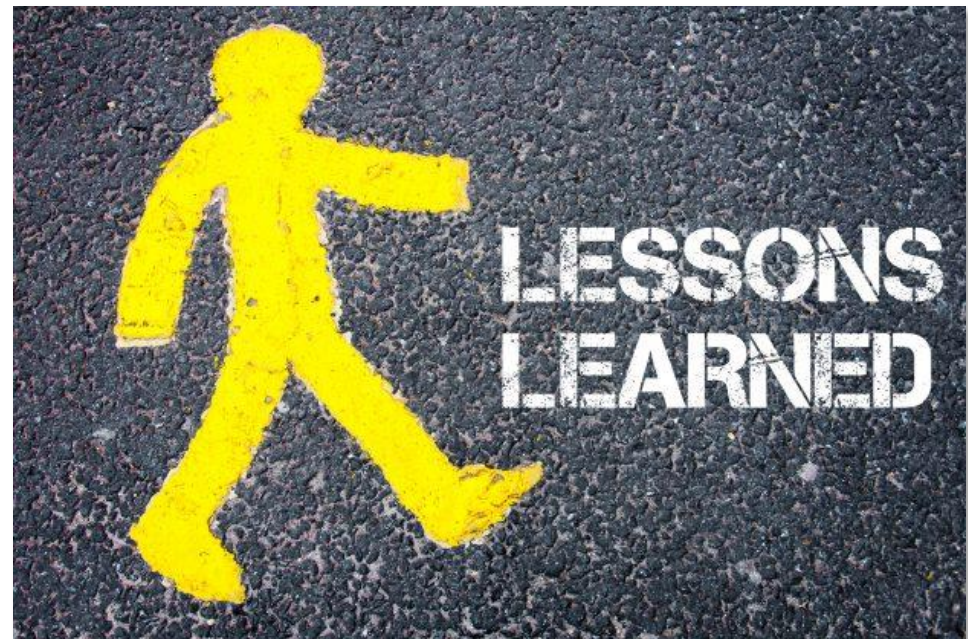


Lessons Learned from Deep Learning in Industry

*Artificial Intelligence in Industry and Finance
3rd European COST Conference on Mathematics for Industry
Thursday, 6th September 2018*

Thilo Stadelmann



Why?



Why?



arXiv monthly submission rates



Why?



arXiv monthly submission rates



Forbes Billionaires Innovation Leadership Money Consumer Industry Lifestyle

GPU TECHNOLOGY CONFERENCE

EUROPE / 9-11 OKTOBER, 2016
DER WICHTIGSTE EVENT ZU KÜNSTLICHER INTELLIGENZ
Sparen Sie 20% mit Code CM0SZM

25,677 views | Aug 20, 2016, 12:11am

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

Bernard Marr Contributor
Enterprise & Cloud

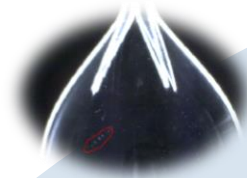
You may have heard about deep learning and felt like it was an area of data science that is incredibly intimidating. How could you possibly get machines to learn like humans? And, an even scarier notion for some, why would we want machines to exhibit human-like behavior? Here, we look at 10 examples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

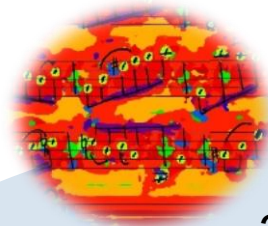
Agenda



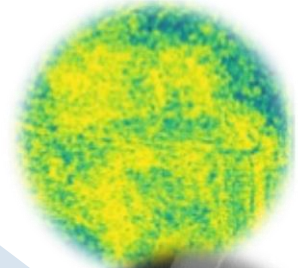
1. Face matching



2. Industrial quality control



3. Music scanning




Lessons learned

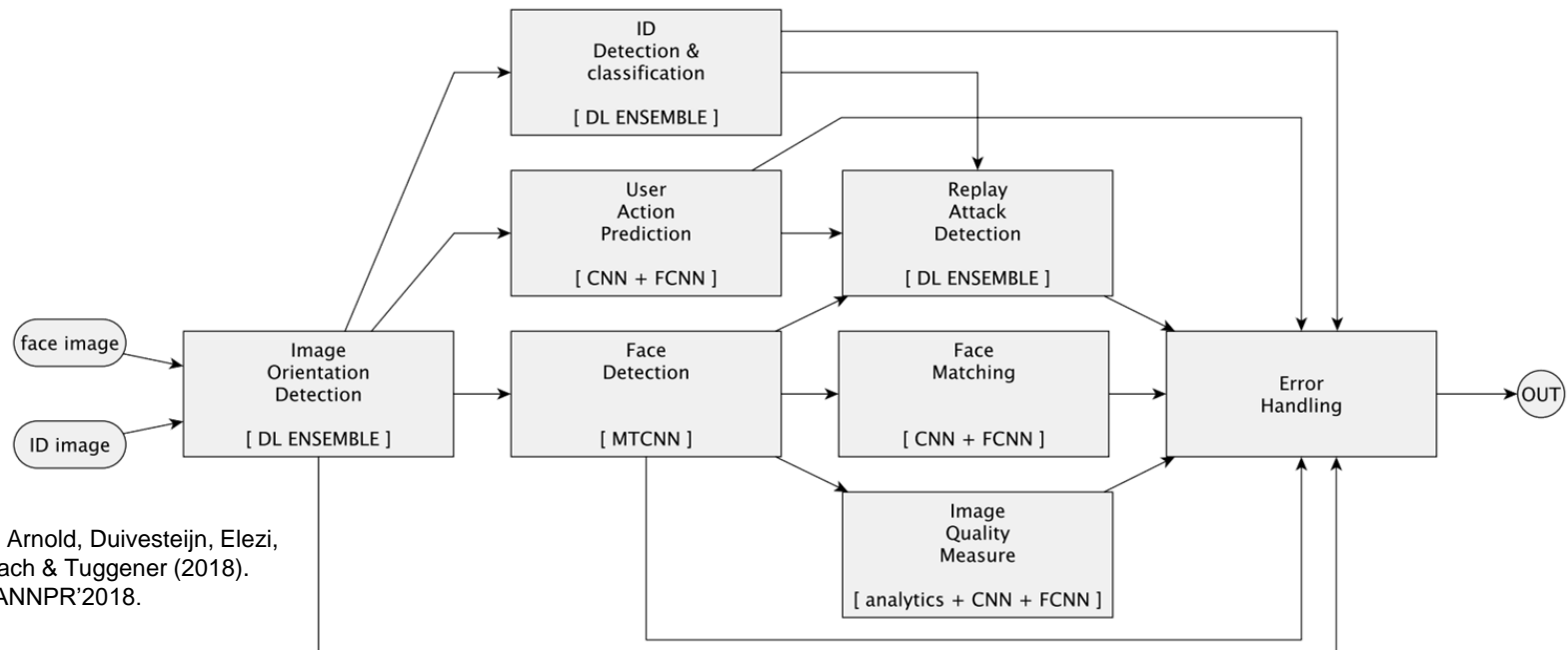
1. Face matching



 **DEEPIIMPACT**

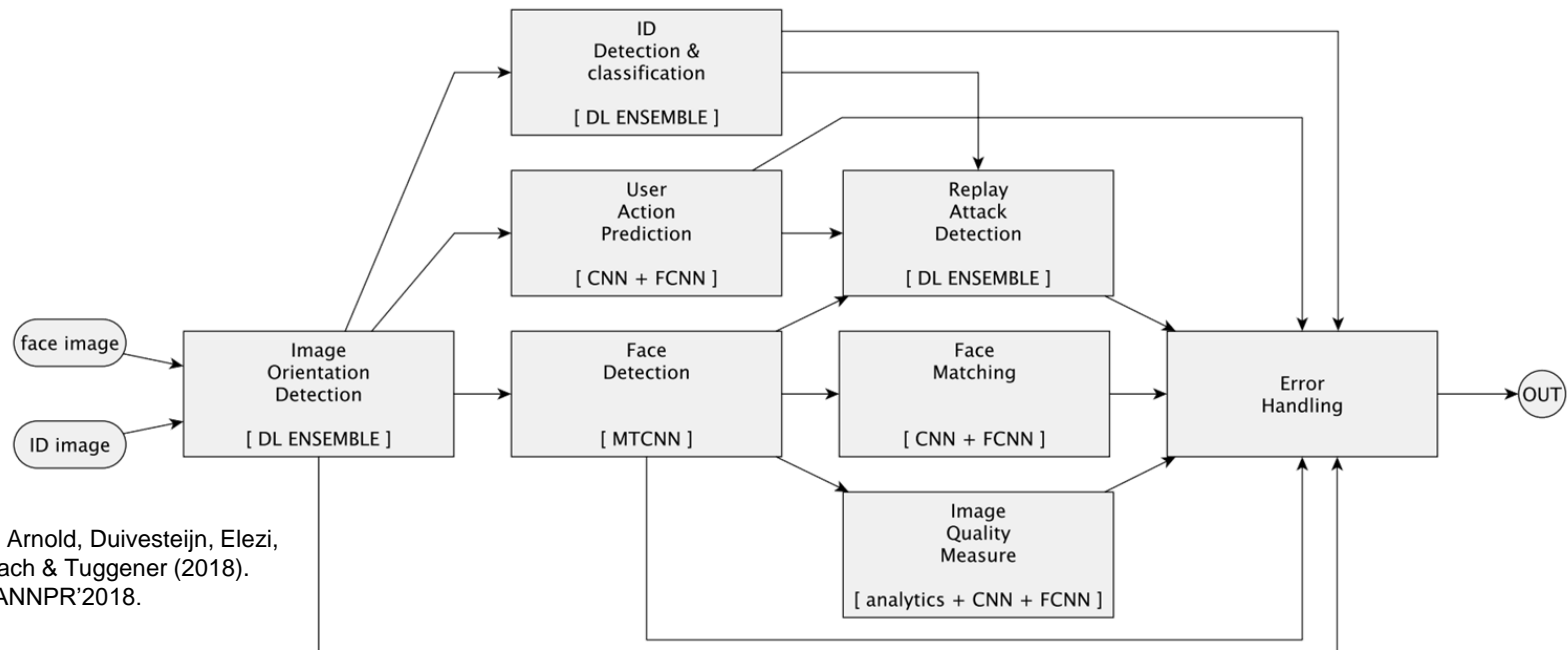
 Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra
Swiss Confederation
Innosuisse – Swiss Innovation Agency

1. Face matching – challenges & solutions



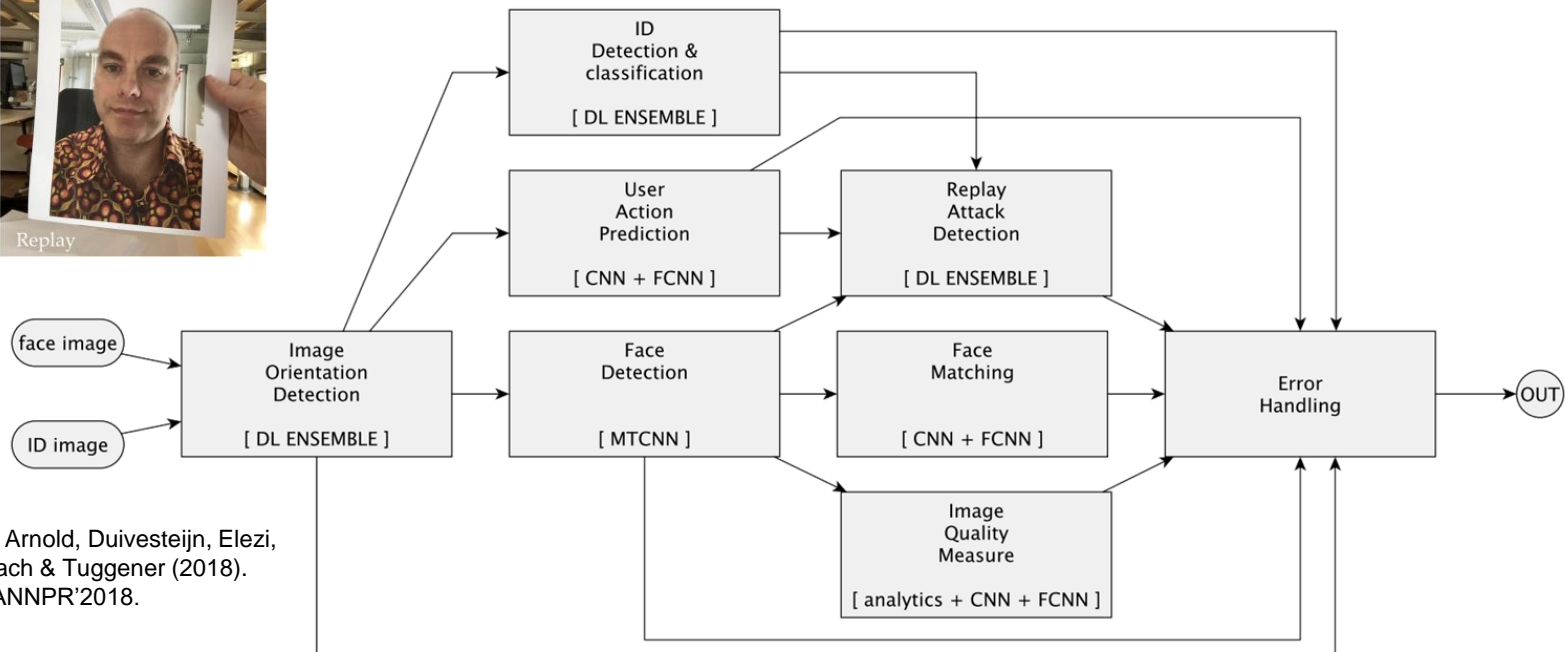
Stadelmann, Amirian, Arabaci, Arnold, Duivesteijn, Elezi, Geiger, Lörwald, Meier, Rombach & Tuggener (2018). «Deep Learning in the Wild». ANNPR'2018.

1. Face matching – challenges & solutions



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

1. Face matching – challenges & solutions

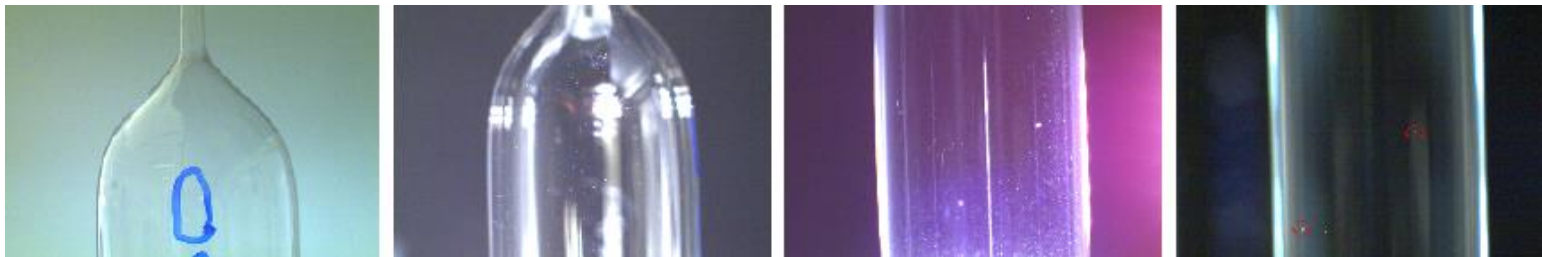


Stadelmann, Amirian, Arabaci, Arnold, Duivesteijn, Elezi, Geiger, Lörwald, Meier, Rombach & Tuggener (2018). «Deep Learning in the Wild». ANNPR'2018.

2. Industrial quality control

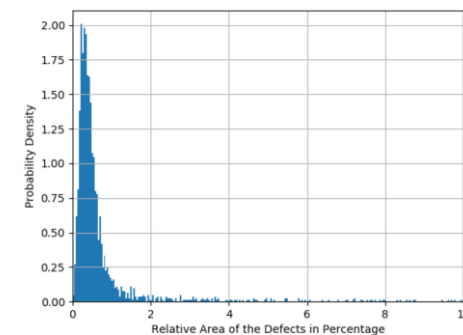
Task

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



Challenges

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

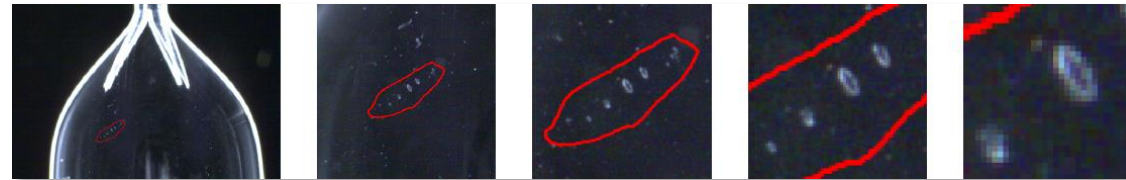


Stadelmann, Duivesteyn, Amirian, Tuggener, Elezi, Geiger & Rombach (2018). «*Deep Learning in the Wild*». ANNPR'2018.

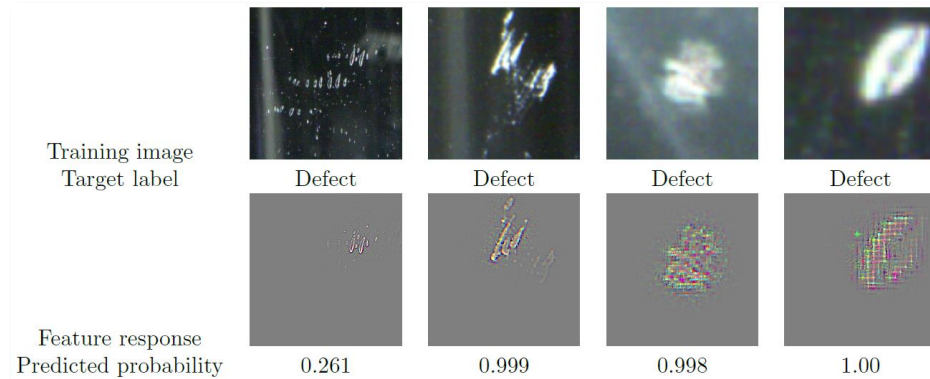
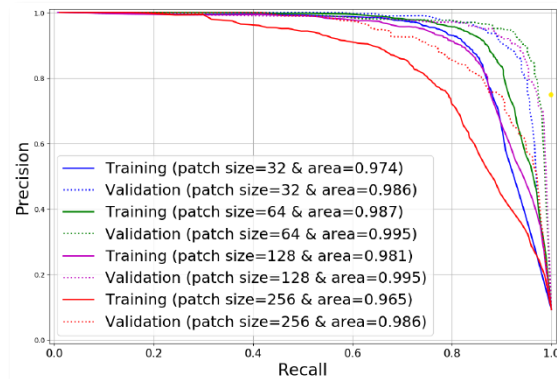
2. Industrial quality control – solutions (Work in progress)

Ingredients

- Weighted loss
- Defect cropping
- Secret sauce



Preliminary results



3. Music scanning

Die Forelle.
Op. 84, No. 2, Schöten.
Für eine Singstimme mit Begleitung des Pianoforte.
Schubert's Werk. Franz Schubert.
Erste Fassung. N° 207

Musik:
Singsstimme:
Pianoforte:



```
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE score-partwise SYSTEM "http://www.musicxml.org/dtd/partwise.dtd" PUBLIC "-//Recordare/DTD MusicML 2.0 Partwise/EN"
- <score-partwise>
- <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="1552.22" 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 (Mollisande on http://www.Musescore.com)</credit words>
- </credit>
- <part-list>
- <score-part id="P1">
- <part name="Ténor">
- <part-abbreviation="Ténor">
- <score-instrument id="P1-13">
- <instrument name="Ténor">
- <score-instrument id="P1-13">
- <midi-channel="1">
- <midi-program="74">
- <volume="78.7402">
- <pan="0">
- </midi-instrument>
- </score-part>
- <part-group type="start" number="1">
- <group-symbol="brace">
- </part-group>
- <score-part id="P2">
- <part name="Piano">
- <score-instrument id="P2-13">
- <instrument name="Piano">
- </score-instrument>
- </part-list>
```



SCOREPAD

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



Die Forelle - Franz Schubert

♩ = 80

Voice

Piano

Vo.

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.

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.

3. Music scanning – challenges & solutions

The image displays a musical score with various annotations. A callout box highlights four specific annotation types:

- (a) accidentalSharp: A sharp symbol (#) placed above a note.
- (b) keySharp: A sharp symbol (#) placed at the beginning of a staff.
- (c) augmentationDot: A red dot placed above a note.
- (d) articStaccatoAbove: A red dot placed above a note.

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

3. Music scanning – challenges & solutions

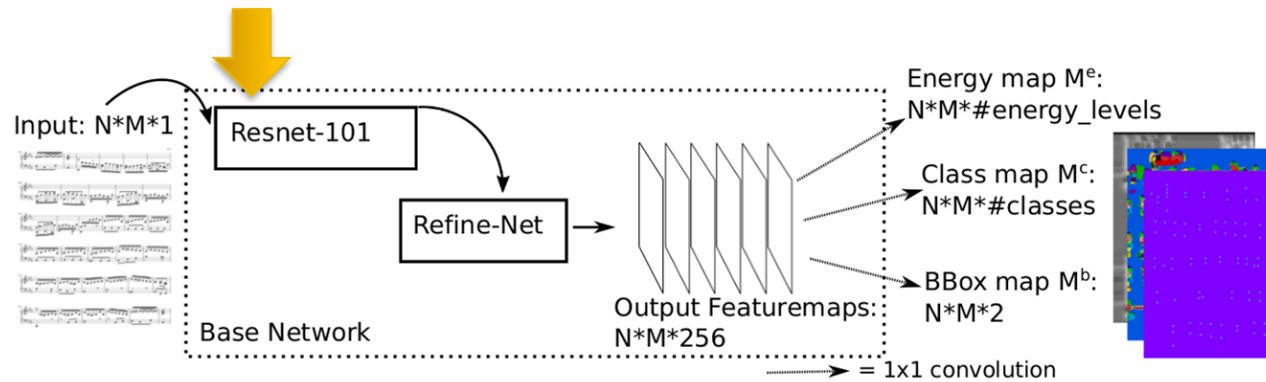
(a) accidentalSharp (b) keySharp

(c) augmentationDot (d) articStaccatoAbove

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

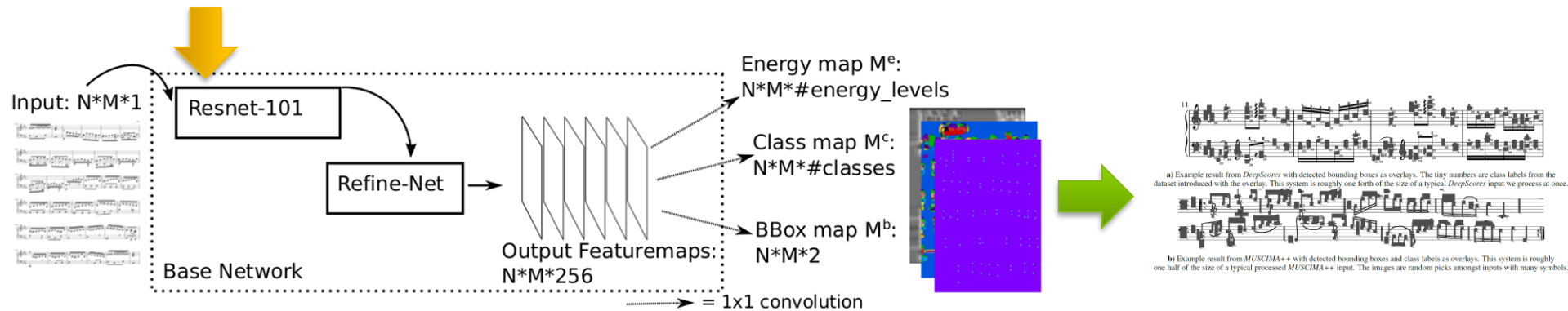
3. Music scanning – challenges & solutions

(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

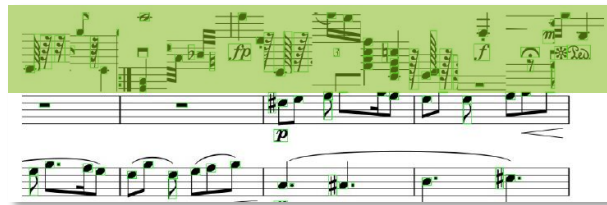


Tuggener, Elezi, Schmidhuber, Pelillo & Stadelmann (2018). «DeepScores – A Dataset for Segmentation, Detection and Classification of Tiny Objects». ICPR'2018.
Tuggener, Elezi, Schmidhuber & Stadelmann (2018). «Deep Watershed Detector for Music Object Recognition». ISMIR'2018.

3. Music scanning – industrialization (Work in progress)

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 mAP** from 16% (on purely synthetic data) **to 73%** on more challenging real-world data set (previous state of the art: 24.8%)

Elezi, Tuggener, Pelillo & Stadelmann (2018). «DeepScores and Deep Watershed Detection: current state and open issues». WoRMS @ ISMIR'2018.

Pacha, Hajic, Calvo-Zaragoza (2018). «A Baseline for General Music Object Detection with Deep Learning». Appl. Sci. 2018, 8, 1488, MDPI.



Lessons learned 1/4

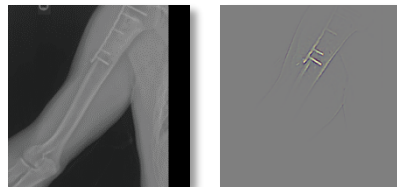
Data


- Acquisition usually **needs much more time** than expected (), yet is the basis for all subsequent success (). Class **imbalance & covariate shift** are usual (, , )



Understanding

- what has been learned** and **how decisions emerge** help both the user and the developer of neural networks to build trust and improve quality (, )






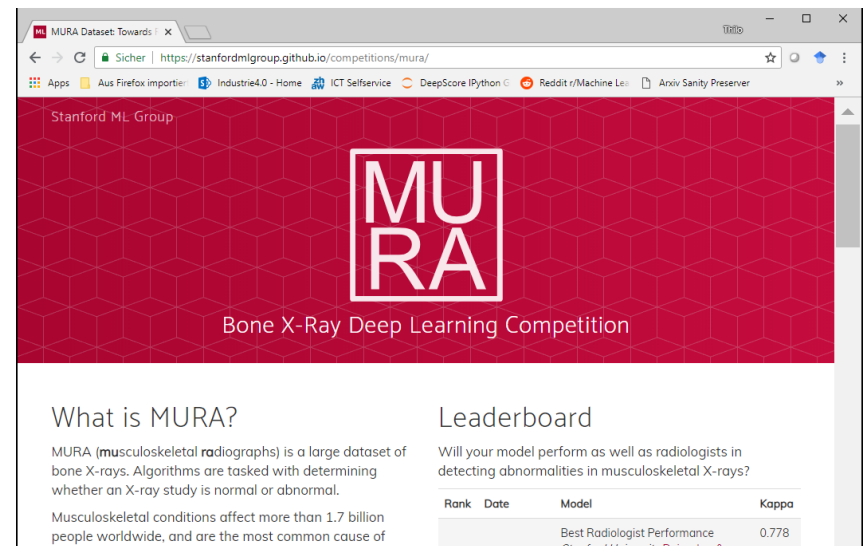
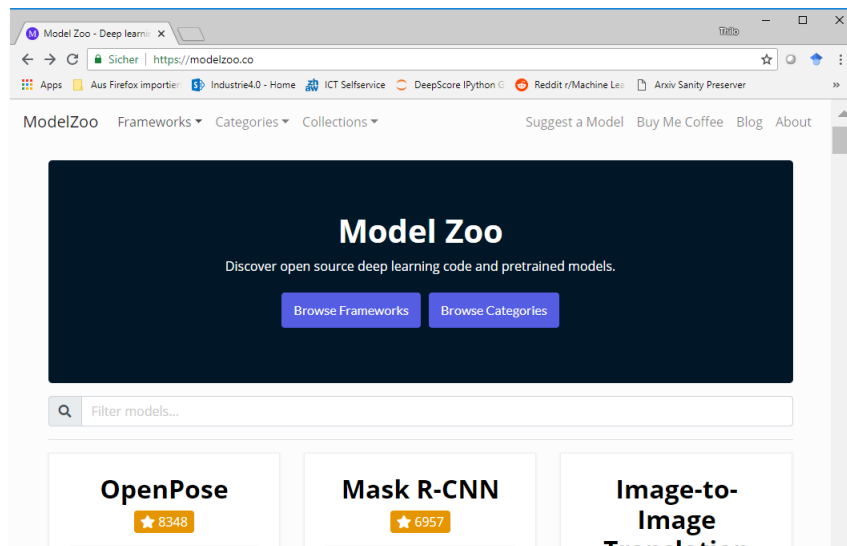
- Operators and business owners** need a basic understanding of used methods to produce usable ground truth and provide relevant subject matter expertise ()

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

Lessons Learned 2/4

Simple baselines

- do a good job in **determining** the **feasibility** as well as the **potential** of the task at hand when final datasets or novel methods are not yet seen ( ,  , )



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


Lessons learned 3/4

Loss shaping

- Usually necessary to enable learning of very complex target functions ()

*“Initially, the training was **unstable** [...] if directly trained on the **combined weighted loss**. Therefore, we now **train** [...] on each of the **three tasks separately**.*

*We further observed that while the network gets trained on the bounding box prediction and classification, the energy level predictions get worse. To avoid this, the network is **fine-tuned only for the energy level loss** [...]. Finally, the network is retrained on the combined task [...] for a few thousand iterations [...].”*


- This includes encoding expert knowledge manually into the model architecture or training setup ().

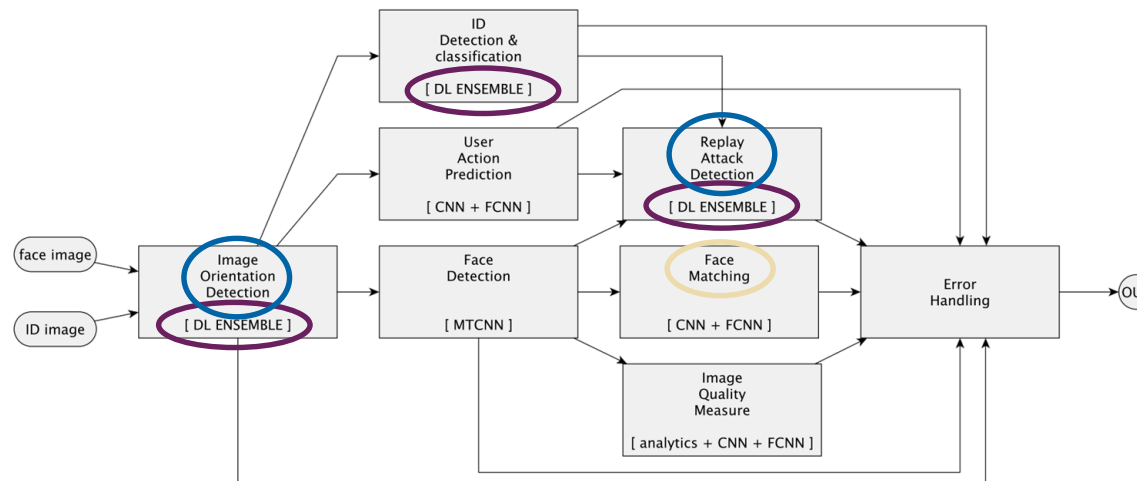
*“The **size of the anomaly** in classifying balloon catheters as good or bad is **quite decisive**. Thus, rescaling the training images is not allowed, and we used a fixed size window around the center of each defect to extract the training images.”*



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

Lessons learned 4/4

Deployment

- Might involve the buildup of up to dozens of **other machine learning** models () to flank the **original core part**.



- **Specialized models** for identifiable sub-problems increase the accuracy in production systems over all-in-one solutions (), and **ensembles** of experts help where no single method reaches adequate performance ().

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

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



On me:

- Prof. AI/ML, head ZHAW Datalab, board SGAICO & 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.

