Zürcher Hochschule
für Angewandte Wissenschaften

Machine Learning for Pattern Recognition @ ZHAW

IDIAP Research Institute, Martigny, July 02, 2018

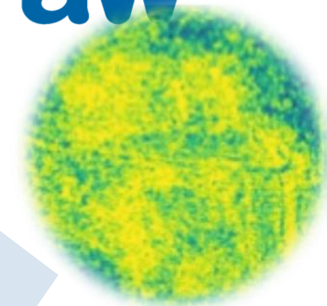
Thilo Stadelmann

data service
Swiss Alliance for
Data-Intensive Services

swiss group for artificial intelligence
and cognitive science
SGAICO

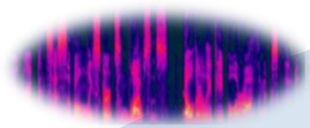
datalab
www.zhaw.ch/datalab

Agenda

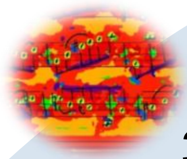


6. Lessons Learned

5. Learning to Cluster



4. Speaker recognition



3. Music scanning

2. Face matching

1. The group

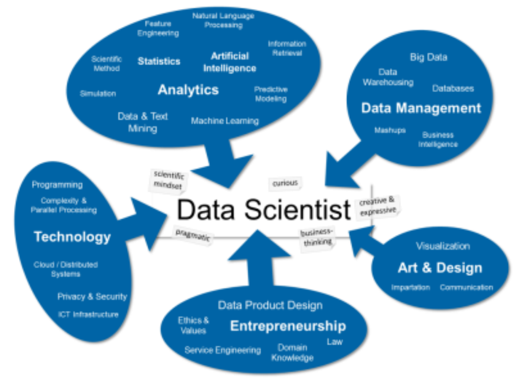
1. ZHAW Zurich University of Applied Sciences, School of Engineering

Switzerland's biggest fully-featured university of applied sciences

- >10'000 students
- >2'600 employees
- >1'000 (associate) professors

School of Engineering emanates from «Technikum Winterthur» (est. 1874)

1. ZHAW Datalab: Est. 2013



Forerunner

- **One of the first** interdisciplinary data science initiatives in Europe
- One of the first interdisciplinary centers at ZHAW

Foundation

- **People**: ca. **70 researchers** from **5 institutes** / **3 departments** opted in
- Vision: **Nationally leading** and **internationally recognized** center of excellence
- Mission: **Generate projects** through critical mass and mutual relationships
- Competency: **Data product design** with structured and unstructured data

Success factors

- **Lean** organization and operation → geared towards projects
- Years of successful **pre-Datalab collaboration**

1. ZHAW Datalab: R&D

Volume

- > 9 Mio. CHF 3rd party funding in first 4 years
- **Overall turnover** of projects up to spring 2017: **> 19.5 Mio**. CHF in **< 4 years**

Spin-offs

- Prognosix – a ZHAW IAS spin-off
- SPINNINGBYTES – a joint spin-off from ZHAW and ETH

Topics: all of digitization

- Industry 4.0 (e.g., CTI project «QualitAI»)
- E-Health (e.g. SystemsX/SNSF project «AneuX»)
- FinTech (e.g., CTI project «DatFrisMo»)
- Mobility (e.g., project «Placebook»)
- Sustainability (e.g., CTI project «EAT-IT CO_2»)
- Technology (e.g., SNSF project «Bio-SODA»)
- …

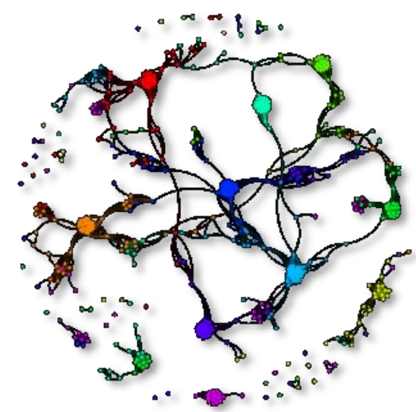


Figure: Visualizing the relationships of all Swiss foundations, based on the similarity of goals as expressed in their statutes. A proud collaboration of InIT and IDP within CTI project «Stiftungsregister SR 2.0»

Swiss Confederation
Commission for Technology and Innovation CTI

1. ML @ Information Engineering Group

Institute of Applied Information Technology, School of Engineering





Machine learning-based Pattern Recognition

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

1. ML @ Information Engineering Group

Institute of Applied Information Technology, School of Engineering

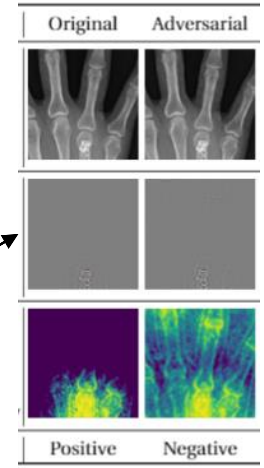
Zürcher Hochschule
für Angewandte Wissenschaften

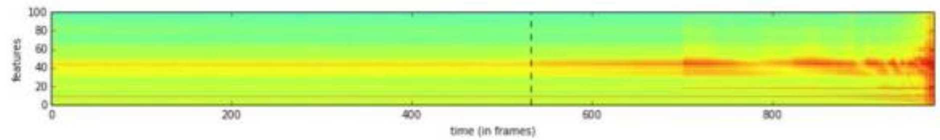




Machine learning-based
Pattern Recognition

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





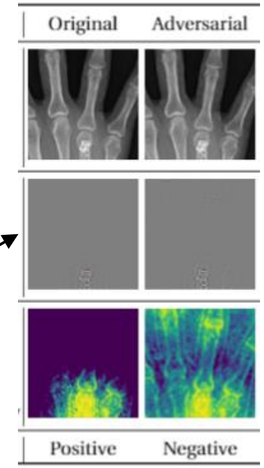
1. ML @ Information Engineering Group

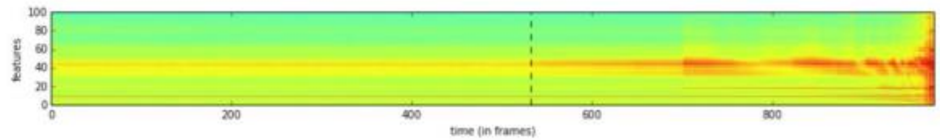
Institute of Applied Information Technology, School of Engineering

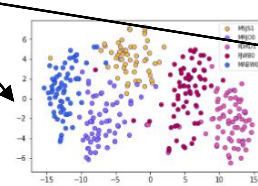
Zürcher Hochschule
für Angewandte Wissenschaften

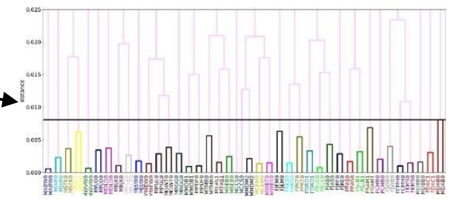
Machine learning-based
Pattern Recognition

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









Zürcher Hochschule
für Angewandte Wissenschaften

1. ML @ Information Engineering Group

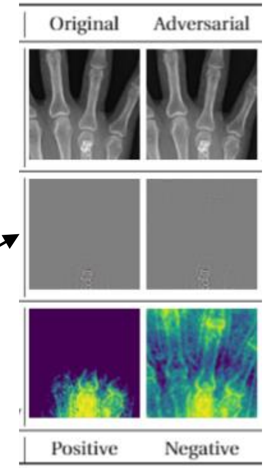
Institute of Applied Information Technology, School of Engineering

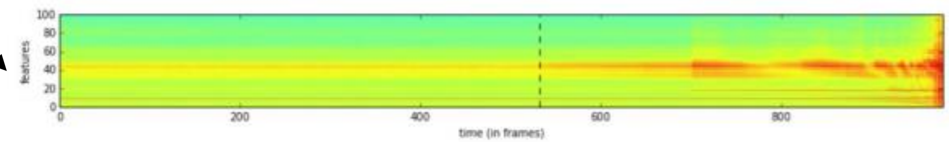


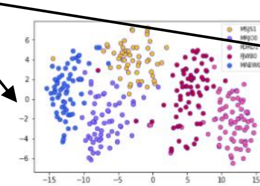


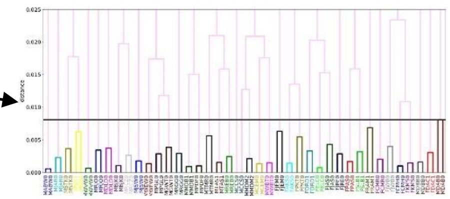
Machine learning-based Pattern Recognition

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





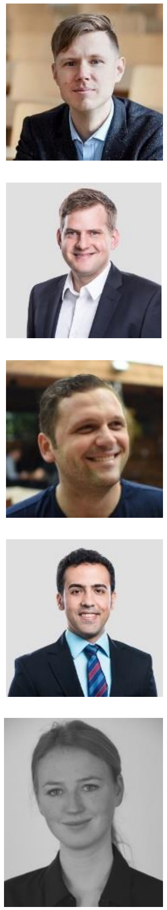




1. ML @ Information Engineering Group

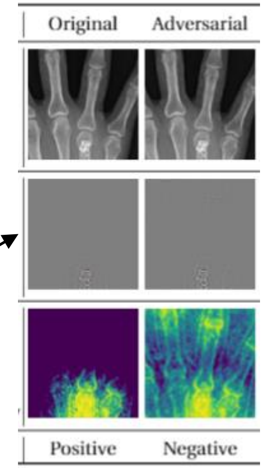
Institute of Applied Information Technology, School of Engineering

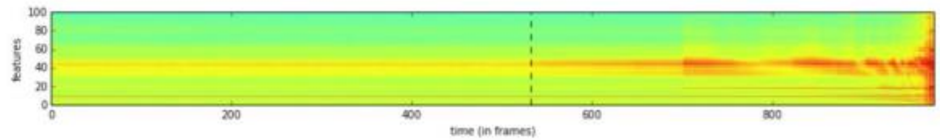
Zürcher Hochschule
für Angewandte Wissenschaften

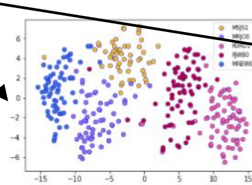


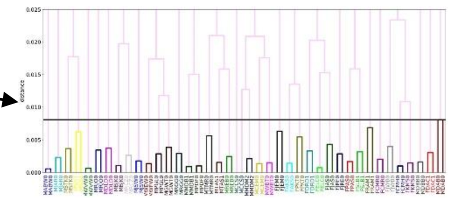
Machine learning-based
Pattern Recognition

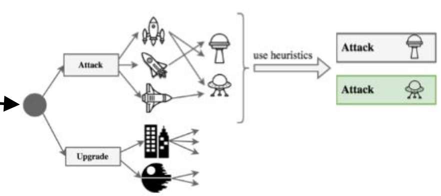
- Robust Deep Learning
- Voice Recognition
- Document Analysis
- Learning to Learn & Control

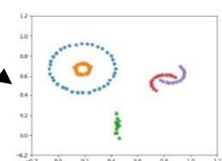


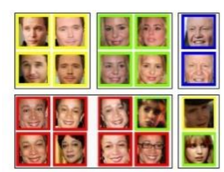












2. Face matching

2. Face matching

2. Face matching – challenges & solutions

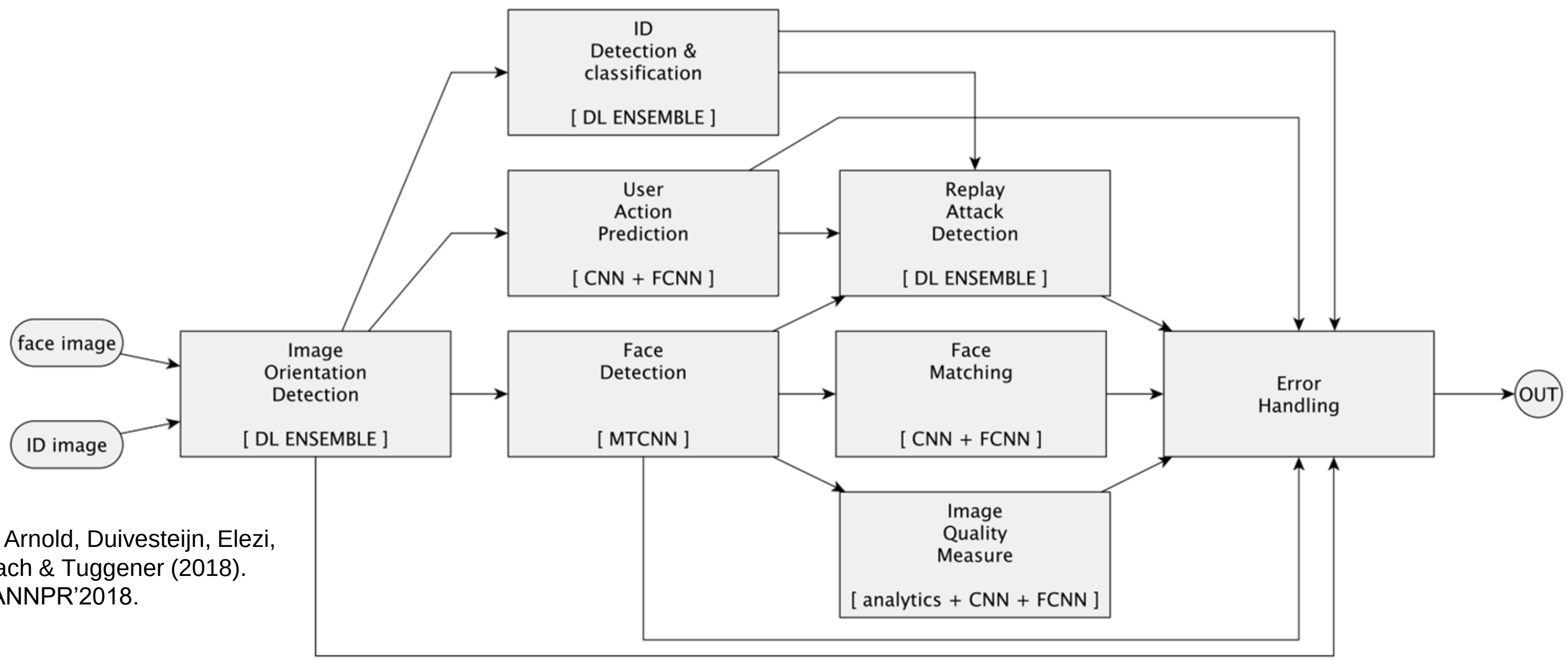
Zürcher Hochschule
für Angewandte Wissenschaften

[!] DEEPIMPACT

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency



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

2. Face matching – challenges & solutions

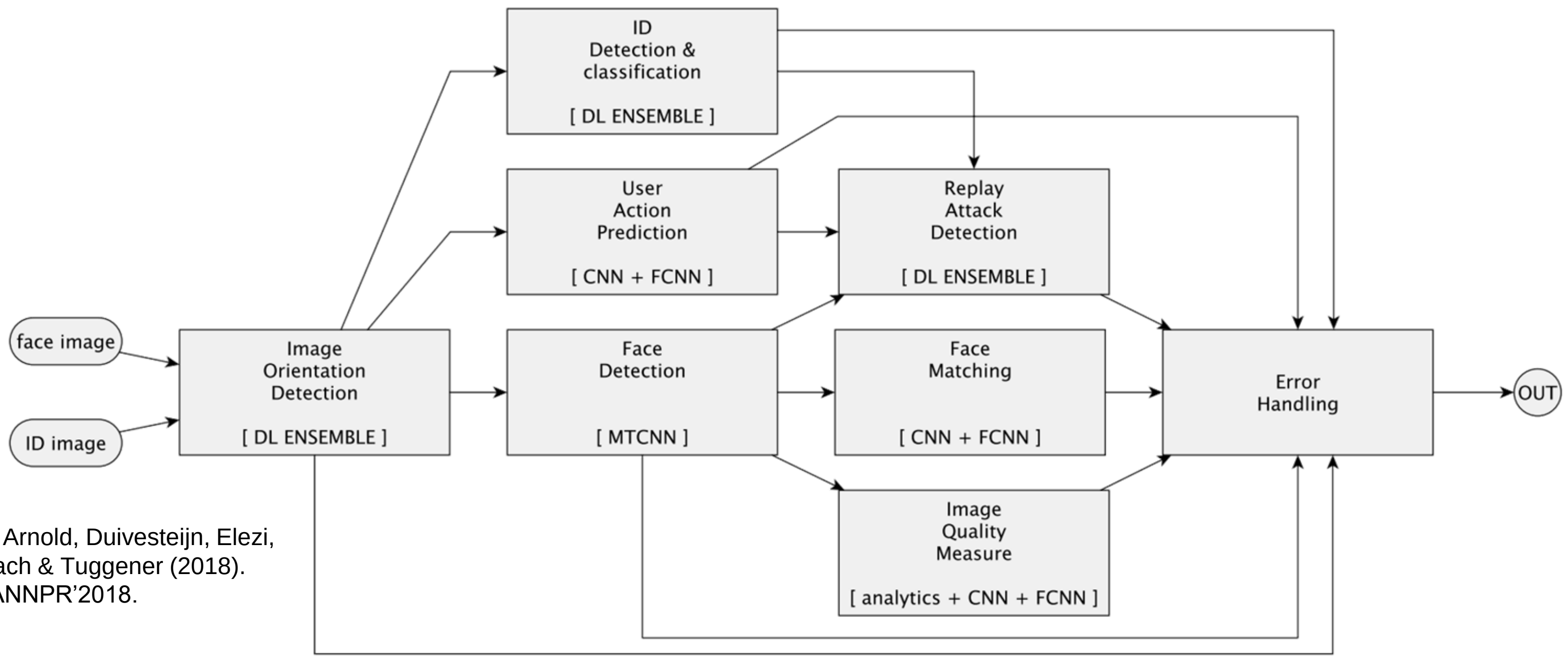
Zürcher Hochschule
für Angewandte Wissenschaften

[!] DEEPIMPACT

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

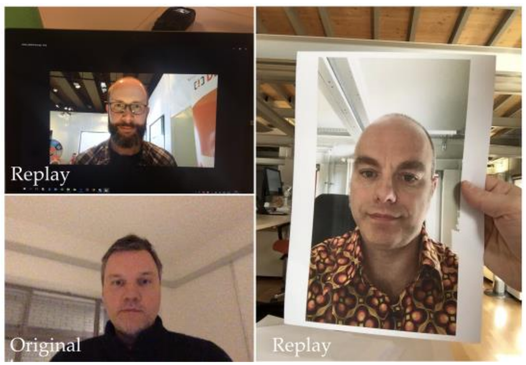
Innosuisse – Swiss Innovation Agency

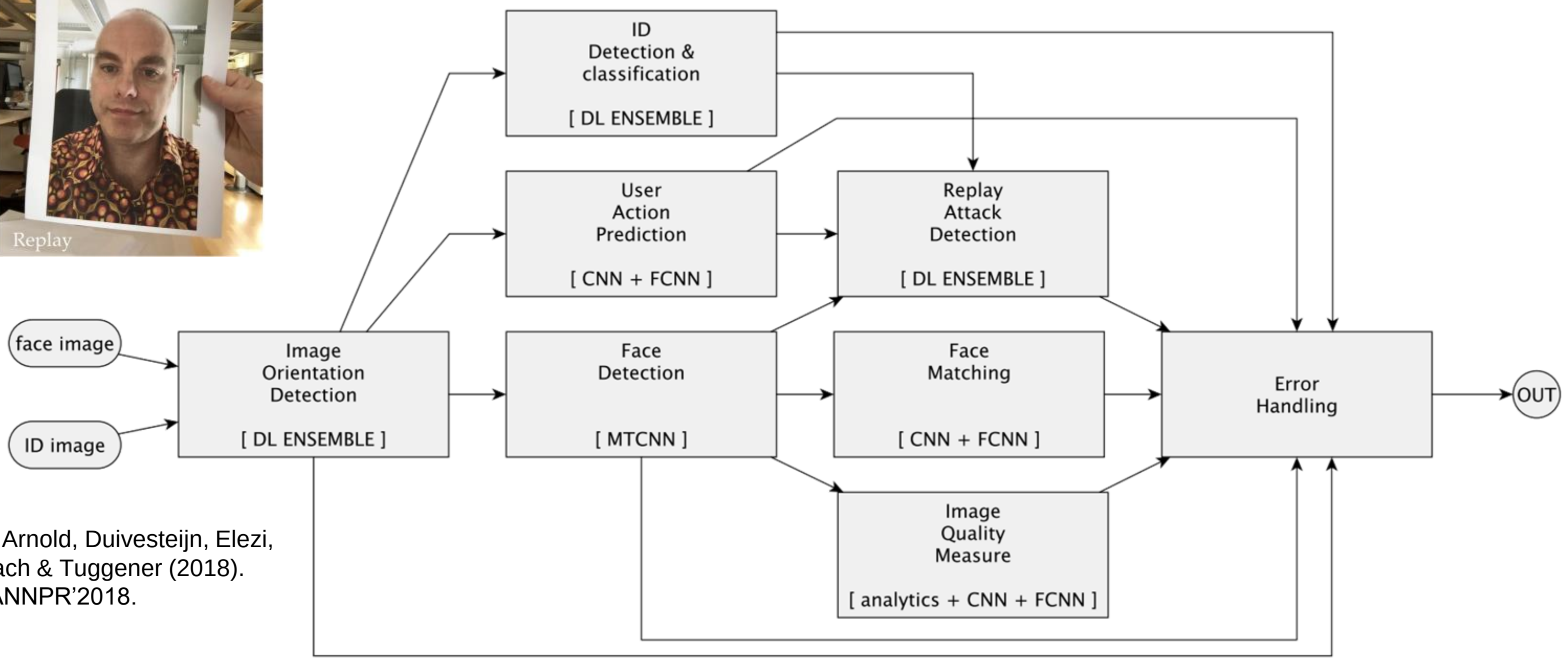


Stadelmann, Amirian, Arabaci, Arnold, Duivesteijn, Elezi, Geiger, Lörwald, Meier, Rombach & Tuggener (2018). *«Deep Learning in the Wild»*. ANNPR'2018.

2. Face matching – challenges & solutions

[!] DEEPIMPACT
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Innosuisse – Swiss Innovation Agency





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

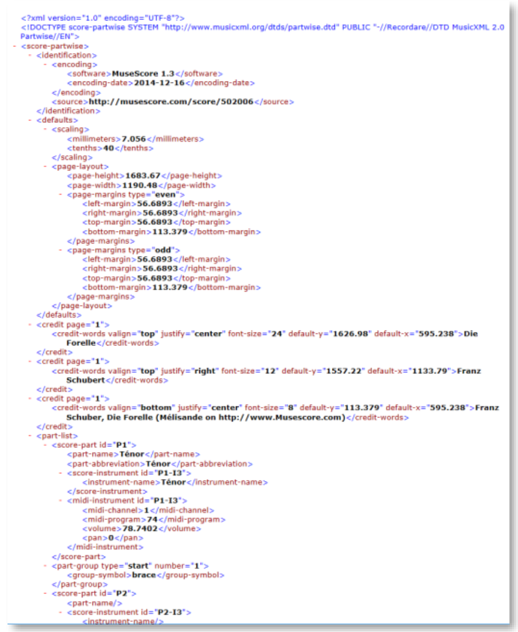
3. Music scanning

Zürcher Hochschule
für Angewandte Wissenschaften

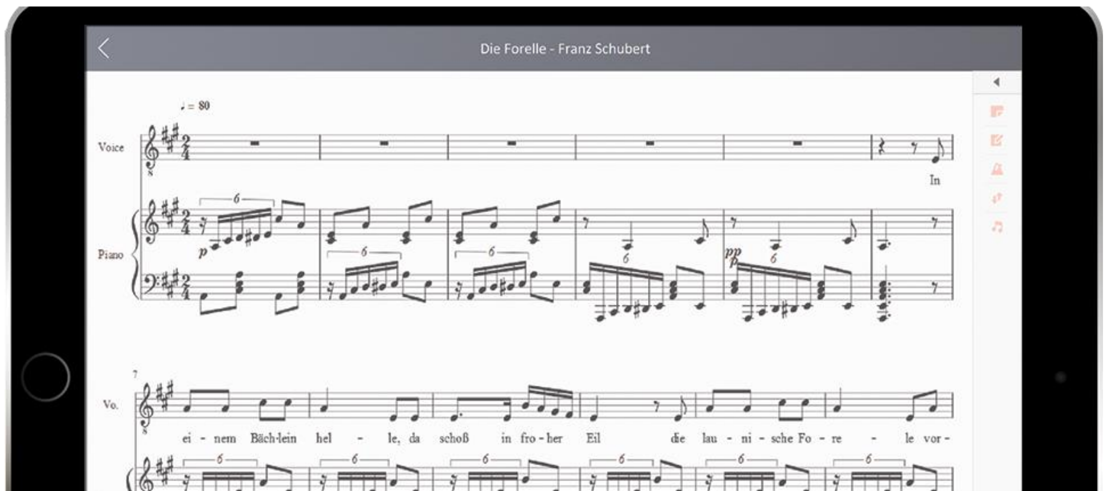
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency



```
<?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/>
```

3. Music scanning – challenges & solutions

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency

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.

3. Music scanning – challenges & solutions

SCOREPAD

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency

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.

3. Music scanning – challenges & solutions

,

SCOREPAD

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency

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.

3. Music scanning – challenges & solutions

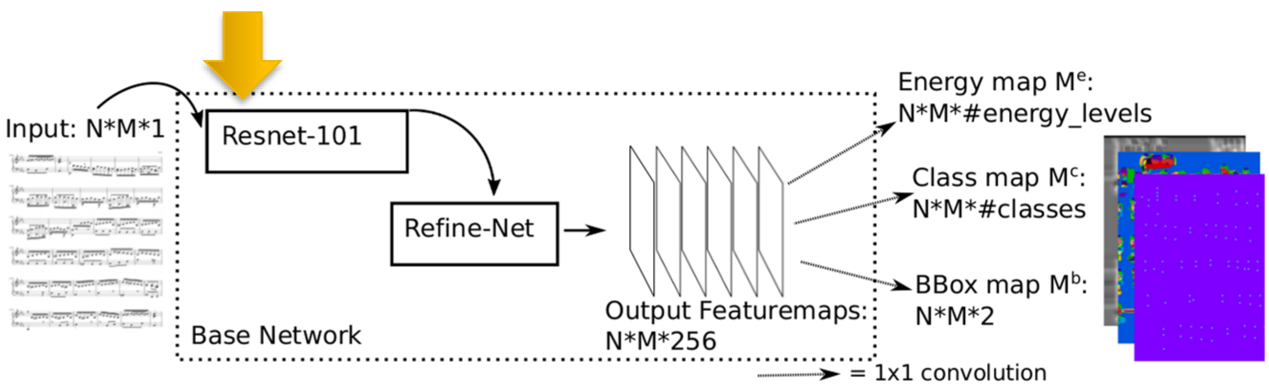
,

SCOREPAD

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency



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.

3. Music scanning – challenges & solutions

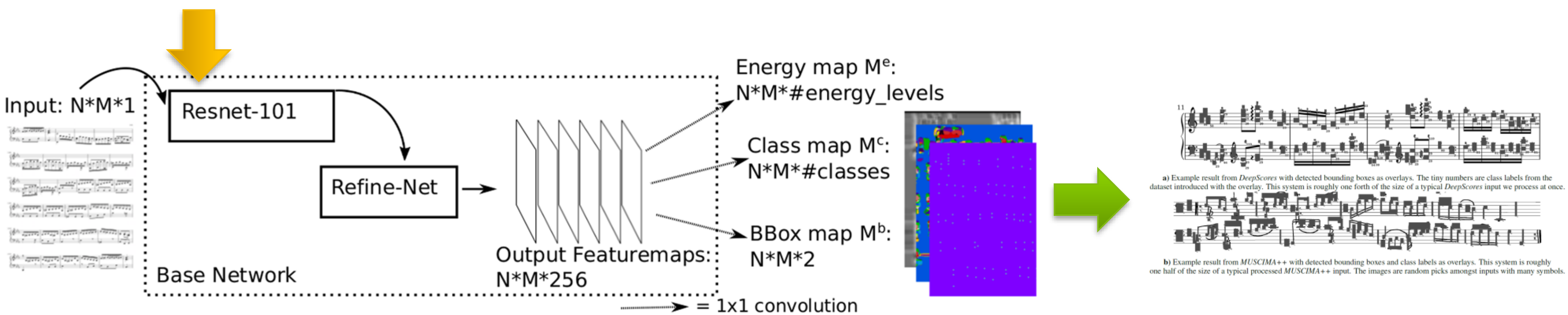
,

SCOREPAD

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

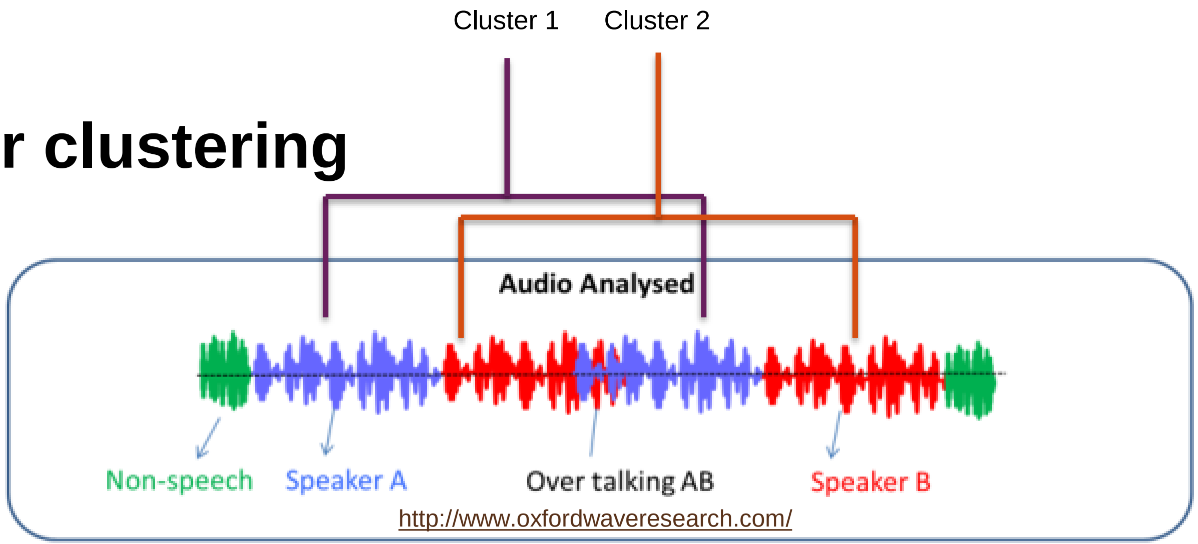
Swiss Confederation

Innosuisse – Swiss Innovation Agency



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.

4. Speaker clustering



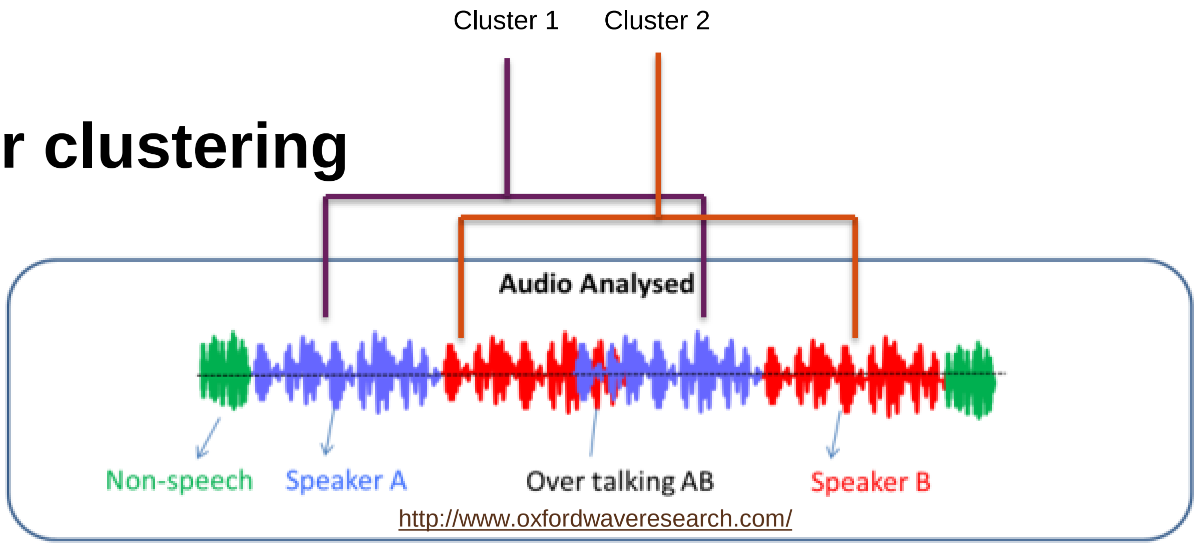
For the 630 training utterances, GMMs with 32 mixtures are built a priori, then an identification experiment is run for the 630 test utterances. It yields a satisfactory 0.5% closed set identification error.

[34]. Evaluations typically concentrate on data sets built from broadcast news/shows and meeting recordings, where diarization error rates ranging from 8% to 24% are reported [28][34][45]. These results are confirmed by more recent

The hypothesis of this paper is: the techniques originally developed for speaker verification and identification are not suitable for speaker clustering, taking into account the escalated difficulty of the latter task. However, the processing chain for speaker clustering is quite large – there are many potential areas for improvement. The question is: *where should improvements be made to improve the final result?*

Stadelmann & Freisleben (2009). *«Unfolding Speaker Clustering Potential: A Biomimetic Approach»*. ACMMM'2009.

4. Speaker clustering



For the 630 training utterances, GMMs with 32 mixtures are built a priori, then an identification experiment is run for the 630 test utterances. It yields a satisfactory 0.5% closed set identification error.

[34]. Evaluations typically concentrate on data sets built from broadcast news/shows and meeting recordings, where diarization error rates ranging from 8% to 24% are reported [28][34][45]. These results are confirmed by more recent

The hypothesis of this paper is: the techniques originally developed for speaker verification and identification are not suitable for speaker clustering, taking into account the escalated difficulty of the latter task. However, the processing chain for speaker clustering is quite large – there are many potential areas for improvement. The question is: *where* should improvements be made to improve the *final* result?

The interpretation of our results has shown that it is the stage of modeling that bears the highest potential: the inclusion of temporal context information among feature vectors is what is crucially missing there. Furthermore, the inclusion

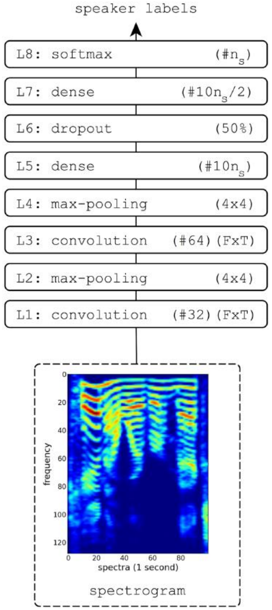
context vector. This corresponds to a syllable length of 130 ms and is found to best capture speaker specific sounds in informal listening experiments over a range of 32–496 ms (in intervals of 16 ms). Our context vector step is one orig-

Stadelmann & Freisleben (2009). *«Unfolding Speaker Clustering Potential: A Biomimetic Approach»*. ACMMM'2009.

4. Speaker clustering – exploiting time information

Zürcher Hochschule
für Angewandte Wissenschaften

CNN (MLSP'16)

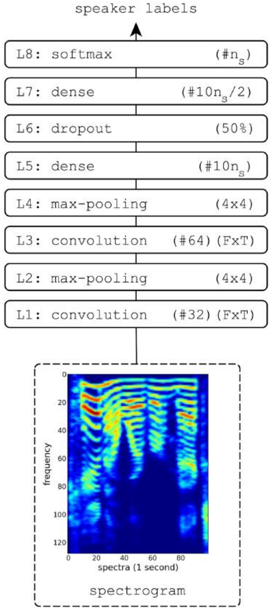


Method	MR	MR (legacy)
RNN /w PKLD	**2.19%** $\left(\frac{1.25\%+2.5\%+1.25\%+3.75\%}{4}\right)$	**4.38%** (average of 4 runs)
CNN /w PKLD [24]	-	**5%**
CNN /w cross entropy [23]	-	**5%**
ν-SVM [40]	6.25%	-
GMM/MFCC [40]	12.5%	-

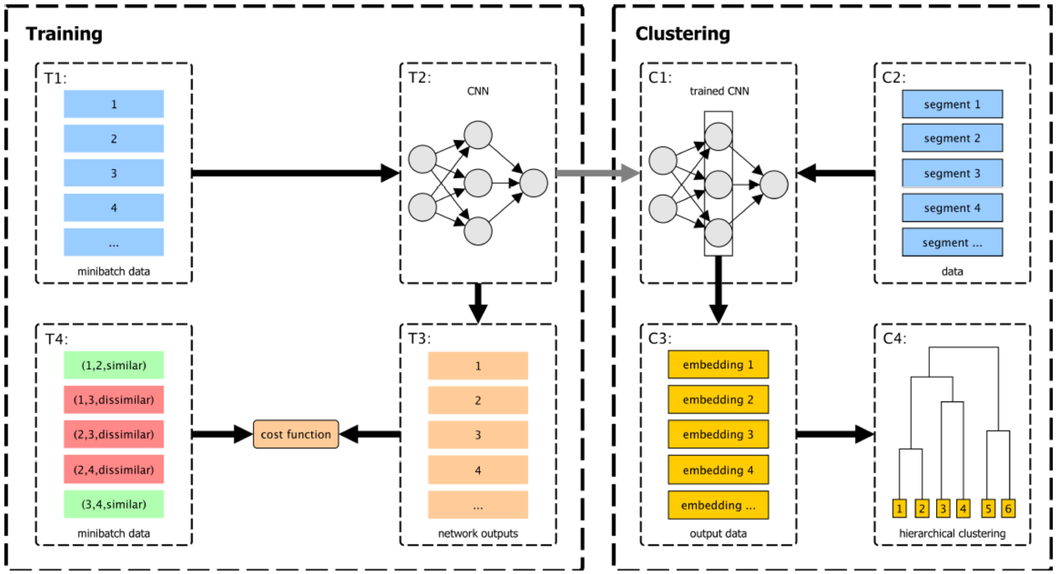
Lukic, Vogt, Dürr & Stadelmann (2016). «Speaker Identification and Clustering using Convolutional Neural Networks». MLSP'2016.
Lukic, Vogt, Dürr & Stadelmann (2017). «Learning Embeddings for Speaker Clustering based on Voice Equality». MLSP'2017.
Stadelmann, Glinski-Haefeli, Gerber & Dürr (2018). «Capturing Suprasegmental Features of a Voice with RNNs for Improved Speaker Clustering». ANNPR'2018.

4. Speaker clustering – exploiting time information

CNN (MLSP'16)



CNN & clustering-loss (MLSP'17)



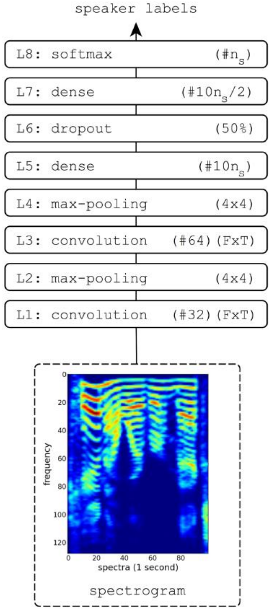
Method	MR	MR (legacy)
RNN /w PKLD	**2.19%** ($\frac{1.25\%+2.5\%+1.25\%+3.75\%}{4}$)	4.38% (average of 4 runs)
CNN /w PKLD [24]	-	**5%**
CNN /w cross entropy [23]	-	**5%**
ν-SVM [40]	6.25%	-
GMM/MFCC [40]	12.5%	-

Lukic, Vogt, Dürr & Stadelmann (2016). «Speaker Identification and Clustering using Convolutional Neural Networks». MLSP'2016.
Lukic, Vogt, Dürr & Stadelmann (2017). «Learning Embeddings for Speaker Clustering based on Voice Equality». MLSP'2017.
Stadelmann, Glinski-Haefeli, Gerber & Dürr (2018). «Capturing Suprasegmental Features of a Voice with RNNs for Improved Speaker Clustering». ANNPR'2018.

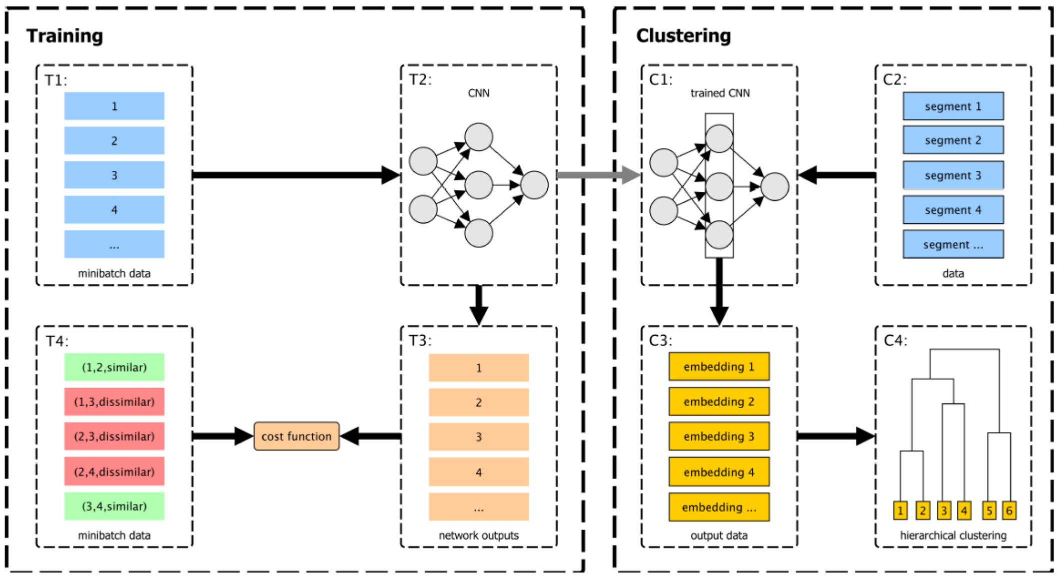
Zürcher Hochschule
für Angewandte Wissenschaften

4. Speaker clustering – exploiting time information

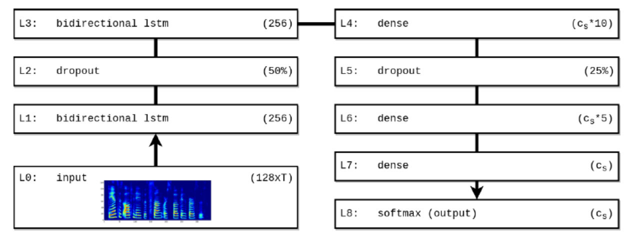
CNN (MLSP'16)



CNN & clustering-loss (MLSP'17)



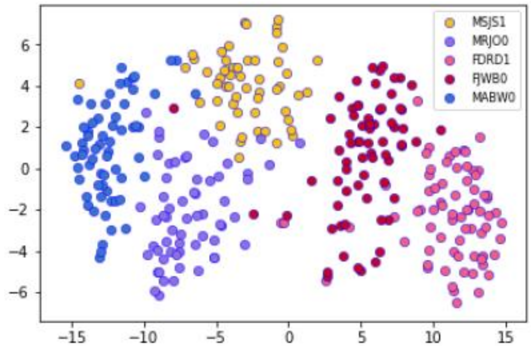
RNN & clustering-loss (ANNPR'18)

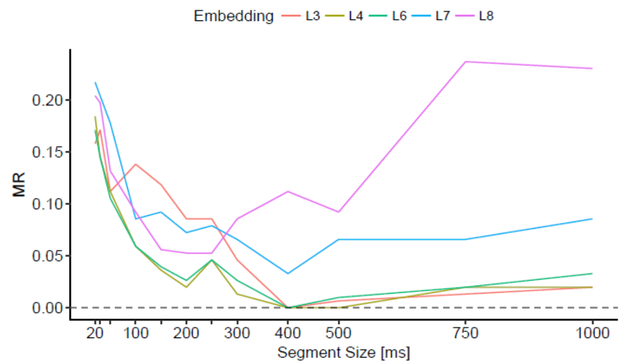


Method	MR	MR (legacy)
RNN /w PKLD	**2.19%** $\left(\frac{1.25\%+2.5\%+1.25\%+3.75\%}{4}\right)$	**4.38%** (average of 4 runs)
CNN /w PKLD [24]	-	**5%**
CNN /w cross entropy [23]	-	**5%**
ν-SVM [40]	6.25%	-
GMM/MFCC [40]	12.5%	-

Lukic, Vogt, Dürr & Stadelmann (2016). «Speaker Identification and Clustering using Convolutional Neural Networks». MLSP'2016.
Lukic, Vogt, Dürr & Stadelmann (2017). «Learning Embeddings for Speaker Clustering based on Voice Equality». MLSP'2017.
Stadelmann, Glinski-Haefeli, Gerber & Dürr (2018). «Capturing Suprasegmental Features of a Voice with RNNs for Improved Speaker Clustering». ANNPR'2018.

4. Speaker clustering – learnings & future work





«Pure» voice modeling seem largely solved

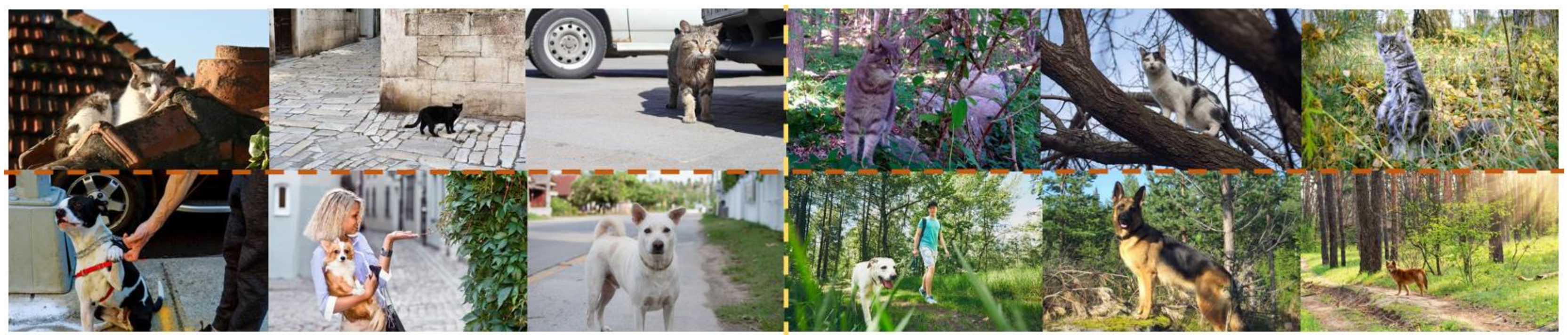
- RNN **embeddings work well** (see t-SNE plot of single segments)
- RNN model robustly exhibits *the predicted* **«sweet spot» for** the used **time information**
- Speaker clustering on clean & reasonably long input works **an order of magnitude better** (*as predicted*)
- Additionally, using a smarter clustering algorithm on top of embeddings makes **clustering on TIMIT as good as identification** (see ICPR'18 paper on dominant sets)

Future work

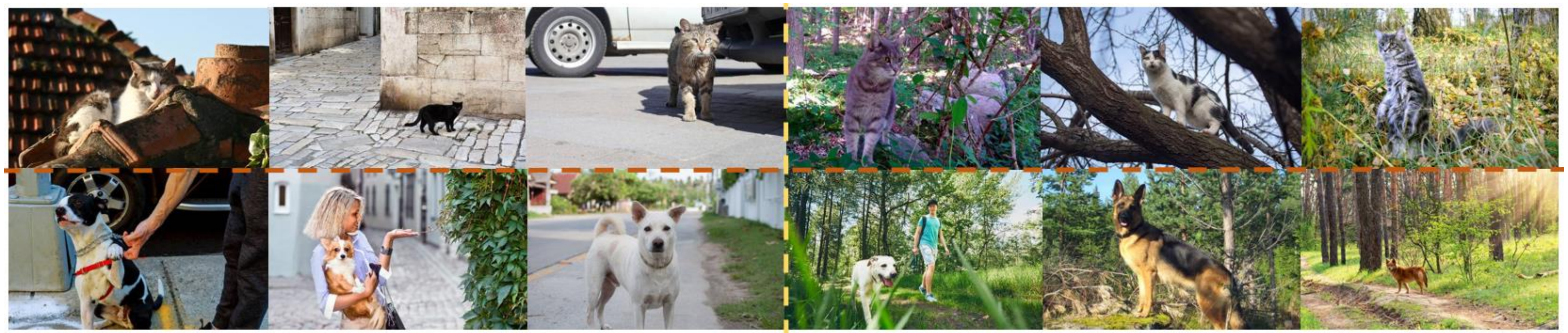
- Make models robust on **real-worldish data** (noise and more speakers/segments)
- Exploit findings for robust reliable **speaker diarization**
- **Learn** embeddings and the clustering algorithm **end to end**

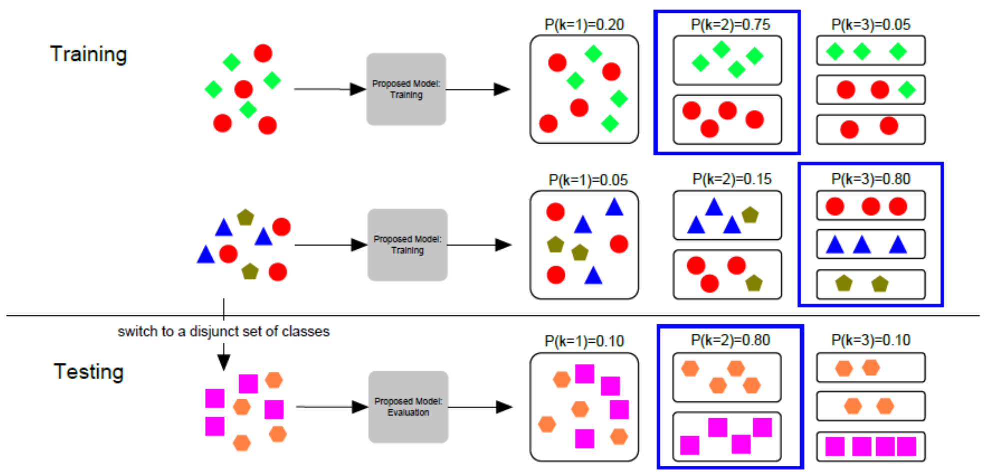
Hibraj, Vascon, Stadelmann & Pelillo (2018). «Speaker Clustering Using Dominant Sets». ICPR'2018.
Meier, Elezi, Amirian, Dürr & Stadelmann (2018). «Learning Neural Models for End-to-End Clustering». ANNPR'2018.

5. Learning to cluster

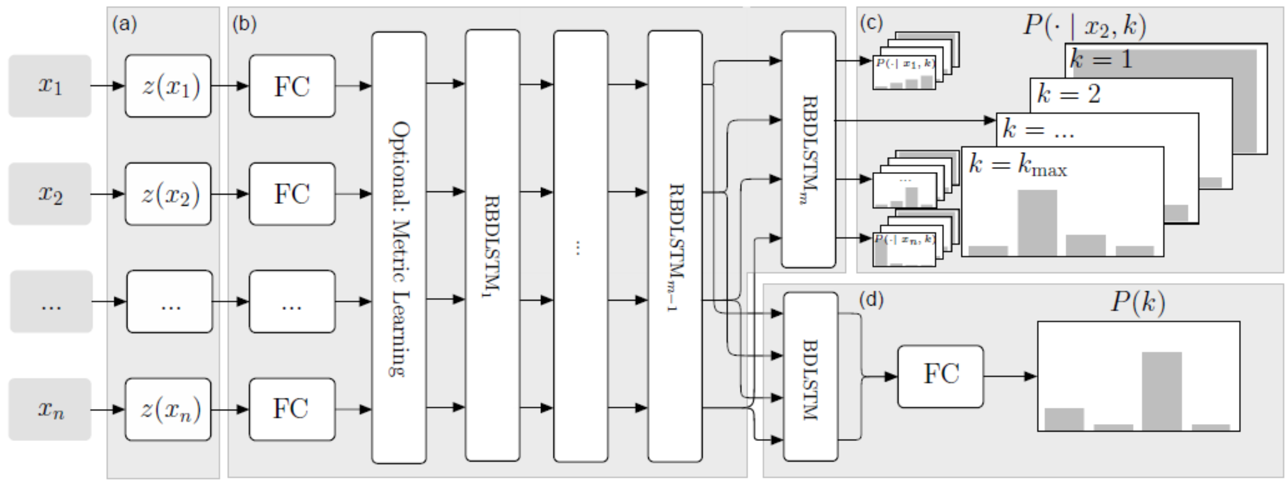


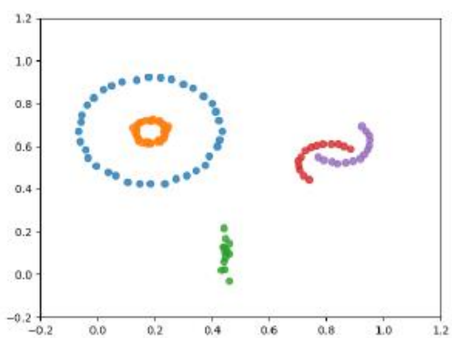
5. Learning to cluster

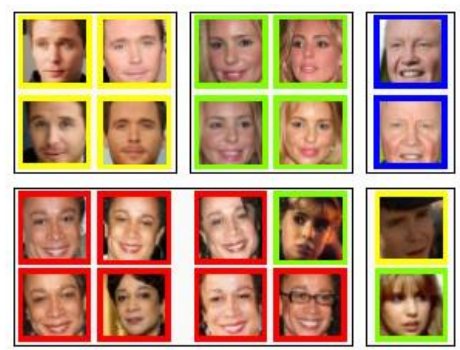




5. Learning to cluster – architecture & examples







Meier, Elezi, Amirian, Dürr & Stadelmann (2018). *«Learning Neural Models for End-to-End Clustering»*. ANNPR'2018.

6. 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)

Robustness is important.

- **Training processes** can be tricky
  → give hints via a unique loss, proper preprocessing and pretraining
- **Risk minimization** instead of error minimization
  → detect all defects at the expense of lower precision




Zürcher Hochschule
für Angewandte Wissenschaften

6. Lessons learned – model interpretability

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

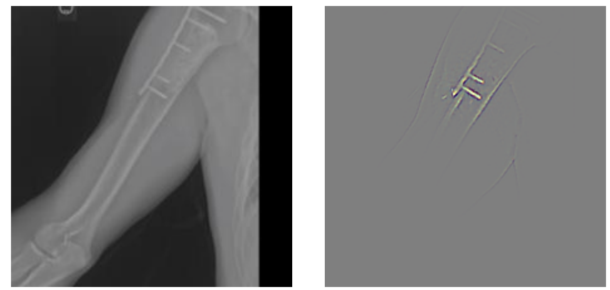
Innosuisse – Swiss Innovation Agency

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



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/

6. Lessons learned – model interpretability

Zürcher Hochschule
für Angewandte Wissenschaften

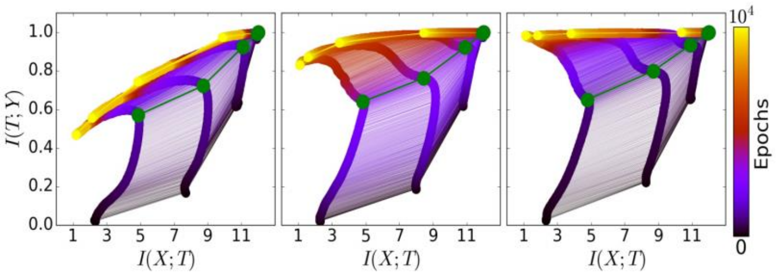
BW-TEC
ENGINEERING AUTOMATION CONSULTING
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Innosuisse – Swiss Innovation Agency

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

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/

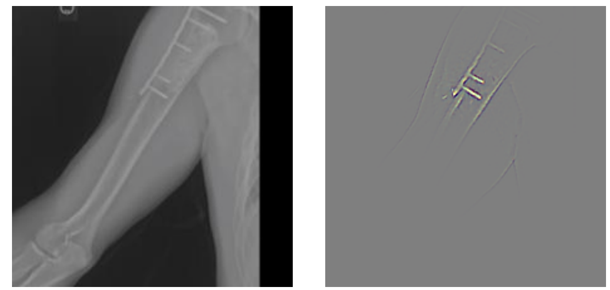
6. Lessons learned – model interpretability

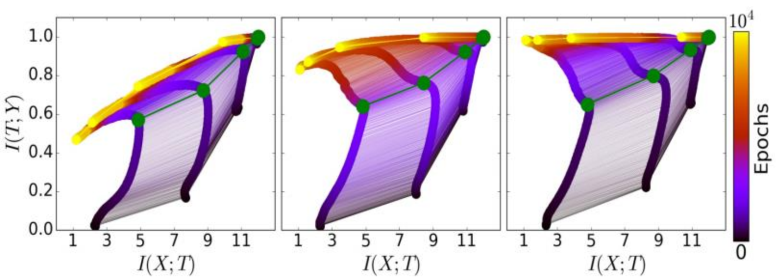
Interpretability is required.

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

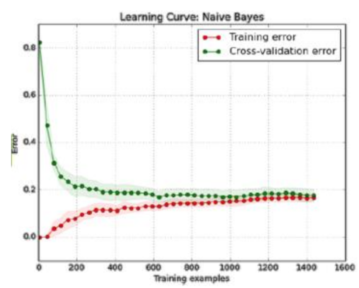
negative X-ray

positive X-ray





DNN training on the Information Plane



a learning curve

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/

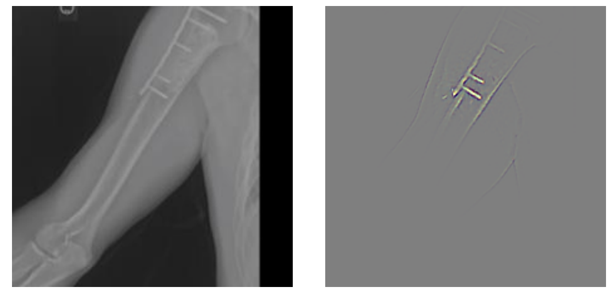
6. Lessons learned – model interpretability

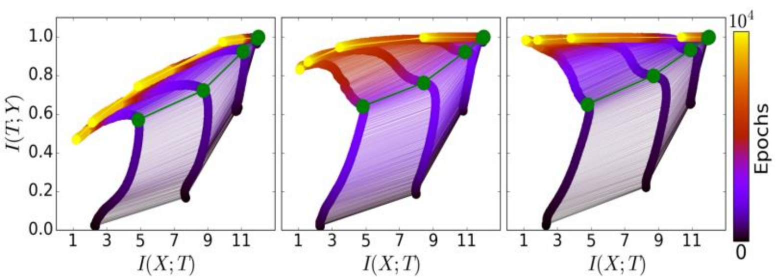
Interpretability is required.

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

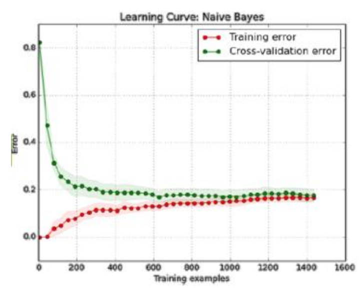
negative X-ray

positive X-ray

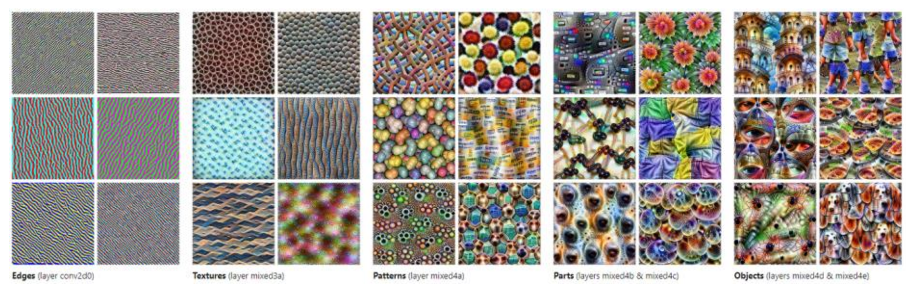




DNN training on the Information Plane



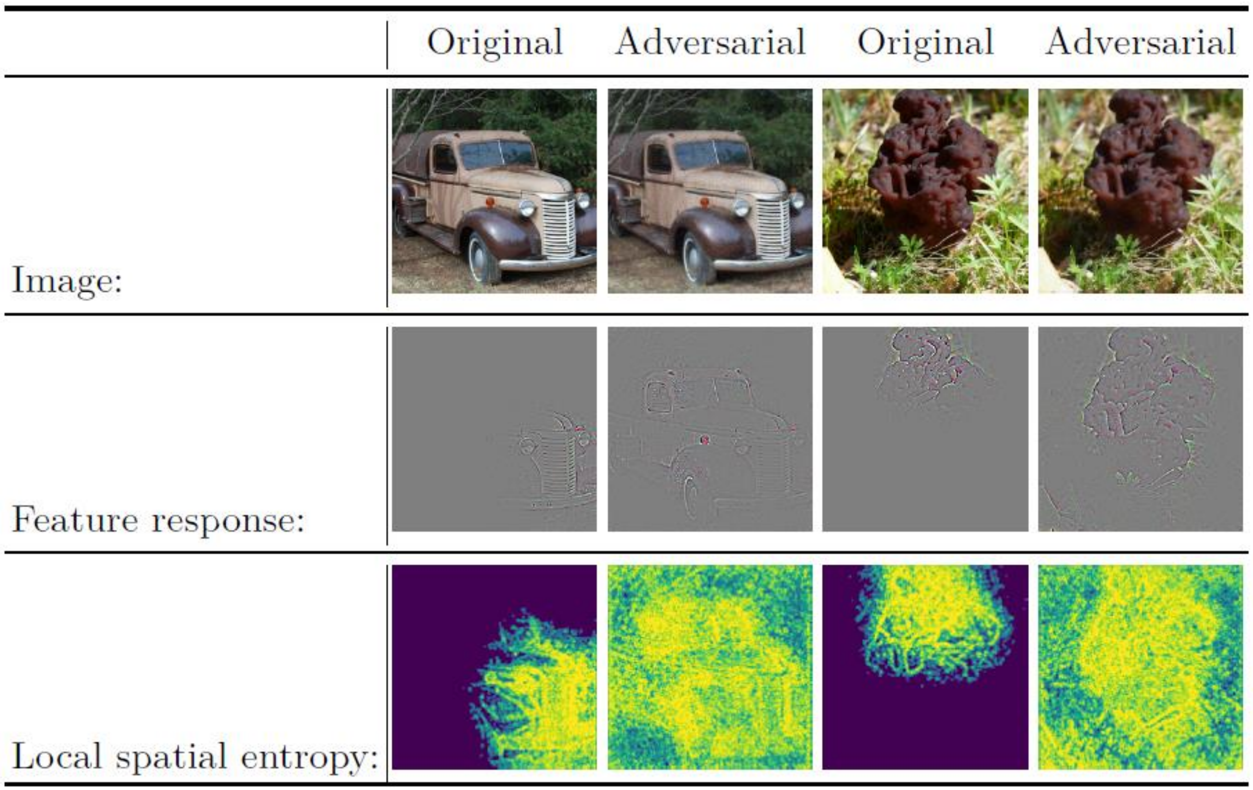
a learning curve



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/

Goody – trace & detect 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.

Conclusions

- Important for DL in practice, and hence target of applied research:
  sample efficiency, **robustness**, **interpretability**
- Future work will include:
  DL-based speaker diarization (dealing with learning transfer and noisyness)
  Novel **object detection** approaches **for many tiny objects**
  ➔ **possibly together?**

datalab
www.zhaw.ch/datalab

swiss group for artificial intelligence and cognitive science

SGAICO

data + service

Swiss Alliance for Data-Intensive Services

On me:

- Head ZHAW Datalab, vice president SGAICO, board Data+Service
- thilo.stadelmann@zhaw.ch
- 058 934 72 08
- https://stdm.github.io/

On the topics:

- AI: https://sgaico.swissinformatics.org/
- Data+Service Alliance: www.data-service-alliance.ch
- Collaboration: datalab@zhaw.ch

➔ Happy to answer questions & requests.

APPENDIX

Swiss Alliance for
Data-Intensive Services

Zürcher Hochschule
für Angewandte Wissenschaften

The Swiss Alliance for Data-Intensive Services provides a significant contribution to **make Switzerland an internationally recognized hub for data-driven value creation**.

In doing so, we rely on **cooperation in an interdisciplinary expert network** of innovative **companies** and **universities** to combine knowledge from different fields into marketable products and services.

Industrial Members

D ONE, la Mobilière, pwc, MIGROS, TRUMPF, Zürcher Kantonalbank, SAP, AWK GROUP, aspaara, bsi, BW-TEC, SKILLUE, ELCA, F&P Personal Robotics, GE, Google, HEXAGON, leanBI, Metrohm, NZZ, record, rockstar recruiting, RAIFFEISEN, SCIGILITY, Swiss Re, geo7, IXC GROUP, wincasa, zazuko, zühlke, comparis.ch, Bystronic, identitas

Academic Members

Bern University of Applied Sciences, EPFL, Hes-so Valais Wallis, Hochschule Luzern, HSR Hochschule für Technik Rapperswil, Université de Fribourg / Universität Freiburg, csem, Université de Genève, Università della Svizzera italiana, n|w University of Applied Sciences and Arts Northwestern Switzerland, University of St.Gallen, USI/SUPSI IDSIA, zhaw

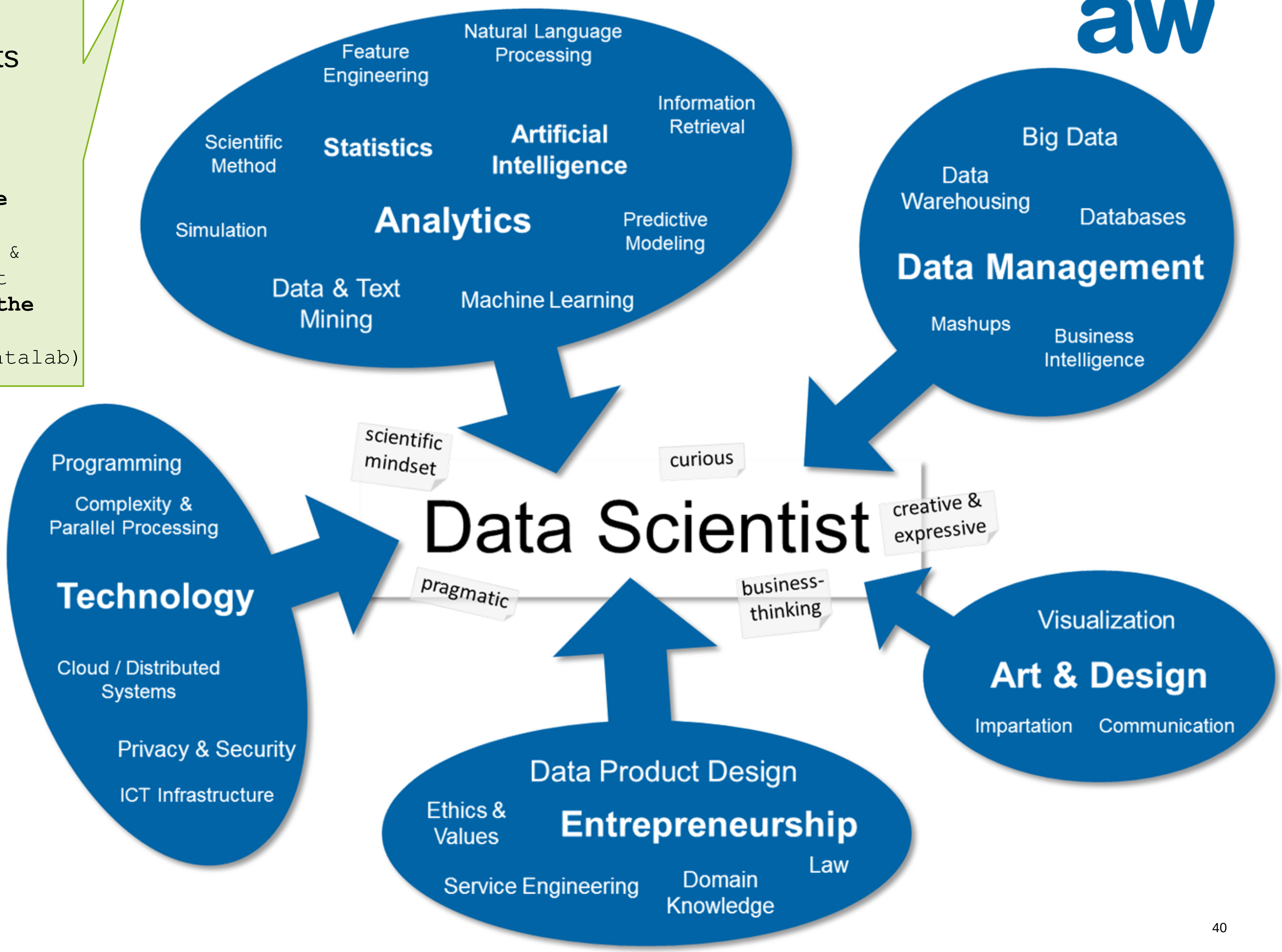
National & International Partners

INDUSTRIE 2025, Swiss Data Alliance, Open Data .ch, Lassalle Institut, Swiss Fintech Innovations, Euresearch, ISSIP

What is Data Science?

Enables Data Products
➔ **Applied** Science
➔ Interdisciplinary

Data Science := "**Unique blend of skills** from analytics, engineering & communication aiming at generating **value from the data** itself […]"
(ZHAW Datalab)



Stadelmann, Stockinger, Braschler, Cieliebak, Baudinot, Dürr and Ruckstuhl (2013). *Applied Data Science in Europe* . ECSS 2013.

PANOPTES – Automated Article Segmentation of Newspaper Pages for "Real Time Print Media Monitoring“

M. Cieliebak & T. Stadelmann, ZHAW

ARGUS MEDIA BASED INTELLIGENCE

zh aw School of Engineering

datalab www.zhaw.ch/datalab

Overview

Partners
Who are we

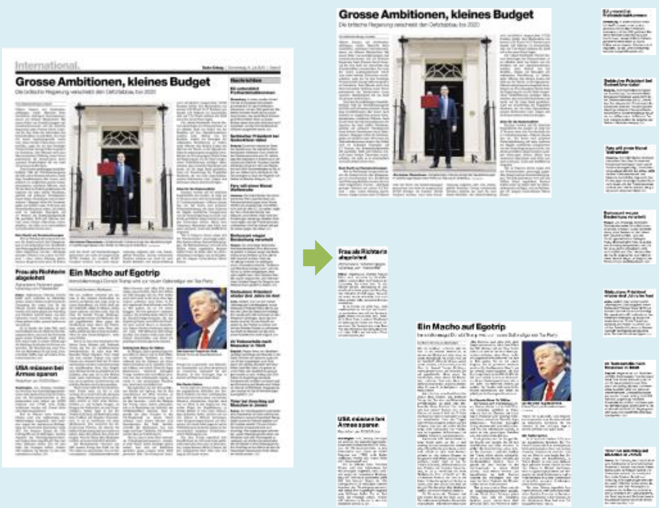
ARGUS der Presse AG
- Switzerland's leading media monitoring and information provider
- Experience of more than 100 years

ZHAW Datalab
- Interdisciplinary research group at Zurich University of Applied Sciences
- Combining the knowledge of different fields related to machine learning

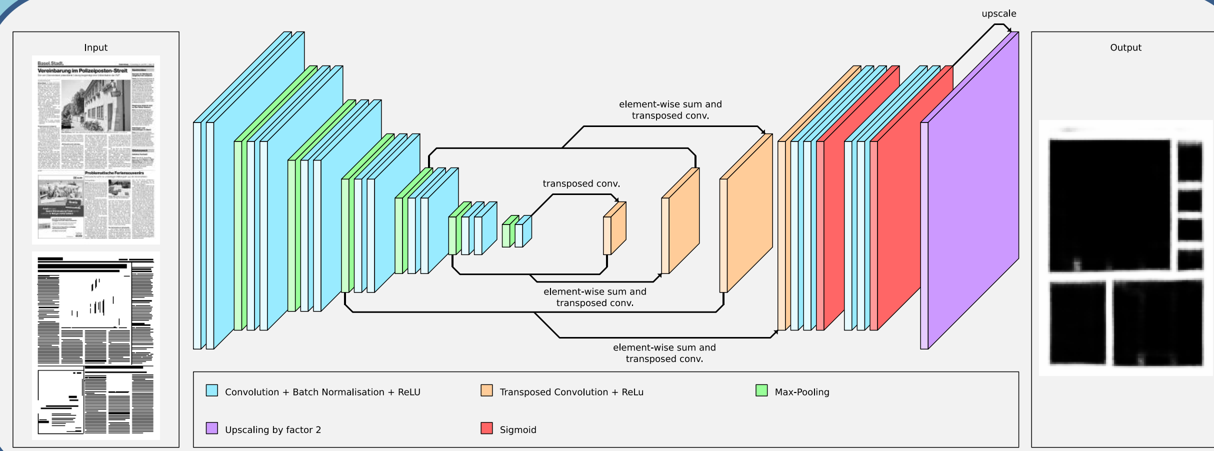
The Project
What do we do

Goal
- Real Time Print Media Monitoring
  - Extraction of relevant articles from newspaper pages
  - Delivering articles to customers

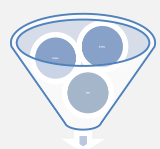
Problem
- Fully automated article segmentation
- Identification of article elements (e.g. title, subtitle, etc.)



Most Successful Approach [3]



Combination
Combination of rules, visual and textual features



Final segmentation

Result

References

[1] D. C. Ciresan, A. Giusti, L. M. Gambardella, and J. Schmidhuber. *Deep neural networks segment neuronal membranes in electron microscopy images.* In *NIPS*, pages 2852–2860, 2012.
[2] T. Mikolov, K. Chen, G. Corrado, and J. Dean. *Efficient Estimation of Word Representations in Vector Space.* In Proceedings of Workshop at *ICLR*, 2013.
[3] B. Meyer, T. Stadelmann, J. Stampfli, M. Arnold, M. Cieliebak. *Fully Convolutional Neural Networks for Newspaper Article Segmentation*. In Proceedings of ICDAR, Kyoto, Japan, 2018.

This project was funded by CTI under project number 17719.1 PFES-ES

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation

Federal Department of Economic Affairs, Education and Research EAER
Commission for Technology and Innovation CTI
Innovation Promotion Agency

swiss group for artificial intelligence and cognitive science

SGAICO

Zürcher Hochschule
für Angewandte Wissenschaften

Bio-SODA: Enabling Complex, Semantic Queries to Bioinformatics Databases through Intuitive Searching over Data

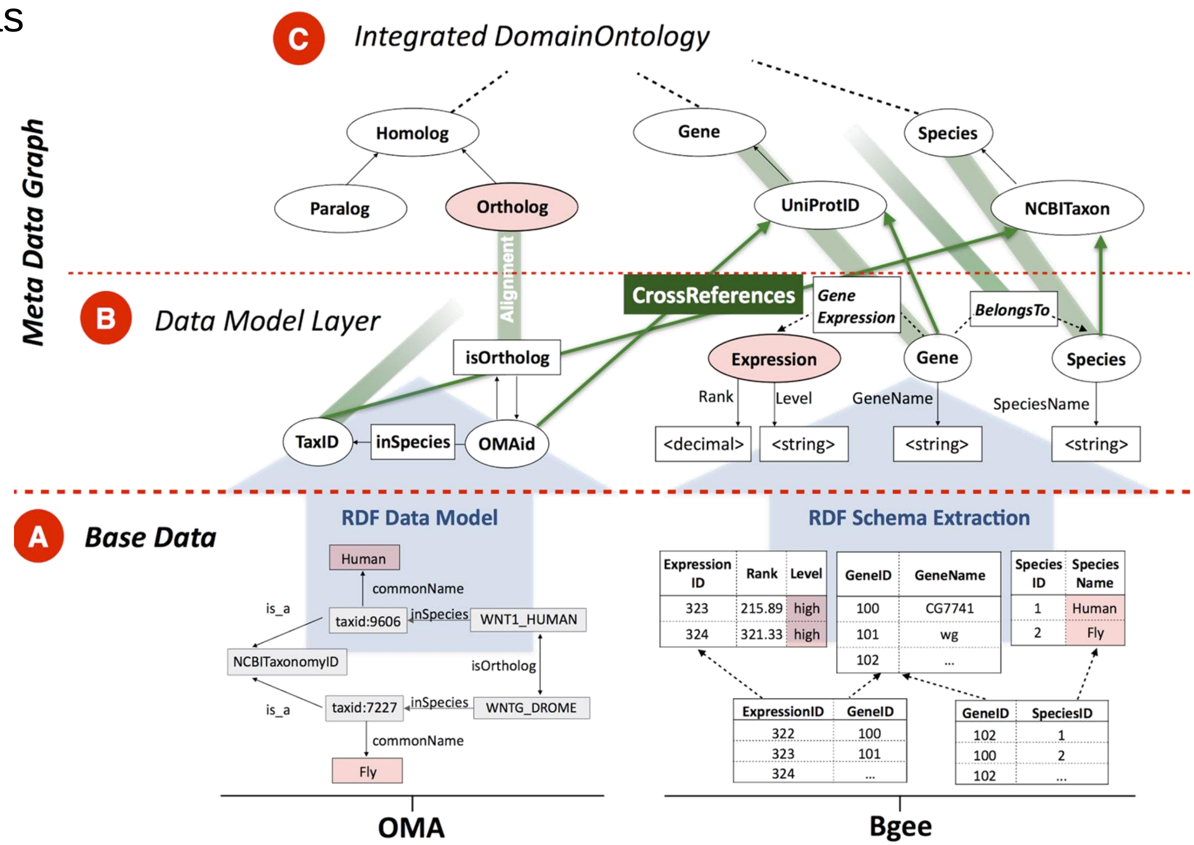
Intuitive exploration

- ✓ without knowing SPARQL, SQL, etc
- ✓ without knowing database schemas
- ✓ large datasets

Impact

- ➢ large bioinformatics user bases
- ➢ future federation of life sciences

Lead: Kurt Stockinger, ZHAW



75 NFP Big Data
Nationales Forschungsprogramm

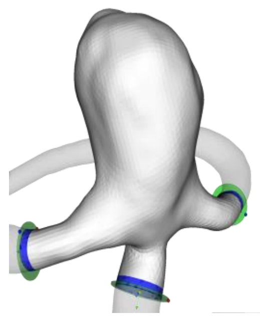
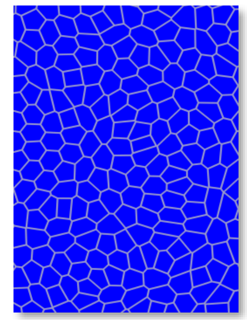
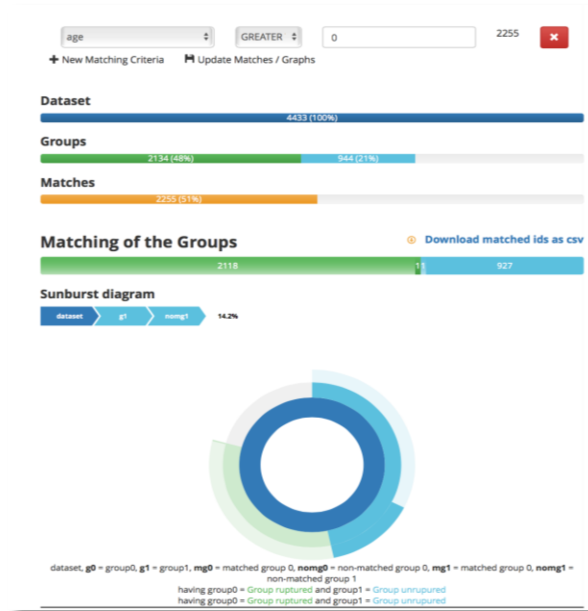
FNSNF
Fonds National Suisse
Schweizerischer Nationalfonds
Fondo Nazionale Svizzero
Swiss National Science Foundation

Zurich University of Applied Sciences zh aw

Unil
UNIL | Université de Lausanne

SIB Swiss Institute of Bioinformatics

AneuX: Ist die Form signifikant für die Gefährdung eines Aneurysmas?

Aneurysm im Röntgenbild (XA)	Isoliertes Aneurysma Zur Formanalyse	Zellen der Gefässwand	Modell der Gefässwand	Webtool für statistische Analyse
				

SystemsX.ch funding: 2M CHF, Begutachtung SNSF

- Morphologische Analyse von Aneurysmen mit Machine Learning
- Biologisch motiviertes Simulationsmodell für Zellwandveränderung
- Aufbau eines Krankheitsmodells für die Behandlungsplanung
- Aufbau einer Datenbank von Aneurysmen
- Erstellung von Werkzeuge zur Analyse der klinischen Daten und Bilddaten

Partner (Co-Antragsteller Sven Hirsch, ZHAW):

AneuX
Shape as Biomarker
for Aneurysm Disease

HUG Hôpitaux Universitaires de Genève

Universität Zürich UZH

UNIVERSITÉ DE GENÈVE

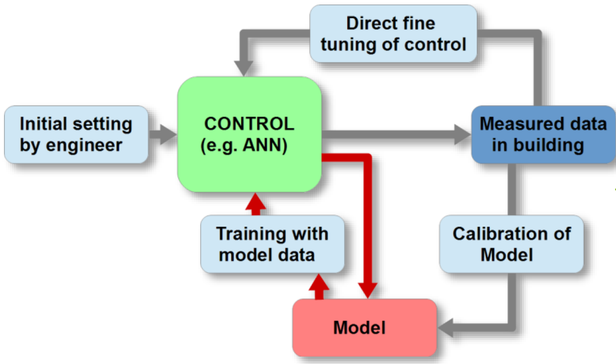
cabmm

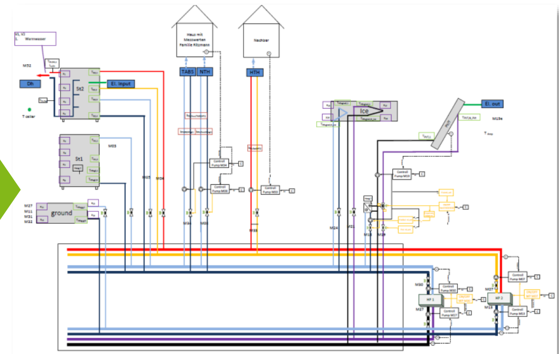
IT'IS FOUNDATION

Swiss Neuro Foundation

ETH zürich

Hydrobus: Simulation-based Optimization





The challenge

- Not enough training data for AI in socio-technological systems

The project

- **Self-adaption** of control to time-varying demands **in a multi-apartment building** using **simulations**
- Combined entropy and **energy optimization** of HVAC-system **based on Model Predictive Control**
- Integrates **renewable energy technology**, social dynamics and scenario-based weather prediction

The upside

- Enables a **Swiss SME** to harvest results from **modern mathematics, data science and AI**
- Gives **science** the opportunity to **test modern approaches on real-world problems**

RINO Electronics AG

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency

Zürcher Hochschule
für Angewandte Wissenschaften

Game playing

GIANTS SOFTWARE

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

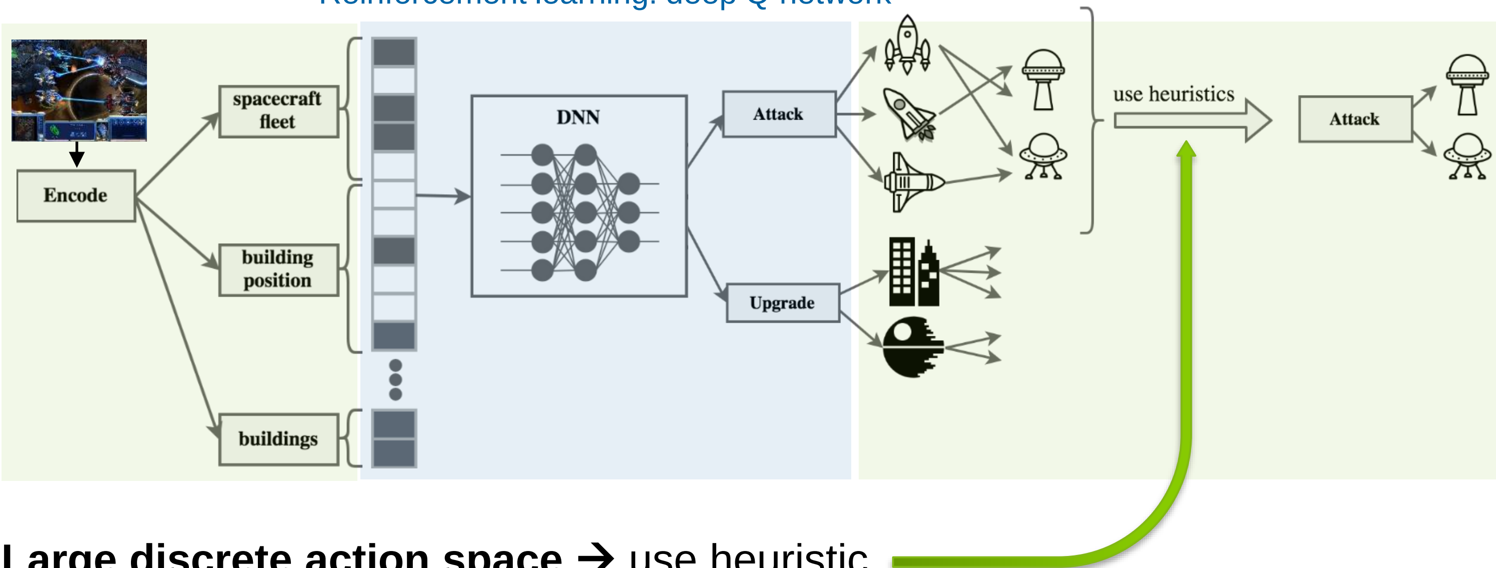
Swiss Confederation

Innosuisse – Swiss Innovation Agency

(symbolic figure)

Game playing – challenges & solutions

Reinforcement learning: deep Q network



Large discrete action space → use heuristic
- makes exploration difficult
- elongates training time

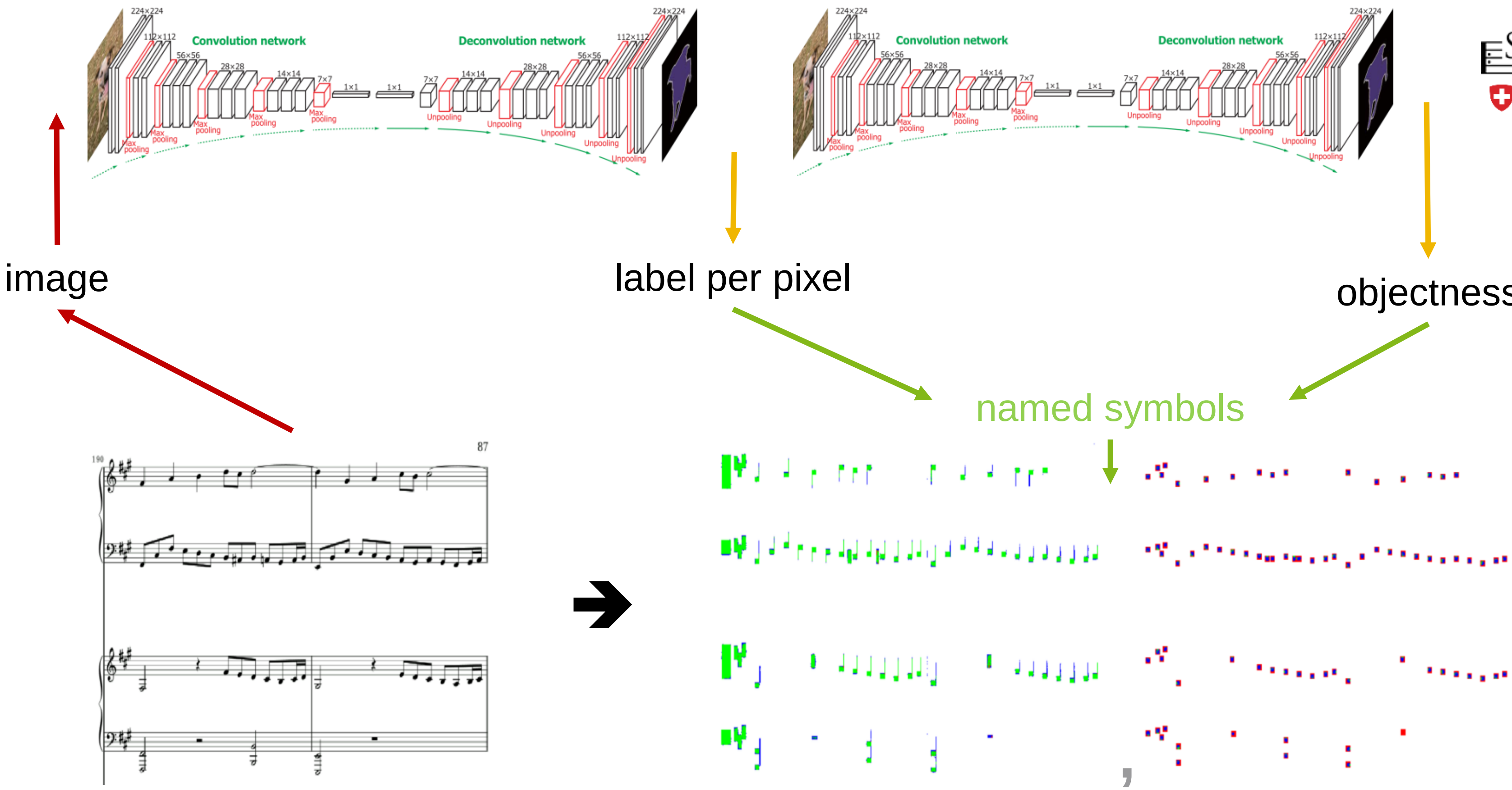
Delayed and sparse reward → do reward shaping
- sequence of actions crucial to get a reward

Distance encoding → use reference points

Transfer Learning → difficult: more complex environment needs other action sequence

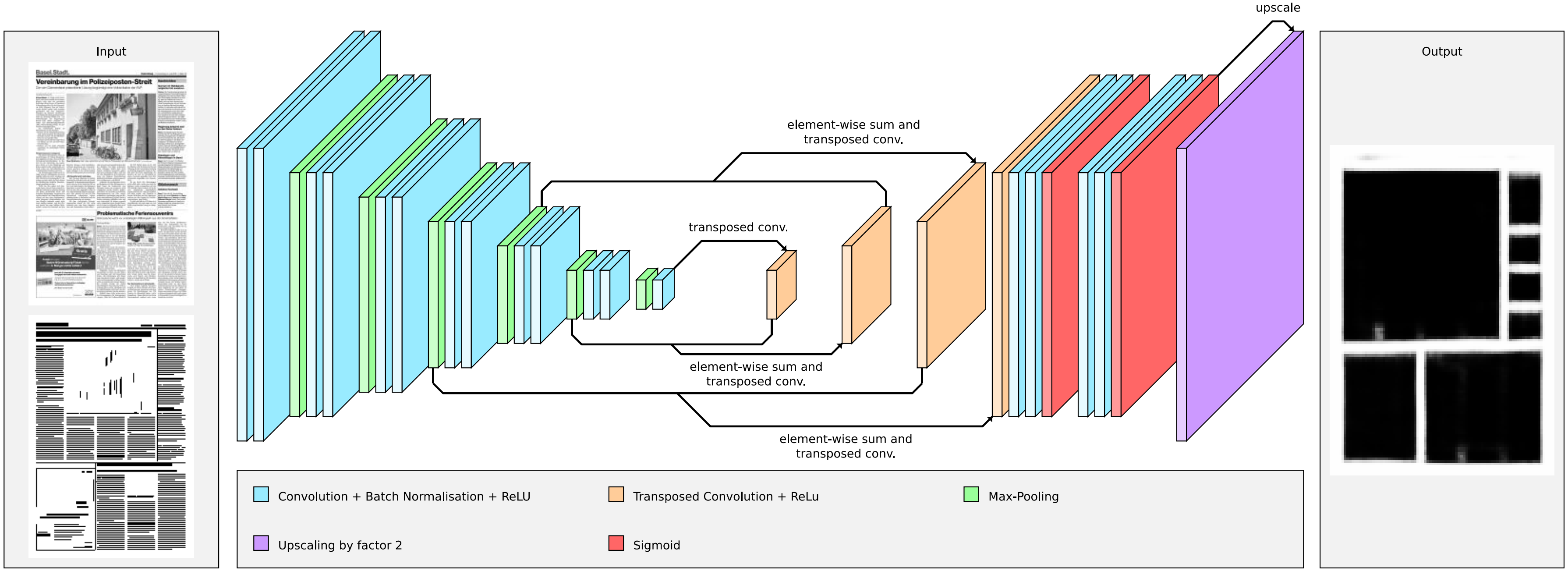
Stadelmann, Amirian, Arabaci, Arnold, Duivesteijn, Elezi, Geiger, Lörwald, Meier, Rombach & Tuggener (2018). *«Deep Learning in the Wild»*. ANNPR'2018.

Optical Music Recognition
Foundation of digitization in orchestras and music schools



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.

Segmentation of newspaper articles
Semi-automatic print media monitoring



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.

Zürcher Hochschule für Angewandte Wissenschaften

Condition monitoring
Maintaining machines on predicted failure only

Mechmine data science technologies

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

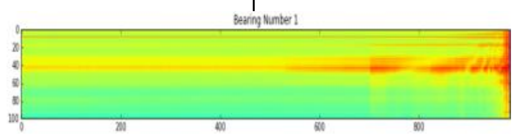
Swiss Confederation

Innosuisse – Swiss Innovation Agency

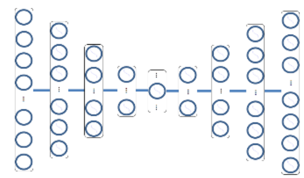
We use machine learning approaches for anomaly detection to learn the normal state of each machine and deviations of it purely from observed sensor signals; the approach combines classic and industry-proven features with e.g. deep learning auto-encoders.

vibration sensors

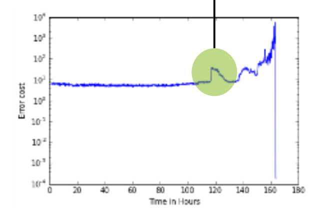
feature extraction



e.g., RNN autoencoder



early detection of fault



Stadelmann, Tolkachev, Sick, Stampfli & Dürr (2018). «*Beyond ImageNet - Deep Learning in Industrial Practice*». In: Braschler et al., «Applied Data Science», Springer.