

Machine Learning for Pattern Recognition @ ZHAW

IDIAP Research Institute, Martigny, July 02, 2018

Thilo Stadelmann



Swiss Alliance for
Data-Intensive Services

swiss group for artificial intelligence
and cognitive science



datalab

www.zhaw.ch/datalab

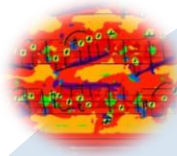
Agenda



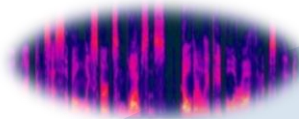
1. The group



2. Face
matching



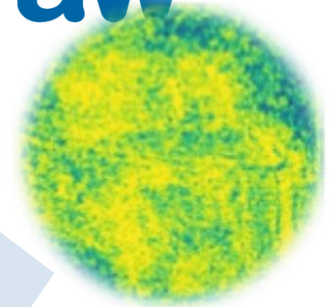
3. Music
scanning



4. Speaker
recognition



5.
Learning
to Cluster



6. Lessons
Learned

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

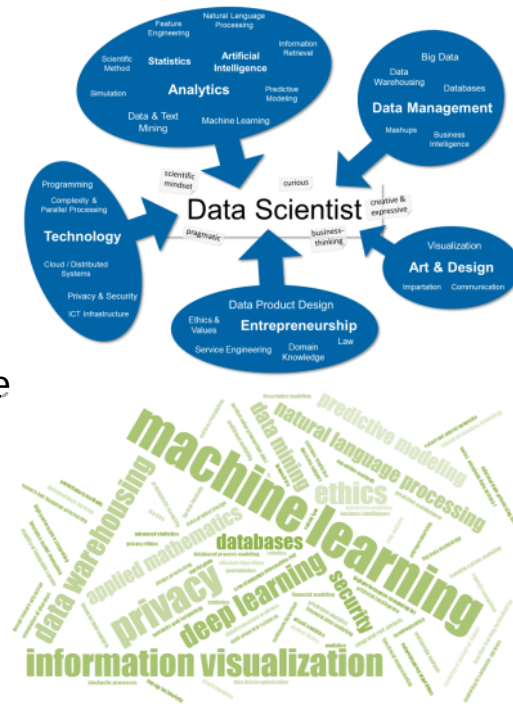
- **One of the first** interdisciplinary data science initiatives in Europe
- One of the first interdisciplinary centers at ZHAW

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

- Prognosis^{IK} – 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 CO₂»)
- Technology (e.g., SNSF project «Bio-SODA»)
- ...

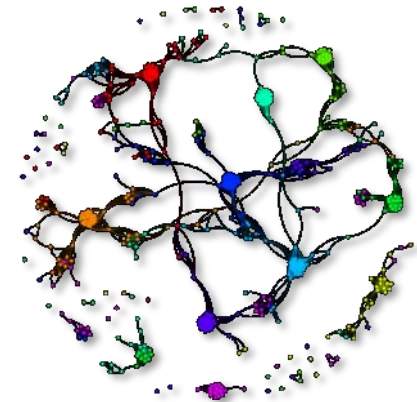
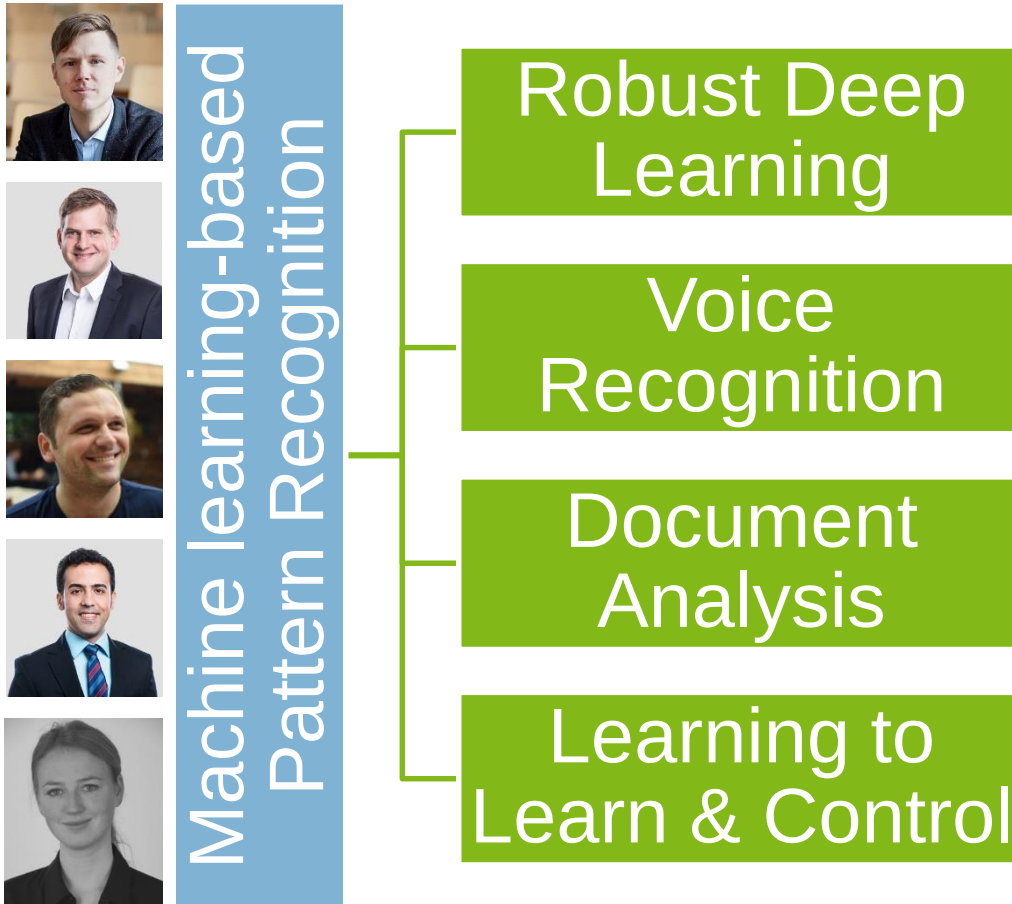


Figure: Visualizing the relationships of all Swiss foundations, based on the similarity of goals as expressed in their statutes. A proud collaboration of InIT and IDP within CTI project «Stiftungsregister SR 2.0»

1. ML @ Information Engineering Group

Institute of Applied Information Technology, School of Engineering



1. ML @ Information Engineering Group

Institute of Applied Information Technology, School of Engineering

Zürcher Hochschule
für Angewandte Wissenschaften

zhaw

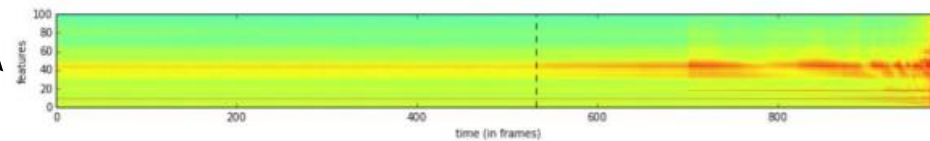
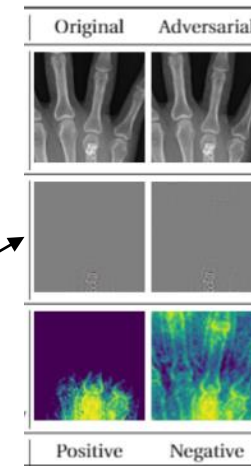
Machine learning-based Pattern Recognition

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



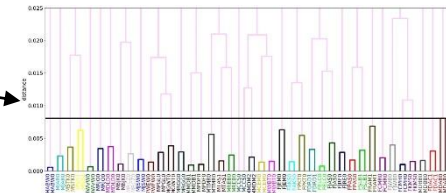
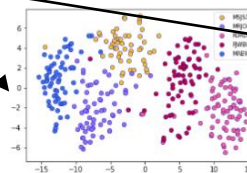
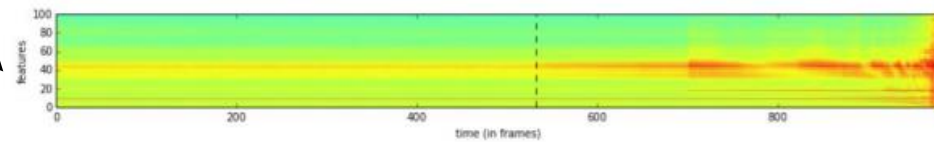
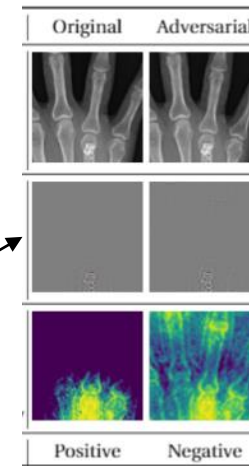
Machine learning-based Pattern Recognition

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

zhaw

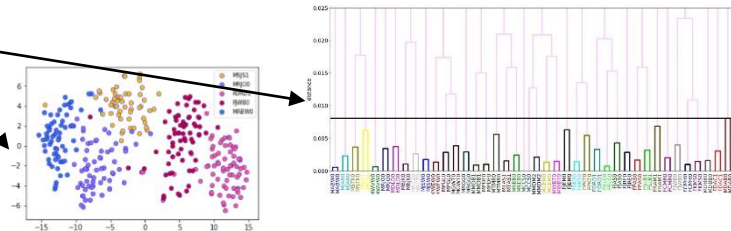
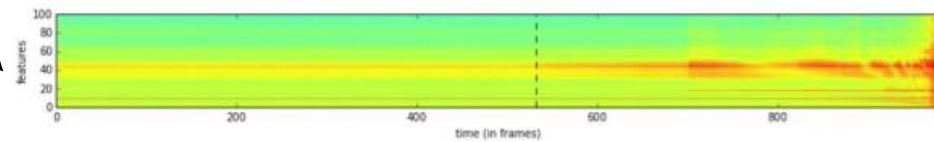
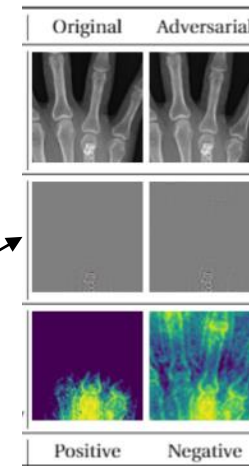
Machine learning-based Pattern Recognition

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

zhaw

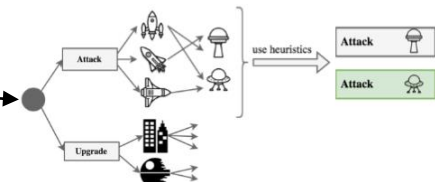
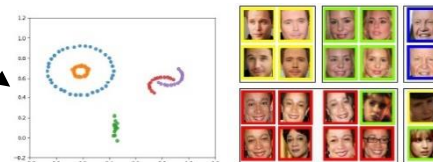
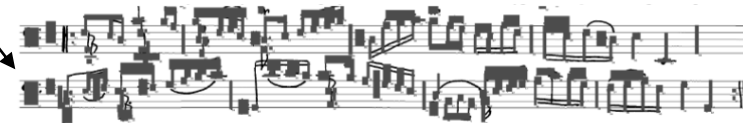
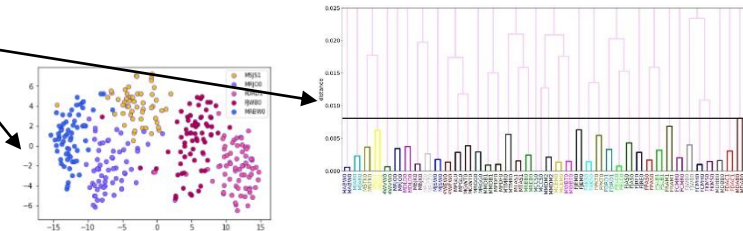
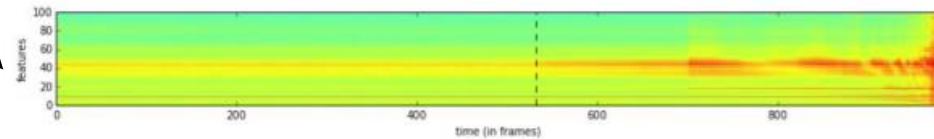
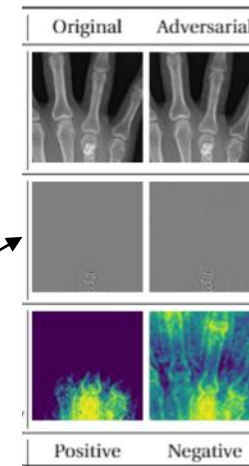
Machine learning-based Pattern Recognition

Robust Deep
Learning

Voice
Recognition

Document
Analysis

Learning to
Learn & Control





[!] DEEPIIMPACT



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Innosuisse – Swiss Innovation Agency

2. Face matching

2. Face matching

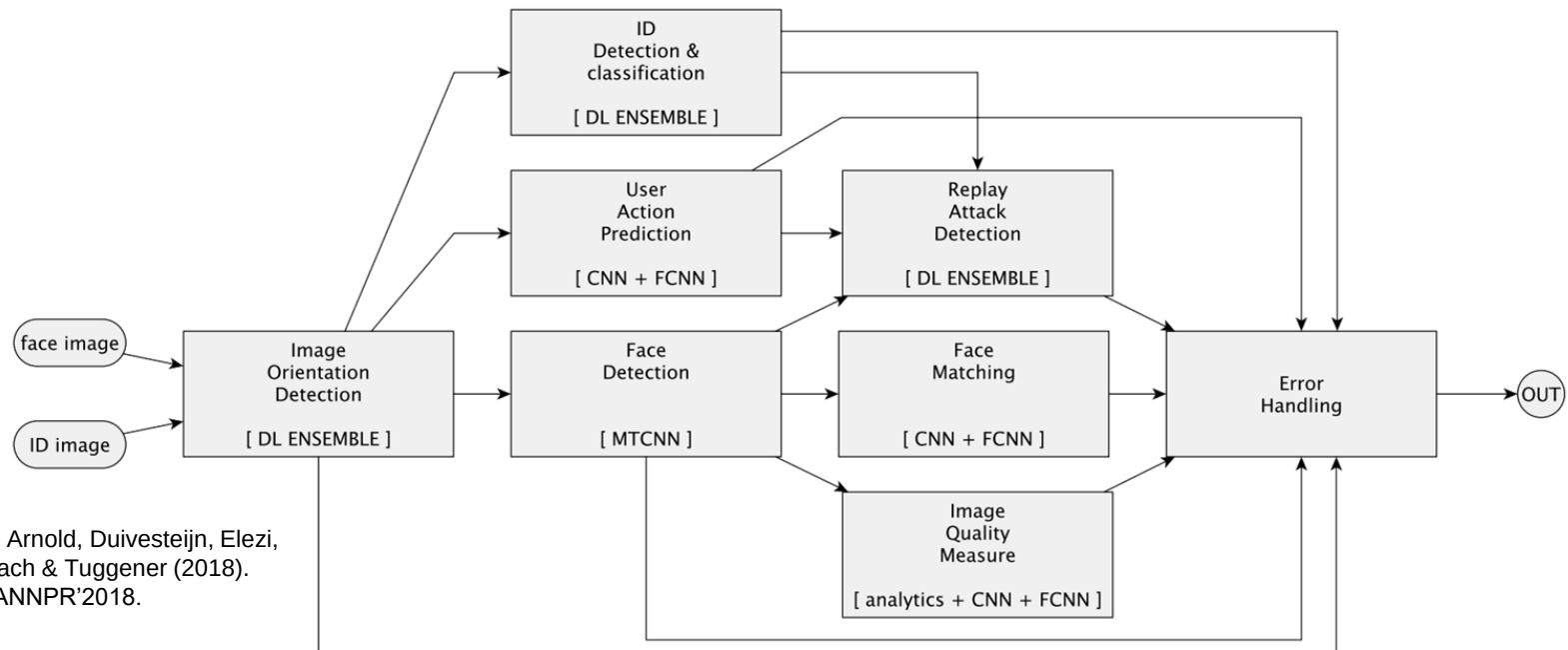


[!] DEEPIMPACT



Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Innosuisse – Swiss Innovation Agency

2. Face matching – challenges & solutions



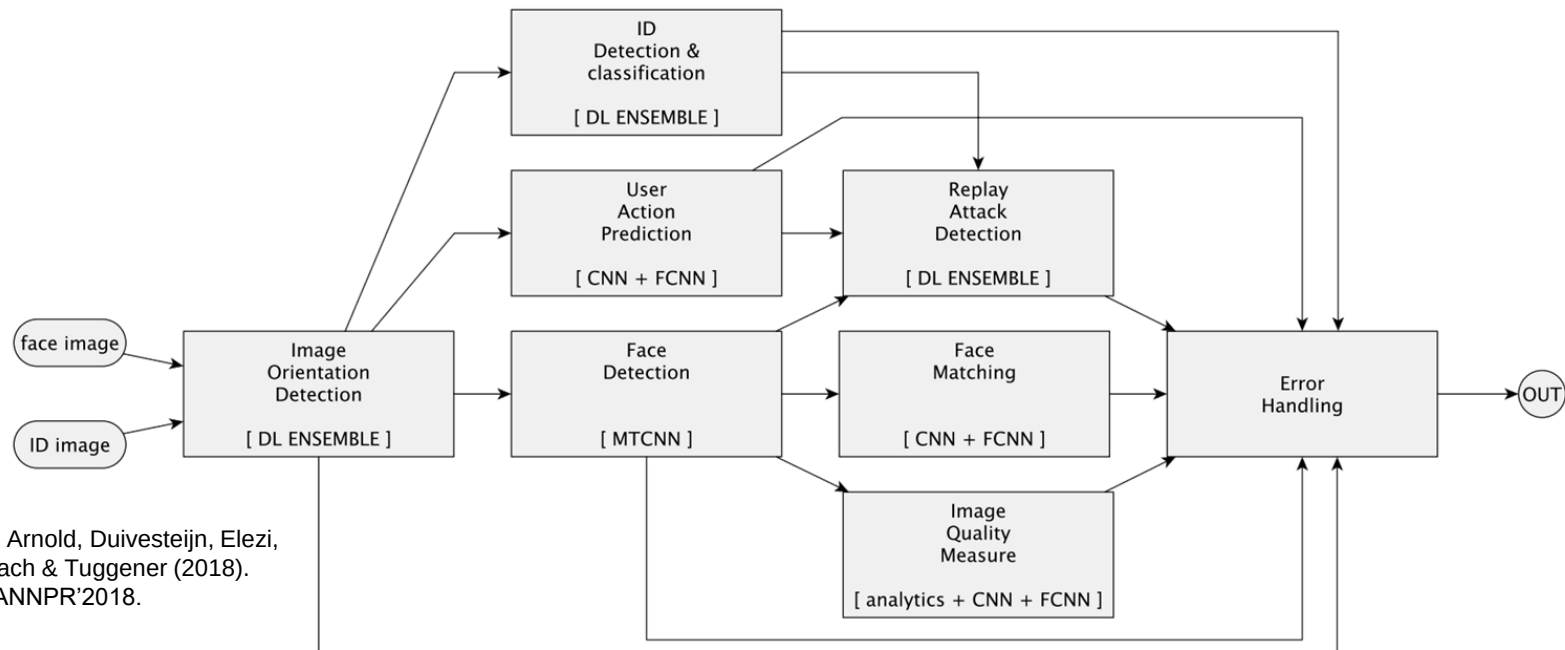
Stadelmann, Amirian, Arabaci, Arnold, Duivesteijn, Elezi, Geiger, Lörwald, Meier, Rombach & Tuggener (2018). «Deep Learning in the Wild». ANNPR'2018.

2. Face matching – challenges & solutions



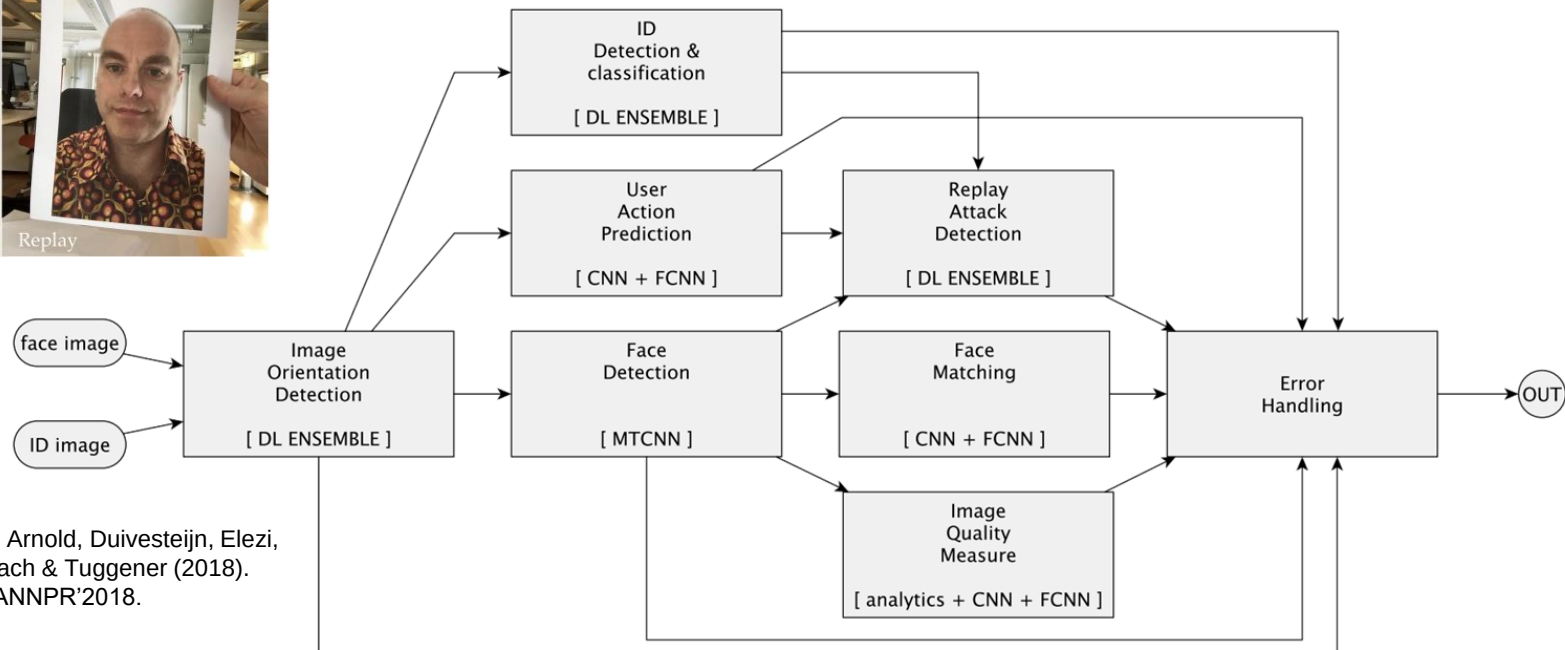
[!] DEEPIIMPACT

 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



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

3. Music scanning

N 152.
Die Forelle.
Gesung von Ch. Fr. D. Schubert.
Für eine Singstimme mit Begleitung des Pianoforte.
Schubert's Werke. compoſirt von
FRANZ SCHUBERT.
Dritte Fassung.
Nº 231

Melod.

Singstimme.

Pianoforte.



```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE score-partwise SYSTEM "http://www.musescore.org/dtd/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>
- <cenths> 40 </cenths>
- <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
Schubert, Die Forelle (Hörsände 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="P1-13">
- <instrument-name> Ténor </instrument-name>
- </score-instrument>
- <midi-instrument id="P1-13">
- <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="P2-13">
- <instrument-name>
```



Die Forelle - Franz Schubert

$\text{♩} = 80$

Voice

Piano

Vo.

ei - nem Büch-lein hel - le, da schoß in fro - her Eil die lau - ni - sche Fo - re - le vor -

3. Music scanning – challenges & solutions



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.

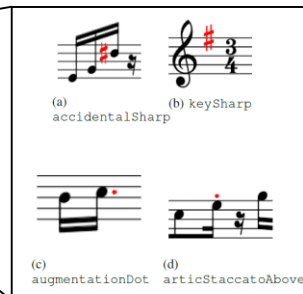
3. Music scanning – challenges & solutions

The image shows a musical score with a callout box highlighting specific annotations. The callout box contains four examples of musical notation with labels:

- (a) accidentalSharp
- (b) keySharp
- (c) augmentationDot
- (d) articStaccatoAbove

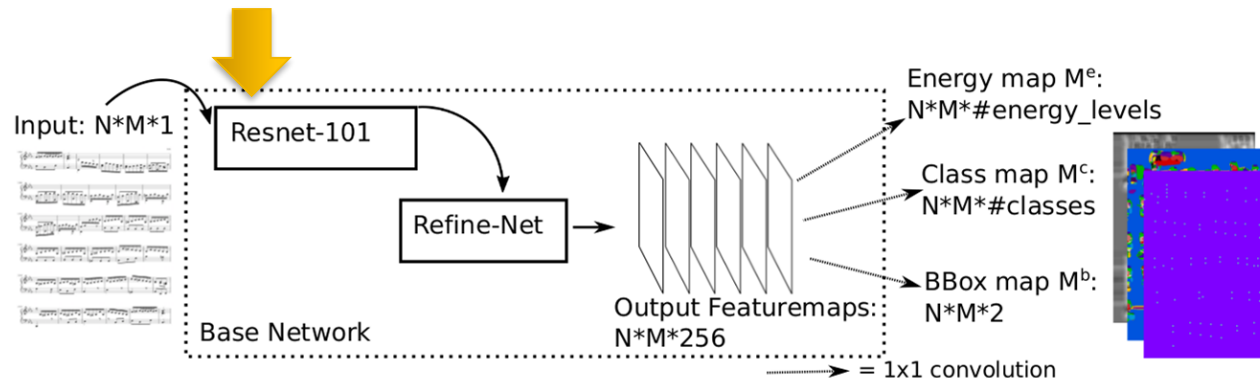
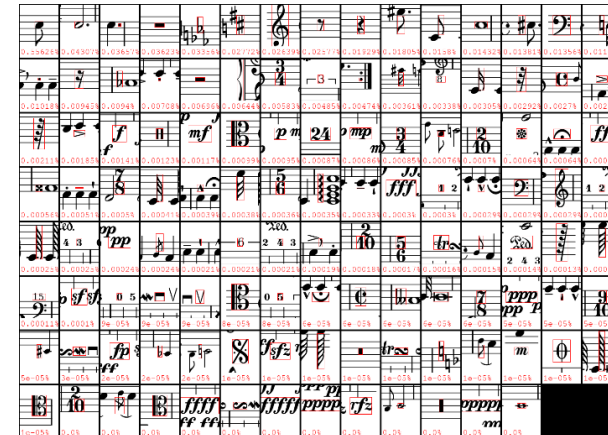
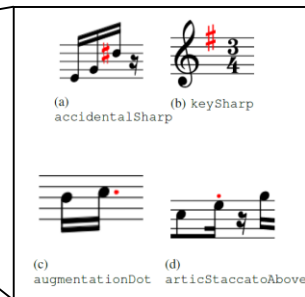
Tuggener, Elezi, Schmidhuber, Pelillo & Stadelmann (2018). «DeepScores – A Dataset for Segmentation, Detection and Classification of Tiny Objects». ICPR'2018.
Tuggener, Elezi, Schmidhuber & Stadelmann (2018). «Deep Watershed Detector for Music Object Recognition». ISMIR'2018.

3. Music scanning – challenges & solutions



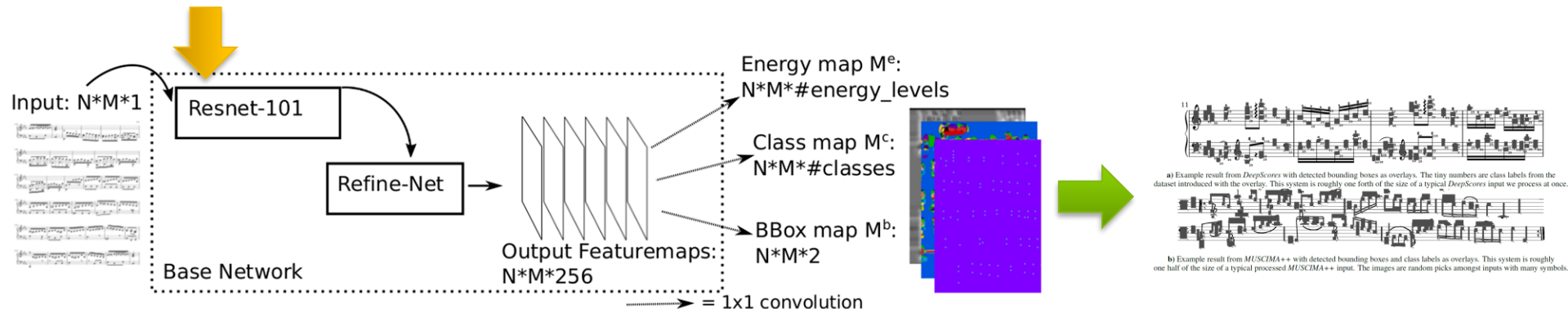
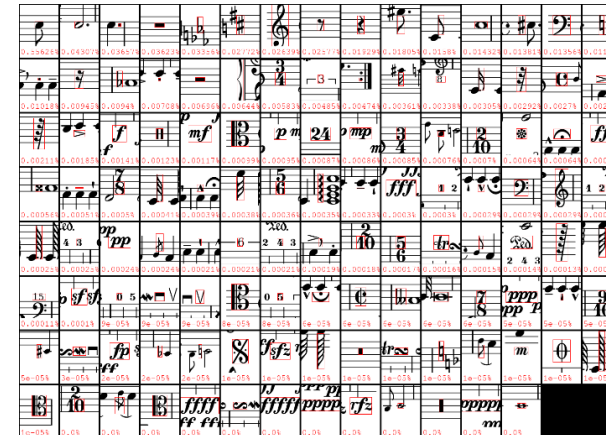
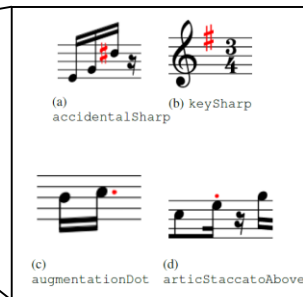
Tuggener, Elezi, Schmidhuber, Pelillo & Stadelmann (2018). «DeepScores – A Dataset for Segmentation, Detection and Classification of Tiny Objects». ICPR'2018.
Tuggener, Elezi, Schmidhuber & Stadelmann (2018). «Deep Watershed Detector for Music Object Recognition». ISMIR'2018.

3. Music scanning – challenges & solutions



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.

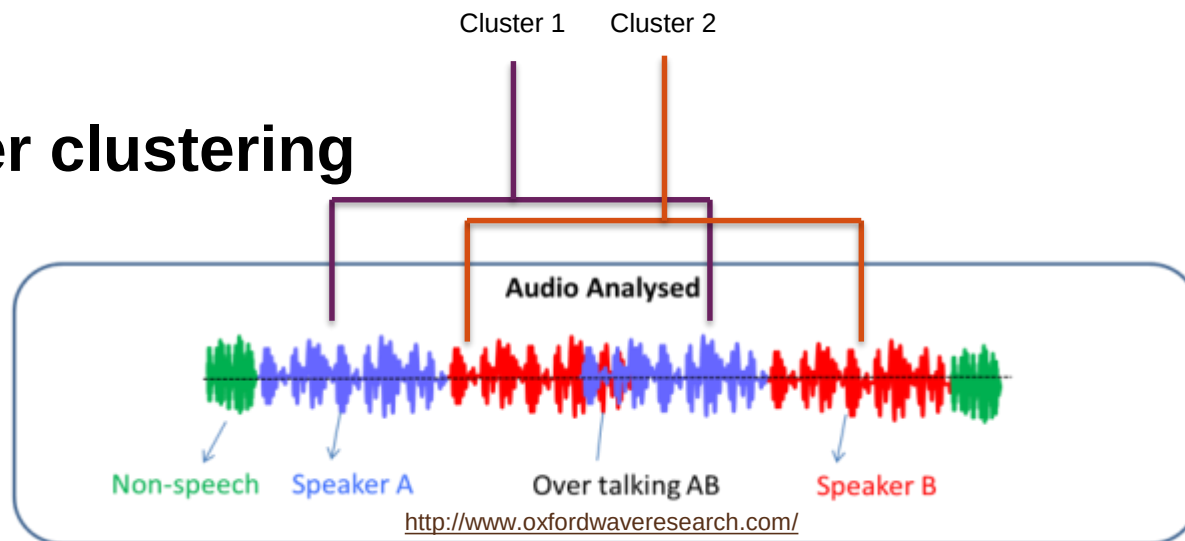
3. Music scanning – challenges & solutions



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

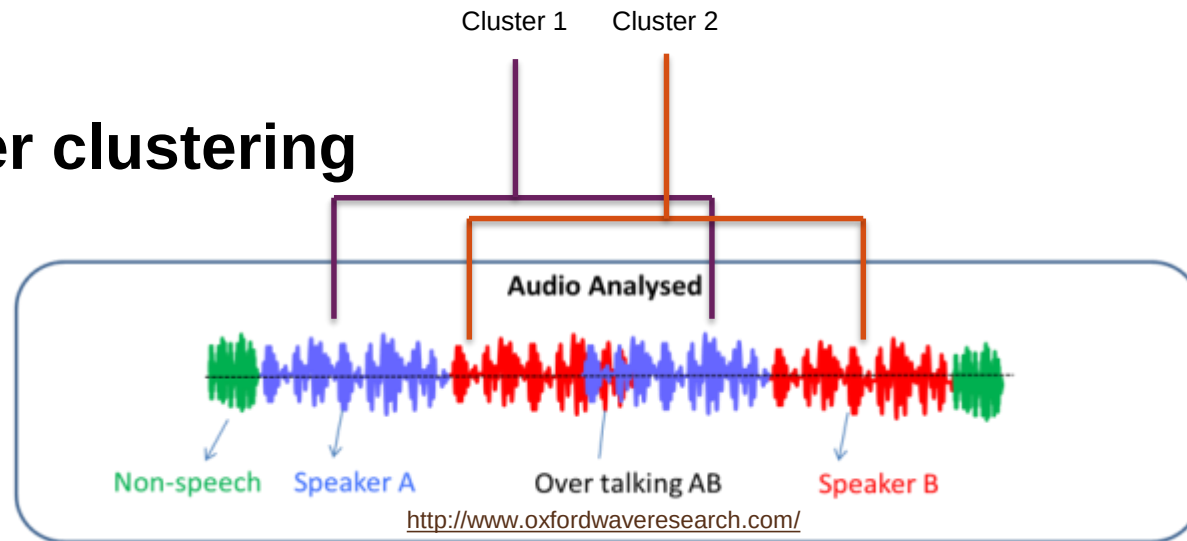


For the 630 training utterances, GMMs with 32 mixtures are built a priori, then an identification experiment is run for the 630 test utterances. It yields a satisfactory 0.5% closed set identification error.

[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?

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?

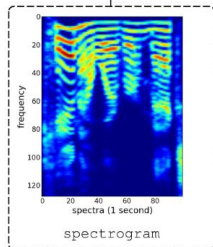
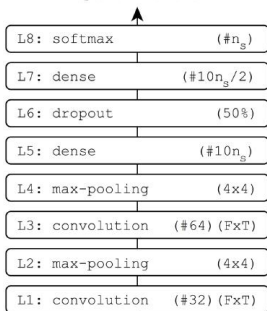
The interpretation of our results has shown that it is the stage of modeling that bears the highest potential: the inclusion of temporal context information among feature vectors is what is crucially missing there. Furthermore, the inclusion

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-

4. Speaker clustering – exploiting time information

CNN (MLSP'16)

speaker labels



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.

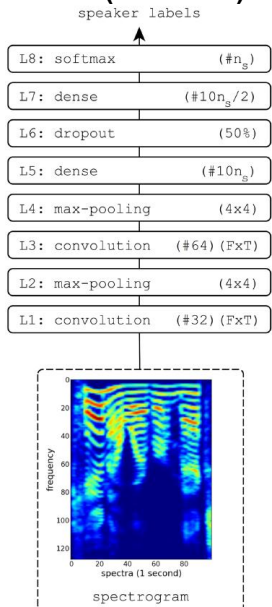
Lukic, Vogt, Dürr & Stadelmann (2017). «Learning Embeddings for Speaker Clustering based on Voice Equality». MLSP'2017.

Stadelmann, Glinski-Haefeli, Gerber & Dürr (2018). «Capturing Suprasegmental Features of a Voice with RNNs for Improved Speaker Clustering». ANNPR'2018.

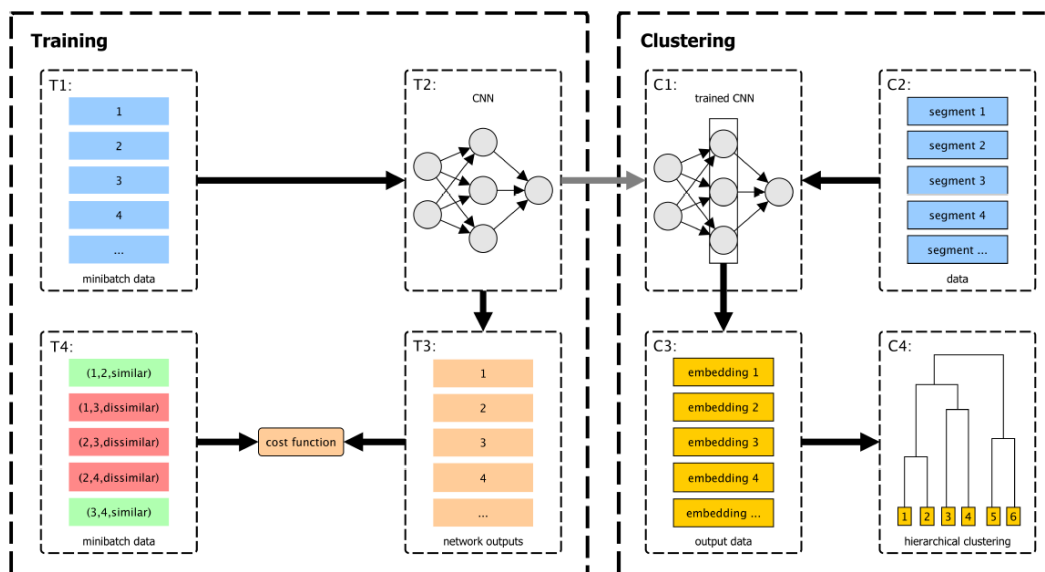
Zürcher Fachhochschule

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.

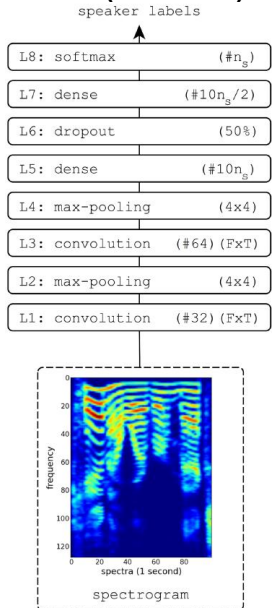
Lukic, Vogt, Dürr & Stadelmann (2017). «Learning Embeddings for Speaker Clustering based on Voice Equality». MLSP'2017.

Stadelmann, Glinski-Haefeli, Gerber & Dürr (2018). «Capturing Suprasegmental Features of a Voice with RNNs for Improved Speaker Clustering». ANNPR'2018.

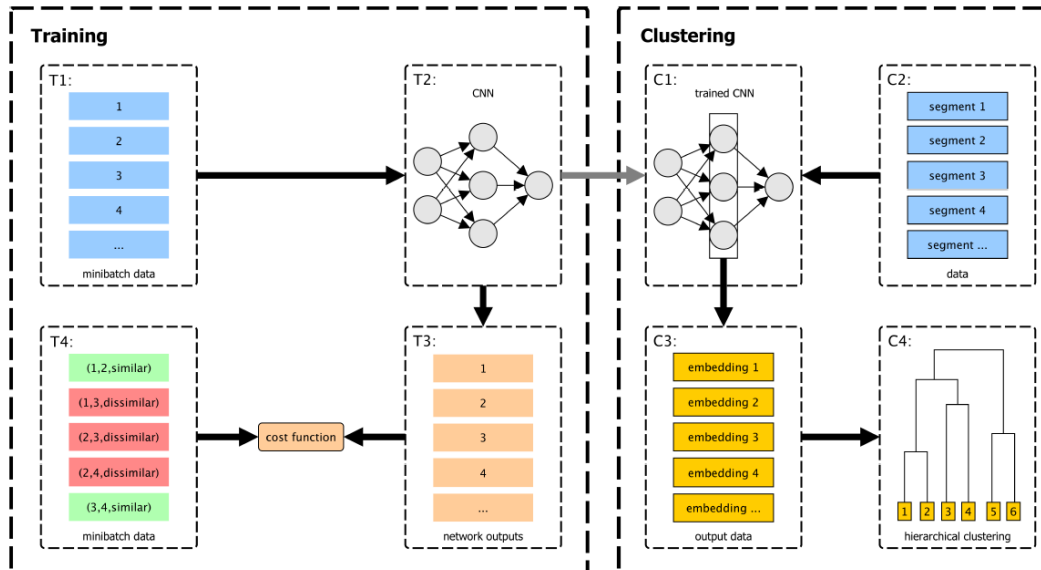
Zürcher Fachhochschule

4. Speaker clustering – exploiting time information

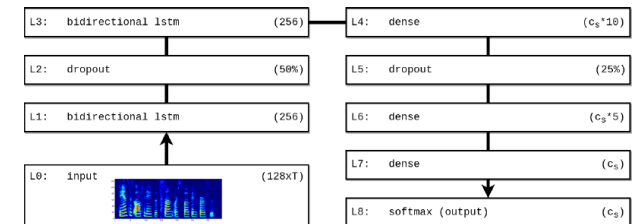
CNN (MLSP'16)



CNN & clustering-loss (MLSP'17)



RNN & clustering-loss (ANNPR'18)



Method	MR	MR (legacy)
RNN /w PKLD	2.19% ($\frac{1.25\%+2.5\%+1.25\%+3.75\%}{4}$)	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%	-

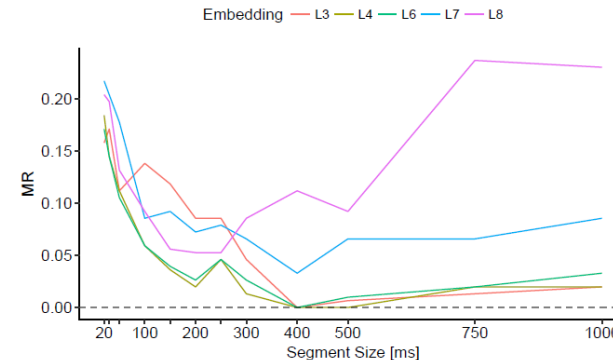
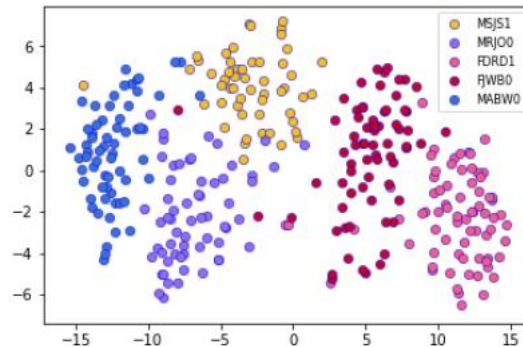
Lukic, Vogt, Dürr & Stadelmann (2016). «Speaker Identification and Clustering using Convolutional Neural Networks». MLSP'2016.

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 Fachhochschule

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)

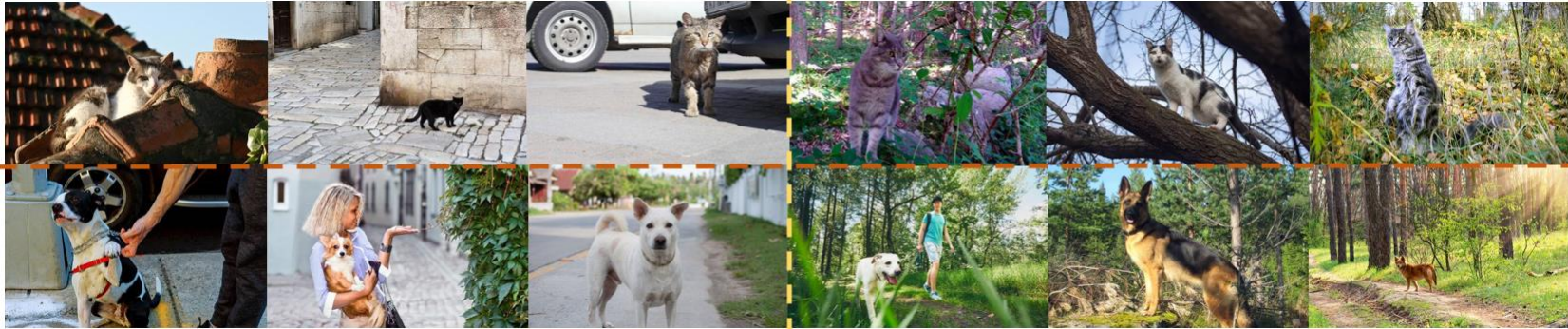
Future work

- 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**

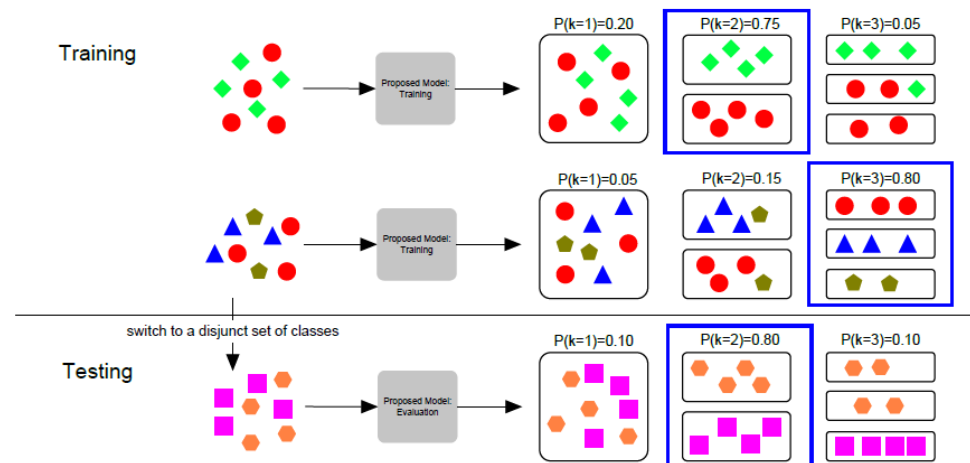
Hibraj, Vascon, Stadelmann & Pelillo (2018). «Speaker Clustering Using Dominant Sets». ICPR'2018.

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

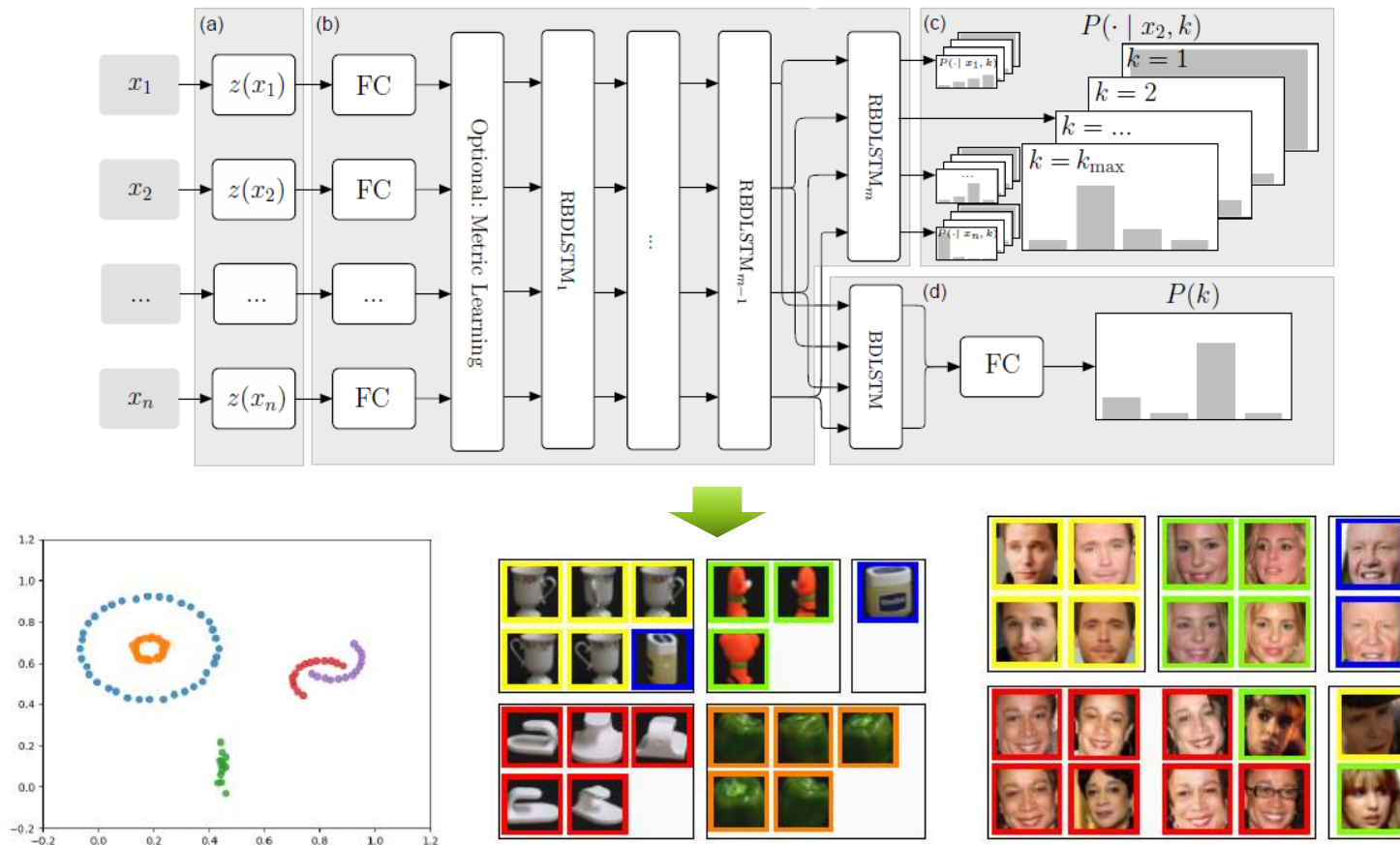
5. Learning to cluster



5. Learning to cluster



5. Learning to cluster – architecture & examples



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



6. Lessons learned – model interpretability

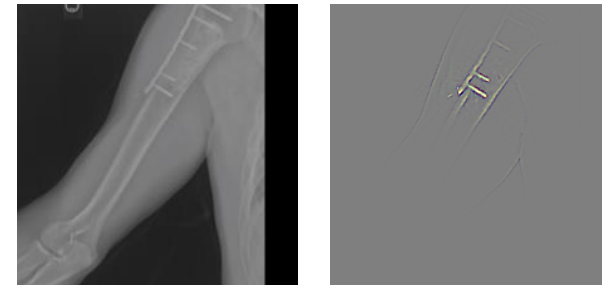
Interpretability is required.

- 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



Stadelmann, Amirian, Arabaci, Arnold, Duivesteyn, 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

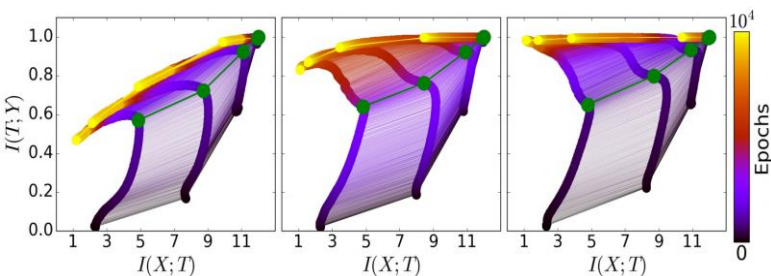
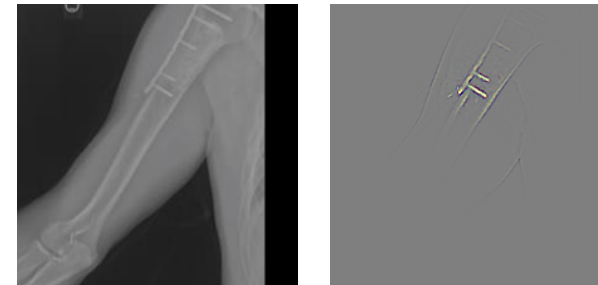
Interpretability is required.

- 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

Stadelmann, Amirian, Arabaci, Arnold, Duivesteyn, 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

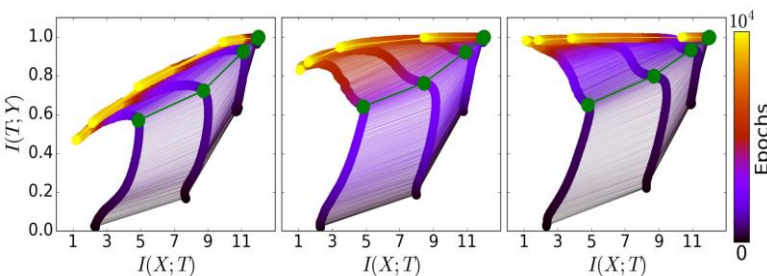
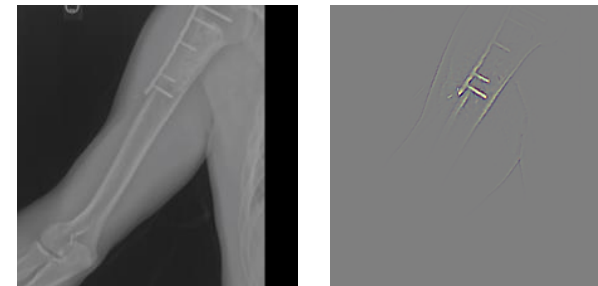
Interpretability is required.

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

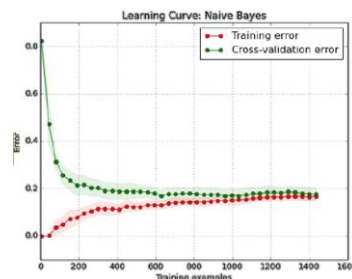
negative X-ray



positive X-ray



DNN training on the Information Plane



a learning curve

Stadelmann, Amirian, Arabaci, Arnold, Duivesteyn, 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

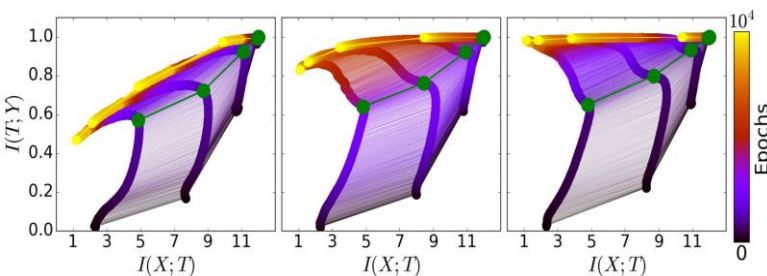
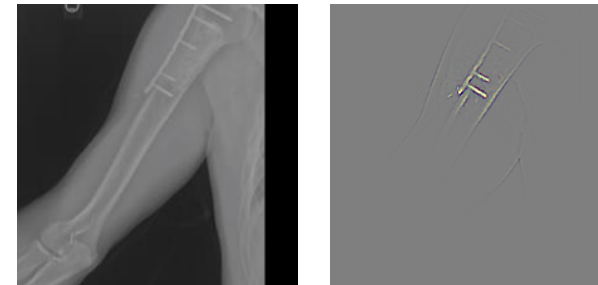
Interpretability is required.

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

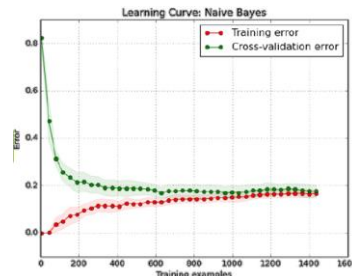
negative X-ray



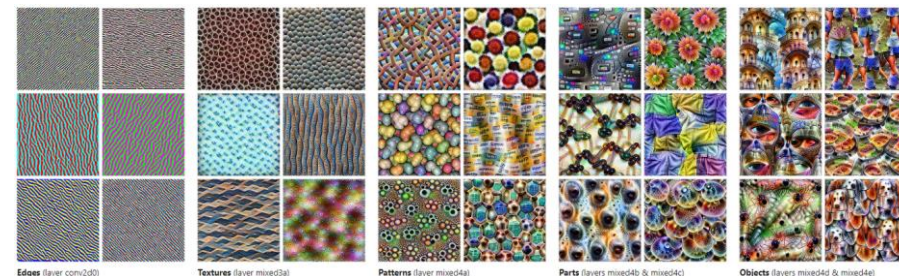
positive X-ray



DNN training on the Information Plane



a learning curve



feature visualization

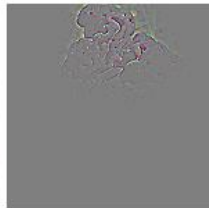
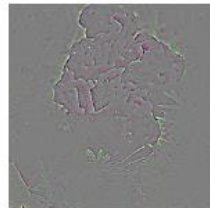
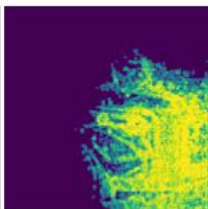
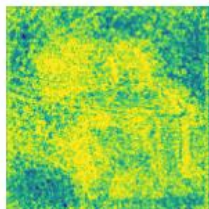
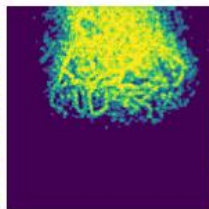
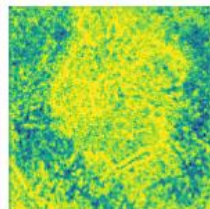
Stadelmann, Amirian, Arabaci, Arnold, Duivesteyn, Elezi, Geiger, Lörwald, Meier, Rombach & Tugener (2018). «Deep Learning in the Wild». ANNPR'2018.

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

	Original	Adversarial	Original	Adversarial
Image:				
Feature response:				
Local spatial entropy:				

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?**



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



Academic Members



National & International Partners



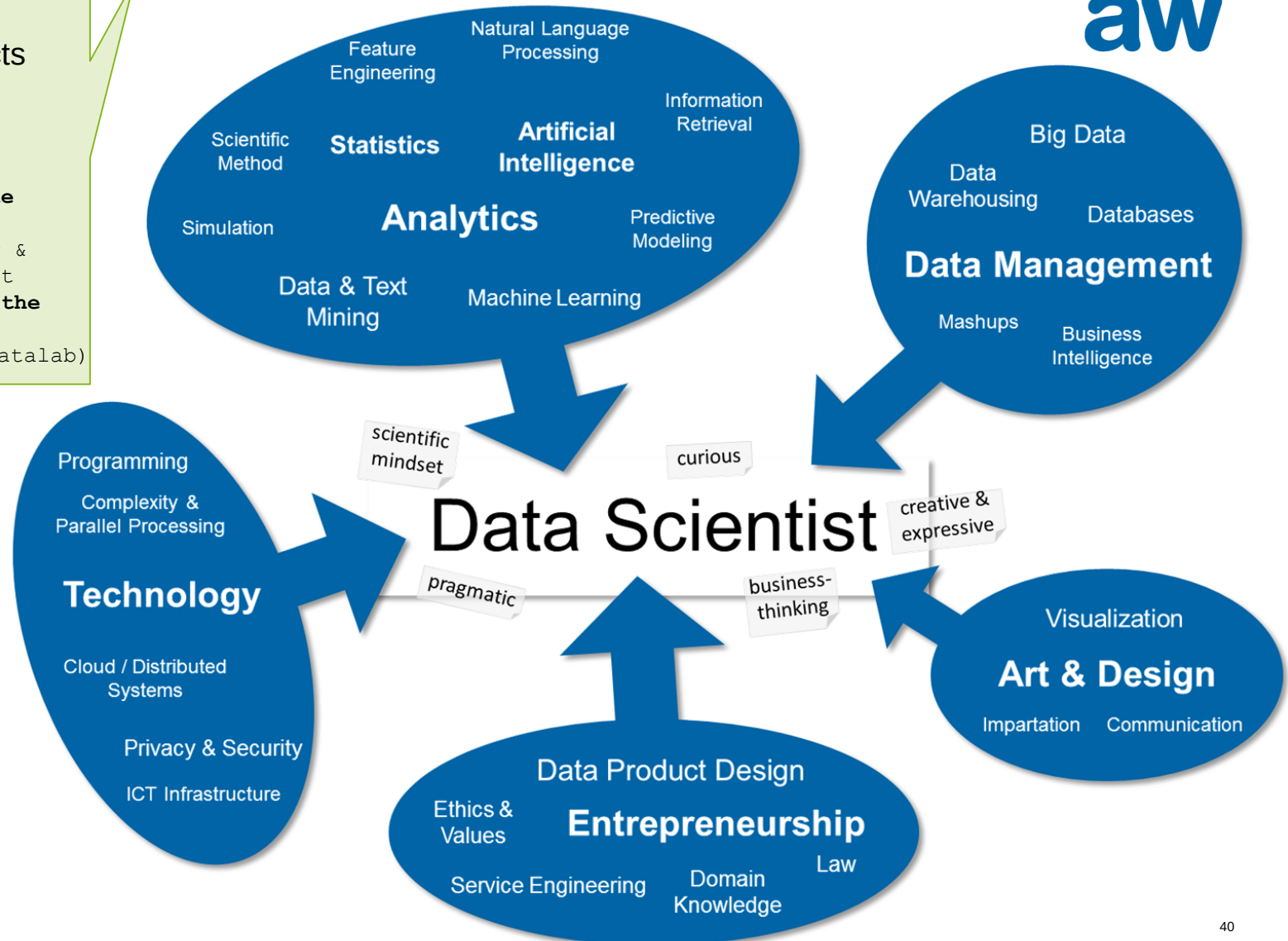
What is Data Science?

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)



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.)

Grosse Ambitionen, kleines Budget



Ein Macho auf Eggenried



Combination

Combination of rules, visual and textual features

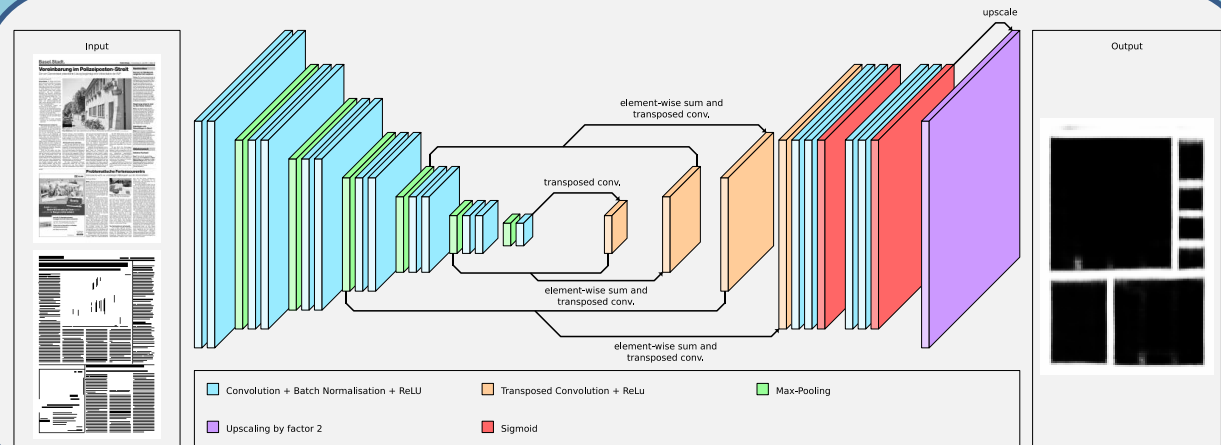


Final segmentation



Result

Most Successful Approach [3]



References

- [1] D. C. Cireşan, 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 ICAR*, Kyoto, Japan, 2018.

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



Big Data
Nationales Forschungsprogramm



FONDS NATIONAL SUISSE
SCHWEIZERISCHER NATIONALFONDS
FONDO NAZIONALE SVIZZERO
SWISS NATIONAL SCIENCE FOUNDATION

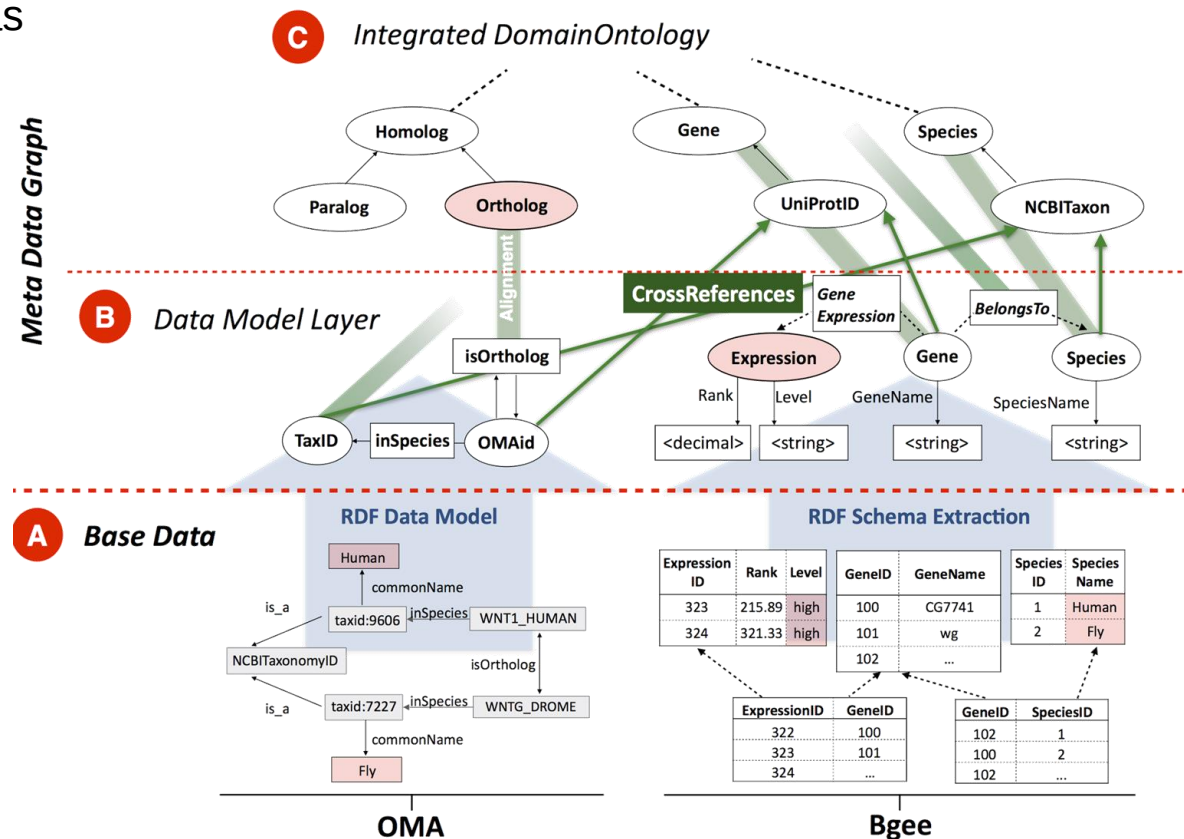
Zürich University
of Applied Sciences



Unil
UNIL | Université de Lausanne



Swiss Institute of
Bioinformatics

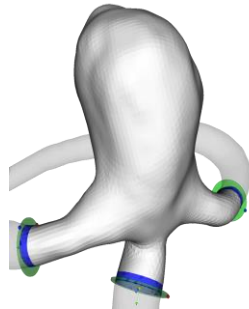


AneuX: Ist die Form signifikant für die Gefährdung eines Aneurysmas?

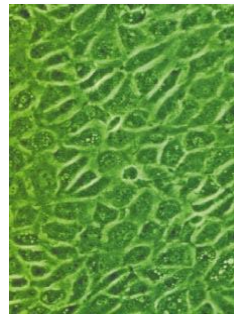
Aneurysm im
Röntgenbild (XA)



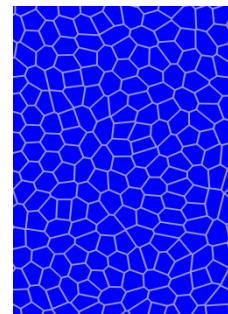
Isoliertes Aneurysma
Zur Formanalyse



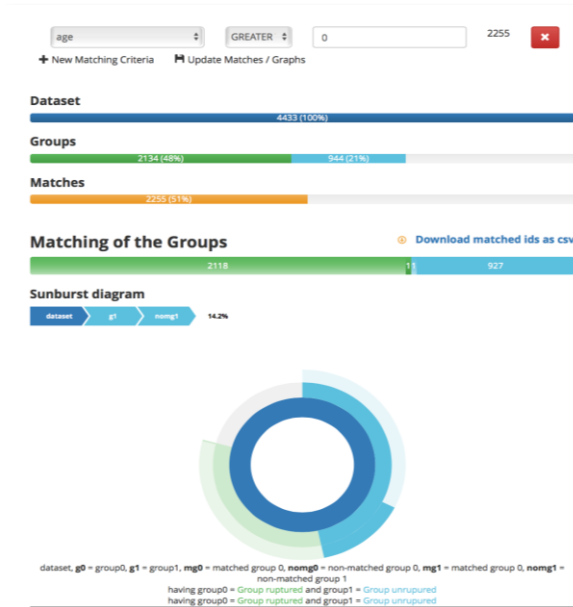
Zellen der
Gefäßwand



Modell der
Gefäßwand



Webtool für statistische Analyse

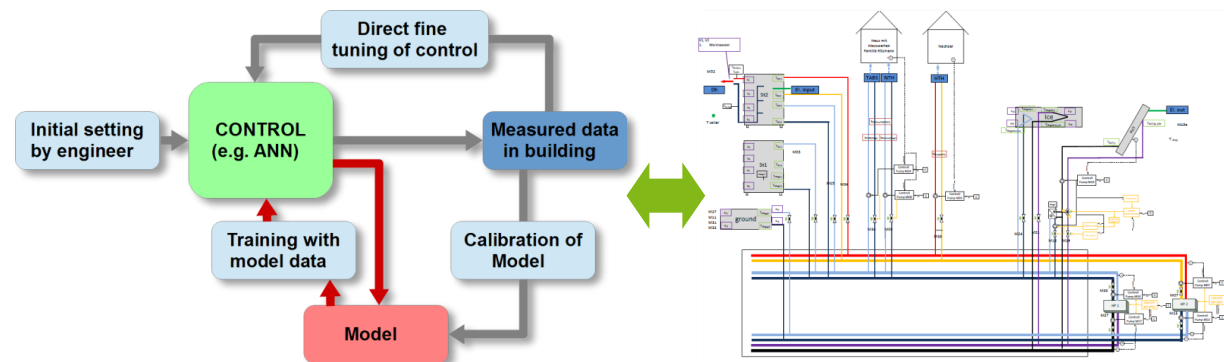


SystemsX.ch funding: 2M CHF, Begutachtung SNSF

- 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 Werkzeugen zur Analyse der klinischen Daten und Bilddaten

Partner (Co-Antragsteller Sven Hirsch, ZHAW):

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**

Game playing

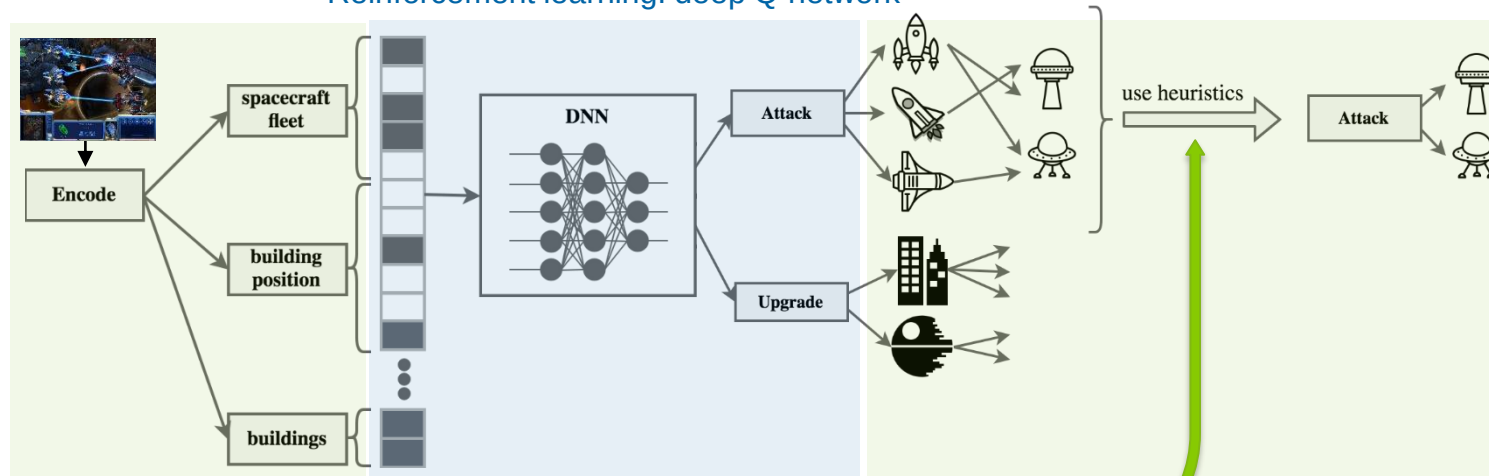


(symbolic figure)



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



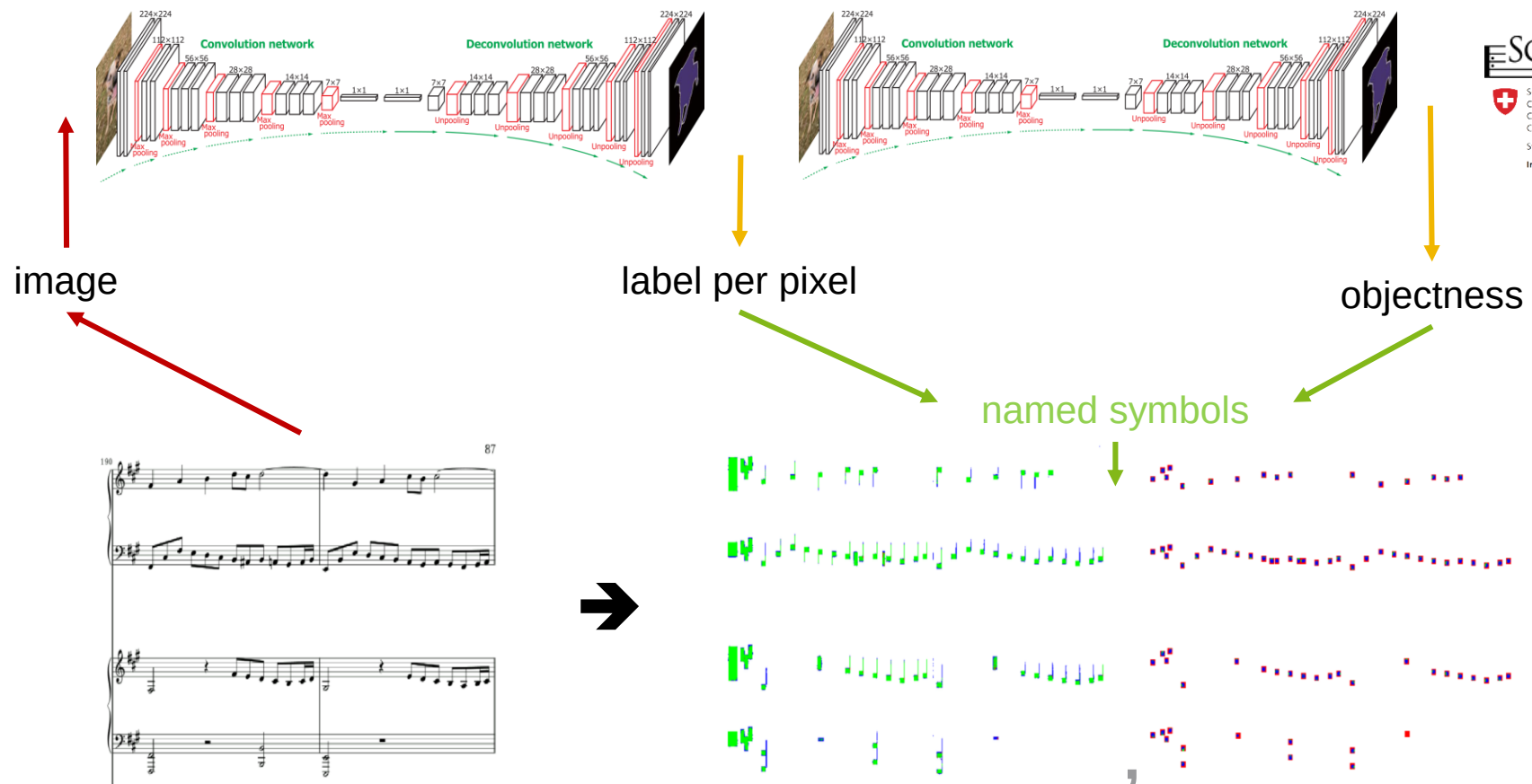
Distance encoding → use reference points

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

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

Optical Music Recognition

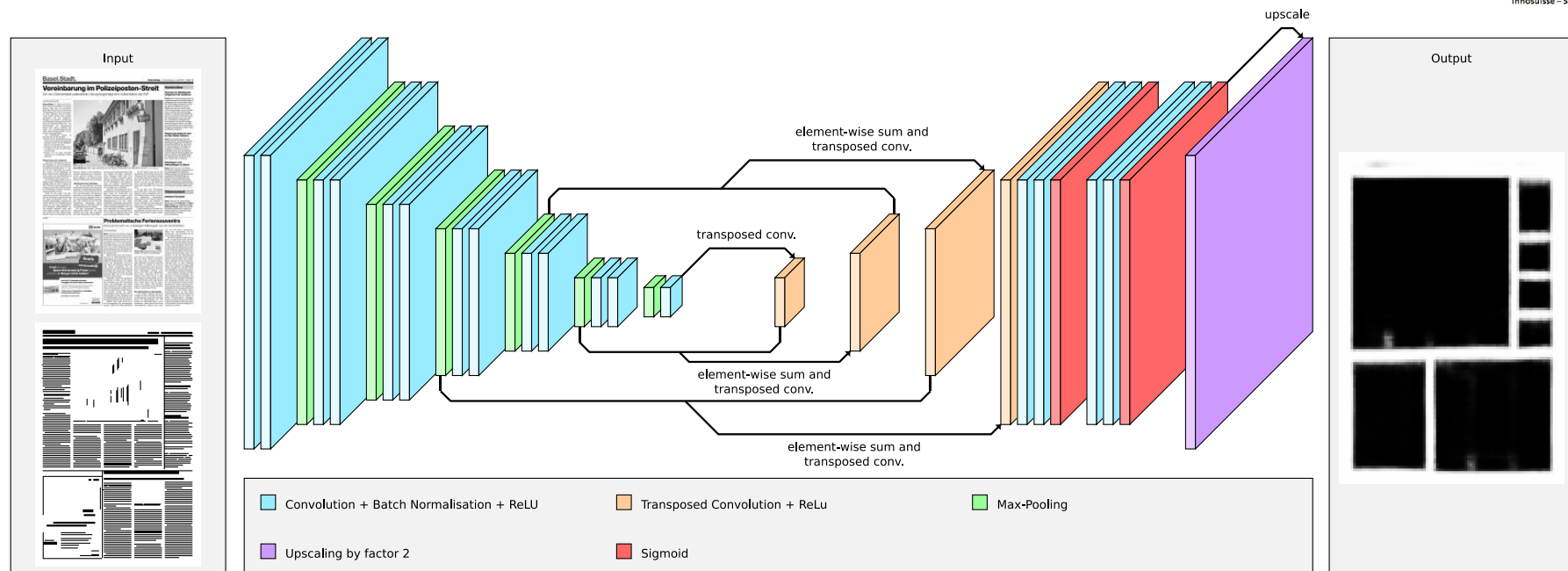
Foundation of digitization in orchestras and music schools



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



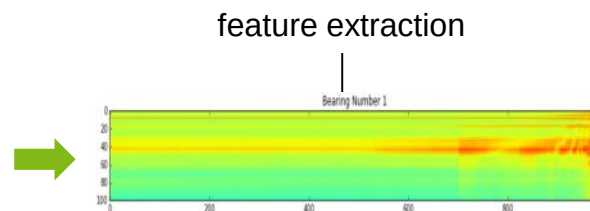
Meier, Stadelmann, Stampfli, Arnold & Cieliebak (2017). «Fully Convolutional Neural Networks for Newspaper Article Segmentation». ICDAR'2017.
Stadelmann, Tolkachev, Sick, Stampfli & Dürr (2018). «Beyond ImageNet - Deep Learning in Industrial Practice». In: Braschler et al., «Applied Data Science», Springer.

Condition monitoring

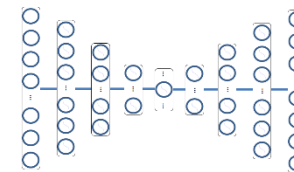
Maintaining machines on predicted failure only

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.

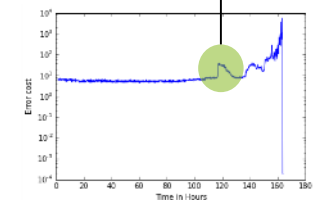
vibration sensors



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.