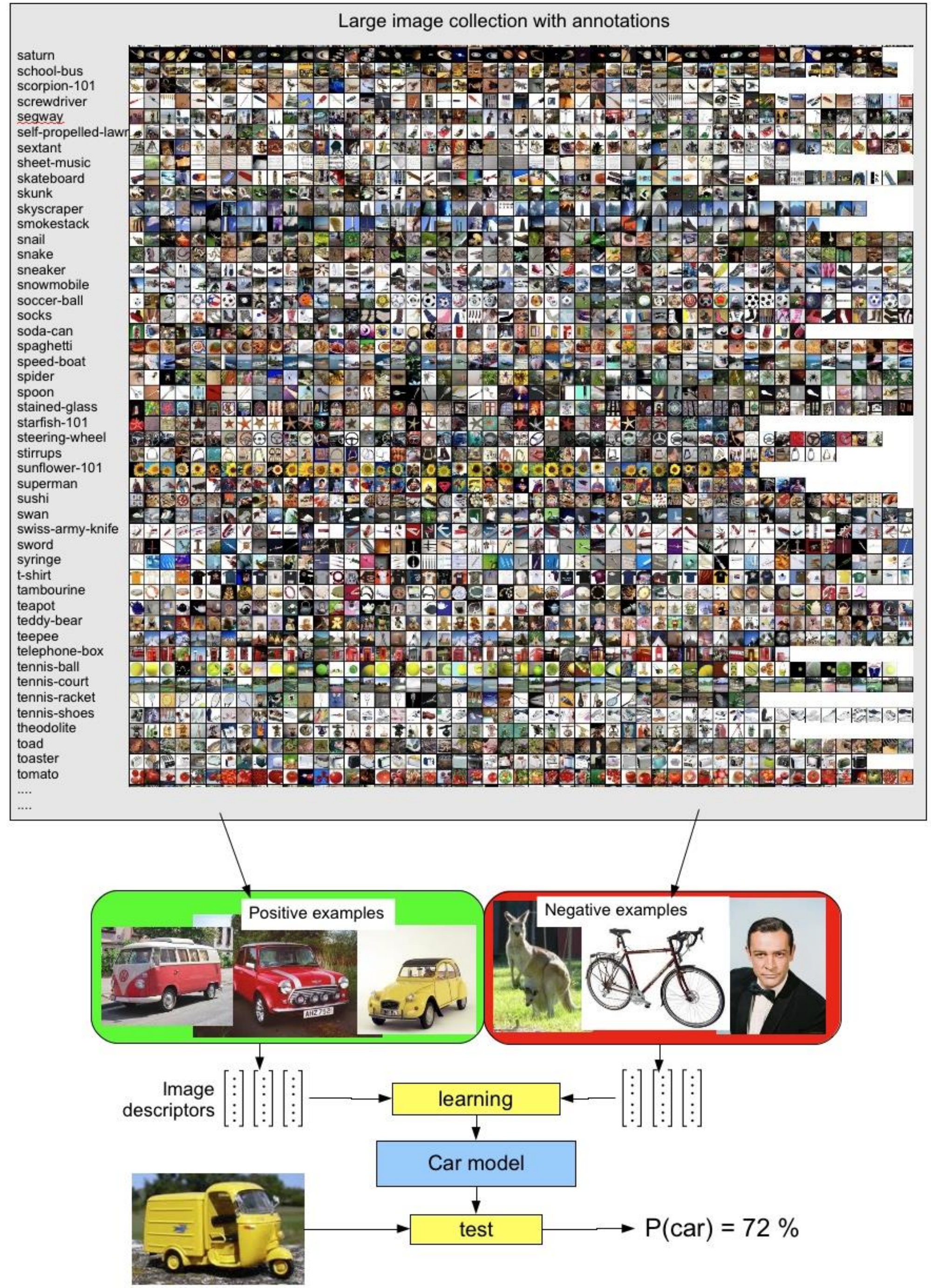
Foundation
Inductive supervised learning

Assumption
- A model fitted to a *sufficiently large* sample of data…
- …will **generalize** to unseen data

Method
- **Searching for optimal parameters of a function**…
- …such that all sample inputs (images) are mapped to the correct outputs (e.g., «car»)

$$f(x) = y$$



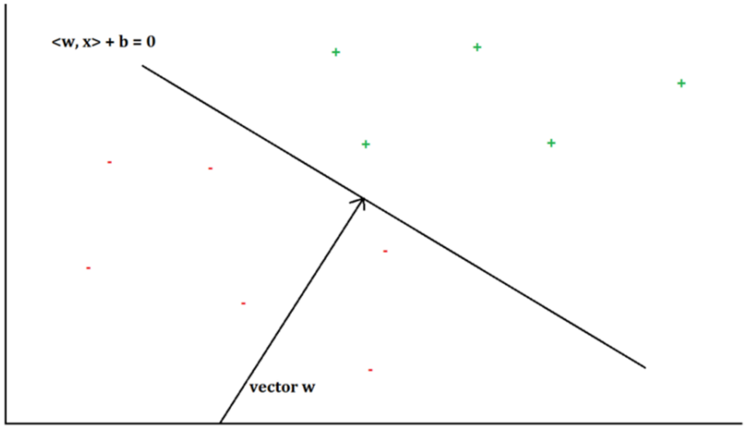


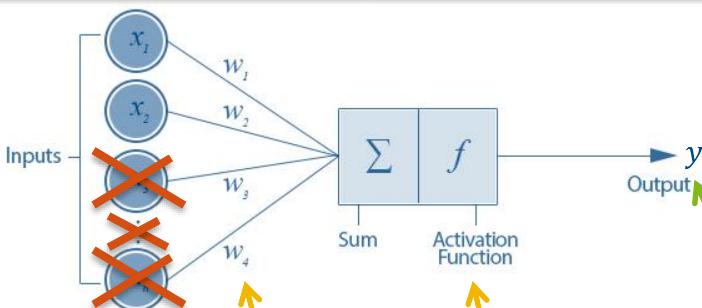
Quelle: http://lear.inrialpes.fr/job/postdoc-large-scale-classif-11-img/attribs_patchwork.jpg

Zürcher Hochschule
für Angewandte Wissenschaften

Search for optimal parameters *of a function*?

neuron





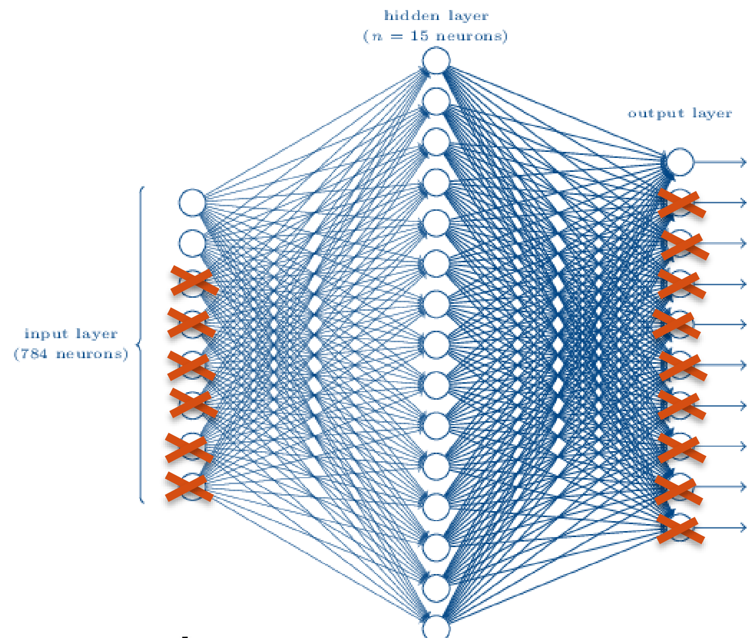
Threshold / decision

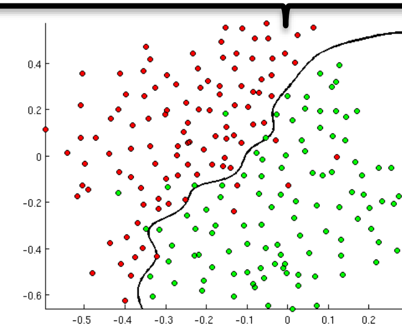
features (e.g., pixel)

Result (interpreted as e.g. «car»)

Tunable parameters

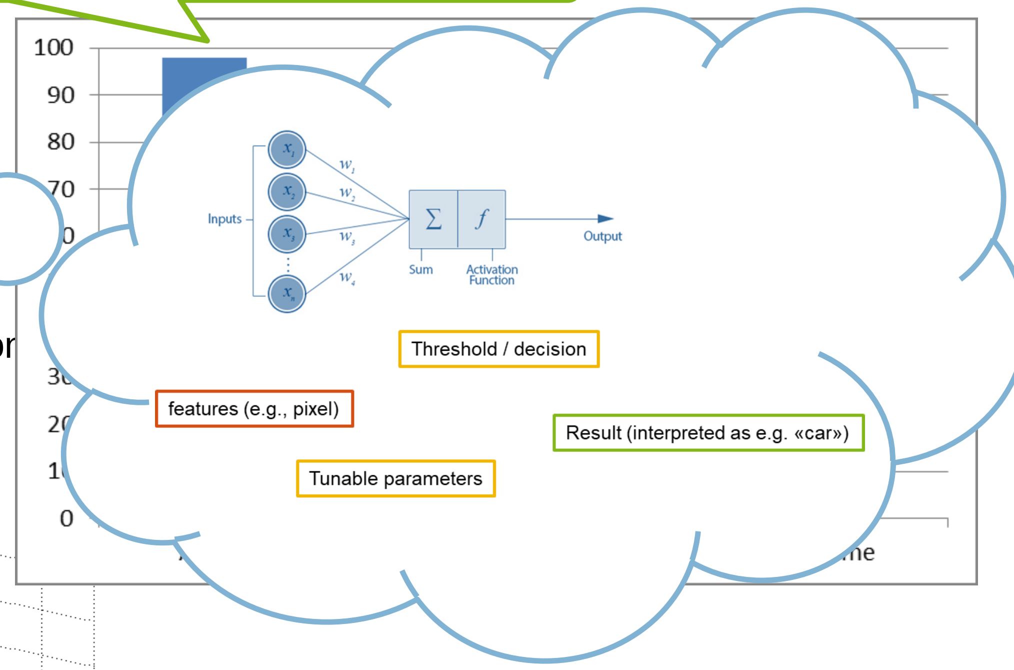
neural net





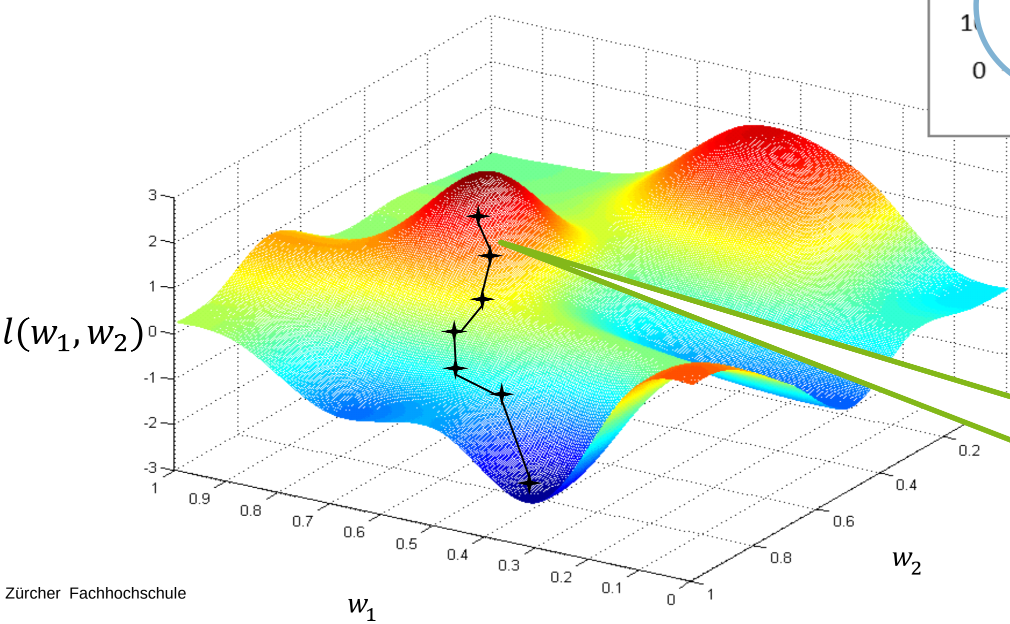
Search for optimal parameters of a function?

Probability [%] for specific event



- Our artificial neural net: $f_W(x) = y$
  with image x, ground truth y and parameters $\boldsymbol{W}$
  ($\boldsymbol{W} = \{w_1, w_2\}$ initialized at random)

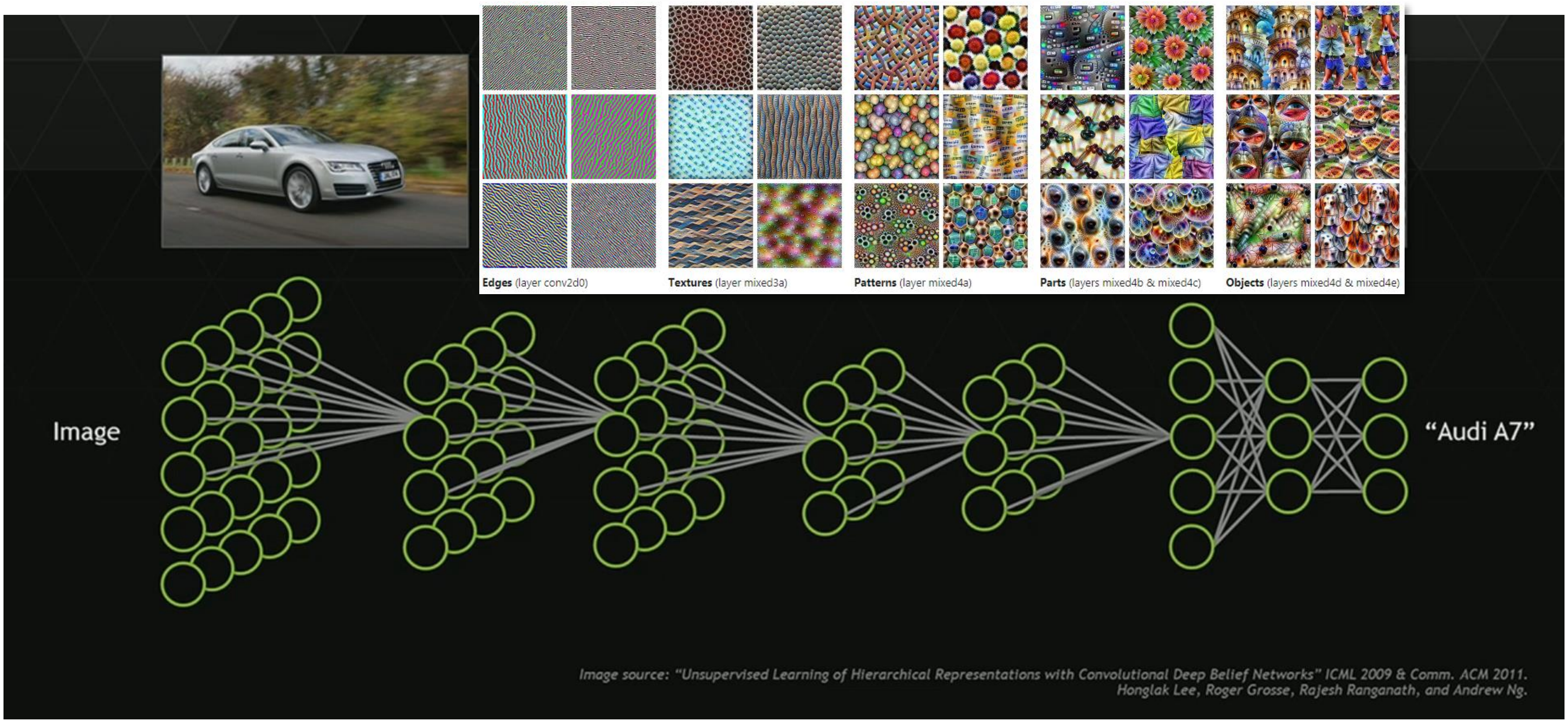
- Error measure: $l(W) = \frac{1}{N}\sum_{i=1}^{N}(f_W(x_i) - y_i)^2$
  Average of (quadratic) difference between prediction and ground truth («loss»)



🡐 error landscape

Method: iterative change of parameters of f in the direction of the steepest descent of J

What does the neural network «see»?
Hierarchy of more complex features



Source: https://www.pinterest.com/explore/artificial-neural-network/
Olah, et al., "Feature Visualization", Distill, 2017, https://distill.pub/2017/feature-visualization/.

Zürcher Hochschule
für Angewandte Wissenschaften

3

Examples of Deep Learning in the Wild

Zürcher Hochschule
für Angewandte Wissenschaften

Print media monitoring

ARGUS DATA INSIGHTS
WISSEN ZUM ERFOLG

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency

Task


International.

Spionage für den Erzfeind Iran

Israelischer Ex-Minister arbeitete als Agent für die Mullahs. Jetzt droht ihm lebenslänglich

Asylbewerber können bleiben

Klage von Le Pen abgelehnt

Nordkoreanischer Diktator zu Besuch in Peking

Challenge

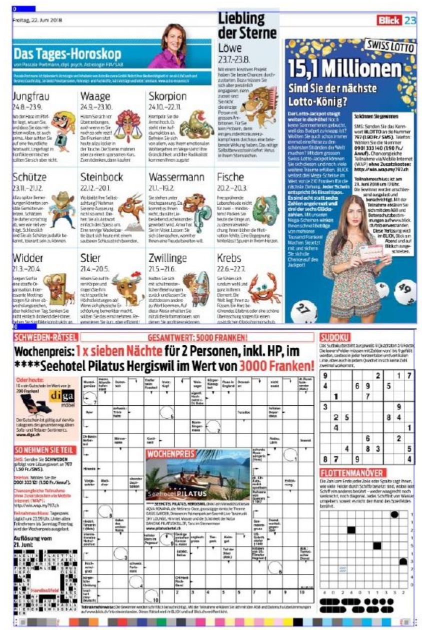

«Steven hat sich alles selber beigebracht»

Verlieren verboten!

Transfer Ticker

Liverpool will Yann Sommer!

Nuisance


Das Tages-Horoskop

Liebling der Sterne

15,1 Millionen

Sind Sie der nächste Lotto-König?

Jungfrau

Waage

Skorpion

Schütze

Steinbock

Wassermann

Fische

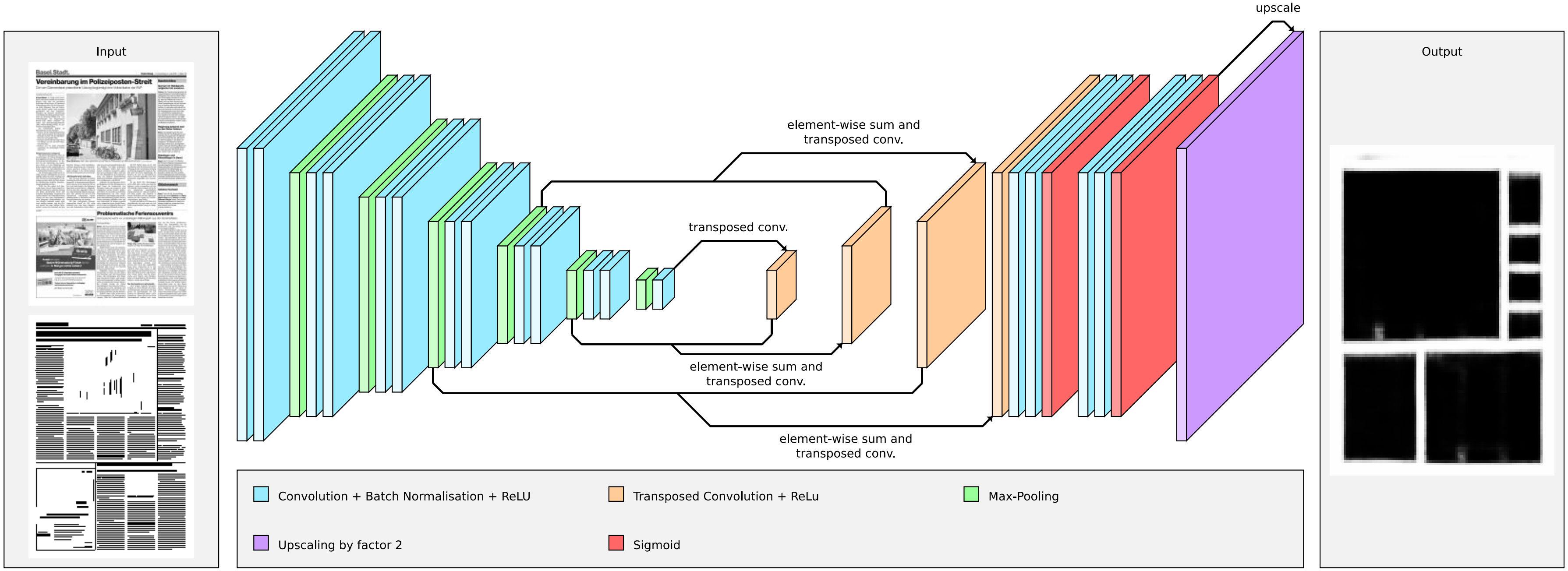
Widder

Stier

Zwillinge

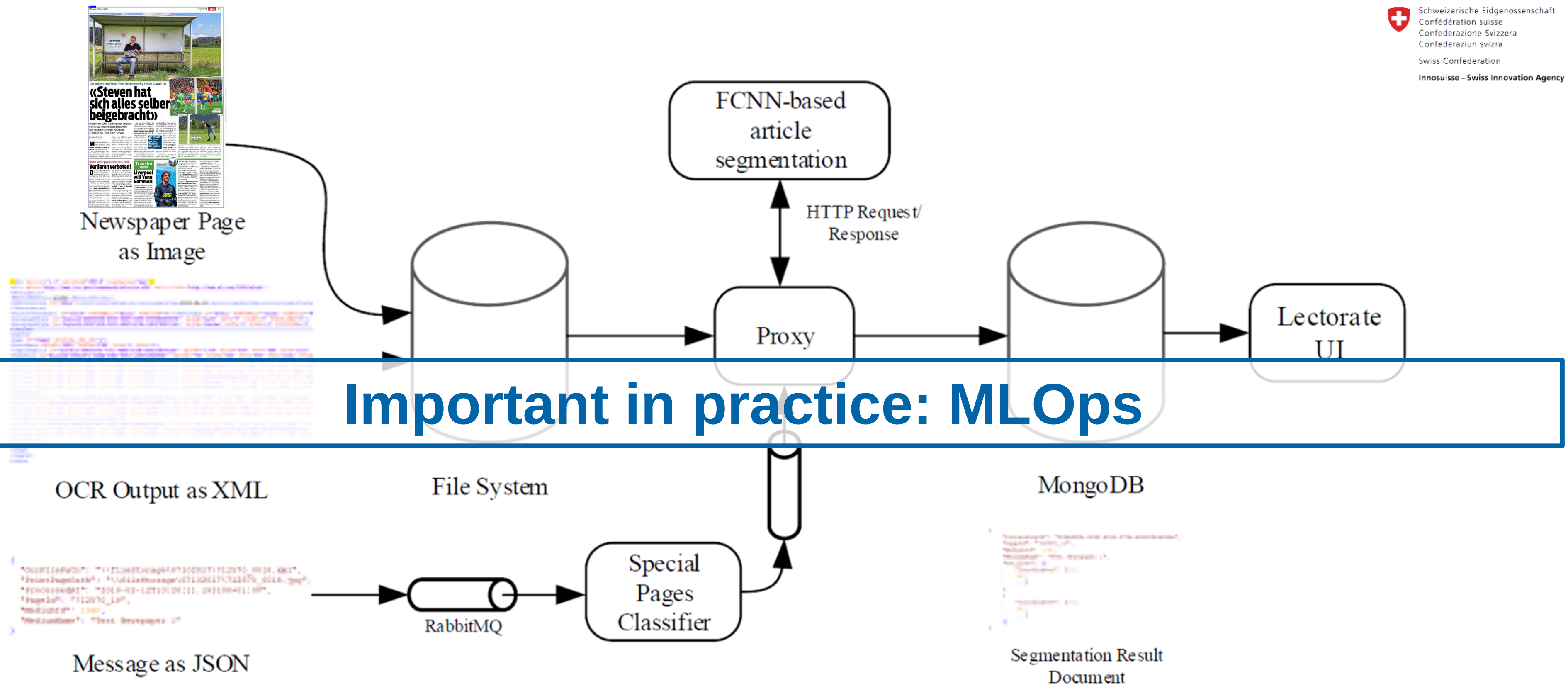
Krebs

Print media monitoring – ML solution



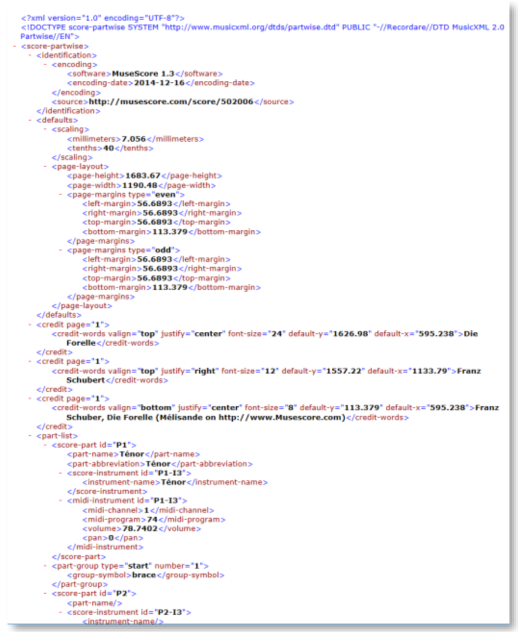
Meier, Stadelmann, Stampfli, Arnold & Cieliebak (2017). «*Fully Convolutional Neural Networks for Newspaper Article Segmentation*». ICDAR'2017.
Stadelmann, Tolkachev, Sick, Stampfli & Dürr (2018). «*Beyond ImageNet - Deep Learning in Industrial Practice*». In: Braschler et al., «Applied Data Science», Springer.

Print media monitoring – deployment



Stadelmann, Amirian, Arabaci, Arnold, Duivesteijn, Elezi, Geiger, Lörwald, Meier, Rombach & Tuggener (2018). *«Deep Learning in the Wild»*. ANNPR'2018.

Symbol detection



```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE score-partwise SYSTEM "http://www.musicxml.org/dtds/partwise.dtd" PUBLIC "-//Recordare//DTD MusicXML 2.0 Partwise//EN">
- <score-partwise>
  - <identification>
    - <encoding>
        <software>MuseScore 1.3</software>
        <encoding-date>2014-12-16</encoding-date>
      </encoding>
      <source>http://musescore.com/score/502006</source>
    </identification>
  - <defaults>
    - <scaling>
        <millimeters>7.056</millimeters>
        <tenths>40</tenths>
      </scaling>
    - <page-layout>
        <page-height>1683.67</page-height>
        <page-width>1190.48</page-width>
      - <page-margins type="even">
          <left-margin>56.6893</left-margin>
          <right-margin>56.6893</right-margin>
          <top-margin>56.6893</top-margin>
          <bottom-margin>113.379</bottom-margin>
        </page-margins>
      - <page-margins type="odd">
          <left-margin>56.6893</left-margin>
          <right-margin>56.6893</right-margin>
          <top-margin>56.6893</top-margin>
          <bottom-margin>113.379</bottom-margin>
        </page-margins>
      </page-layout>
    </defaults>
  - <credit page="1">
      <credit-words valign="top" justify="center" font-size="24" default-y="1626.98" default-x="595.238">Die Forelle</credit-words>
    </credit>
  - <credit page="1">
      <credit-words valign="top" justify="right" font-size="12" default-y="1557.22" default-x="1133.79">Franz Schubert</credit-words>
    </credit>
  - <credit page="1">
      <credit-words valign="bottom" justify="center" font-size="8" default-y="113.379" default-x="595.238">Franz Schuber, Die Forelle (Mélisande on http://www.Musescore.com)</credit-words>
    </credit>
  - <part-list>
    - <score-part id="P1">
        <part-name>Ténor</part-name>
        <part-abbreviation>Ténor</part-abbreviation>
      - <score-instrument id="P1-I3">
          <instrument-name>Ténor</instrument-name>
        </score-instrument>
      - <midi-instrument id="P1-I3">
          <midi-channel>1</midi-channel>
          <midi-program>74</midi-program>
          <volume>78.7402</volume>
          <pan>0</pan>
        </midi-instrument>
      </score-part>
    - <part-group type="start" number="1">
        <group-symbol>brace</group-symbol>
      </part-group>
    - <score-part id="P2">
        <part-name/>
      - <score-instrument id="P2-I3">
          <instrument-name/>
```

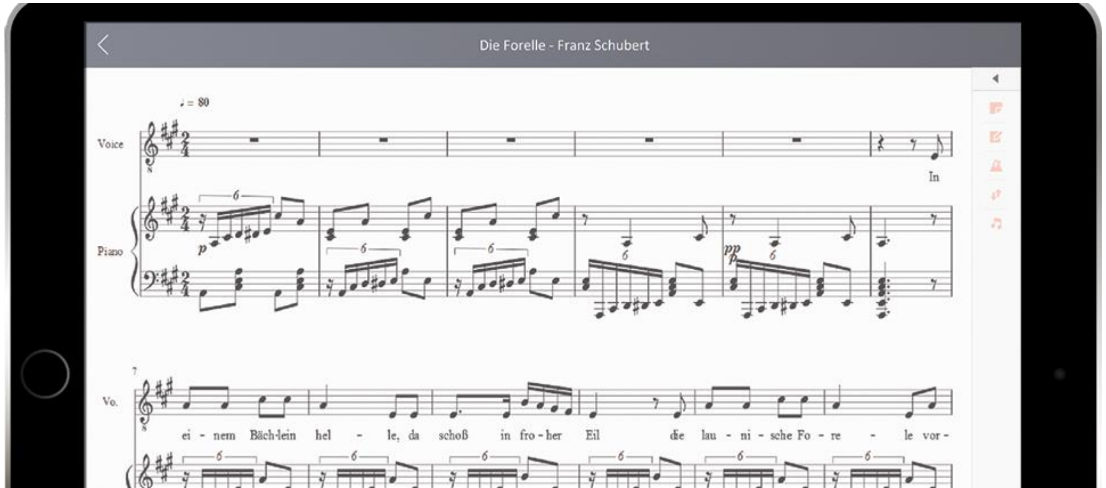

Zürcher Hochschule für Angewandte Wissenschaften

SCOREPAD

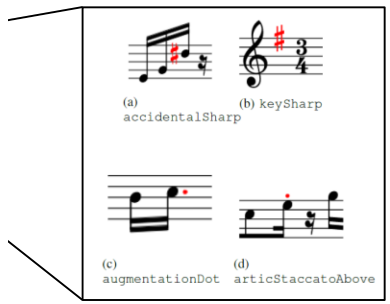
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency



Symbol detection – challenges & solutions

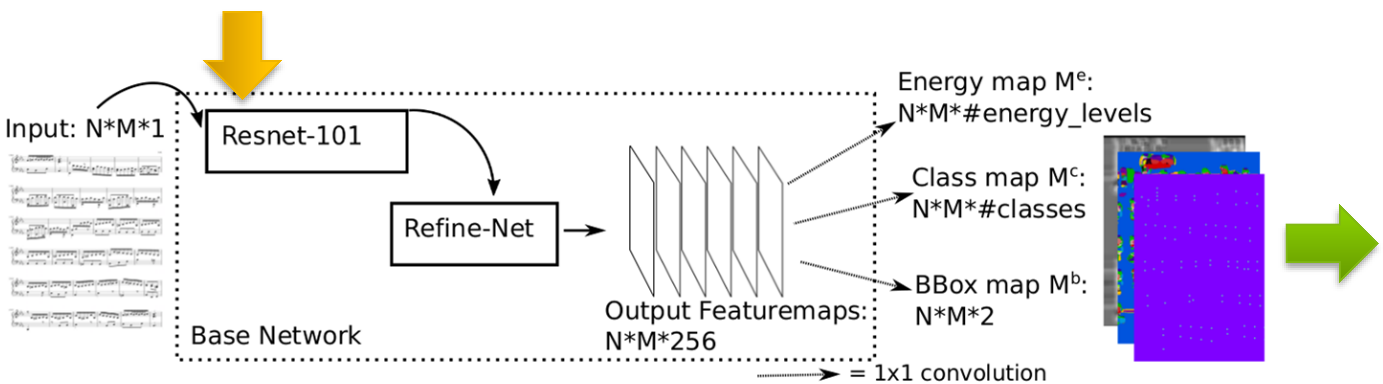


SCOREPAD

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency



a) Example result from *DeepScores* with detected bounding boxes as overlays. The tiny numbers are class labels from the dataset introduced with the overlay. This system is roughly one forth of the size of a typical *DeepScores* input we process at once.

b) Example result from *MUSCIMA++* with detected bounding boxes and class labels as overlays. This system is roughly one half of the size of a typical processed *MUSCIMA++* input. The images are random picks amongst inputs with many symbols.

Tuggener, Elezi, Schmidhuber, Pelillo & Stadelmann (2018). *«DeepScores – A Dataset for Segmentation, Detection and Classification of Tiny Objects»*. ICPR'2018.
Tuggener, Elezi, Schmidhuber & Stadelmann (2018). *«Deep Watershed Detector for Music Object Recognition»*. ISMIR'2018.

Symbol detection – industrialization

Zürcher Hochschule
für Angewandte Wissenschaften

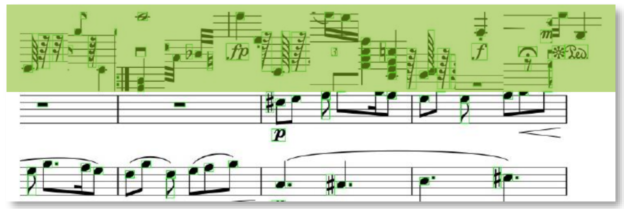
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency

Current results on class imbalance and robustness challenges

1. Added sophisticated **data augmentation** in every page's margins



2. Put additional effort (and compute) into hyperparameter **tuning** and **longer training**
3. Trained also on scanned (more **real-worldish**) scores

Sufficient condition: lots of tuning

➔ **Improved** our **mAP** from 16% (on purely synthetic data) **to 73%** on more challenging real-world data set (additionally, using Pacha et al.'s evaluation method as a 2nd benchmark: SotA from 24.8% to 47.5%)

Elezi, Tuggener, Pelillo & Stadelmann (2018). *«DeepScores and Deep Watershed Detection: current state and open issues»*. WoRMS @ ISMIR'2018.
Pacha, Hajic, Calvo-Zaragoza (2018). *«A Baseline for General Music Object Detection with Deep Learning»*. Appl. Sci. 2018, 8, 1488, MDPI.

Lessons learned

Data is key

- Many real-world projects miss the required **quantity & quality** of data
  → even though «big data» is not needed
- **Class imbalance** needs careful dealing
  → special loss, resampling (also in unorthodox ways)
- **Unsupervised** methods need to be used creatively
- Users & label providers need to be **trained**

Prerequisite: stable data acquisition pipeline
Learning from (raw) data is powerful, yet one is fully dependent on what is in that data
Important in practice: MLOps

Robustness is important

- **Training processes** can be tricky
  → give hints via a unique loss, proper preprocessing and pretraining

Sufficient condition: lots of tuning
Deep learning is no silver bullet

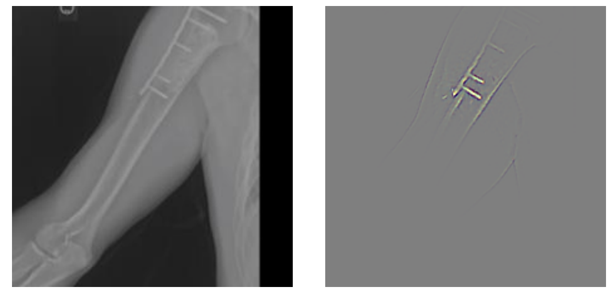
Lessons learned – model interpretability

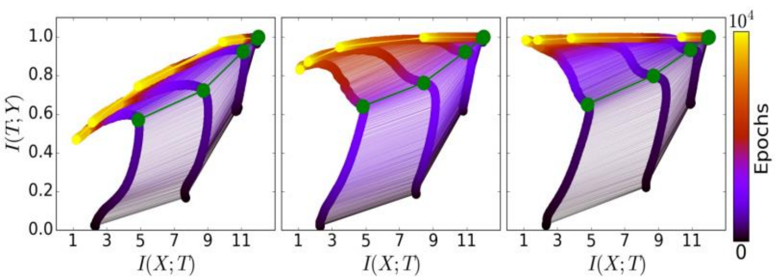
Interpretability is required.

- Helps the developer in «debugging», needed by the user to trust
  → visualizations of learned features, training process, learning curves etc. should be «always on»

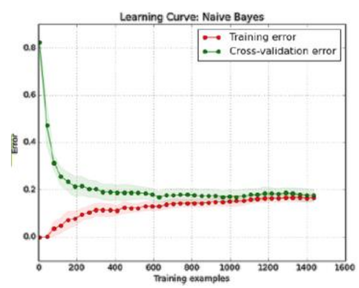
negative X-ray

positive X-ray

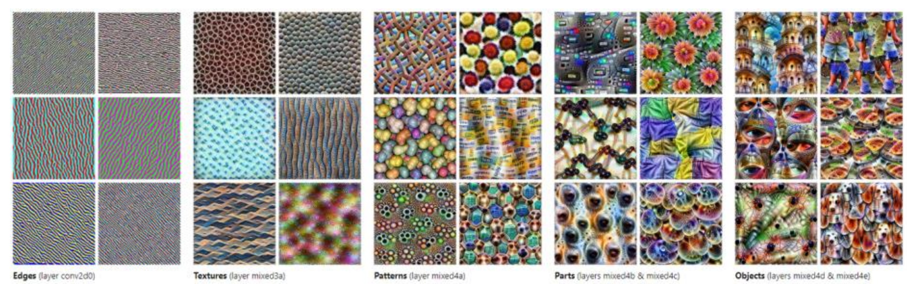




DNN training on the Information Plane



a learning curve



feature visualization

Stadelmann, Amirian, Arabaci, Arnold, Duivesteijn, Elezi, Geiger, Lörwald, Meier, Rombach & Tuggener (2018). *«Deep Learning in the Wild»*. ANNPR'2018.
Schwartz-Ziv & Tishby (2017). *«Opening the Black Box of Deep Neural Networks via Information»*.
https://distill.pub/2017/feature-visualization/, https://stanfordmlgroup.github.io/competitions/mura/

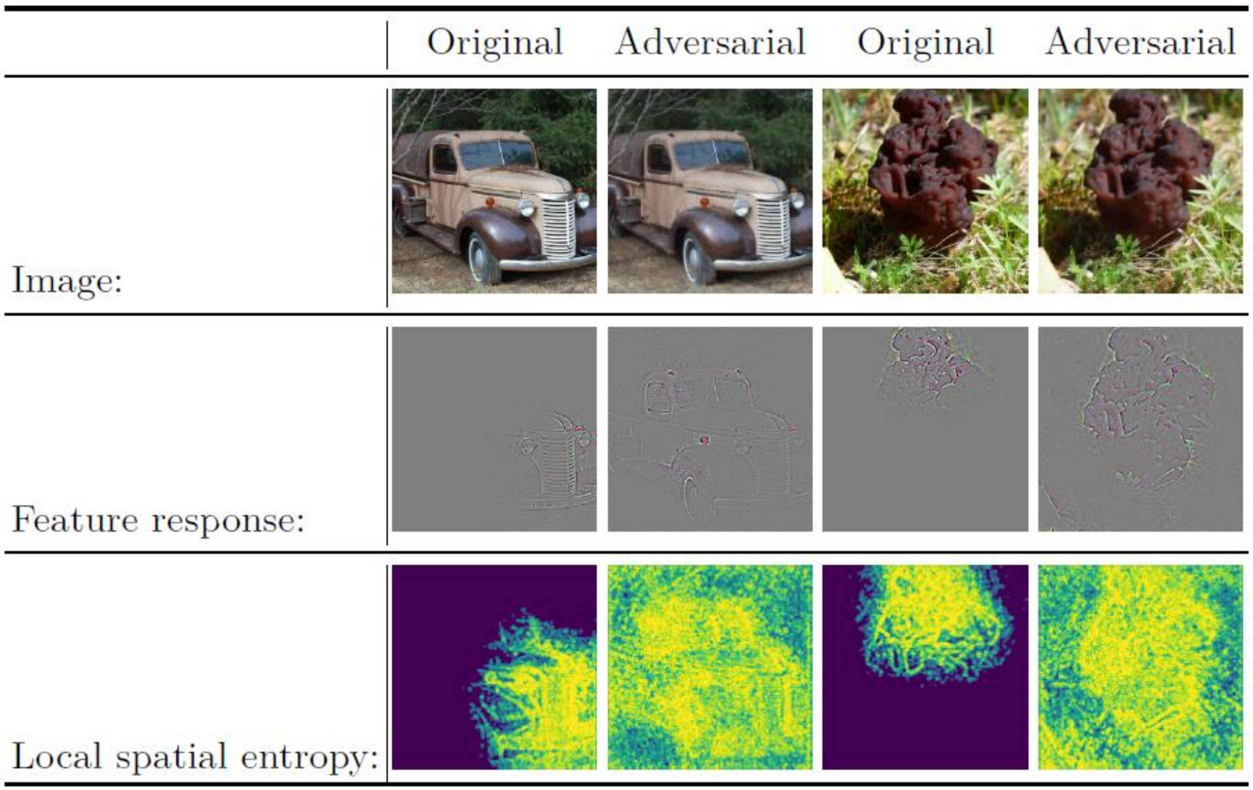
Zürcher Hochschule
für Angewandte Wissenschaften

Lessons learned – detecting adversarial attacks
…using average local spatial entropy of feature response maps

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency

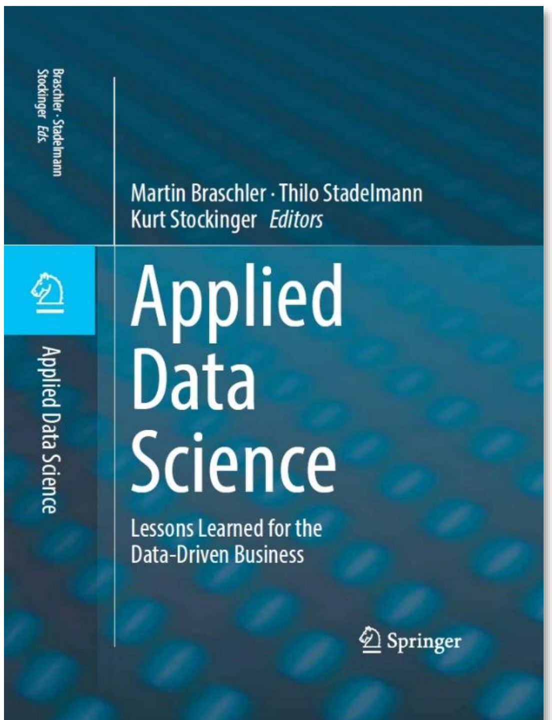


Amirian, Schwenker & Stadelmann (2018). *«Trace and Detect Adversarial Attacks on CNNs using Feature Response Maps»*. ANNPR’2018.

Zürcher Hochschule
für Angewandte Wissenschaften

Conclusions

- Deep learning **is applied** and deployed in «normal» businesses (non-AI, SME)
- It does not need big-, but some **data** (**effort usually underestimated**)
- DL **training** for new use cases **can be tricky** (→ needs thorough experimentation)
- New **theory and visualizations** help to debug & understand
  → *the training process*
  → *individual results*



About me:

- Prof. AI/ML, scientific director ZHAW digital
- Email: stdm@zhaw.ch
- Phone: +41 58 934 72 08
- Web: https://stdm.github.io/
- Twitter: @thilo_on_data
- LinkedIn: thilo-Stadelmann

Further contacts:

- Collaboration: datalab@zhaw.ch

→ Happy to answer questions & requests.

datalab
www.zhaw.ch/datalab

data service
Swiss Alliance for Data-Intensive Services