Zürcher Hochschule
für Angewandte Wissenschaften

Industrielle Anwendungsmöglichkeiten für Deep Learning-basierte Künstliche Intelligenz

Endress+Hauser Technologieforum, Sternenhof Auditorium, Reinach BL
01. Februar 2019

Thilo Stadelmann

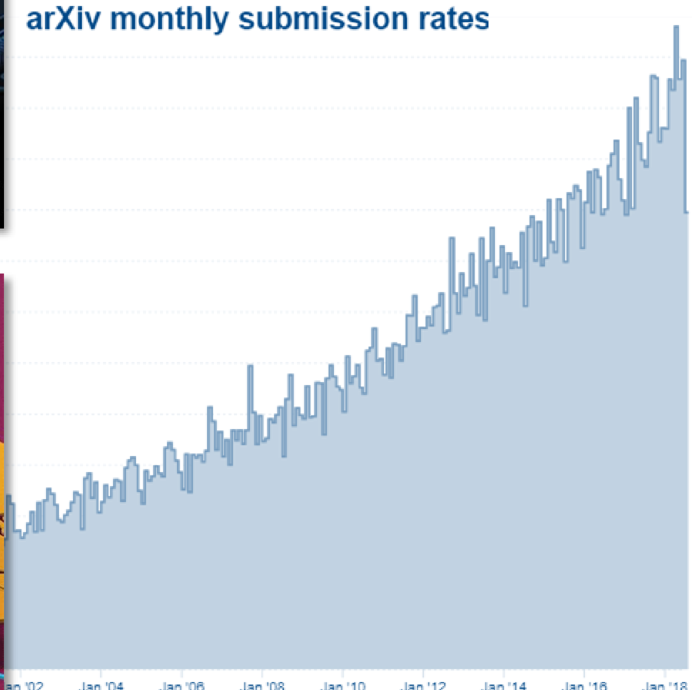
Why?


nature
At last — a computer program that can beat a champion Go player PAGE 484
ALL SYSTEMS GO
SONGBIRDS À LA CARTE
SAFEGUARD TRANSPARENCY
WHEN GENES GOT 'SELFISH'

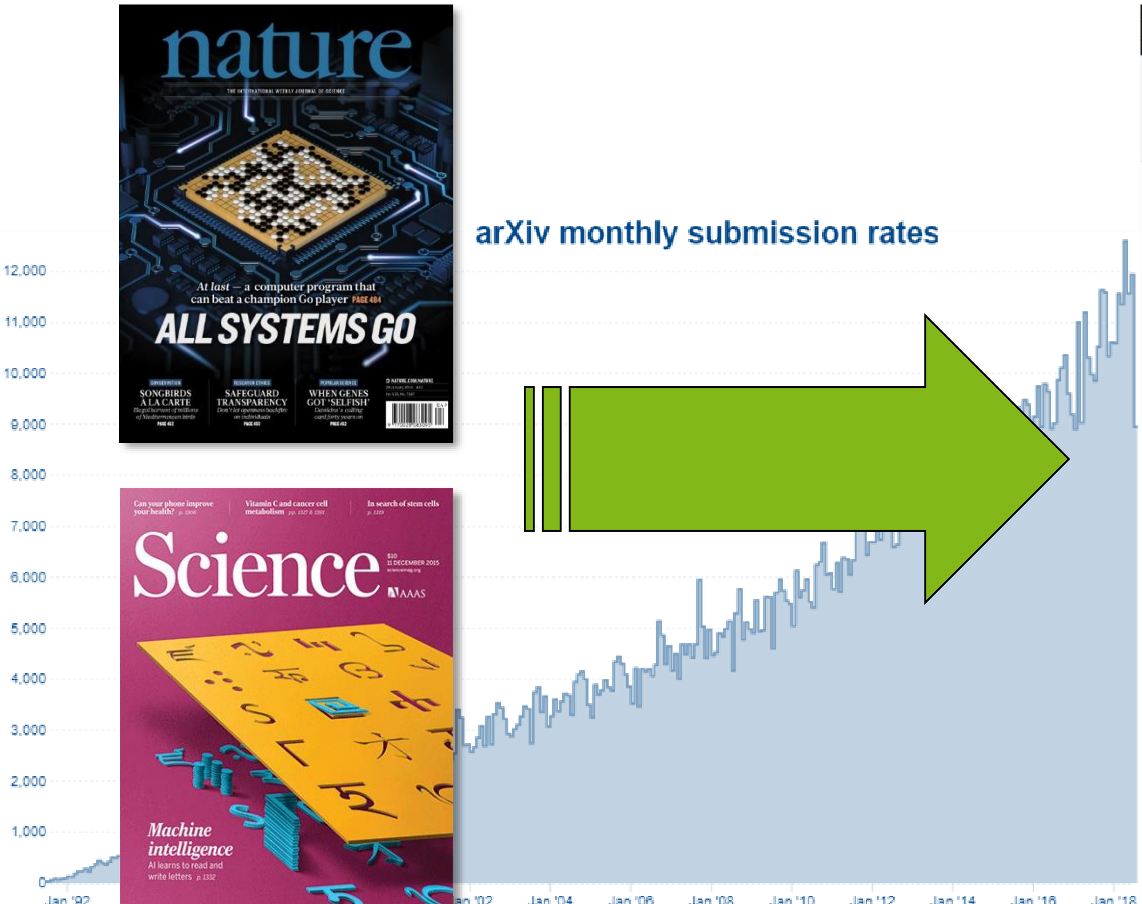

Can your phone improve your health?
Vitamin C and cancer cell metabolism
In search of stem cells
Science
11 DECEMBER 2015
AAAS
Machine intelligence
AI learns to read and write letters p. 1332

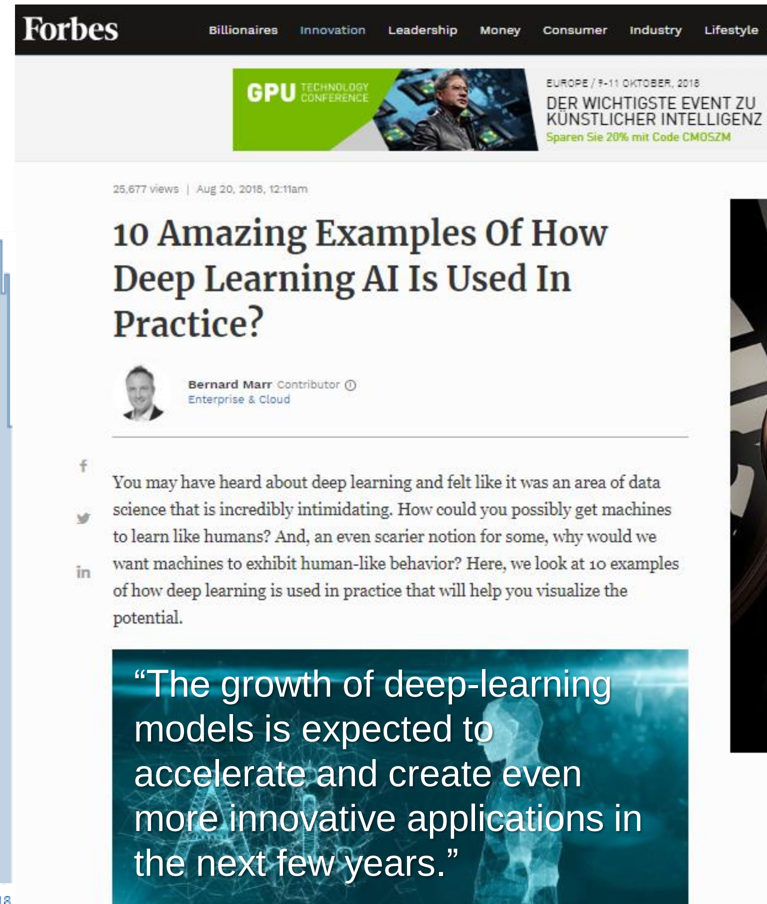
Why?

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

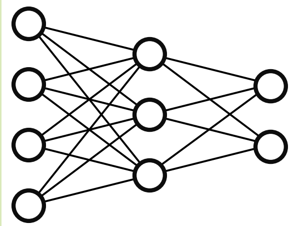
Classical image processing

Feature extraction
(SIFT, SURF, LBP, HOG, etc.)

(0.2, 0.4, …)

(0.4, 0.3, …)

Classification
(SVM, neural network, etc.)

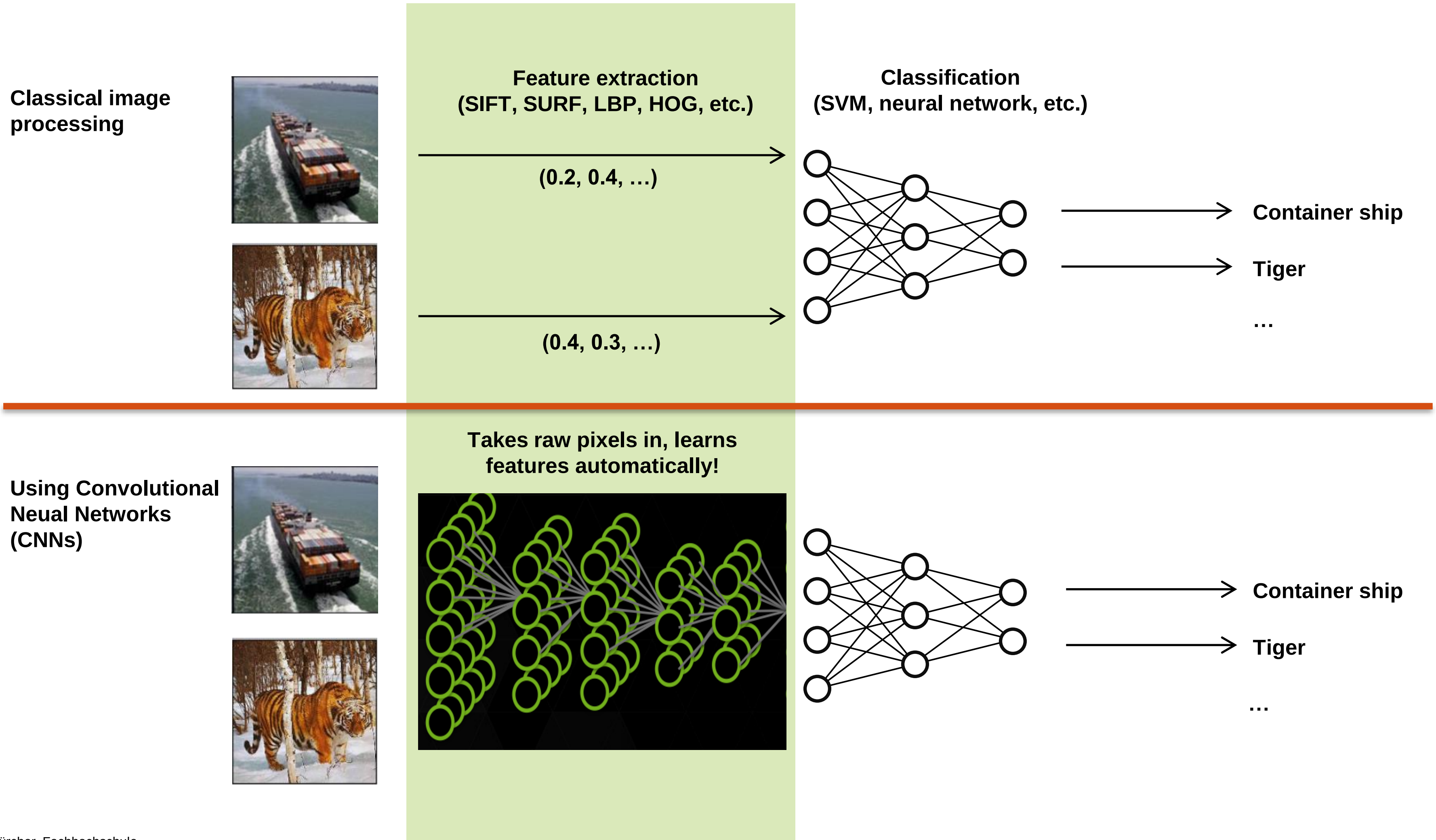


Container ship

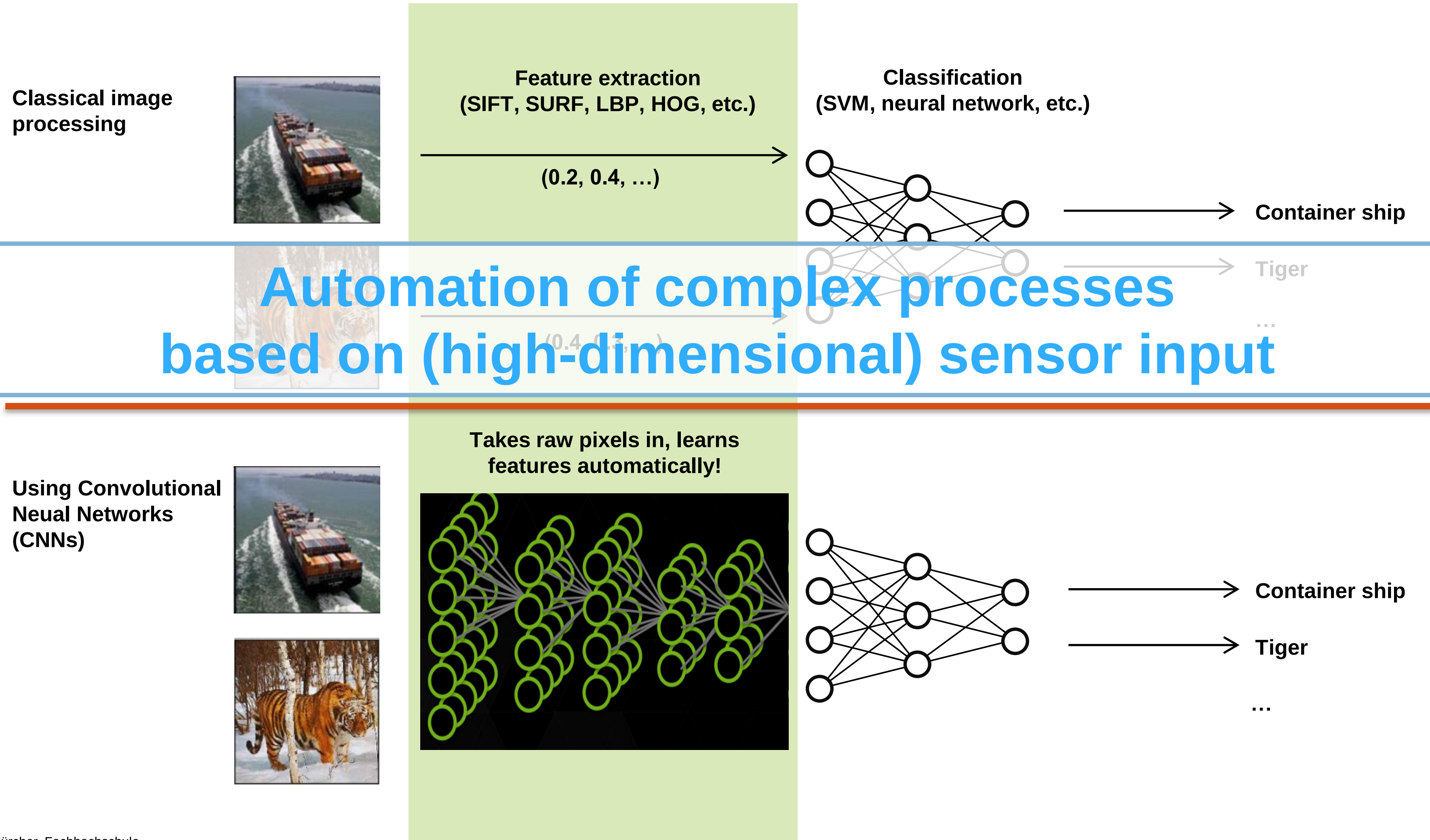
Tiger

…

Idea: Add depth to learn features automatically



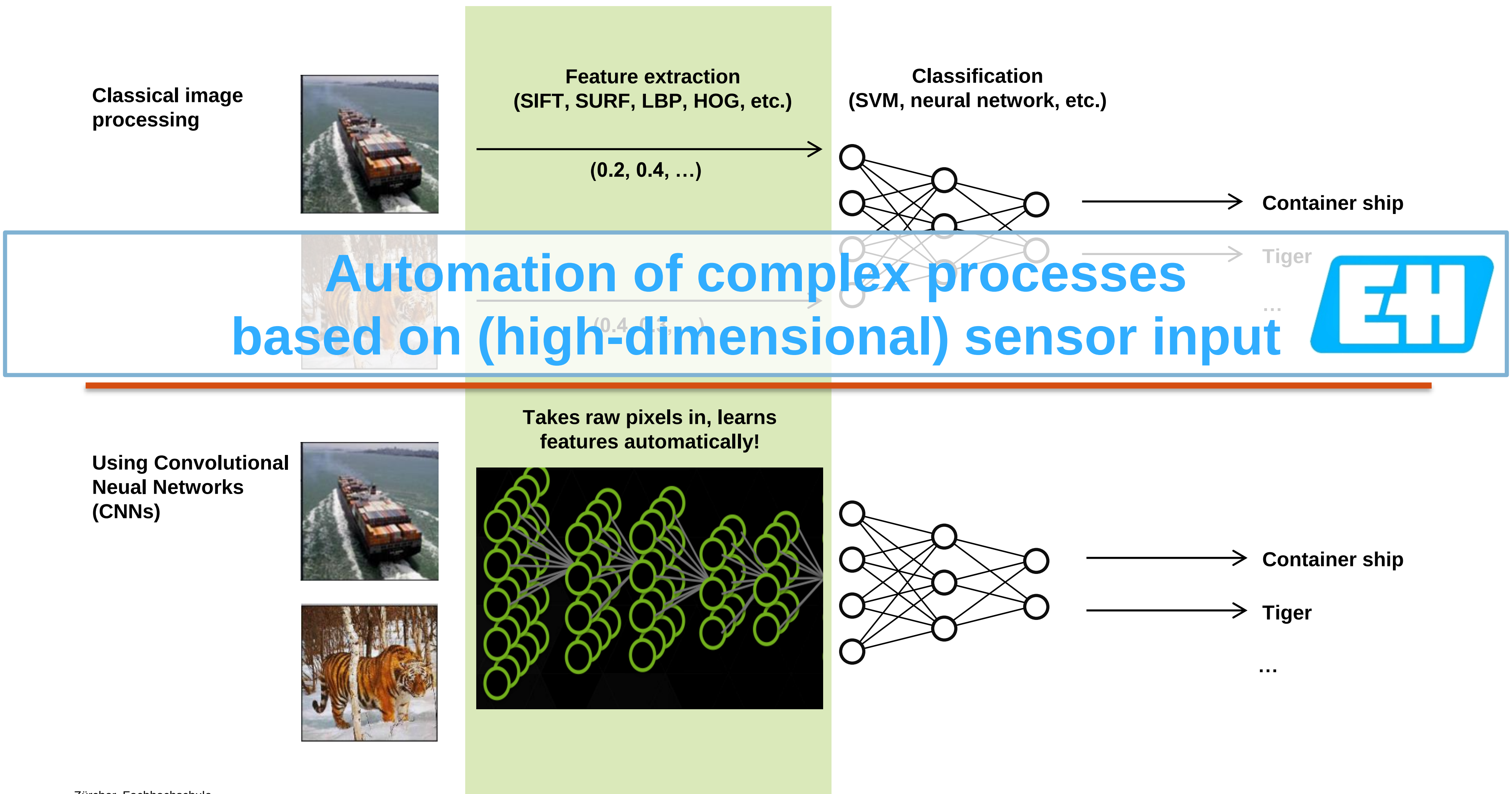
Idea: Add depth to learn features automatically



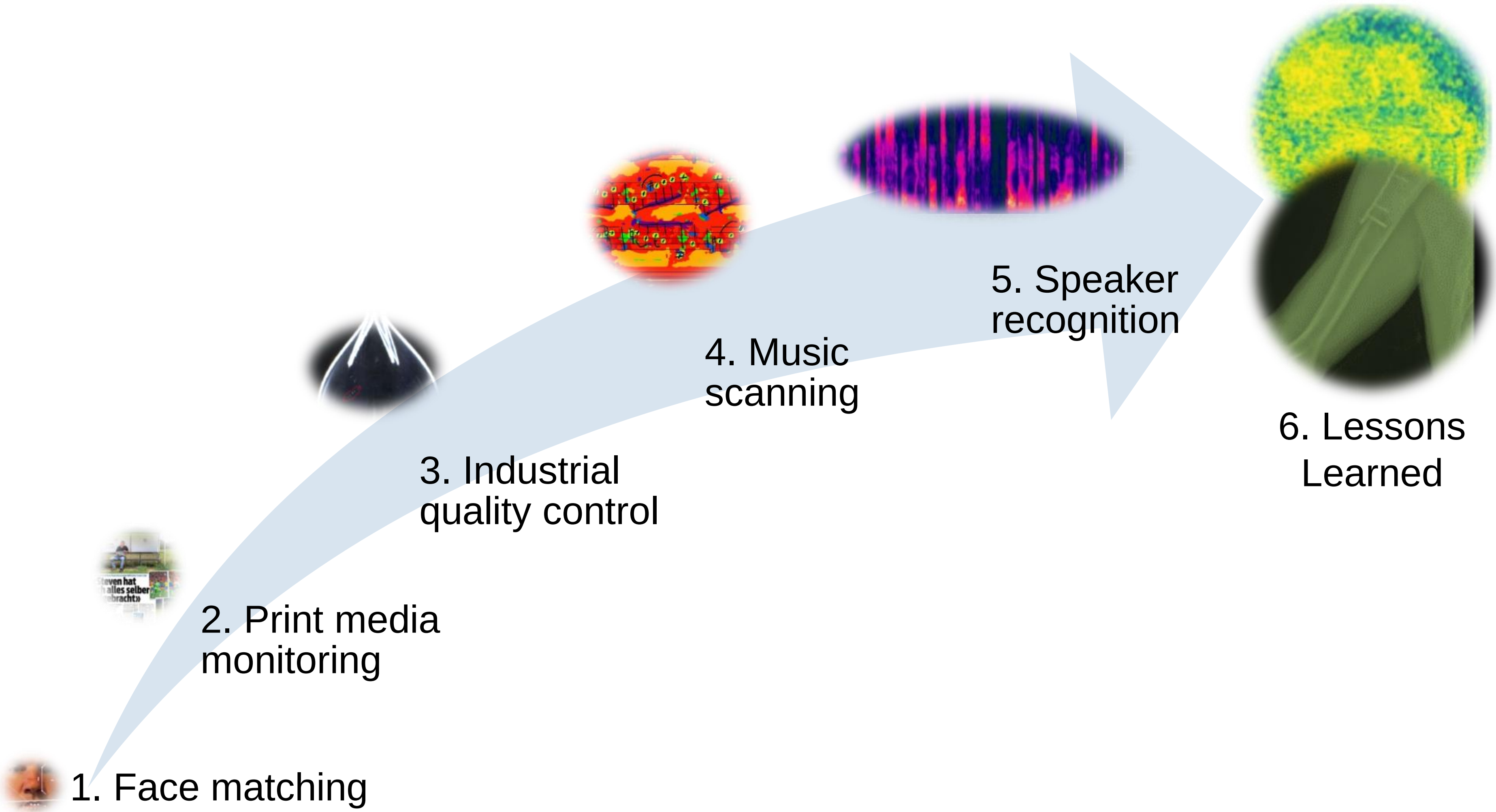
Automation of complex processes based on (high-dimensional) sensor input

Zürcher Hochschule
für Angewandte Wissenschaften

Idea: Add depth to learn features automatically



Agenda


1. Face matching
2. Print media monitoring
3. Industrial quality control
4. Music scanning
5. Speaker recognition
6. Lessons Learned

1. Face matching

1. Face matching

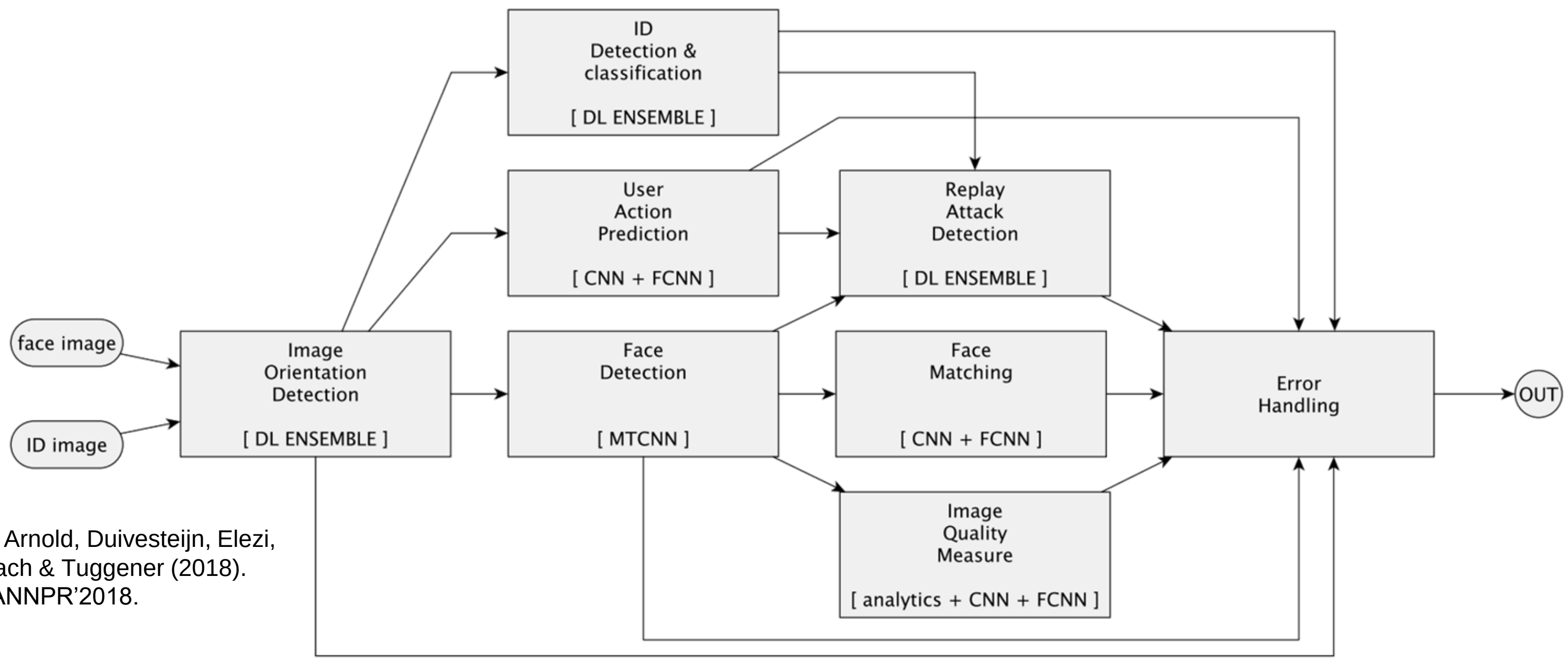
1. Face matching – challenges & solutions

DEEPIMPACT

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

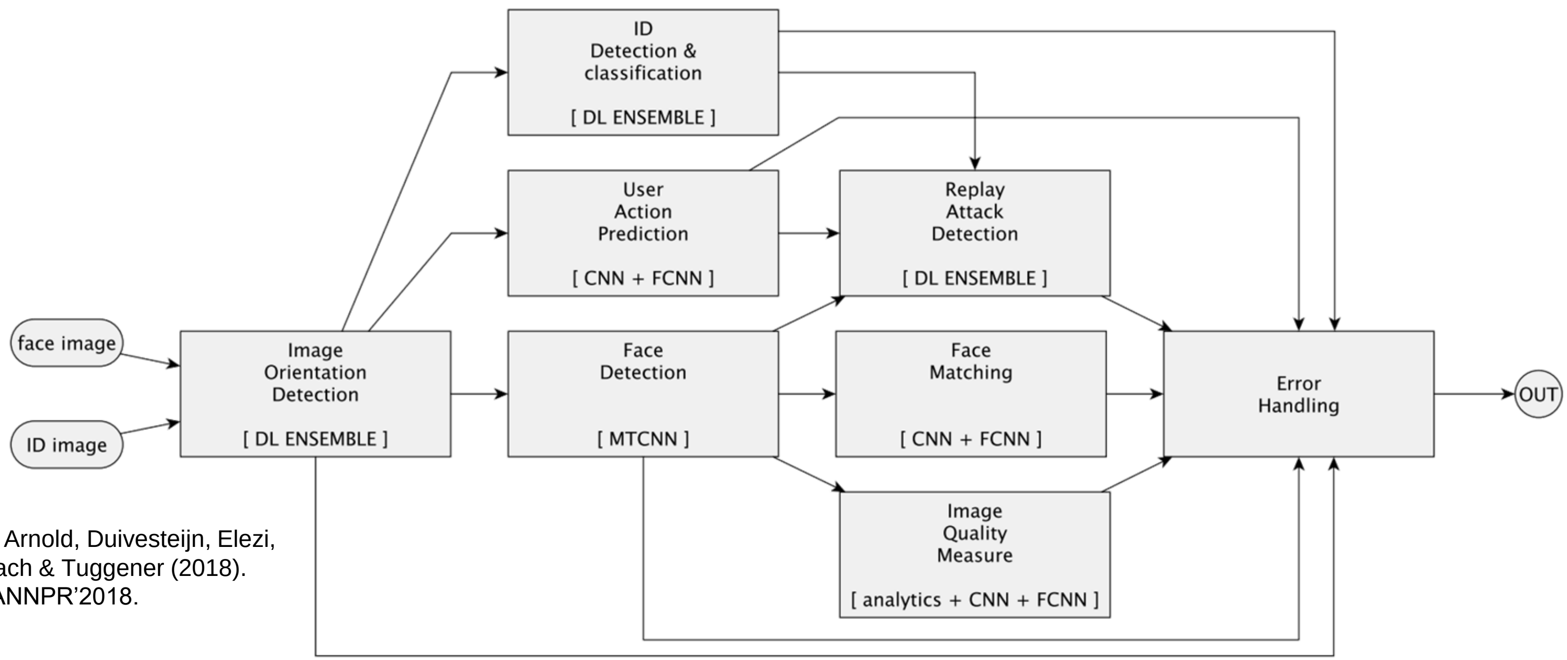
Innosuisse – Swiss Innovation Agency



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

1. Face matching – challenges & solutions





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

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

Zürcher Hochschule
für Angewandte Wissenschaften

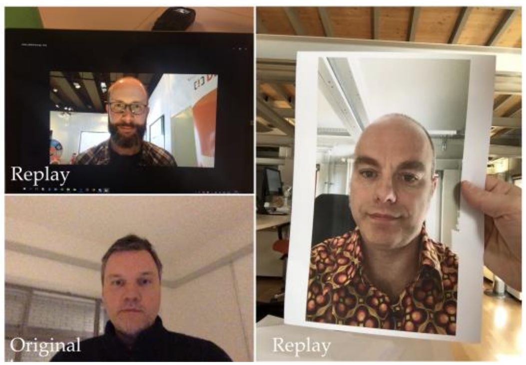
1. Face matching – challenges & solutions

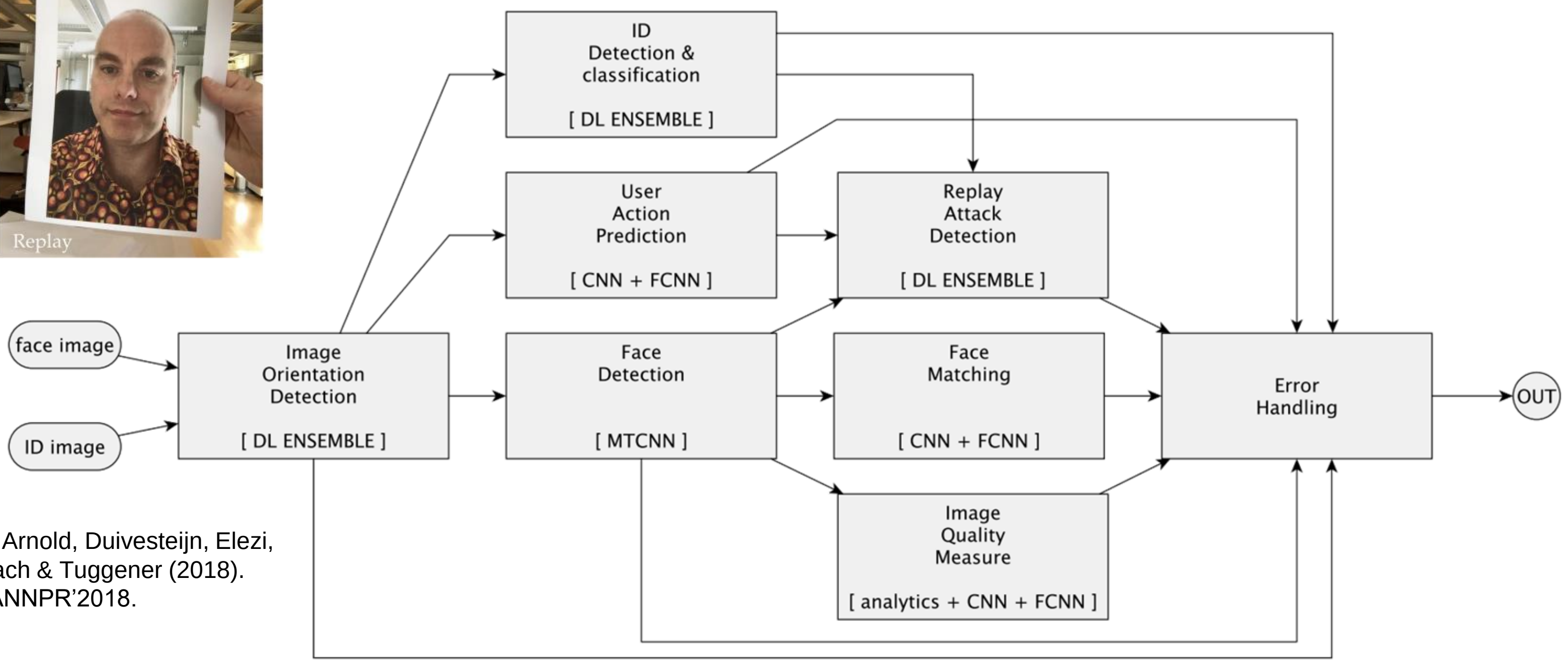
[!] DEEPIMPACT

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.

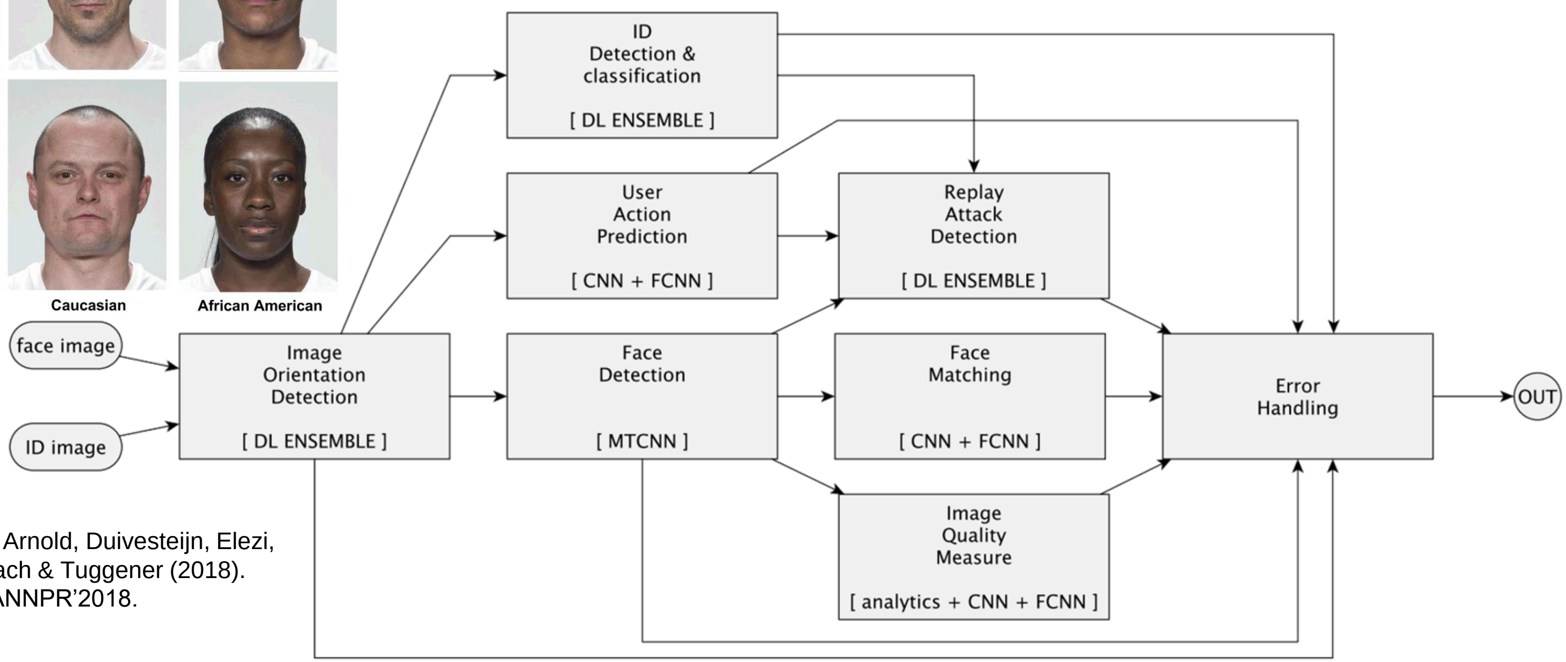
1. Face matching – challenges & solutions

[!] 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.

2. 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
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
Wochenpreis: 1 x sieben Nächte für 2 Personen, inkl. HP, im ****Seehotel Pilatus Hergiswil im Wert von 3000 Franken!

2. Print media monitoring – ML solution

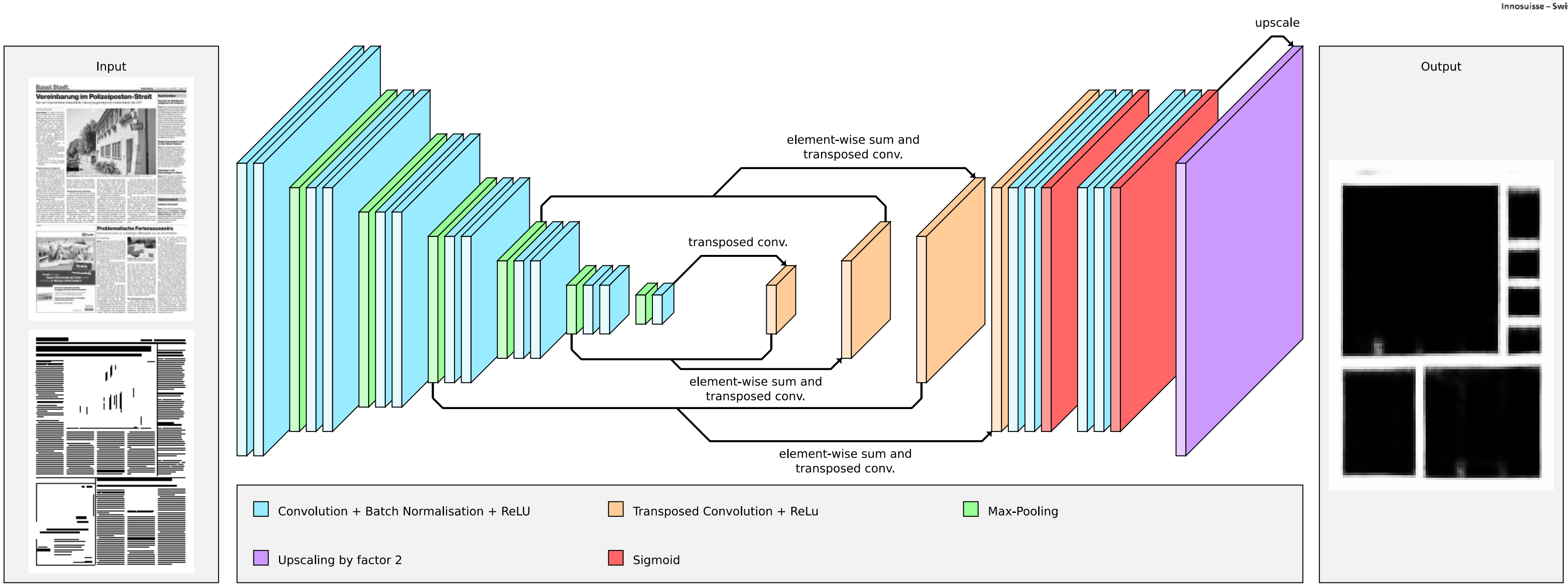
Zürcher Hochschule
für Angewandte Wissenschaften

ARGUS DATA INSIGHTS
WISSEN ZUM ERFOLG

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

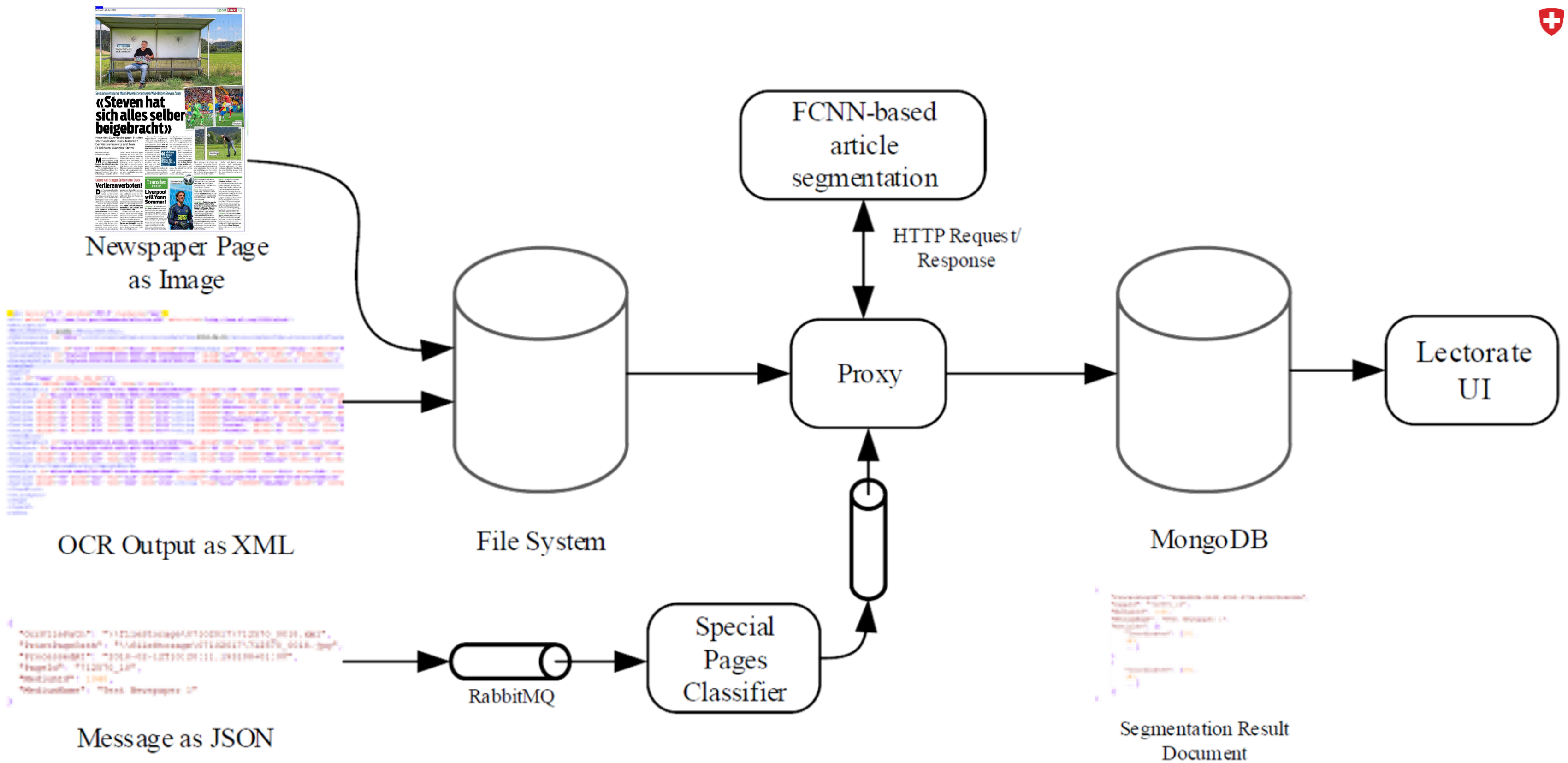
Swiss Confederation

Innosuisse – Swiss Innovation Agency



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.

2. Print media monitoring – deployment



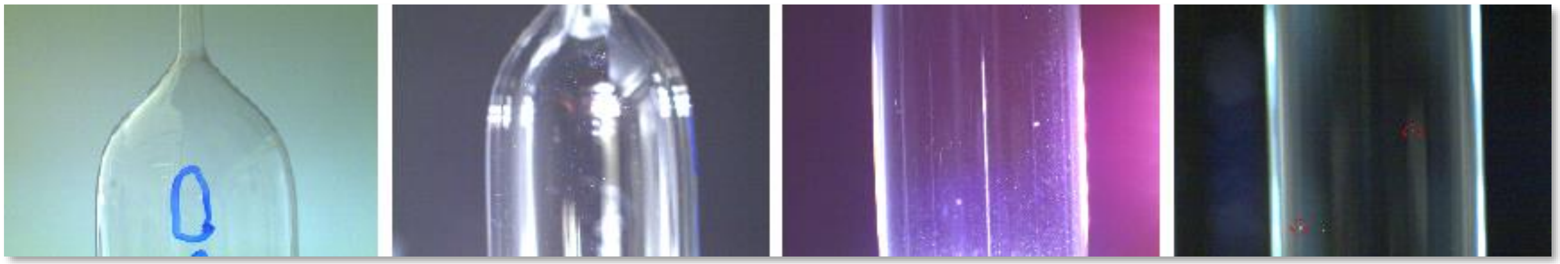
Stadelmann, Amirian, Arabaci, Arnold, Duivesteijn, Elezi, Geiger, Lörwald, Meier, Rombach & Tuggener (2018). *«Deep Learning in the Wild»*. ANNPR'2018.

3. Industrial quality control

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

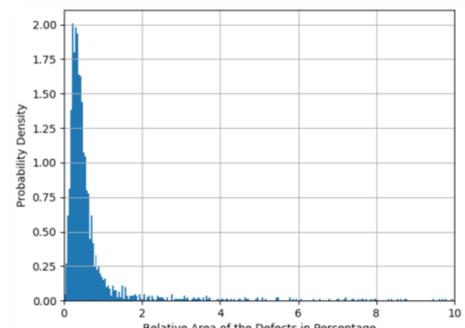
Task

- Reliably **sort out faulty balloon catheters** in image-based production quality control



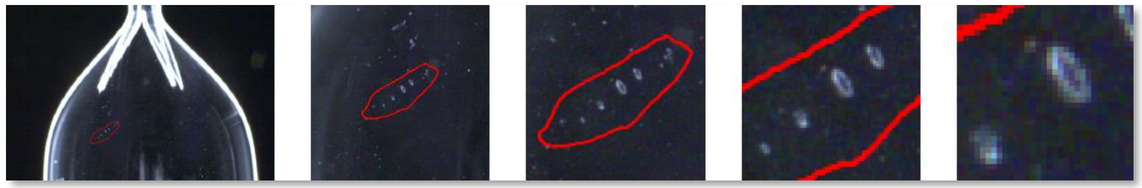
Challenges

- **Non-natural** image source, class **imbalance**, **optic**al conditions, **variation** in defect size & shape

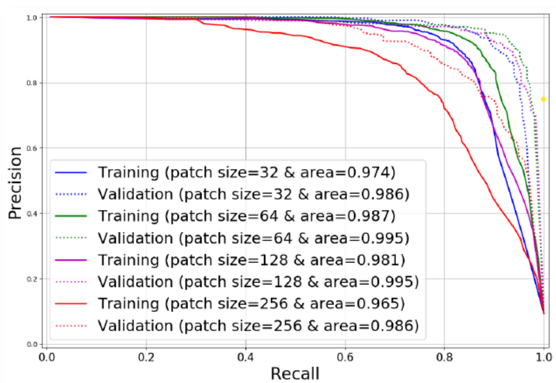


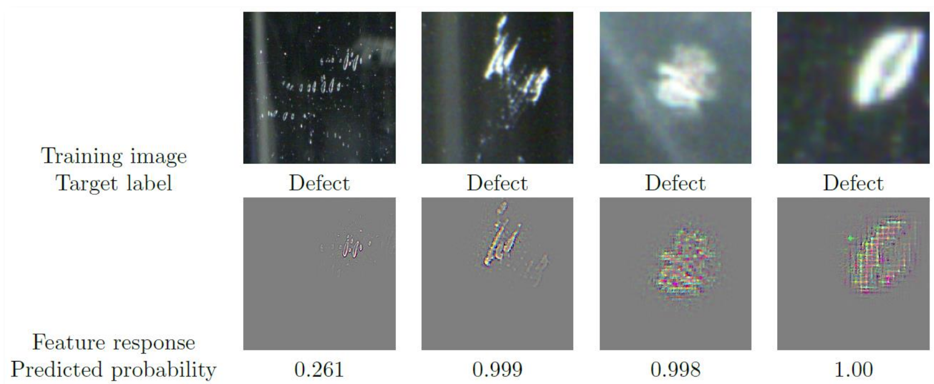
3. Industrial quality control – baseline results

Ingredients
- Weighted loss
- Defect cropping
- Careful customization



Interm results





3. Industrial quality control – recent results
(Work in progress)

- Human performance isn't flawless

3. Industrial quality control – recent results
(Work in progress)

- Human performance isn't flawless

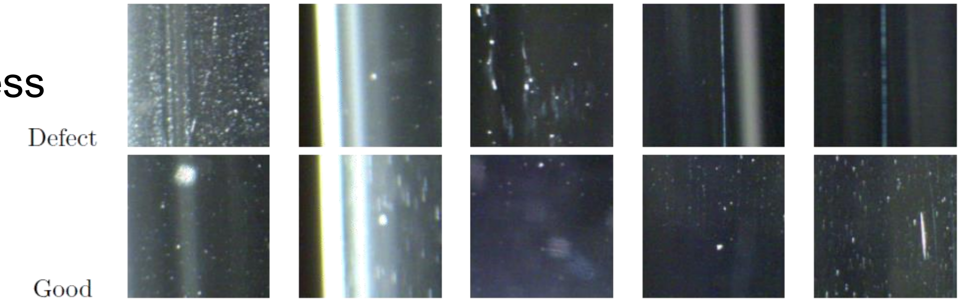


Figure 2: Samples of failure cases in classification. The shown *defect* samples in the table are not recognized as a defects, and the *good* images are misclassified as defects.

3. Industrial quality control – recent results
(Work in progress)

Zürcher Hochschule
für Angewandte Wissenschaften
zh
aw

- Human performance isn't flawless
- Tailoring pays off

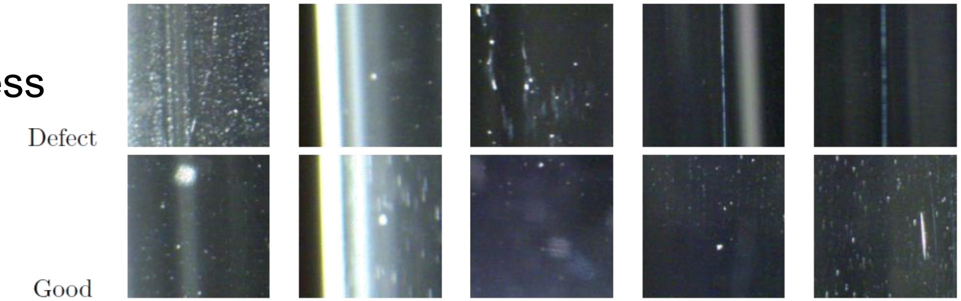
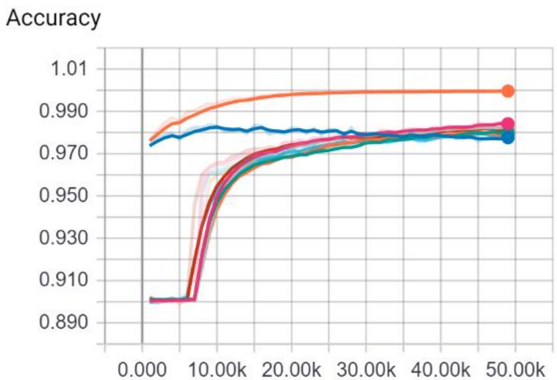


Figure 2: Samples of failure cases in classification. The shown *defect* samples in the table are not recognized as a defects, and the *good* images are misclassified as defects.

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



Name	Smoothed	Value	Step	Time	Relative
QualitAI_VGG19_Full_Pretrained\train	0.9996	0.9996	49.00k	Tue Jan 22, 02:32:13	8h 30m 56s
QualitAI_VGG19_Full_Pretrained\validation	0.9776	0.9783	49.00k	Tue Jan 22, 02:32:24	8h 30m 56s
QualitAI_VGG19_Full_Random\train	0.9841	0.9841	49.00k	Thu Jan 24, 19:28:02	10h 29m 2s
QualitAI_VGG19_Full_Random\validation	0.9798	0.9798	49.00k	Thu Jan 24, 19:28:14	10h 29m 2s
QualitAI_VGG19_Half\train	0.9827	0.9835	49.00k	Thu Jan 24, 13:01:47	4h 9m 12s
QualitAI_VGG19_Half\validation	0.9792	0.9798	49.00k	Thu Jan 24, 13:01:54	4h 9m 11s
QualitAI_VGG19_Quarter\train	0.9817	0.9823	49.00k	Thu Jan 24, 10:53:52	2h 17m 21s
QualitAI_VGG19_Quarter\validation	0.9791	0.9806	49.00k	Thu Jan 24, 10:53:56	2h 17m 21s

3. Industrial quality control – recent results
(Work in progress)

Zürcher Hochschule
für Angewandte Wissenschaften
zh
aw

- Human performance isn't flawless
- Tailoring pays off
- Data shortage may be outsmarted

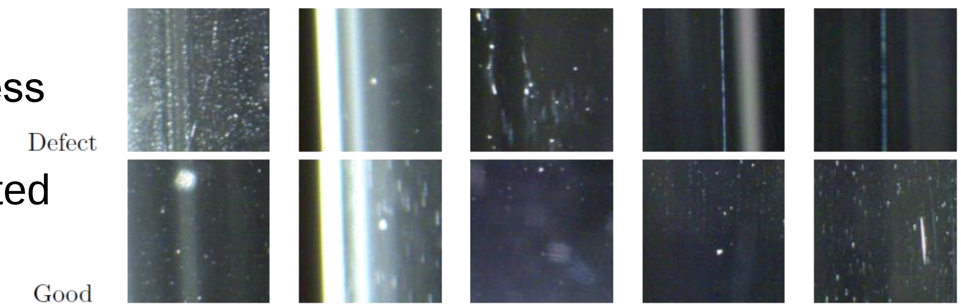
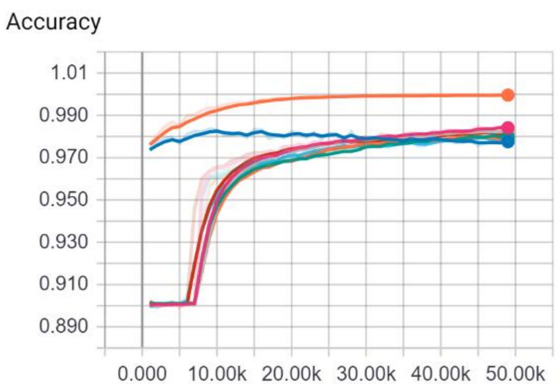
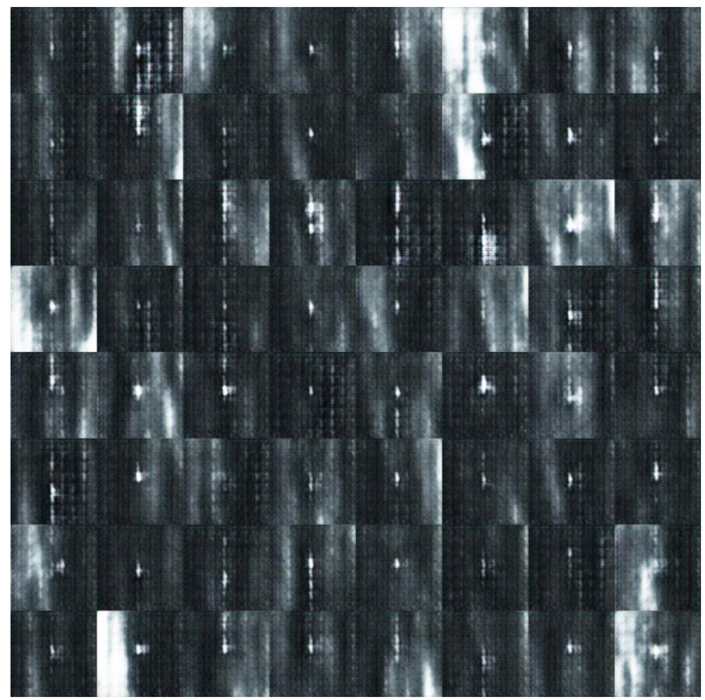


Figure 2: Samples of failure cases in classification. The shown *defect* samples in the table are not recognized as a defects, and the *good* images are misclassified as defects.

BW-TEC
Engineering Automation Consulting
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Innosuisse – Swiss Innovation Agency



Name	Smoothed	Value	Step	Time	Relative
QualitAI_VGG19_Full_Pretrained\train	0.9996	0.9996	49.00k	Tue Jan 22, 02:32:13	8h 30m 56s
QualitAI_VGG19_Full_Pretrained\validation	0.9776	0.9783	49.00k	Tue Jan 22, 02:32:24	8h 30m 56s
QualitAI_VGG19_Full_Random\train	0.9841	0.9841	49.00k	Thu Jan 24, 19:28:02	10h 29m 2s
QualitAI_VGG19_Full_Random\validation	0.9798	0.9798	49.00k	Thu Jan 24, 19:28:14	10h 29m 2s
QualitAI_VGG19_Half\train	0.9827	0.9835	49.00k	Thu Jan 24, 13:01:47	4h 9m 12s
QualitAI_VGG19_Half\validation	0.9792	0.9798	49.00k	Thu Jan 24, 13:01:54	4h 9m 11s
QualitAI_VGG19_Quarter\train	0.9817	0.9823	49.00k	Thu Jan 24, 10:53:52	2h 17m 21s
QualitAI_VGG19_Quarter\validation	0.9791	0.9806	49.00k	Thu Jan 24, 10:53:56	2h 17m 21s



4. Music scanning


```
<?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/>
```

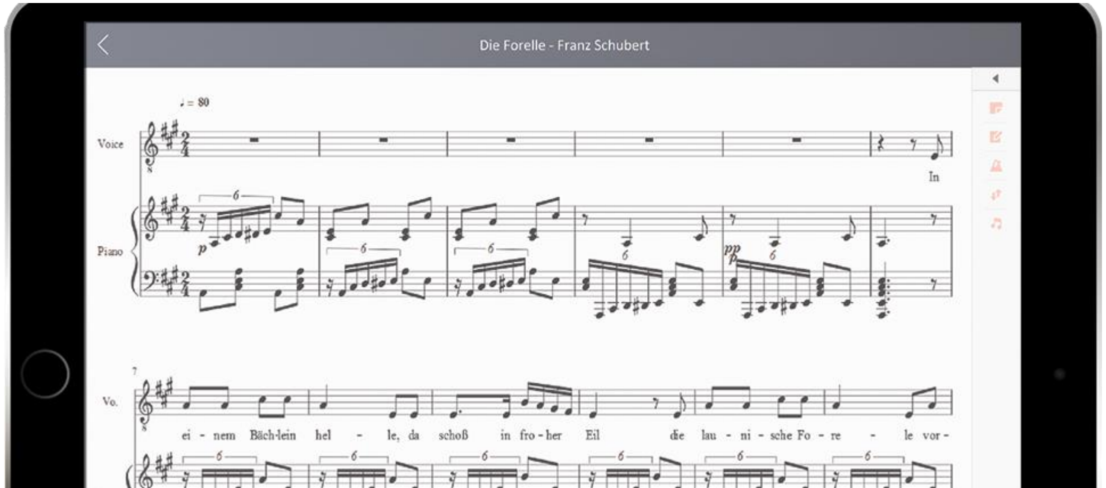

Zürcher Hochschule für Angewandte Wissenschaften

SCOREPAD

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

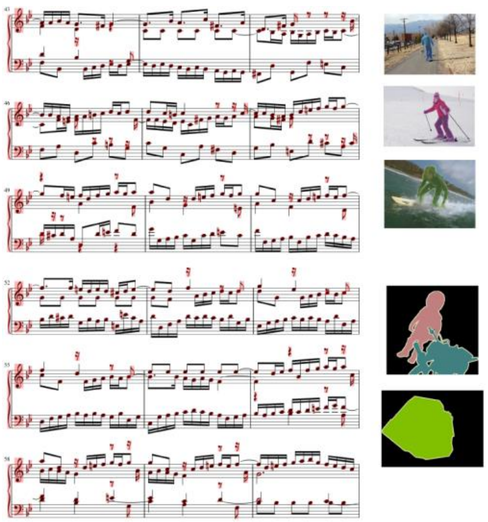
Innosuisse – Swiss Innovation Agency



4. Music scanning – challenges & solutions

Tuggener, Elezi, Schmidhuber, Pelillo & Stadelmann (2018). *«DeepScores – A Dataset for Segmentation, Detection and Classification of Tiny Objects»*. ICPR’2018.

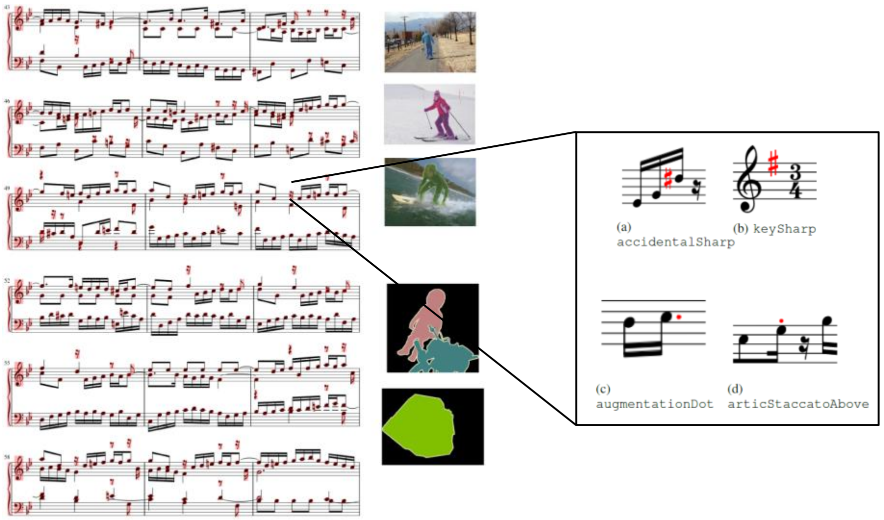
4. Music scanning – challenges & solutions



Tuggener, Elezi, Schmidhuber, Pelillo & Stadelmann (2018). *«DeepScores – A Dataset for Segmentation, Detection and Classification of Tiny Objects»*. ICPR’2018.

Zürcher Hochschule
für Angewandte Wissenschaften

4. Music scanning – challenges & solutions



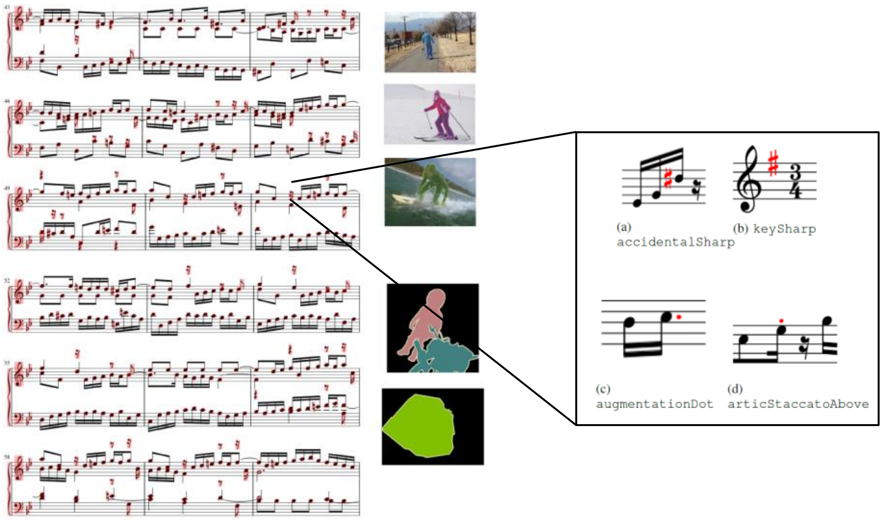
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

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.

4. Music scanning – challenges & solutions



,

SCOREPAD

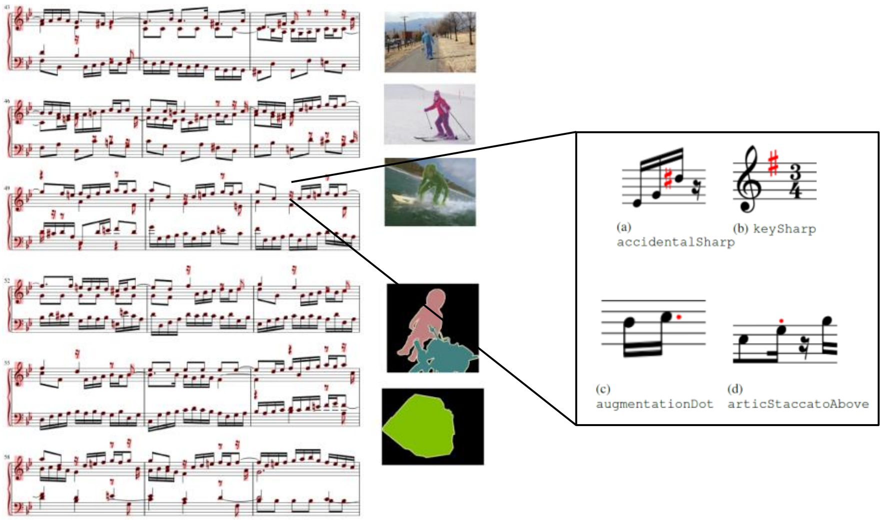
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

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.

4. Music scanning – challenges & solutions



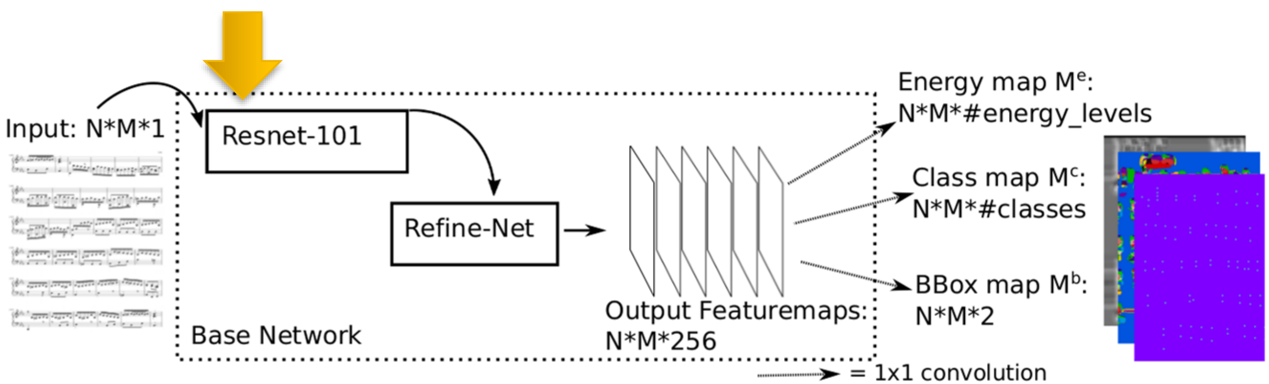
,

SCOREPAD

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

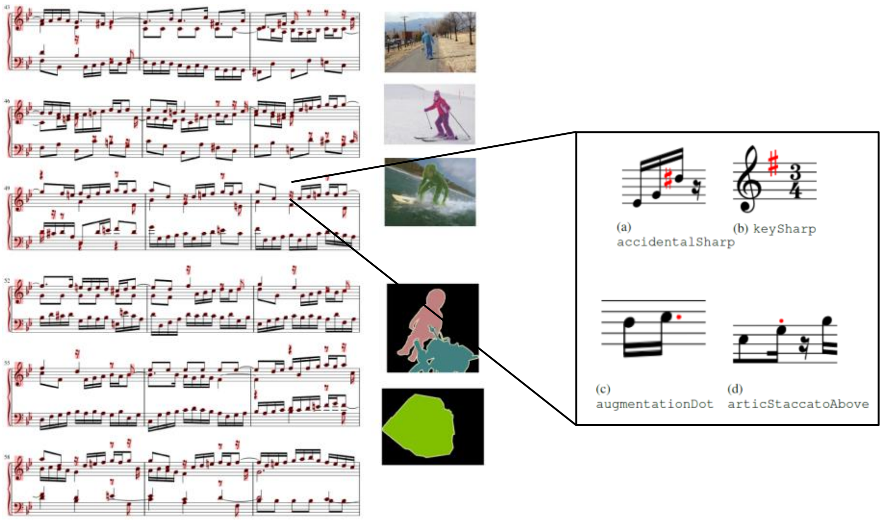
Swiss Confederation

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. Music scanning – challenges & solutions

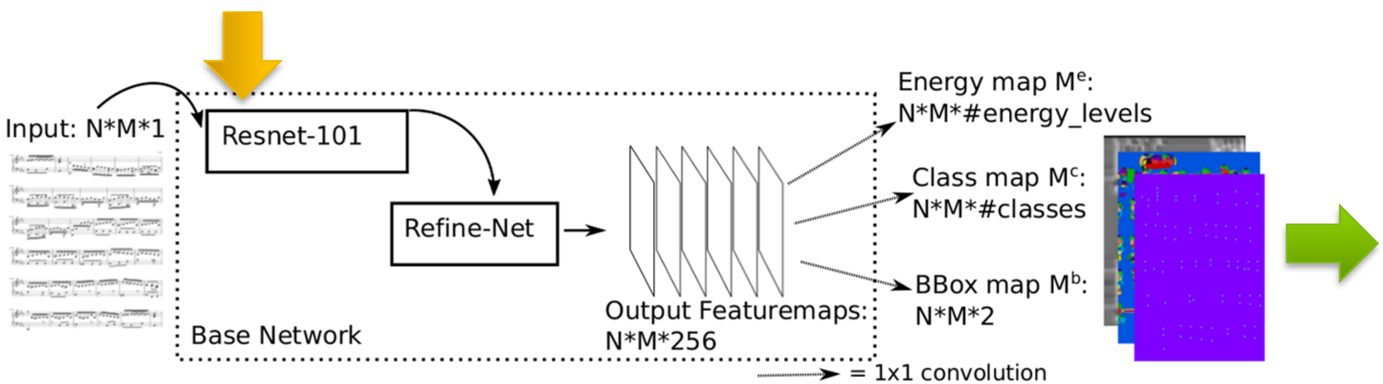


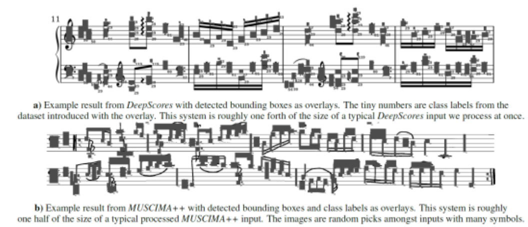
,

Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Innosuisse – Swiss Innovation Agency



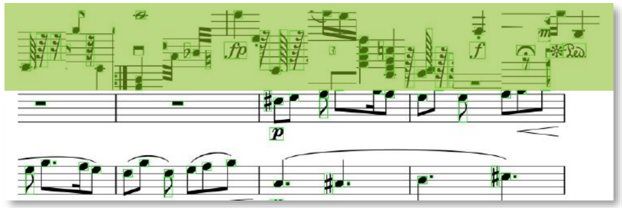


Tuggener, Elezi, Schmidhuber, Pelillo & Stadelmann (2018). *«DeepScores – A Dataset for Segmentation, Detection and Classification of Tiny Objects»*. ICPR'2018.
Tuggener, Elezi, Schmidhuber & Stadelmann (2018). *«Deep Watershed Detector for Music Object Recognition»*. ISMIR'2018.

4. Music scanning – industrialization

Recent results on class imbalance and robustness challenges

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

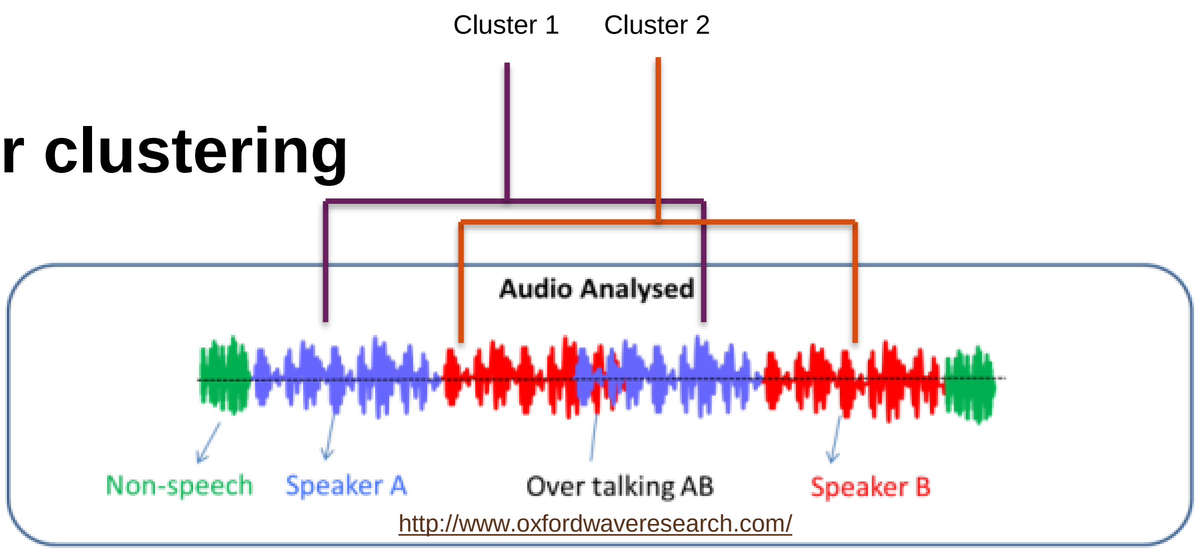


2. Put additional effort (and compute) into hyperparameter **tuning** and **longer training**
3. Trained also on scanned (more **real-worldish**) scores

➔ **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: 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.

5. Speaker clustering



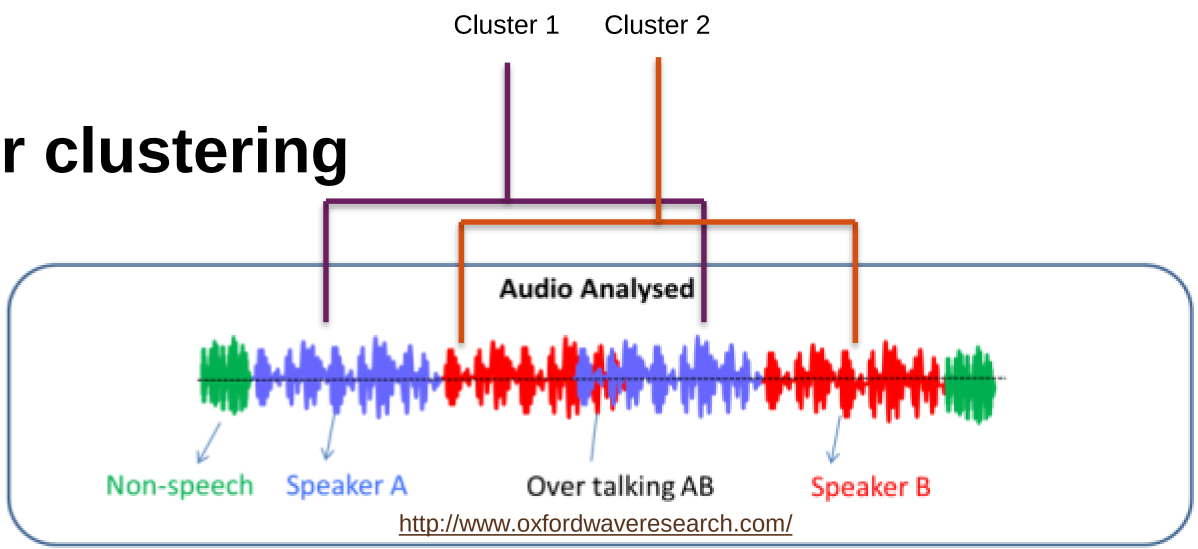
For the 630 training utterances, GMMs with 32 mixtures are built a priori, then an identification experiment is run for the 630 test utterances. It yields a satisfactory 0.5% closed set identification error.

[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.

5. 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?*

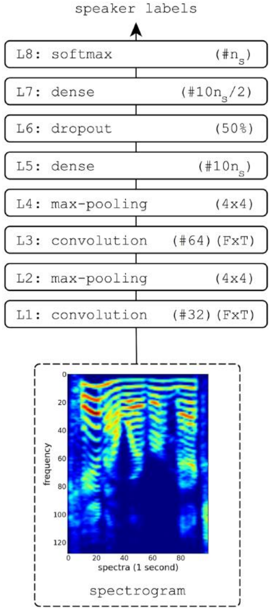
The interpretation of our results has shown that it is the stage of modeling that bears the highest potential: the inclusion of temporal context information among feature vectors is what is crucially missing there. Furthermore, the inclusion

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.

5. Speaker clustering – exploiting time information

CNN (MLSP'16)

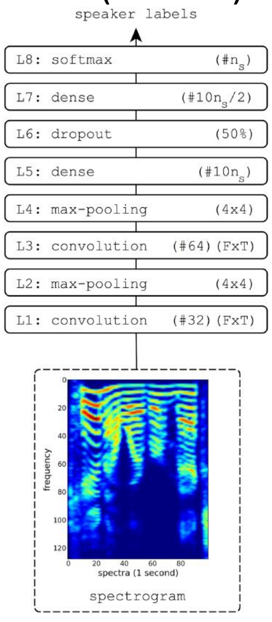


Method	MR	MR (legacy)
RNN /w PKLD	**2.19%** $\left(\frac{1.25\%+2.5\%+1.25\%+3.75\%}{4}\right)$	**4.38%** (average of 4 runs)
CNN /w PKLD [24]	-	**5%**
CNN /w cross entropy [23]	-	**5%**
ν-SVM [40]	6.25%	-
GMM/MFCC [40]	12.5%	-

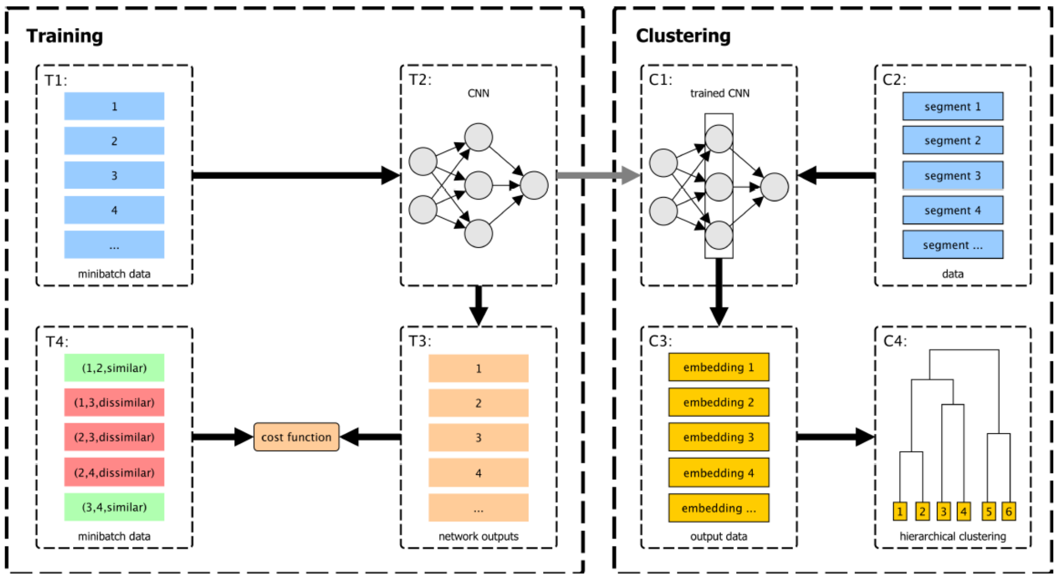
Lukic, Vogt, Dürr & Stadelmann (2016). «Speaker Identification and Clustering using Convolutional Neural Networks». MLSP'2016.
Lukic, Vogt, Dürr & Stadelmann (2017). «Learning Embeddings for Speaker Clustering based on Voice Equality». MLSP'2017.
Stadelmann, Glinski-Haefeli, Gerber & Dürr (2018). «Capturing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

5. Speaker clustering – exploiting time information

CNN (MLSP'16)



CNN & clustering-loss (MLSP'17)

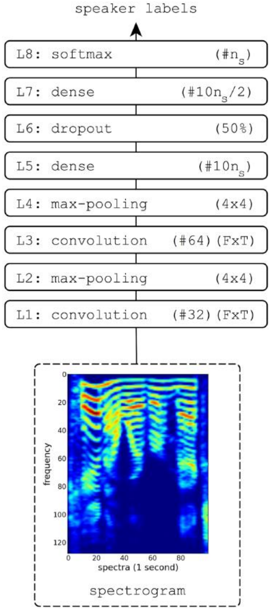


Method	MR	MR (legacy)
RNN /w PKLD	**2.19%** $\left(\frac{1.25\%+2.5\%+1.25\%+3.75\%}{4}\right)$	4.38% (average of 4 runs)
CNN /w PKLD [24]	-	5%
CNN /w cross entropy [23]	-	5%
ν-SVM [40]	6.25%	-
GMM/MFCC [40]	12.5%	-

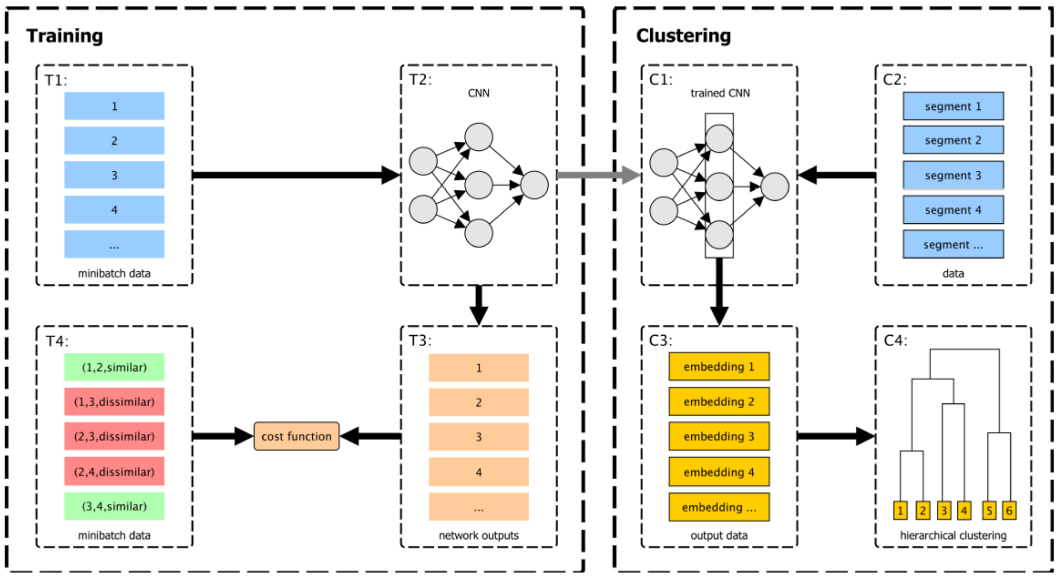
Lukic, Vogt, Dürr & Stadelmann (2016). «Speaker Identification and Clustering using Convolutional Neural Networks». MLSP'2016.
Lukic, Vogt, Dürr & Stadelmann (2017). «Learning Embeddings for Speaker Clustering based on Voice Equality». MLSP'2017.
Stadelmann, Glinski-Haefeli, Gerber & Dürr (2018). «Capturing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

5. Speaker clustering – exploiting time information

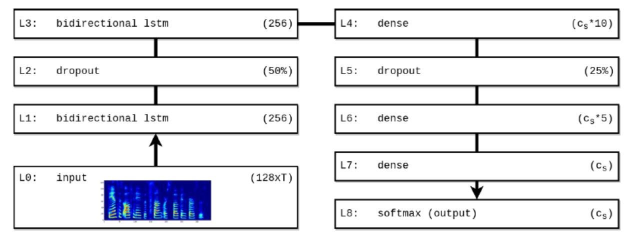
CNN (MLSP'16)



CNN & clustering-loss (MLSP'17)



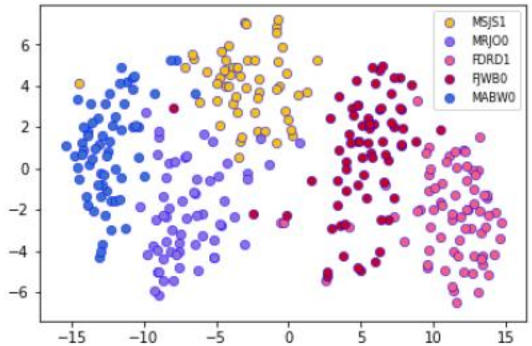
RNN & clustering-loss (ANNPR'18)

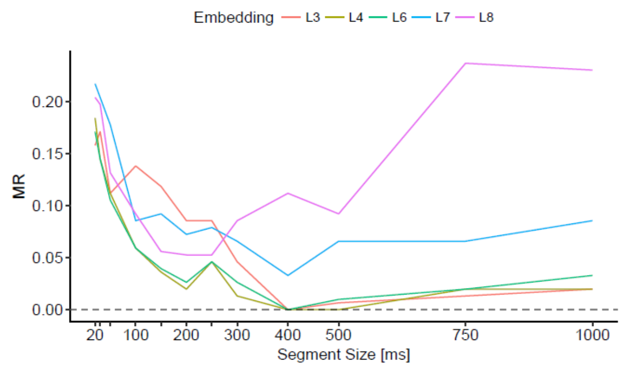


Method	MR	MR (legacy)
RNN /w PKLD	**2.19%** $\left(\frac{1.25\%+2.5\%+1.25\%+3.75\%}{4}\right)$	**4.38%** (average of 4 runs)
CNN /w PKLD [24]	-	**5%**
CNN /w cross entropy [23]	-	**5%**
ν-SVM [40]	6.25%	-
GMM/MFCC [40]	12.5%	-

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.

5. Speaker clustering – learnings & future work





«Pure» voice modeling seems 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.

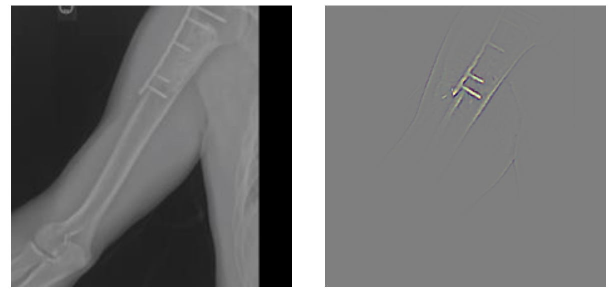
6. Lessons learned – model interpretability

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, 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

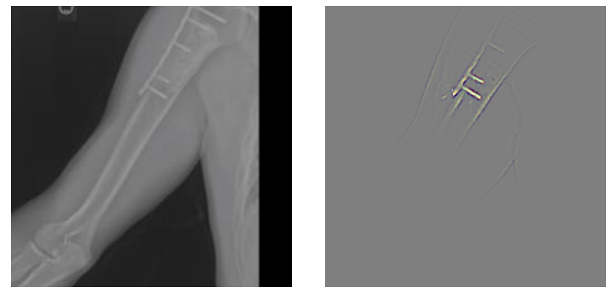
Zürcher Hochschule
für Angewandte Wissenschaften

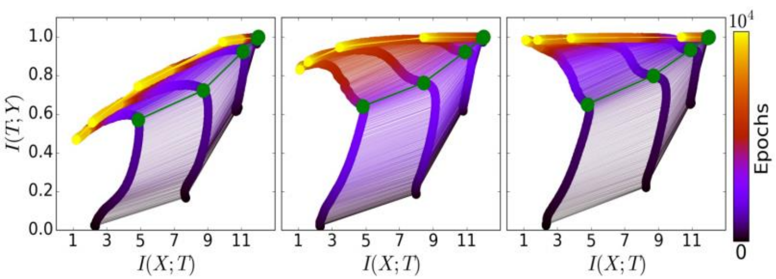
Interpretability is required.

- 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/

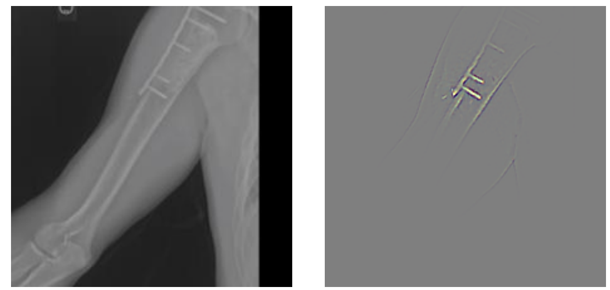
6. Lessons learned – model interpretability

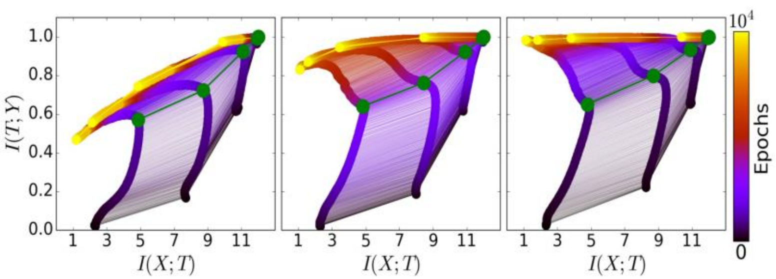
Interpretability is required.

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

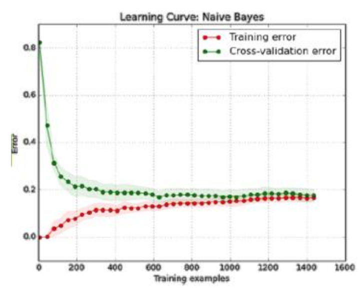
negative X-ray

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

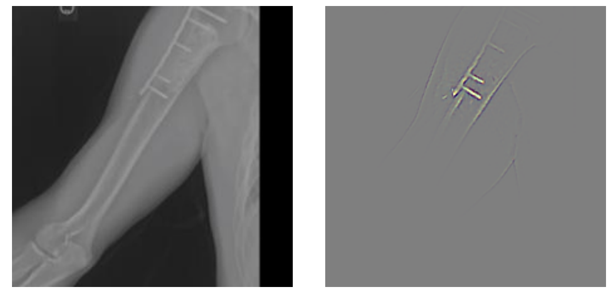
Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Innosuisse – Swiss Innovation Agency

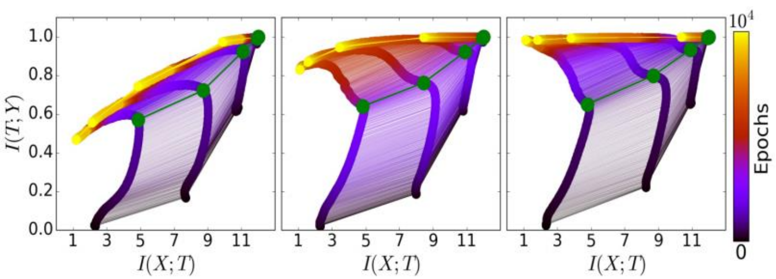
Interpretability is required.

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

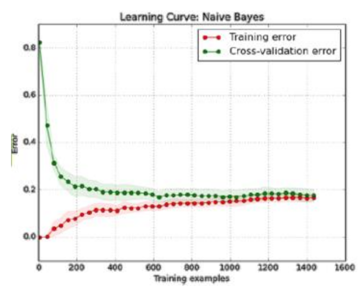
negative X-ray

positive X-ray

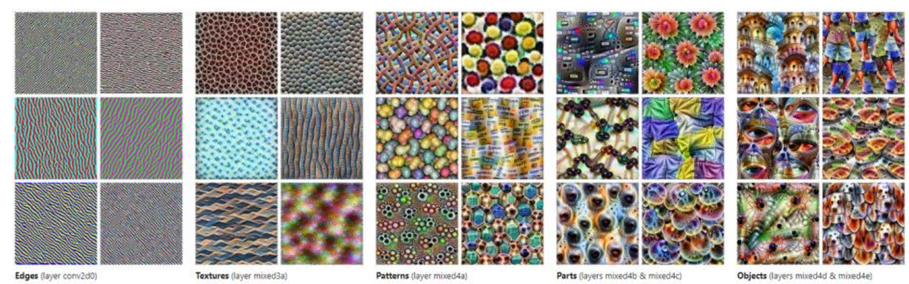




DNN training on the Information Plane



a learning curve

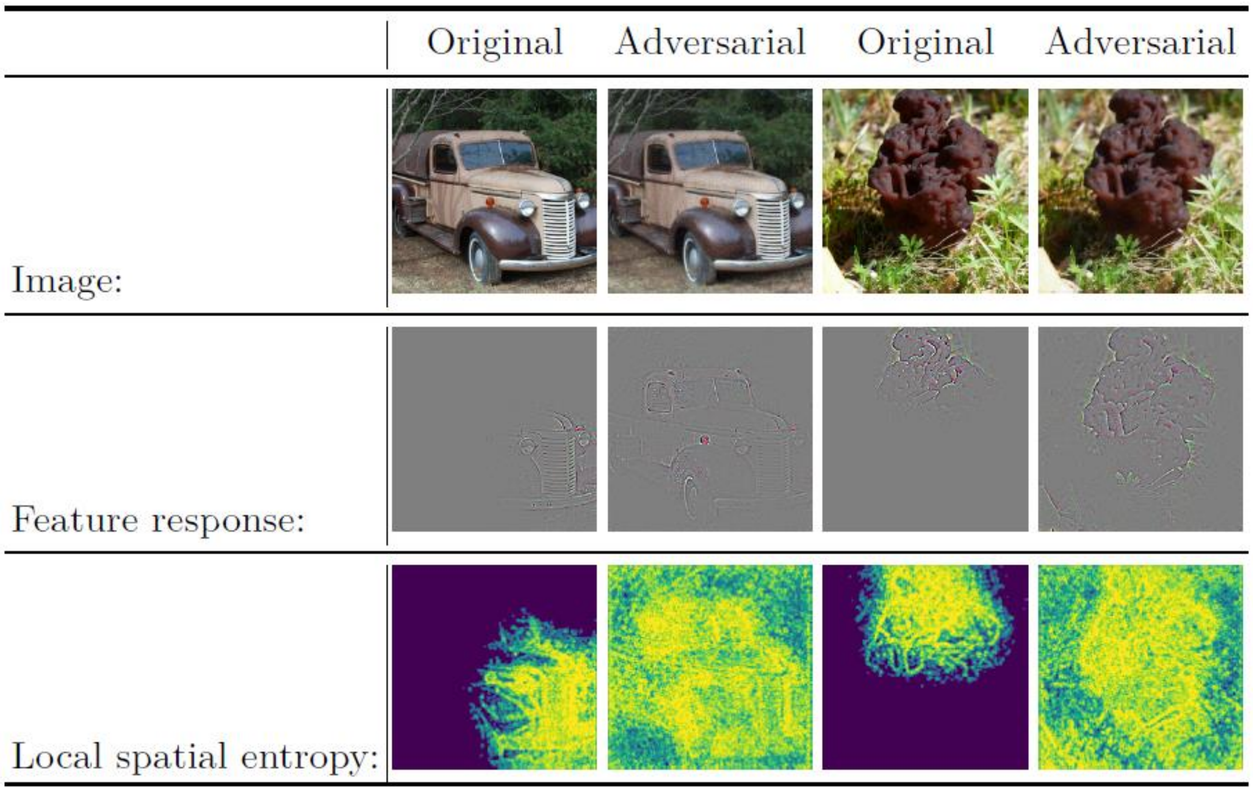


feature visualization

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/

6. 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



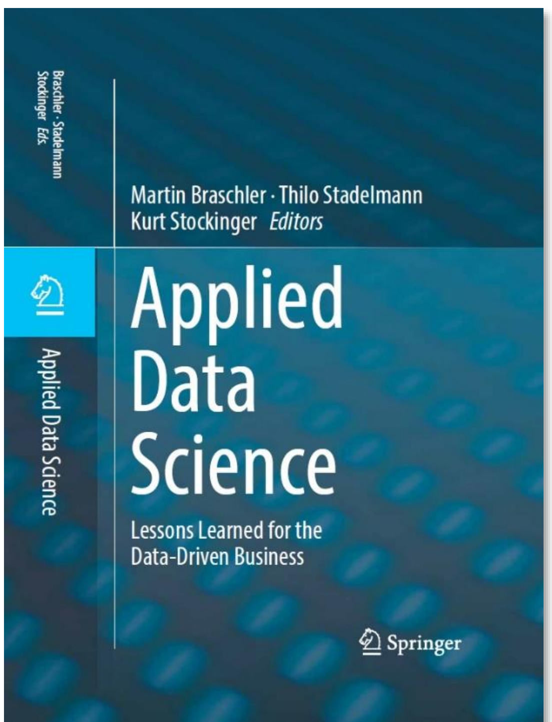
Amirian, Schwenker & Stadelmann (2018). *«Trace and Detect Adversarial Attacks on CNNs using Feature Response Maps»*. ANNPR’2018.

Zürcher Hochschule
für Angewandte Wissenschaften

zh
aw

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/RL **training** for new use cases **can be tricky** (→ needs thorough experimentation)
- New **theory and visualizations** help to debug & understand
  → *the training process*
  → *individual results*



On me:

- Prof. AI/ML, scientific director ZHAW digital, head ZHAW Datalab, board Data+Service
- thilo.stadelmann@zhaw.ch
- 058 934 72 08
- @thilo_on_data
- https://stdm.github.io/

Further contacts:

- Data+Service Alliance: www.data-service-alliance.ch
- Collaboration: datalab@zhaw.ch

➔ Happy to answer questions & requests.

datalab
www.zhaw.ch/datalab

data service
Swiss Alliance for
Data-Intensive Services