Segmentation of the paper structure in historical prints

Master research project (5 ECTS)

Paper structure

- In the early modern period, paper was still scooped by hand
- Thus, analyzing the paper structure of historical documents can reveal information about the origin of the prints

Orange lines: warp wires

Image source: H-4763, Germanisches Nationalmuseum Nürnberg

Task

- Automatically segment the warp wires in the paper structure
- Compute some statistics regarding the lines (e.g. varying distance between the wires)
- Implementation in Python

- Master research project (5 ECTS)
- In cooperation with the Germanisches Nationalmuseum and TH Köln

Contact:
Aline Sindel
Room 10.138
aline.sindel@fau.de

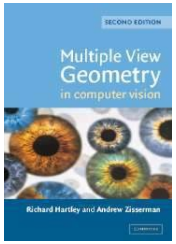
Computer Vision Project Assignments

Jul 22nd 2019

André Aichert
Pattern Recognition Lab (CS 5)
Friedrich-Alexander-Universität-Erlangen-Nürnberg

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Bundle-Adjustment for Calibration of FD-CT Scanners.

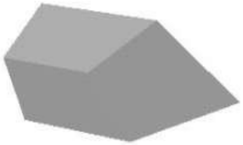


Multiple View Geometry in Computer Vision
Richard Hartley and Andrew Zisserman
Cambridge University Press, March 2004.

Step 1: Factorization

- Factorization: Recover geometry AND projection
  → Given a set of images with the same points
  → Without additional information
  → Up to a single 3D projective transformtion

- Optional: Stratification
  → Add knowledge: angles
  → Undistort projective solution



15 DOF

7 DOF

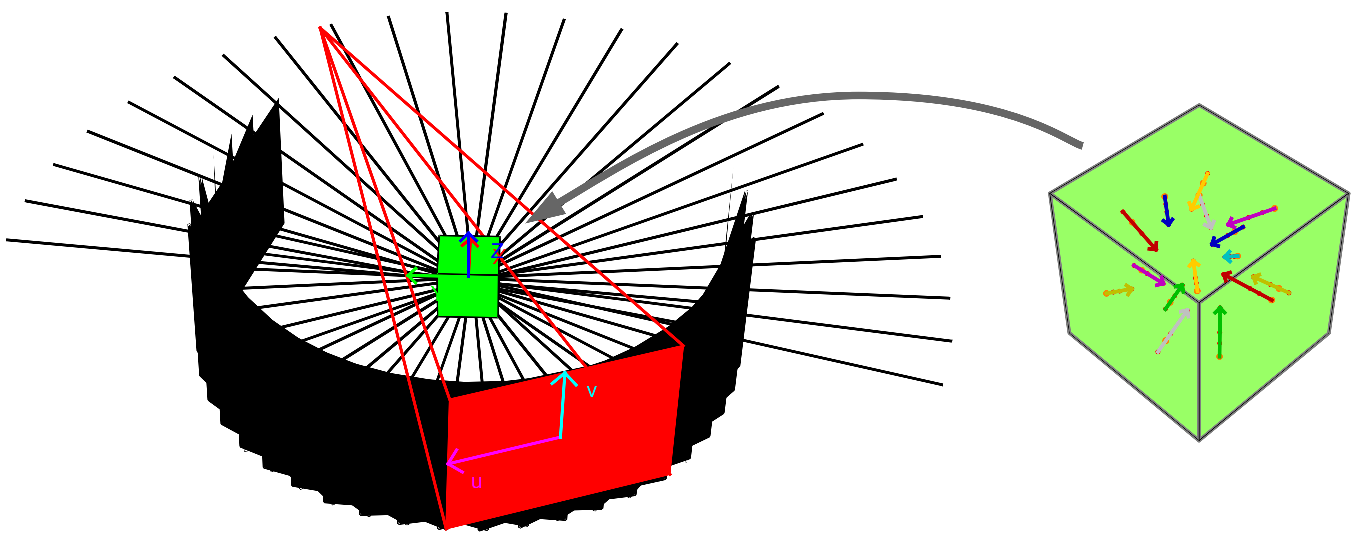
Step 2: Bundle Adjustment

- Recover geometry AND projection

  → Given initial guess of projection and points

  → Determine optimal metric reconstruction and projection

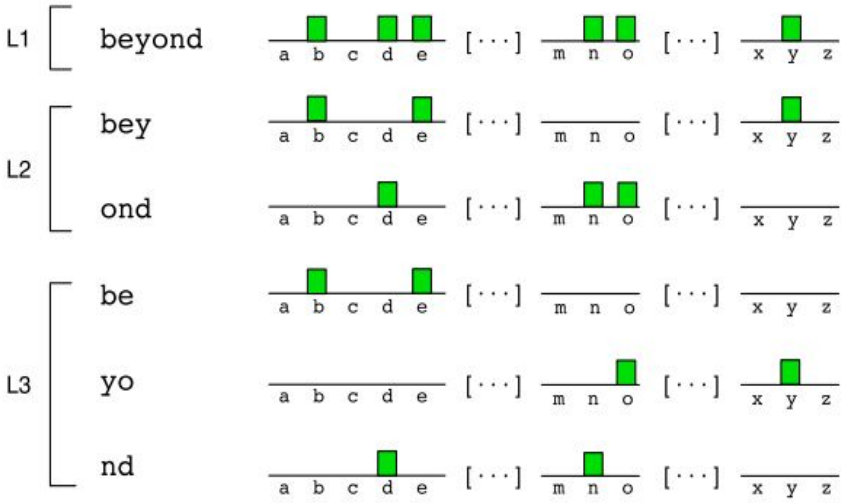
  → Application: Computed tomography scanner calibration.



Topics

Compressing PHOC-like representations

- PHOC-like: An vector representation of a string that can be generated from word-images
- Can we compress them? Will standard compression techniques work?
- Do they preserve their their joint image-string searchability



Contact: ***anguelos.nikolaou@fau.de***

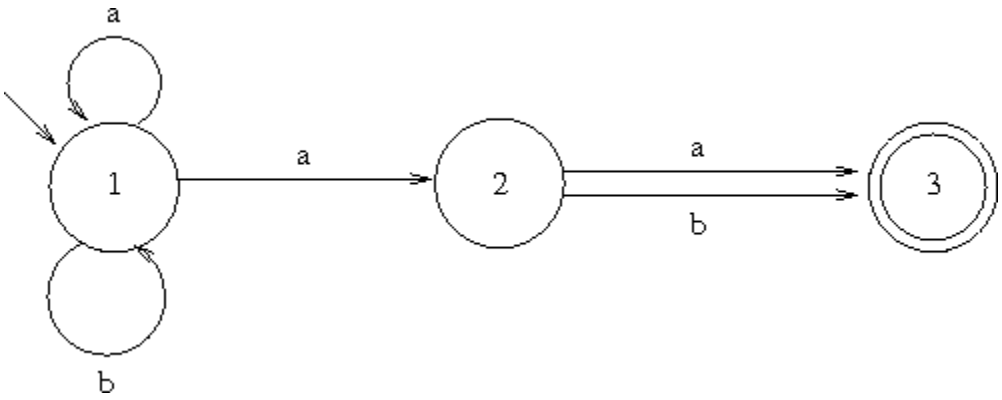
ECTS: 5/10/MT (depends)

Deep regular expressions

- Regular expressions are easily compiled to NFA (Non-discrete Finite-state Automata)
- Typically regular expression engines are implemented by compiling NFA to larger DFA (Discrete Finite-state Automata)
- Can we work directly on NFA?
- Can we use it on top of a Deep Neural Network?
- What are the benefits?

Contact: ***anguelos.nikolaou@fau.de***

ECTS: 5/10/MT (depends)



Decomposing 2D Convolutions

- 2D Convolutions complexity: N^2
- Two consecutive 1D Convolutions complexity N*2
- How much do we lose if we train on 2D and do inference on 2x1D?
- How much do we lose if we train on 2x1D and do inference on 2x1D?

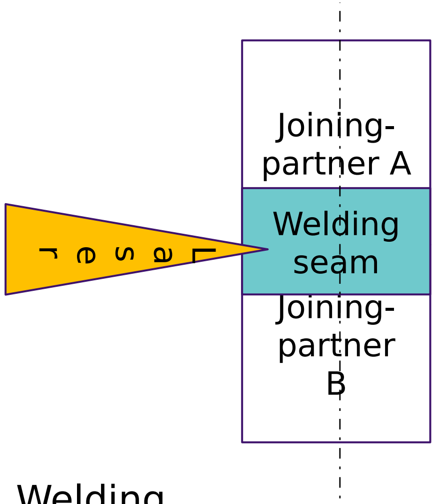
Contact: ***anguelos.nikolaou@fau.de***

ECTS: 5

Industrial Segmentation

5 / 10 ECTS Research/Master Project
Contact: vincent.christlein@fau.de

Procedure for metallographic grinding:



Welding

Separatic

Embedding and grinding in plastic

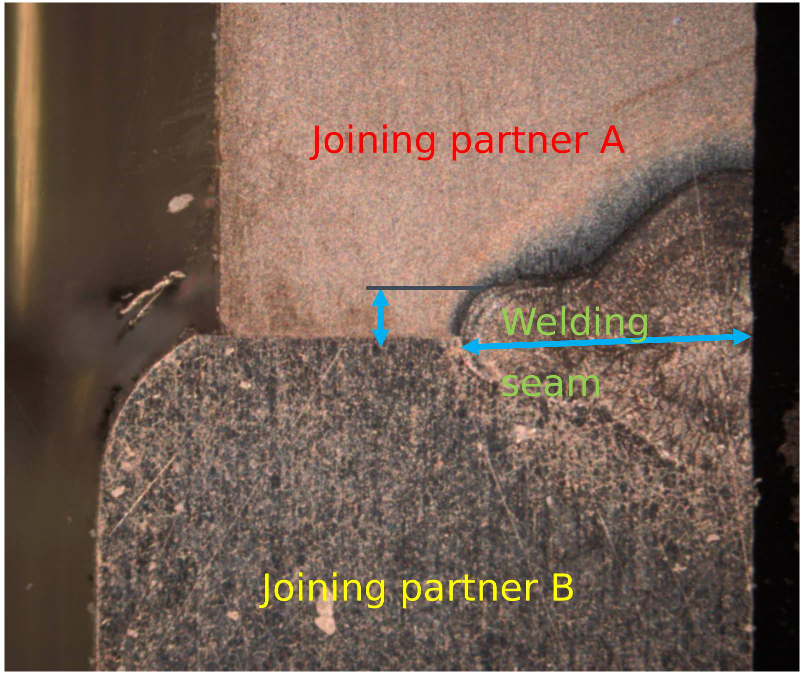
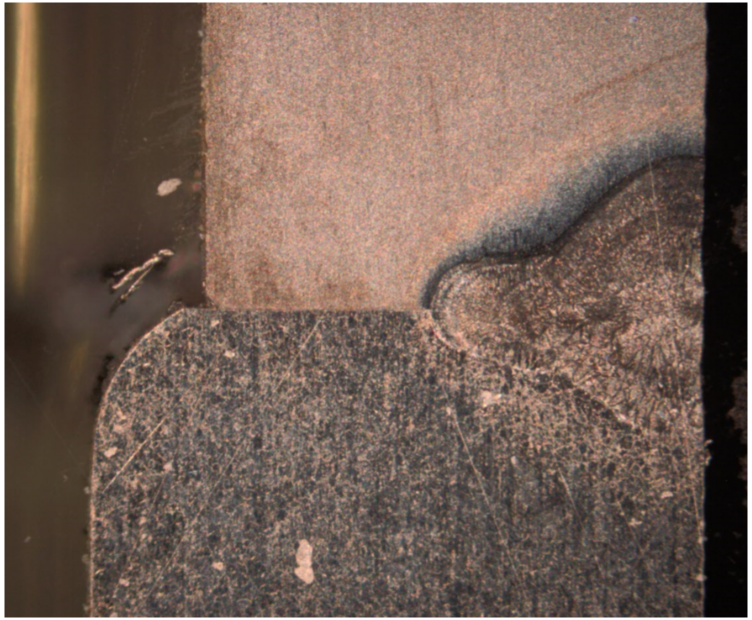


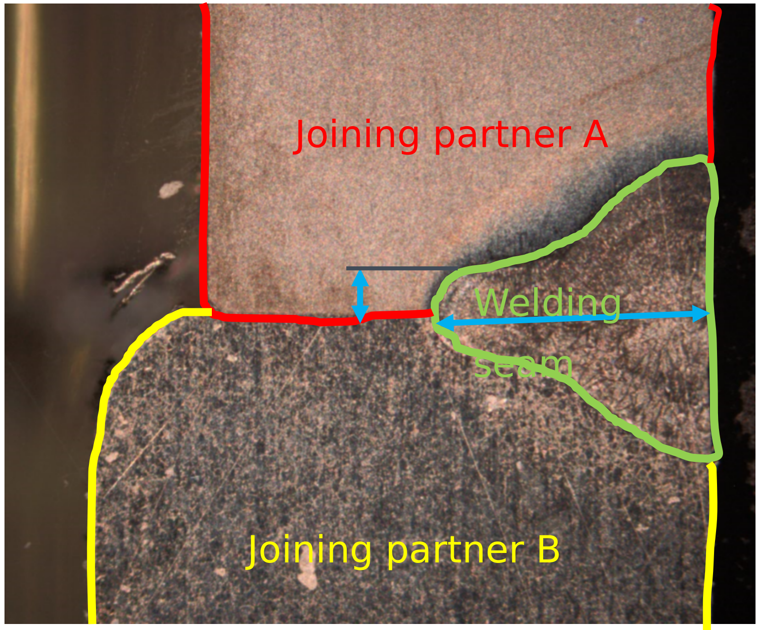
Image acquisition with measuring microscope and manual measurement

- Manual measurement error-prone
- Measure seam at only a few points
- time expenditure

Standard test methods for welds:
Benefits for a large part of the manufacturing industry

Computer Vision Task on Laser Welds



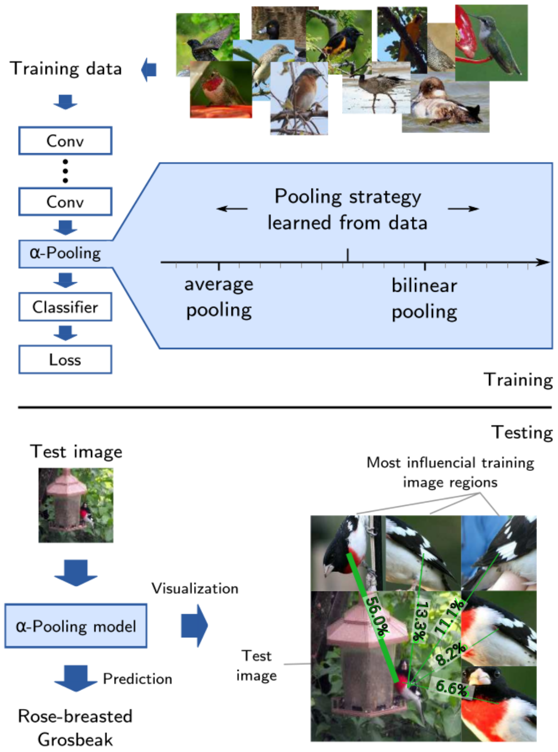


- Segmentation of the joint into joining partner and welding seam
- Dimensioning the seam at selected points
- Structure of a learning process using the example of one or two welding seams
- Representation of the detection reliability

Advanced Pooling

- Implementation and evaluation of advanced pooling techniques (e.g. alpha pooling, etc.)

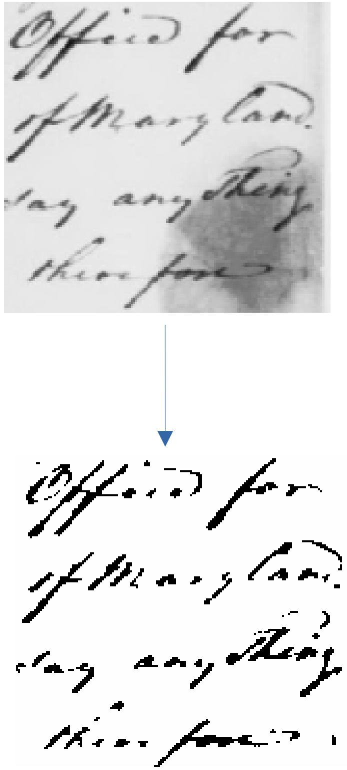
- 5 ECTS
- Contact: vincent.christlein@fau.de
- Implementation in pytorch



Improved Binarization

- Content integration for binarization based on U-Net variation for binarization

- 5 ECTS
- Contact: vincent.christlein@fau.de
- Implementation in pytorch

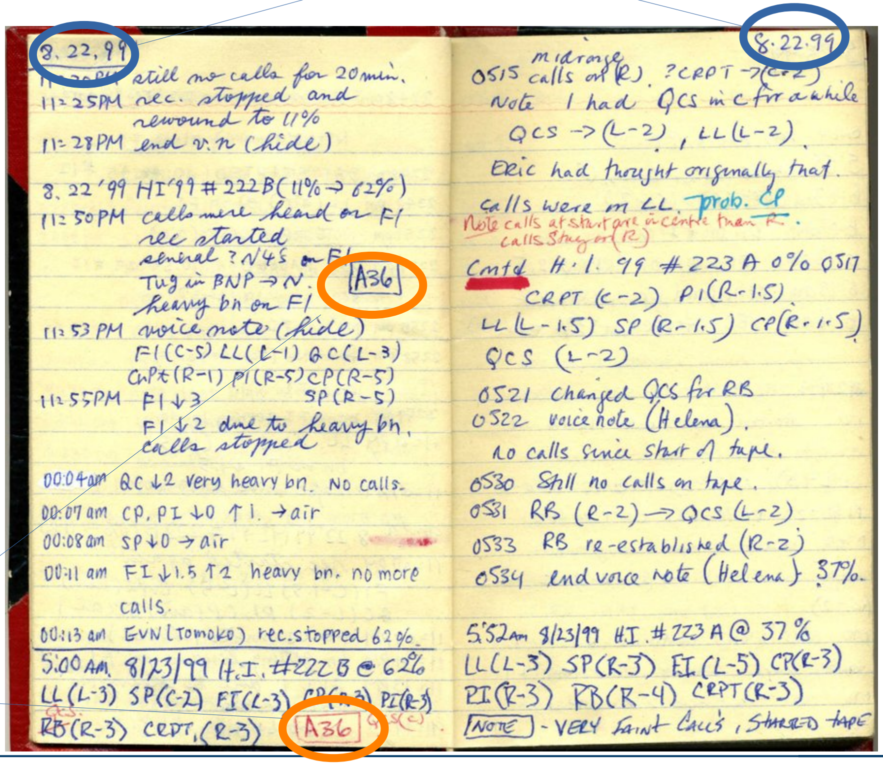


Keyword Spotting in Orca Sighting Journals

- HTR-based Keywpord Spotting
- Extract Date and Orca Family

- 5/10 ECTS Project / MT
- Contact: vincent.christlein@fau.de
- Implementation in pytorch

Date

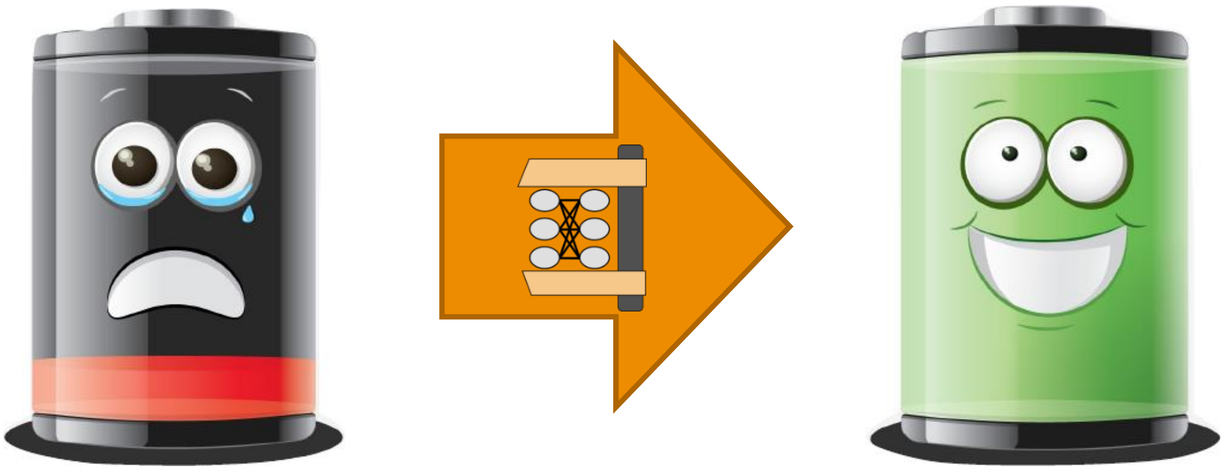


Family-ID

AUDIO LABS

DNN Optimization in Audio

Axel Plinge



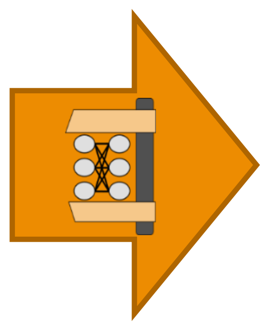
FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

Fraunhofer IIS

DNN Optimization in Audio

Motivation

- Training DNNs requires Graphical Processing Units (GPUs)
- They still need considerable resources (energy) at run-time
- Applications should run on embedded devices in real-time!
- It can be done: AlexNet (244MB) ➔ SqueezeNet (5MB)



[I16] Iandola, F. N., Moskewicz, M. W., et al. “SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <1MB model size” arXiv 1602.07360

DNN Optimization in Audio
Fraunhofer IIS

- Fraunhofer IIS in Erlangen is the “home of mp3”
- 250+ employees working on audio, video, multimedia, virtual reality and more

DNN Optimization in Audio

Master Thesis Topics

We want YOU to optimize our Applications!

Apply, investigate and develop state-of-the-art deep compression methods to one of the following:

i. Speaker localization with microphone arrays & CNN
ii. Speech separation using (B-)LSTM
iii. Language modelling by RNN for natural language interfaces
iv. Speaker verification with ResNet-like architecture

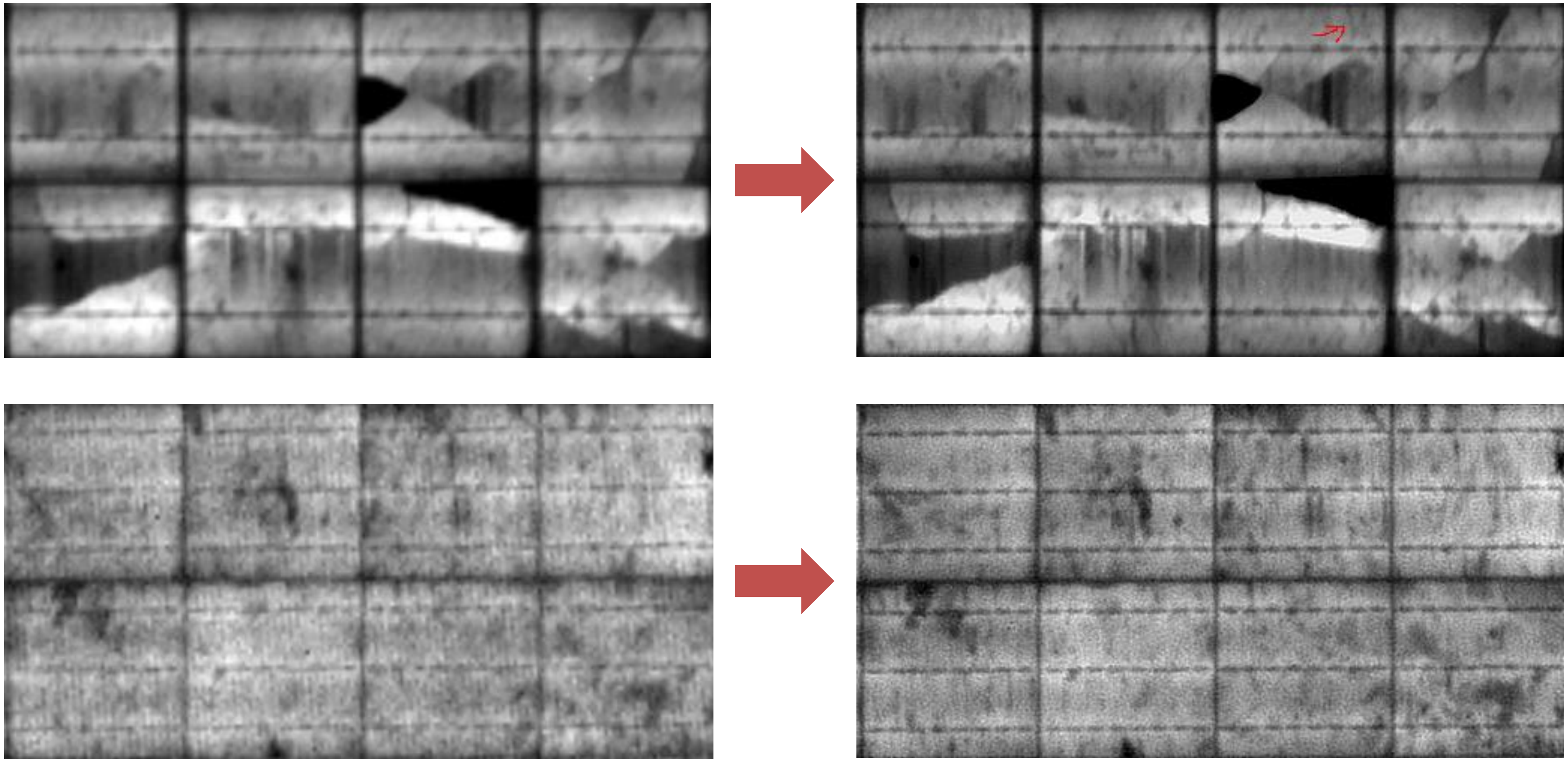
[H15] S. Han, H. Mao, et al., 2015, "Deep Compression: Compressing Deep Neural Networks with Pruning, trained Quantization and Huffman coding." ArXiv:1510.00149

Multi-frame superresolution for defect detection on solar cells

Master project

Idea

- **Low**-resolution-images of solar modules → **high** resolution images
- Enables detection of more defects



Principle

Registration

- Estimate parameters of pinhole camera
- Warp pixels into common grid

↓

Interpolation

- Interpolate with higher resolution
- Well suited for **CUDA** implementation

↓

Restoration

- Invert camera PSF by deconvolution

Caught you attention?

- Implement a classic CV pipeline
- Have **fun coding** C++ and CUDA
- Method is **known to work**

Contact:
Mathis Hoffmann (09.153)
mathis.hoffmann@fau.de

HI ERN
Helmholtz Institute
Erlangen-Nürnberg

Quality Control of Solarparks - Failure detection and analysis using statistical methods

Thema – Projektarbeit – Bachelorarbeit - Masterarbeit

Mai 2019 ||| Dr.-Ing. Claudia Buerhop ||| High Throughput Methods in Photovoltaics

Quality control of solarparks

Inspection using imaging techniques, e.g. thermography

benefit:

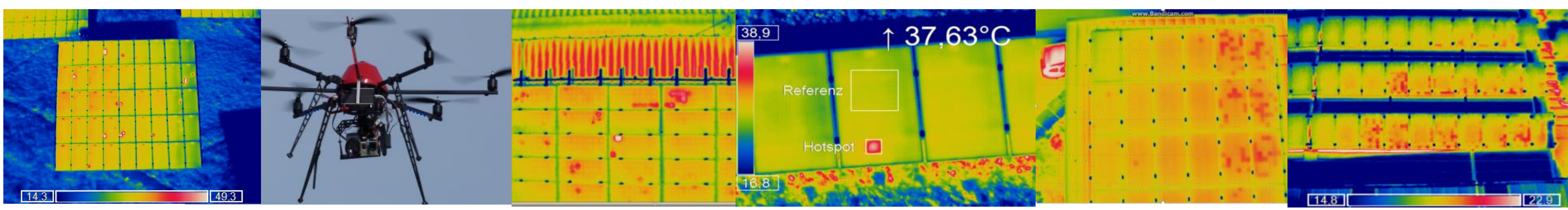
Fast

Contactless – without operation interruption

During real operating conditions – during sunshine

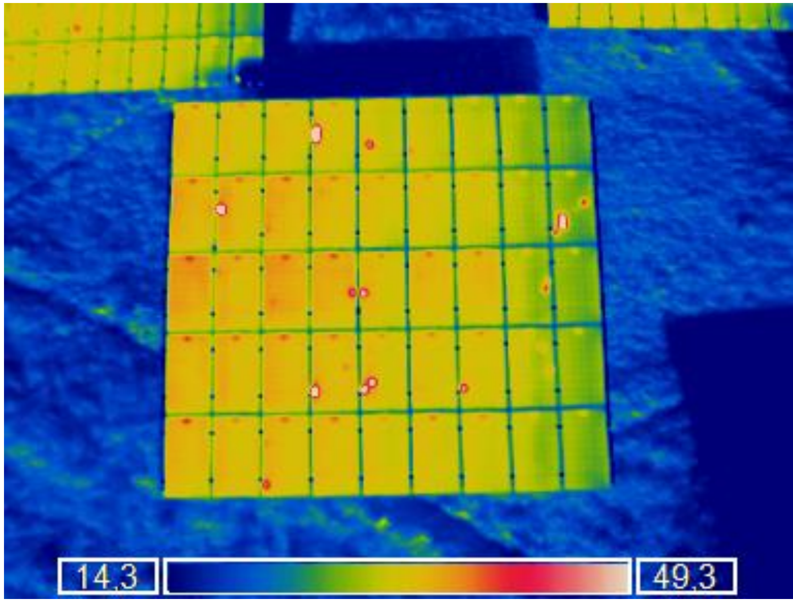
Quality check on module level

Predicting the power the basing on IR-images is advantageous because time-consuming electrical measurements are avoided and no operating interruption is necessary.



Task:

- detection of thermal anomalies
- Identification of malperforming PV-modules inbetween mostly well-performing PV-modules
- Prediction of the module power basing on IR-images



TODOs:

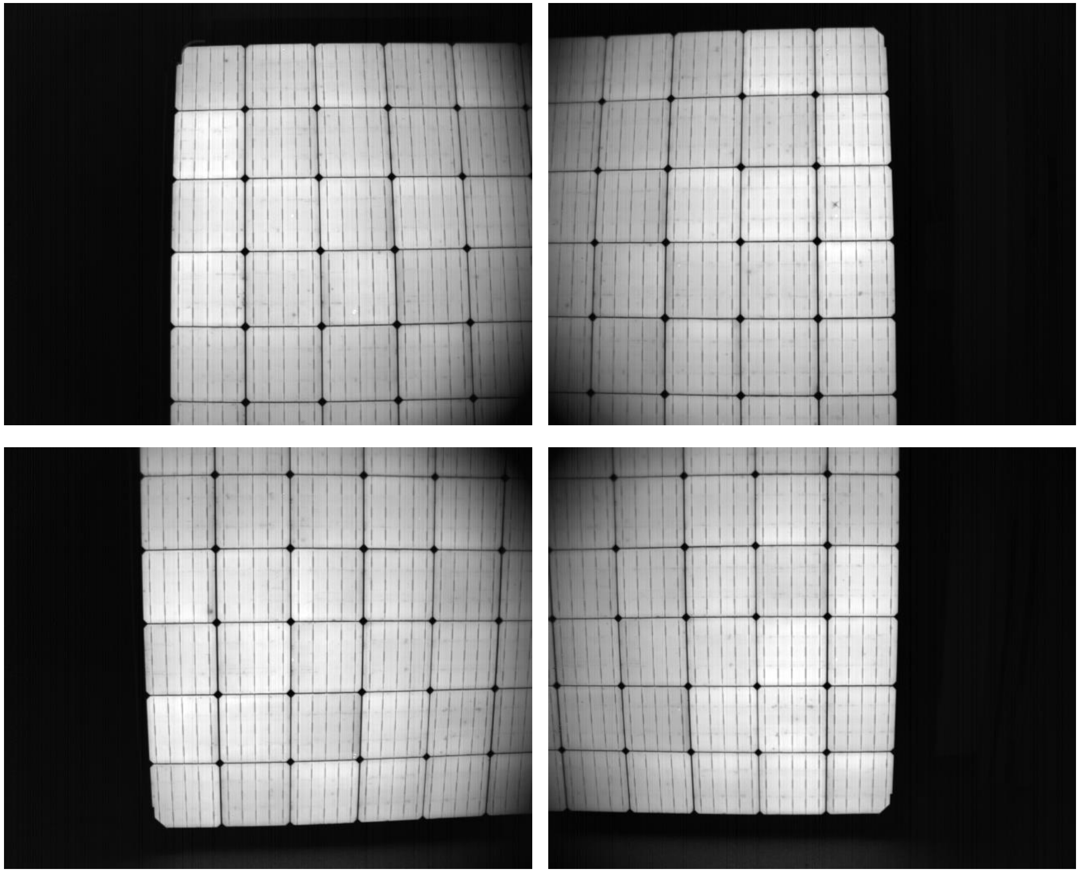
- Machine learning techniques for power prediction, deep learning
- Processing the recorded IR-movies and -images of PV-systems recorded at field conditions
- Training a deep learning model on modules with known power
- Ensuring that it generalizes to unknown data under varying conditions

Stitching of solar modules

Master project

Idea

- **Partial views** of a solar module → image of the complete module
- Obtain a higher resolution per cell



→ ?

Steps

- Detection:
  - Extend existing module detection pipeline
  - Alternatively: Code your own
- Match keypoints between images
- Compute stitched image

Caught you attention?

- Find a creative solution
- Code in whatever language you prefer
- Get 5-10 ECTS

Contact:
Mathis Hoffmann (09.153)
mathis.hoffmann@fau.de

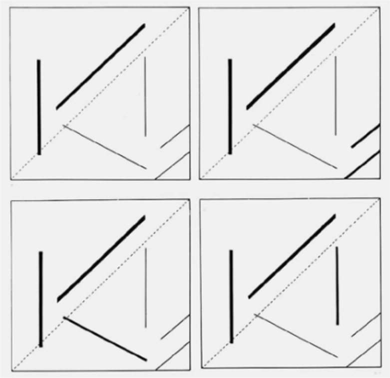
Understanding the structure / geometry in art images

Masters' project (10 ECTS)

May 28, 2019

Idea

- Using the **gaze** and **pose** estimates, find the underlying structure within paintings.
- In the below shown two images, the left image is the input, while the right image is the expected output.



Steps

- **Gaze Detection** in Art images
- **Pose Estimation** in Art images
- Combine the information from the above two steps to suggest the **underlying structure** within a paiting.
- Programming skills required : python (preferrably)

Did you find it interesting?

Contact : Prathmesh Madhu (09.156)
Email : prathmesh.madhu@fau.de

Pattern Recognition Lab

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

FACULTY OF ENGINEERING

Adversarial Examples for Emotion Analysis

Research Project (Bachelor/Master) 5/10 ECTS

Computer Vision Group, Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg

Motivation

Can we generate realistic images that can elicit emotions like real images, with people as the main object of interest? [1]

(a) Image depicting the emotions: *Pleasure and Happiness*

(b) *Adversarial Image that should show similar emotions...*

[1] Song, Yang, et al. "Constructing unrestricted adversarial examples with generative models." Advances in Neural Information Processing Systems. 2018.

Outline

1. **Dataset Building** Set up the emotion-based data to work with
2. **Methods** GANs (Generative Adversarial Networks), CycleGANs, more?
3. **I**mplementation, **A**nalysis and **C**onclusion

Interested?
Contact for more information/discussion:

Ronak Kosti (Room: 10.136)
ronak.kosti@fau.de

Pattern Recognition Lab

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG

FACULTY OF ENGINEERING

Emotion detection in Art

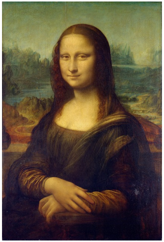
Research Project (Bachelor/Master) 5/10 ECTS

Computer Vision Group, Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg

Motivation

It is challenging to detect emotions of people in art paintings:

(a) This person is in the state of *Anger*



(b) What can be said about the emotion of this person?

Outline

Using current emotion recognition pipelines, modify various deep networks to detect emotions in Art images (or paintings in digital format).

1. **Current Research** Reviewing current state-of-art methods for emotion detection of people in images
2. **Data** Choosing an appropriate dataset for training (or already chosen!?)
3. **Implementation** Evaluate the performance of different models on the collected data
4. **A**nalysis and **C**onclusion

Interested?
Contact for further information/discussion:

Ronak Kosti (Room: 10.136)
ronak.kosti@fau.de

Saliency detection for Emotions

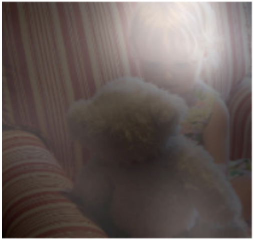
Master Thesis

Computer Vision Group, Pattern Recognition Lab, Friedrich-Alexander University Erlangen-Nürnberg

Motivation

Detecting the regions of image that are salient for emotion recognition **AND/OR** sentiment elicitation [1]

(a) Source Image



(b) *Expected Salient Region*

Figure: An image has lot of information. Which regions have more significance for emotion analysis?

[1] Fan, Shaojing, et al. "Emotional attention: A study of image sentiment and visual attention." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

Outline

Using the background research/models, find the salient regions (objects, people, stuff) which elicits emotions - *Saliency as a bridge between low and high level vision.*

1. **Literature review** Emotion Recognition *AND/OR* Sentiment Analysis
2. **Data** Mining and building Datasets/Resources
3. **Methods** Attention Models, Context Analysis, etc
4. **I**mplementation, **A**nalysis and **C**onclusion

Interested?
Contact for further information/discussion:

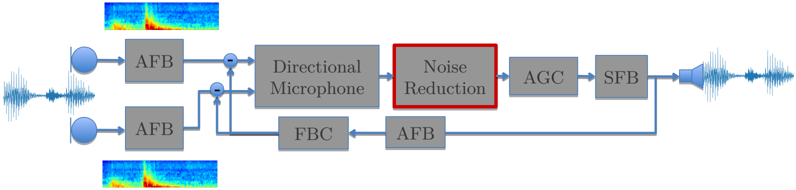
Ronak Kosti (Room: 10.136)
ronak.kosti@fau.de

Deep Learning based Noise Reduction for Hearing Aids

Hendrik Schröter
Speech Processing Group, Friedrich-Alexander University of Erlangen-Nürnberg
July 22th 2019

Hearing Aid Pipeline

Replace conventional noise reduction algorithms with deep learning based approach:



AFB: Analysis Filterbank
SFB: Synthesis Filterbank
AGC: Automatic Gain Control
FBC: Feedback Canceler

Figure: Typical hearing aid pipeline[1].

[1] Figure from: Ehrensperger, Kai, “Deep Learning-based Noise Reduction for Hearing Instrument Applications”, MA thesis (Friedrich-Alexander University Erlangen-Nürnberg, 2018)

Denoising using Deep Learning

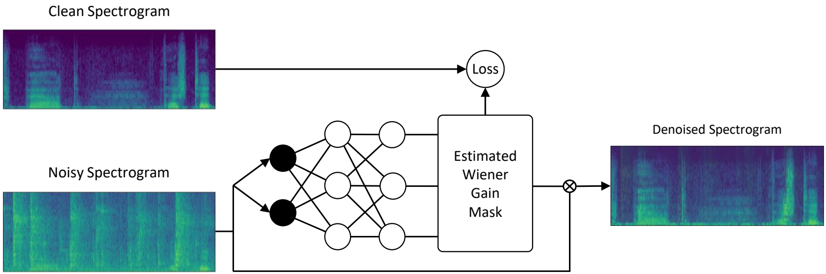


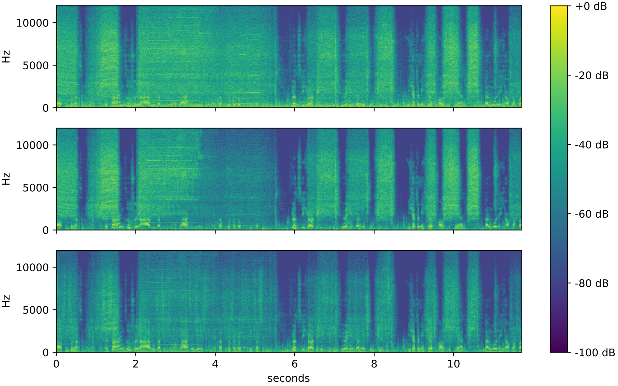
Figure: Simplified schematic figure of the neural network training.

Example: Denoising using Deep Learning

(▶ Noisy)

(▶ Conventional)

(▶ DL based)



the hearing company

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

FACULTY OF ENGINEERING

Distillation Learning for Noise Reduction

Research Project Master (10 ECTS) / Master Thesis

Hendrik Schröter
Speech Processing Group, Friedrich-Alexander University of Erlangen-Nürnberg
WS 2019/20

Distillation Learning

Improve an already existing deep-learning based noise reduction and reduce the number of parameters using distillation learning.

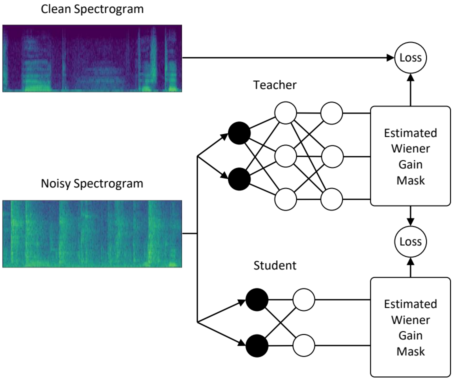
Concept distillation learning (or student/teacher networks):

- A powerful teacher network is trained on the data with hard labels.
- The student is trained to model the teacher's output distribution.
- I.e. the student does not try to predict the hard labels, but rather should learn to imitate the output of the teacher.

sivantos
the hearing company

FAU FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
FACULTY OF ENGINEERING

Distillation Learning



Teacher network:

- Deeper network, more parameters
- “Easier” input, i.e. higher SNR
- Relaxed real-time constraints

Requirements:

- Deep learning basics
- Signal processing basics (complex numbers, Fourier transform)

Contact:

Hendrik Schröter (Room 10.138)

☎ +49 9131 85 27882

✉ hendrik.m.schroeter@fau.de

Deep Learning based Beamforming for Hearing Aids

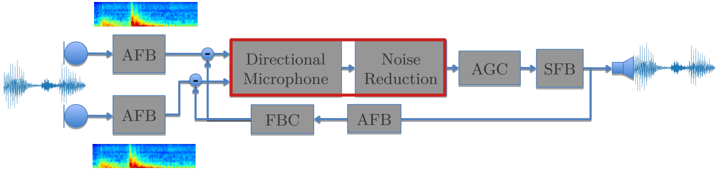
Master Thesis

Hendrik Schröter
Speech Processing Group, Friedrich-Alexander University of Erlangen-Nürnberg
WS 2019/20

Hearing Aid Pipeline

Improve an already existing deep-learning based noise reduction using multi-channel signals, which enables to exploit directional information.

Using this, we want to replace traditional directional signal processing and noise reduction with deep learning based approach:



AFB: Analysis Filterbank

AGC: Automatic Gain Control

SFB: Synthesis Filterbank

FBC: Feedback Canceler

Figure: Typical hearing aid pipeline[1].

[1] Figure from: Ehrensperger, Kai, "Deep Learning-based Noise Reduction for Hearing Instrument Applications", MA thesis (Friedrich-Alexander University Erlangen-Nürnberg, 2018)

Data

- Multi-channel noise signals from hearing aids
- Clean speech signals, transformed with HRTFs (Head-related transfer function)

Beamforming

a) Use multiple channels to estimate a multi-channel Wiener filter

b) Use multiple channels and positional information of the microphones to estimate beamforming coefficients

Requirements:

- Deep learning basics
- Signal processing basics (complex numbers, Fourier transform)

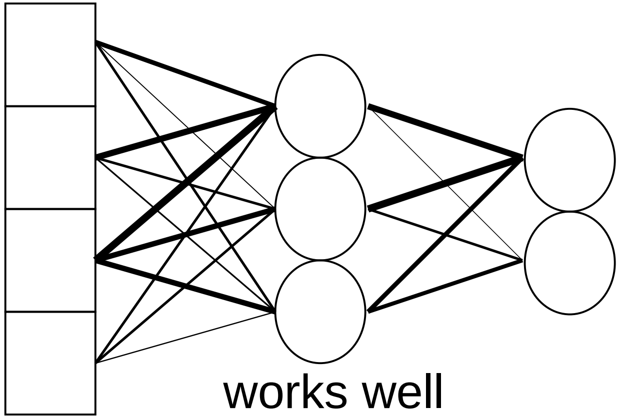
Contact:

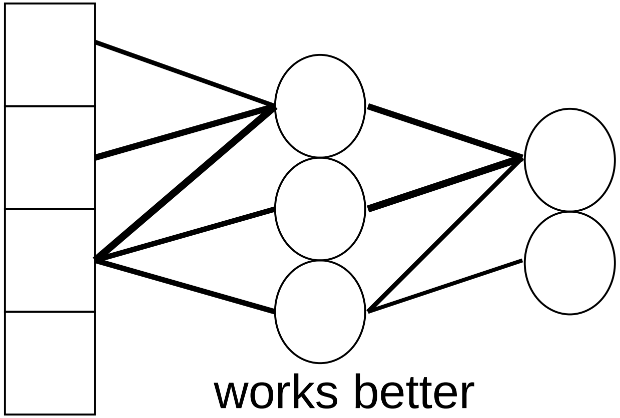
Hendrik Schröter (Room 10.138)

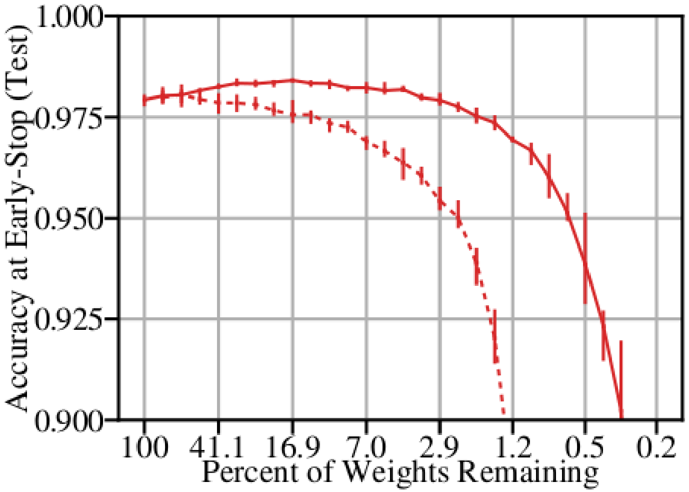
☎ +49 9131 85 27882

✉ hendrik.m.schroeter@fau.de

The Lottery Ticket Hypothesis
Finding Sparse, Trainable Neural Networks







- New network pruning approach

- Removing up to 80 % of weights

- Produces networks as good/better

- Are resulting structures consistent?

- Is transfer learning possible?

- Can layers be resized?

… many other open questions

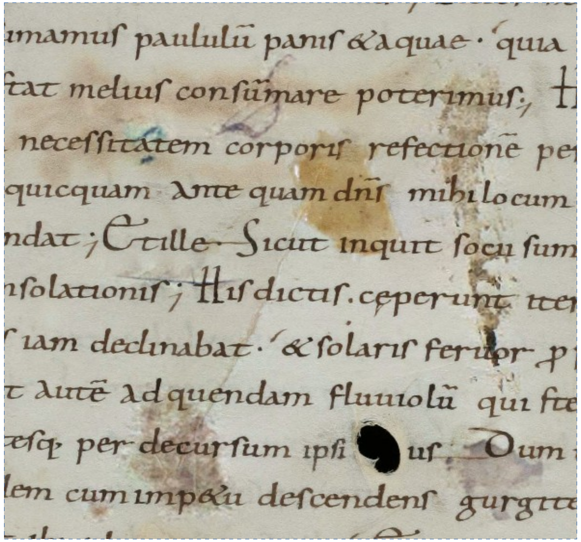
5-10 ECTS project – contact: mathias.seuret@fau.de

Based on a paper by Jonathan Frankle, Michael Carbin

Gradient-domain Data Augmentation : Degradation Model

Color domain

Gradient domain



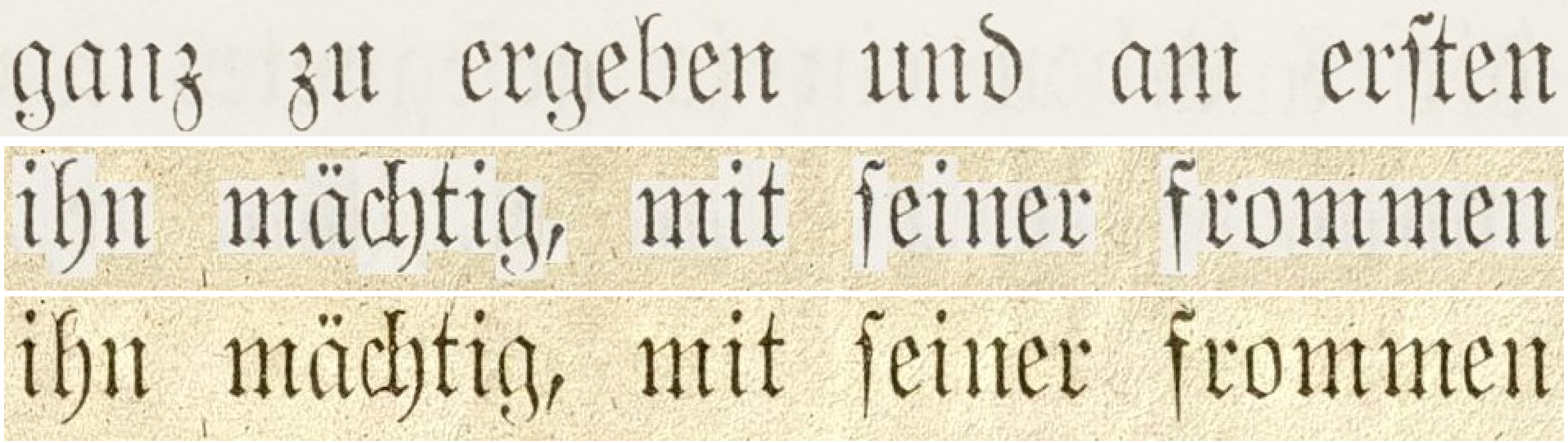
- Data augmentation method
- Paste gradients of stains
- Pixels reconstructed from gradients
- "Fools" human experts

- Random noise location: unrealistic
- Fingerprints in margins, water stains top/bottom
- Degradation location, big (unlabeled) data

➡ location probability, more realistic results

5-10 ECTS project – contact: mathias.seuret@fau.de

Mimicking Typesetting & Printing



- OCR for ancient documents: open problem
- Synthetic data needed
- Gradient-domain approach
- Toy-example proof of concept

- Automatic character & baseline extraction
- Character- & forme-level augmentation (GAN?)
- ‘’Print’’ pages with multiple fonts
- Evaluation through OCR

From 5 ECTS to Master project – contact: mathias.seuret@fau.de

Weakly supervised multimodel lesion detection and classification in mammogram & ultrasound

Master's Project (10 ECTS)

Motivation

- Mulitmodel breast image analysis for malignancy **detection** and **classification**

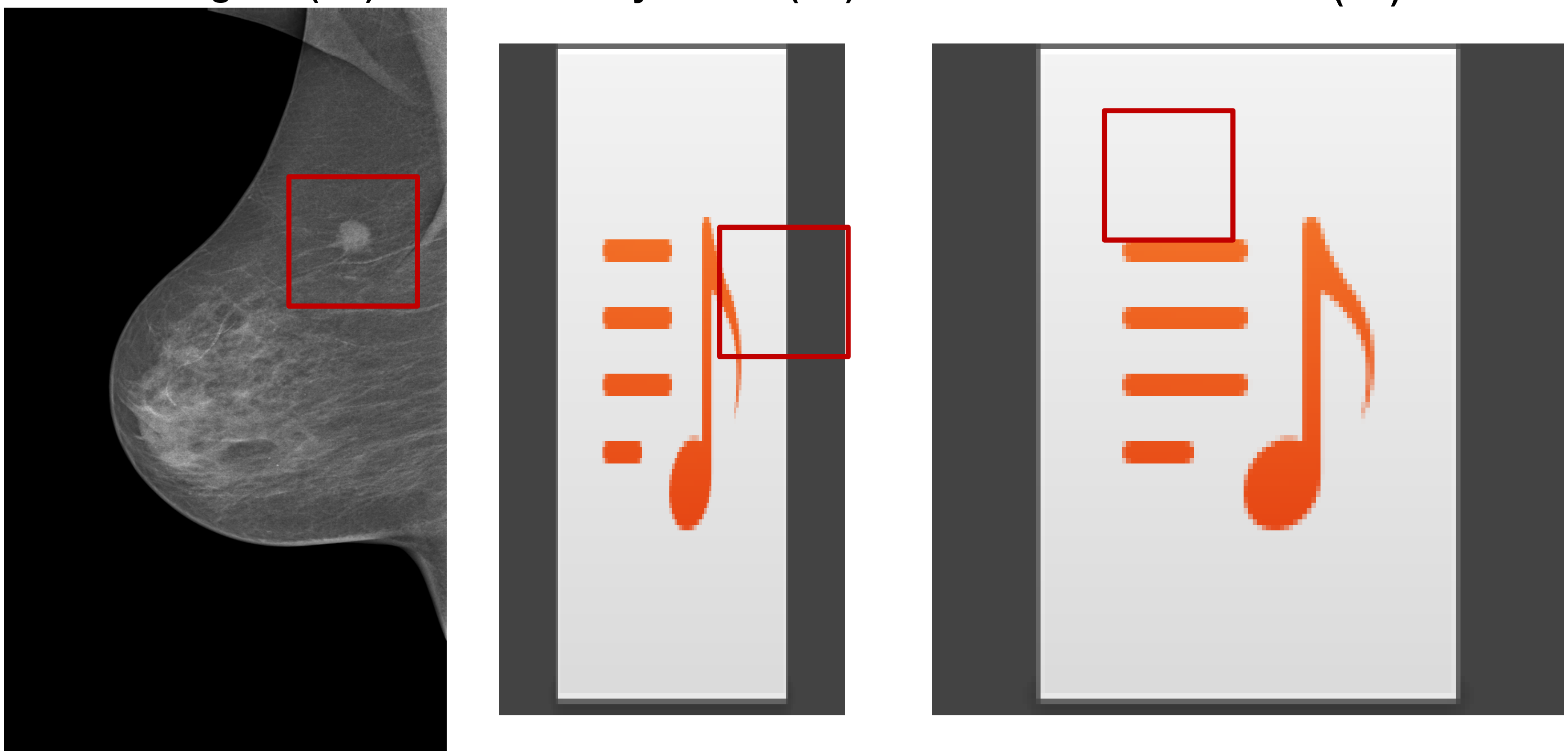


Figure 1: A 55 years old patient with a malignant lesion in left-side breast, diagnosed with BI-RADS 5

Steps

- Design weakly supervised **multimodal learning** method using cross-modality fusion
  - Feature learning level
  - Classifier/decision-making level
  - For learning: no manual annotation, but pathology label

- **Requirements:**
  - Programming skills: Python + Keras/TensorFlow
  - Deep understanding of volumetric/high-dimensional data

Contact for further information/discussion:
Sulaiman Vesal M.Sc. (Room: 10.136)
Sulaiman.vesal@fau.de

Left ventricle quantification using spatiotemporal feature learning

Master's Project (10 ECTS)

Motivation

- Assessing the **heart's function**, the **left ventricle (LV) function**, morphology and temporal dynamics is of clinical interest
  - Cavity and myocardium size
  - Cavity dimension
  - Regional wall thicknesses
  - Heart phase (systole or diastole)

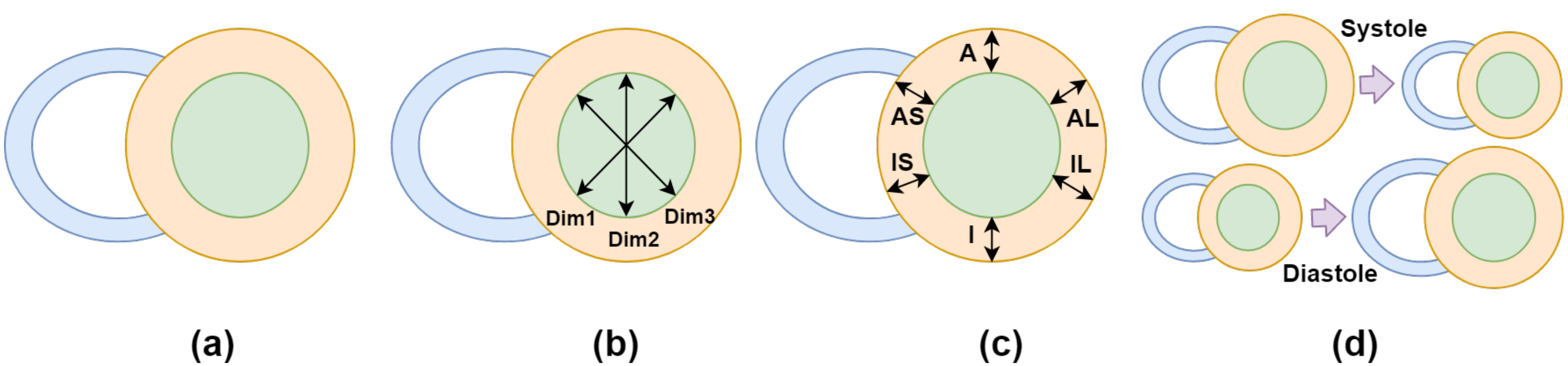


Figure 1: Illustration of LV indices to be quantified for short-axis cardiac image.(a) Cavity (green) and myocardium (yellow) areas. (b) directional dimensions of cavity (black arrows). (c) Regional wall thicknesses (black arrows). (d) Phase(systole or diastole)

Steps

- Develop effective machine learning models that can estimate a set of clinically significant LV indices
  - Supervised localization of LVs in short-axis cine MR images
  - Investigate the use of **spatiotemporal convolutions**
  - **Multi-task learning** for both cardiac phase detection and LV indices estimation

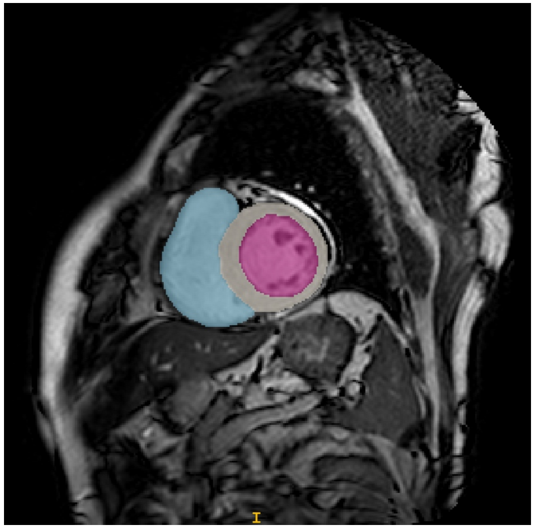


Figure 2: Cine-MR image with segmented left ventricle and myocardium

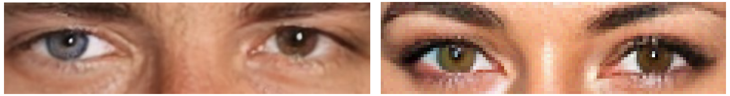
- **Requirements:**
  - Programming skills: Python + Keras/TensorFlow
  - Deep understanding of volumetric/high-dimensional data

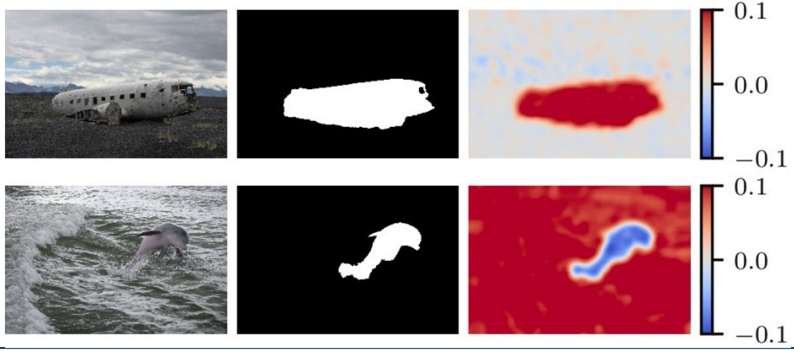
Contact for further information/discussion:
Sulaiman Vesal M.Sc. (Room: 10.136)
Sulaiman.vesal@fau.de

Multimedia Security Group

- **Image enhancement:**
  - Superresolution of compressed data
- **Image/video forensics:**
  - Has an image been retouched?
  - Is part of a video computer-generated?

How "dangerous" is GAN-generated CGI?





What traces leave manipulations in the compression container?

How can we learn to detect manipulated faces from few training examples?

Example Open Projects or Theses

Guess characters on unreadable licence plates	Statistical video manipulation detection	Physics-based image manipulation detection	How easily can DL-based forgery detectors be fooled?
-> CNN to deal with strongly compressed video frames of licencse plates	-> Deep anomaly detector / device parameter regressor	-> Learning-based methods for classical vision tasks, e.g., shadow segmentation	-> Can we construct a counter-forensics adversarial example image laundry just from "innocent" JPEG settings?

Who to talk to

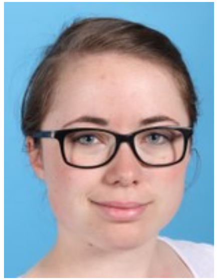
- We run projects between the Pattern Recognition Lab, the Computer Graphics Lab, and the IT Security Infrastructures Lab
- Group Members

Amir Davari

Benjamin Hadwiger

Benedikt Lorch

Patrick Mullan



Franziska Schirrmacher

- For concrete Projects or Theses:
  Contact Franziska Schirrmacher, franziska.schirrmacher@fau.de