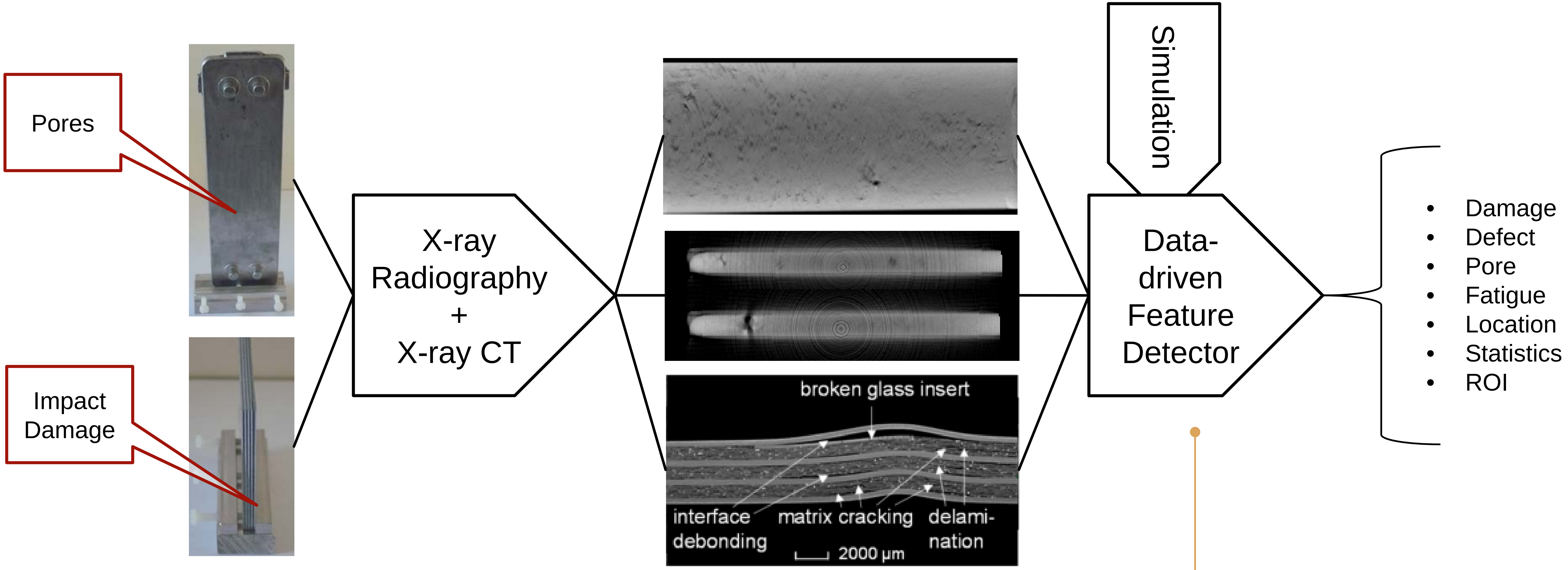
STEFAN BOSSE[1,2*]
DIRK LEHMHUS[3]

[1]Fachbereich Mathematik und Informatik
Universität Bremen
sbosse@uni-bremen.de www.edu-9.de
[2]Fakultät Maschinenbau
Universität Siegen
[3]Fraunhofer IFAM, Bremen



AUTOMATED X-RAY-BASED DAMAGE DETECTION AND CHARACTERISATION IN MATERIALS BY DATA-DRIVEN ANOMALY PREDICTOR MODELS TRAINED BY FUSION OF REAL AND SIMULATED X-RAY DATA

Stefan Bosse

FEMS Euromat 2023 - 07.09.2023

Universität Bremen

CONTENT

Universität Bremen

INTRODUCTION: NON-DESTRUCTIVE TESTING

- Spatially resolved Inspection and Testing of structures requires **image-based measuring methods**
- **Non-destructive testing** (NDT) of metal-based structures can exploit different imaging methods, mainly:
  - **X-ray** Radiography (single projection) and Computer Tomography (CT, multi-projection)
  - Guided **Ultrasonic Waves** (GUW) and Ultrasonic Sonography
- **Detection of hidden damages, defects, and impurities (e.g., pores) is still a challenge!**

Primary Goal. Automated Damage, Defect, and Impurity Detection in materials and structures including composites using single X-ray projection images (from LowQ/MidQ devices) and data-driven feature marking models (Convolutional Neural Networks).

INTRODUCTION: EXPERIMENTS AND GOALS

Different specimens, structure geometries, materials, and defects are considered in this work! They pose different coincidence between material and image features.

1. **Homogeneous Aluminum Die Casting Plates (150x40 mm) with pore defects**
2. **Composite Fibre Metal Laminate plates (FML, aluminum and PREG layers, 50 x 50 mm) with impact damages posing layer delaminations, deformation, cracks, and kissing bond defects.**

Secondary Goal. Migration from laboratory (High-Q/Mid-Q) to in-field (Low-Q) measuring techniques and devices.

INTRODUCTION: AUTOMATED FEATURE DETECTION

- Automated feature detection and marking in measuring images can occur on different levels:
  - **Region-of-Interest Search**
  - **Feature marking and Maps**
  - **Damage and defect classification**
  - **Damage and defect localization**
  - **Global statistical aggregates (e.g., pore density, distribution)**
- Either classical numerical and model-based algorithms (e.g., edge detection using a Soebel filter or Canny detectors) or data-driven models are used for feature marking („Machine Learning“)

Data-driven models require data! Data must contain a sufficient statistical variance and distribution of features to be detected. That‘s the first issue with most engineering data! Additionally, supervised data modelling requires accurately labelled strong feature examples, commonly not available, and being the second issue and downfall in data-driven modelling.

DATA-DRIVEN NDT FRAMEWORK

Radiography

One Projection

Transmission

MidQ and LowQ Devices

Input data for inference with data-driven predictor models

CT

Multiple Projections

MidQ and HighQ Devices

Used for detailed material and defect characterisation

Input for CAD models and simulatio

ML

Data-driven Modelling

Feature/ROI Marking

Pore Analysis

Damage Detection

Anomaly Detection

Supervised / Unsupervised

Semi CT

Some Projections

LowQ Devices

Reconstruction or Feature Extraction by data-driven CNN ML models

Simulation

2D/3D

Raytracing

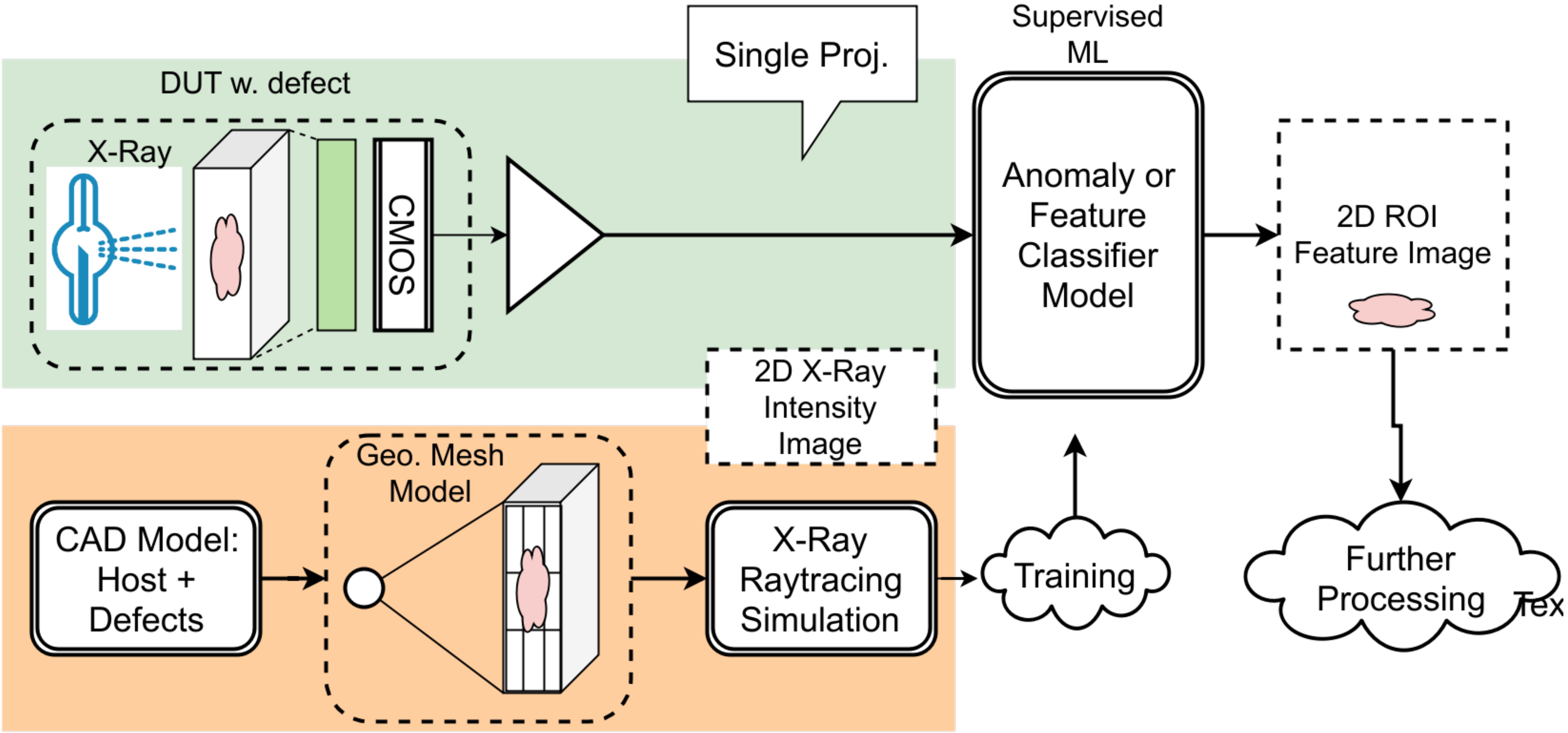
Transmission

(Reflection and Diffraction neglected)

Input data for training of data-driven predictor models

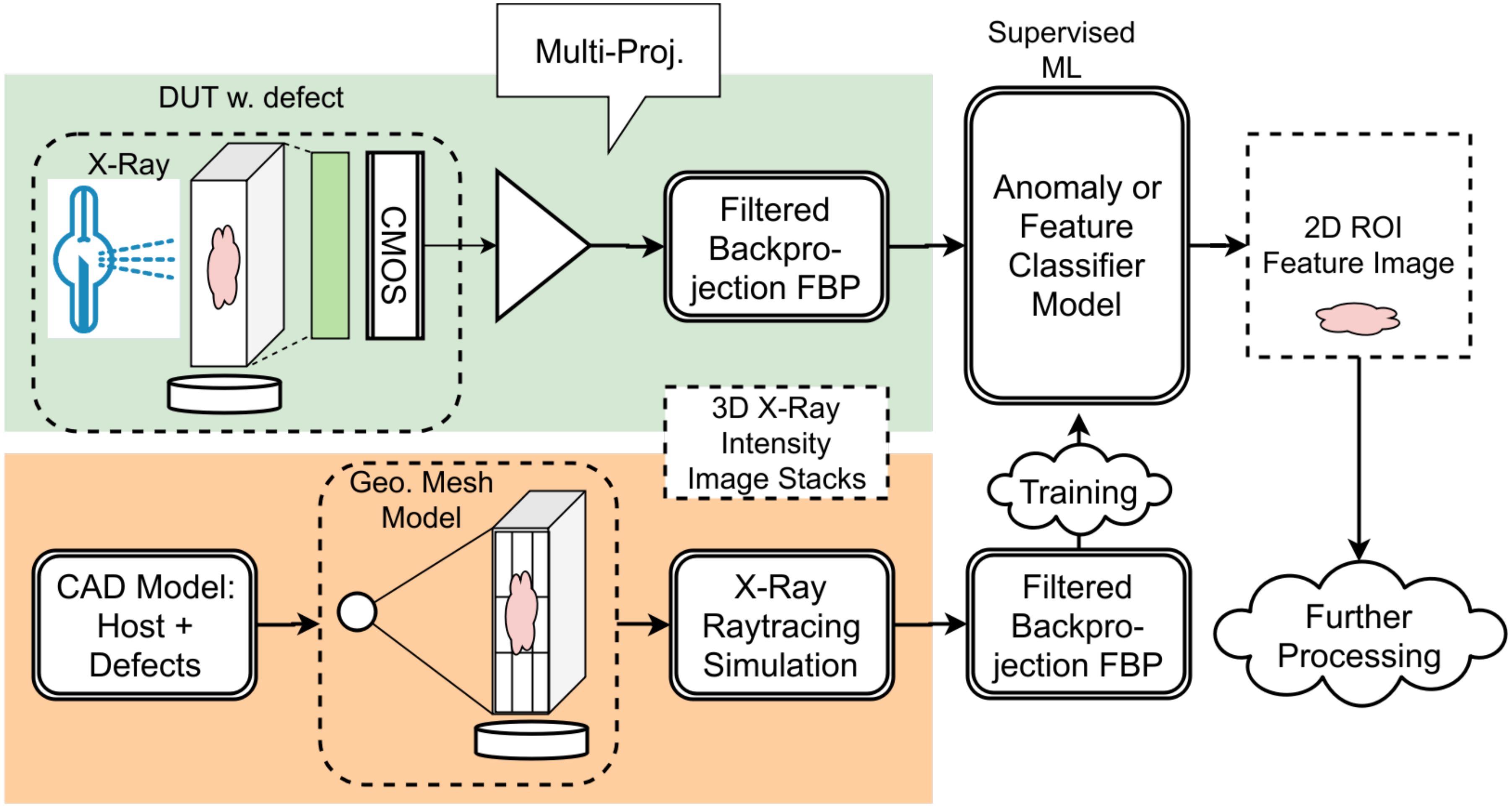
Data and Meth. Fusion

PRINCIPLE CONCEPT (1) : X-RAY RADIOGRAPHY



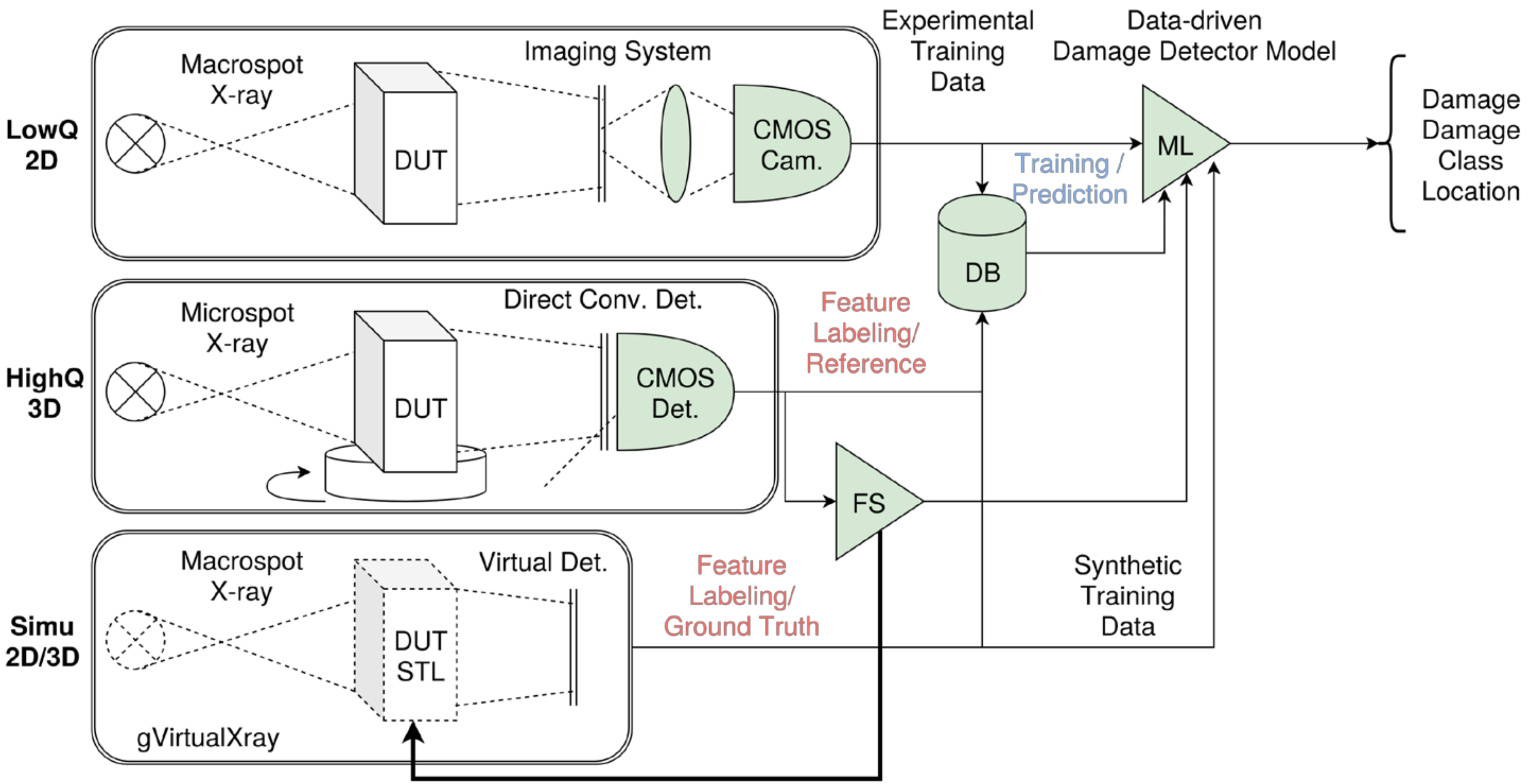
Universität Bremen

PRINCIPLE CONCEPT (2): X-RAY COMPUTER TOMOGRAPHY



Universität Bremen

ADVANCED CONCEPT

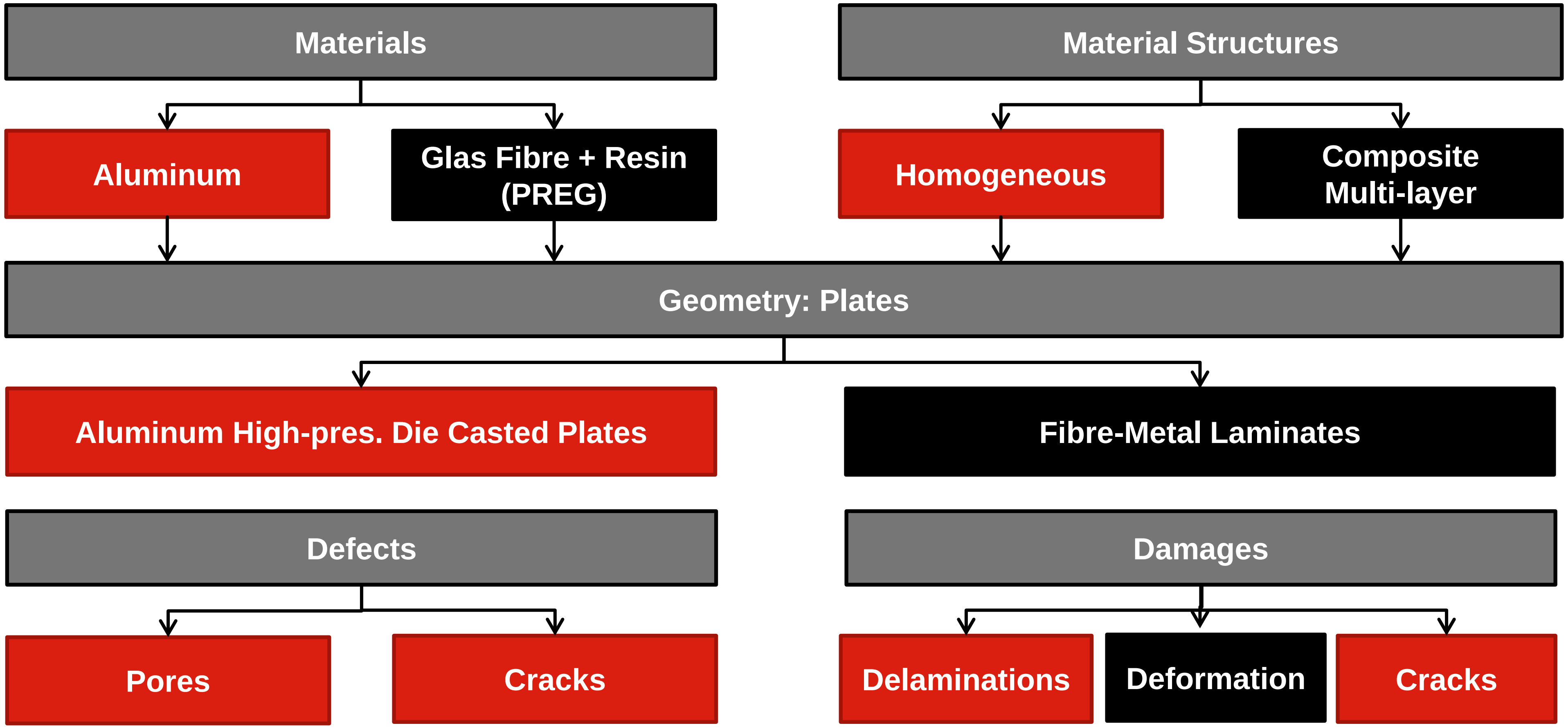


Universität
Bremen

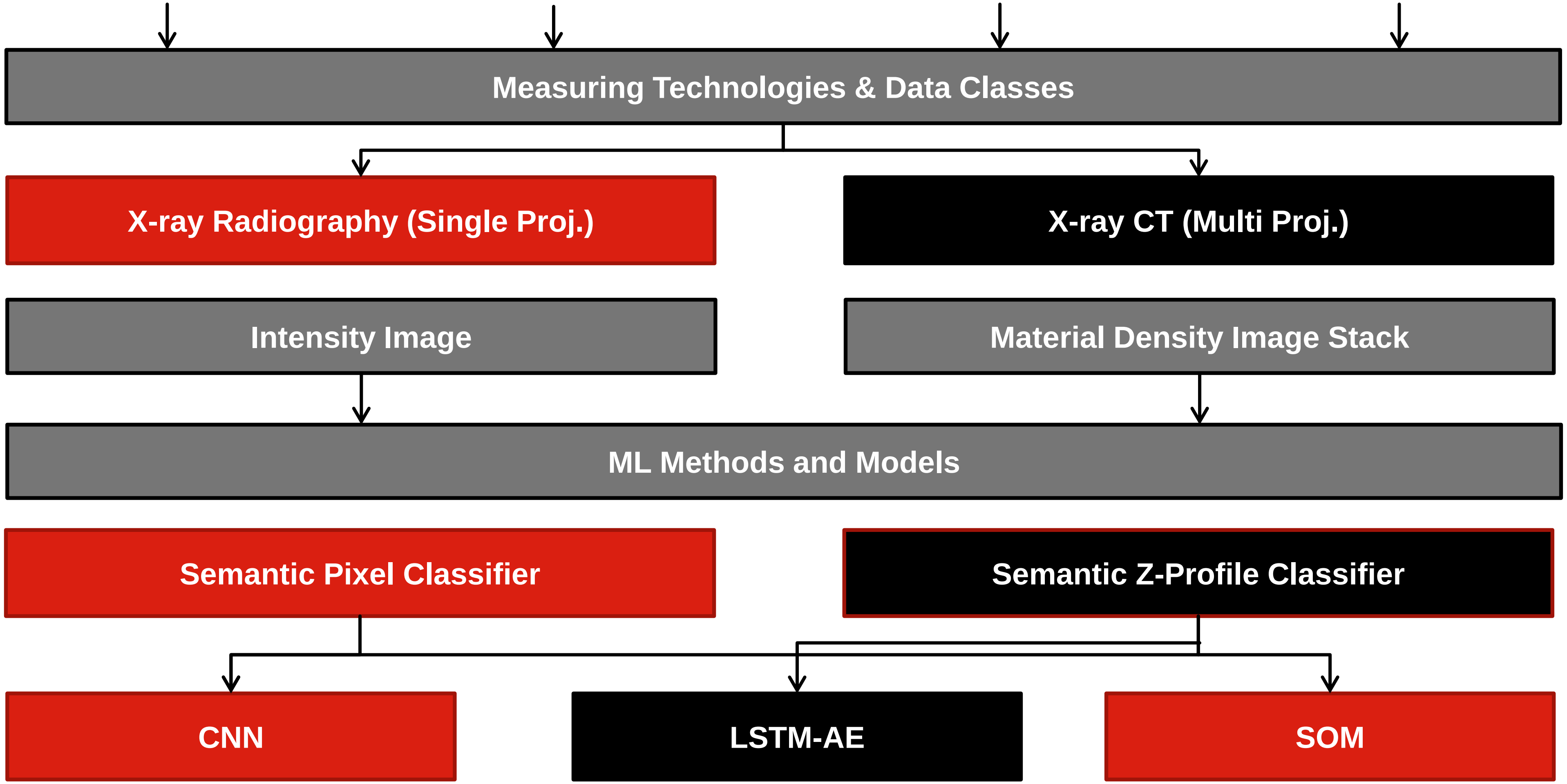
DEVICE CLASSES

	High-Q	Mid-Q	Low-Q
Single Projection	✓	✓	✓
Mult-Projection (Rotation)	✓	✓	✓
X-ray Tube Focal Diameter	5µm	0.8mm	0.8mm
X-ray Voltage/Current	-120 kV/2 mA	-120 kV/10 mA	-70 kV/1 mA
Detector	2000x2000 20 µm Screen / Microsc.	1000x1000 200 µm Direct Sci.	2000x1000 3/40 µm Screen/Imaging
Digital Resolution [Bits]	16	16	12
Sampling Time	500 ms-10 s	10 ms-1 s	5 s
Distance Object/Source	5-10 cm	10-50 cm	10-30 cm
Costs	1000 k€ (Zeiss)	500 k€ (IFAM)	1 k€ (Bosse)

TAXONOMY

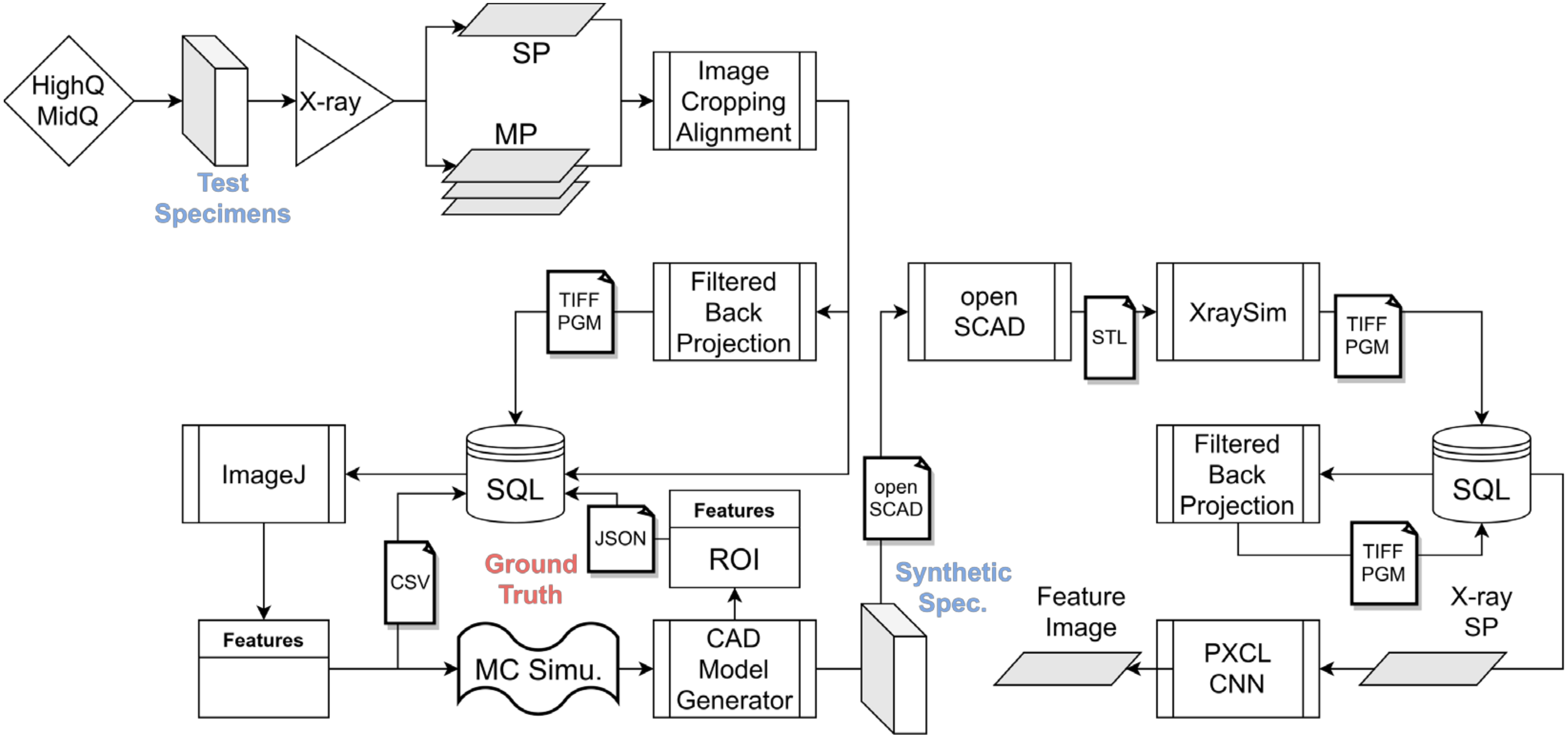

Materials
Material Structures
Aluminum
Glas Fibre + Resin (PREG)
Homogeneous
Composite Multi-layer
Geometry: Plates
Aluminum High-pres. Die Casted Plates
Fibre-Metal Laminates
Defects
Damages
Pores
Cracks
Delaminations
Deformation
Cracks

TAXONOMY

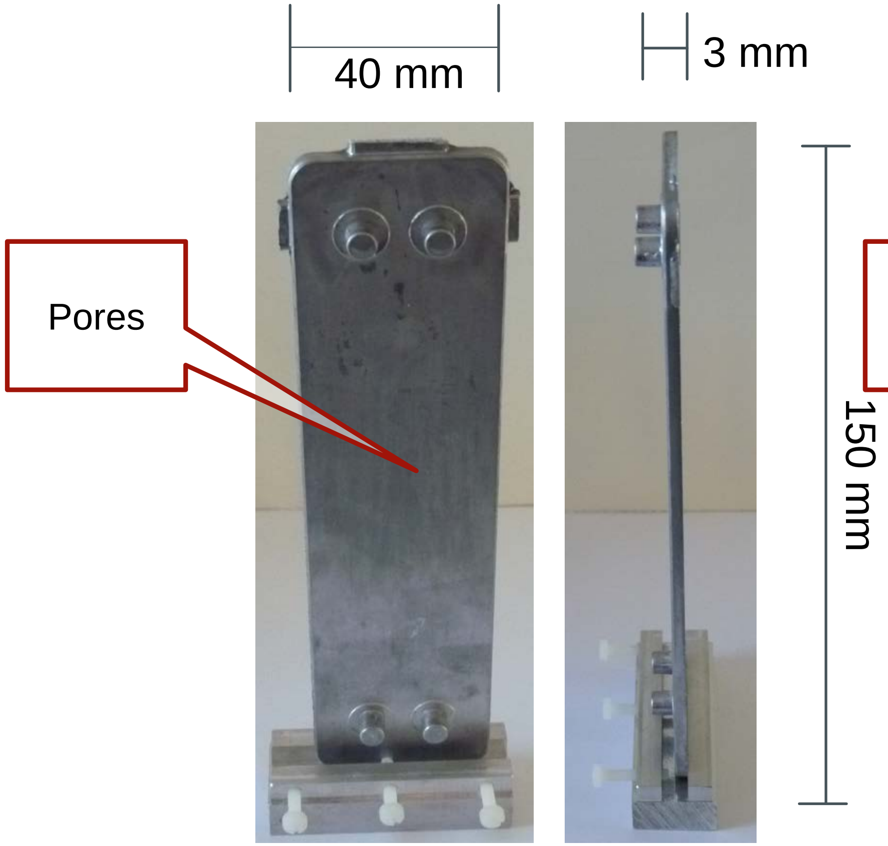


Universität Bremen

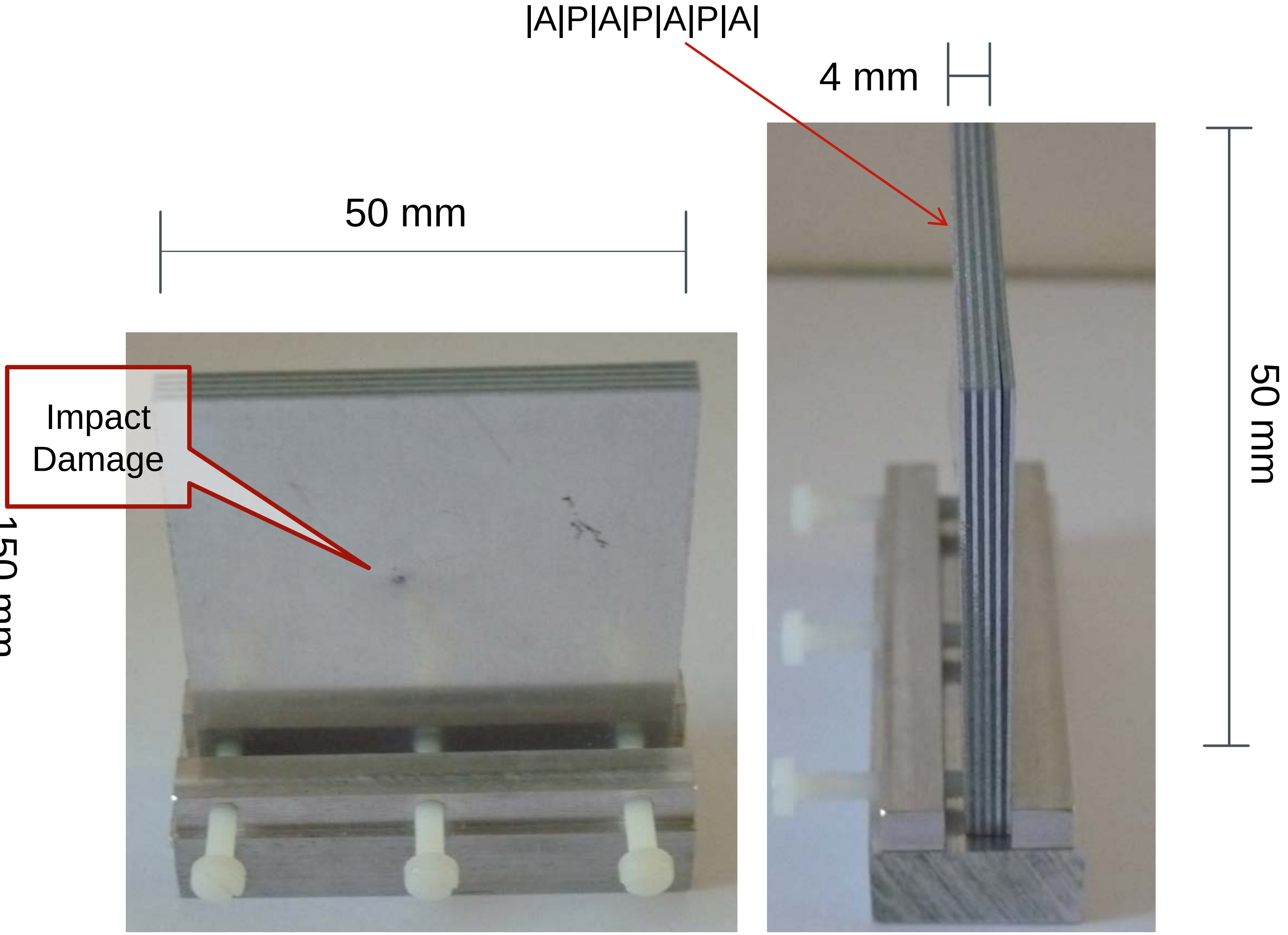
WORKFLOW


HighQ
MidQ
Test
Specimens
X-ray
SP
MP
Image
Cropping
Alignment
Filtered
Back
Projection
TIFF
PGM
open
SCAD
STL
XraySim
TIFF
PGM
ImageJ
SQL
JSON
Features
ROI
CSV
Ground
Truth
Features
MC Simu.
CAD
Model
Generator
open
SCAD
Synthetic
Spec.
Feature
Image
Filtered
Back
Projection
SQL
TIFF
PGM
X-ray
SP
PXCL
CNN

SPECIMENS

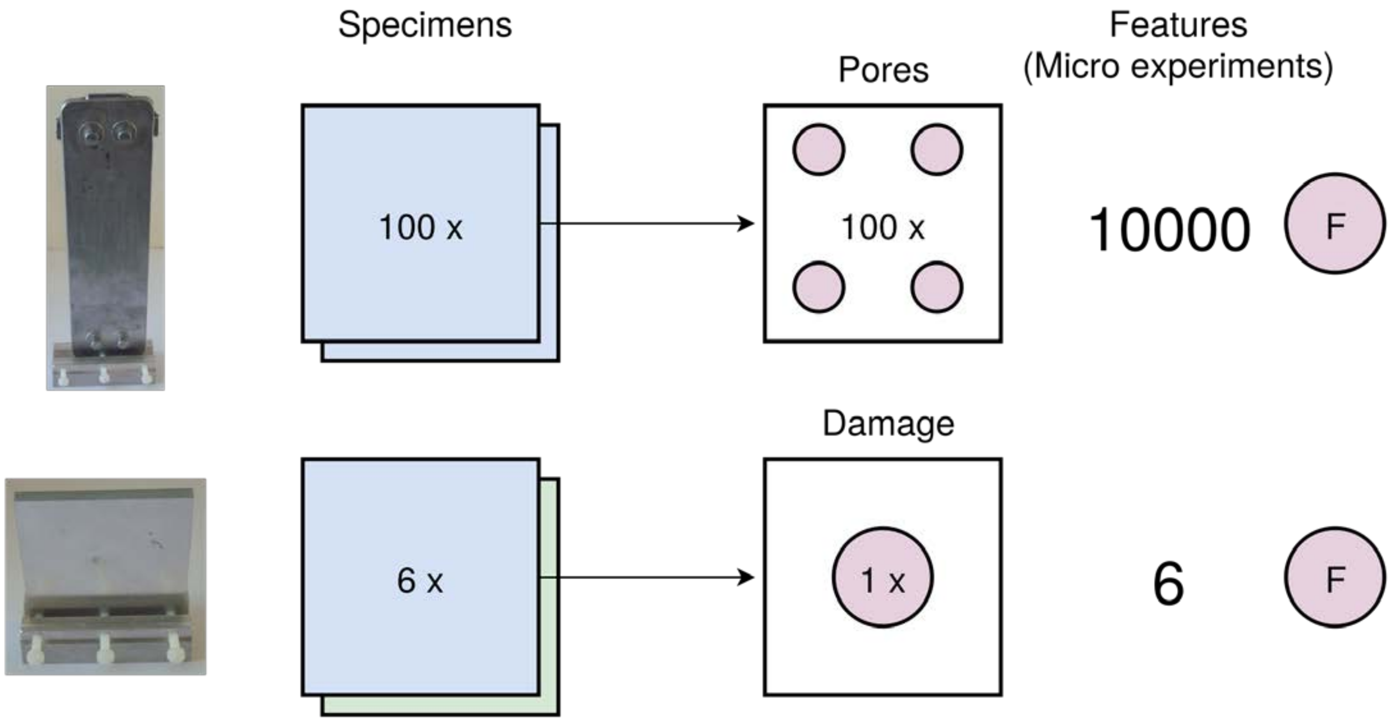


1. Aluminum Die Casting Plate (IFAM)



2. GLARE FML Plate (DFG FOR 3022)

DATA VARIANCE: THE FIRST CHALLENGE



In this work a semantic pixel classifier is used for feature marking. From the model point of view, each pixel (and neighbour pixels) of an X-ray image or volex of a CT image stack is a sample instance!

METHODS AND ALGORITHMS

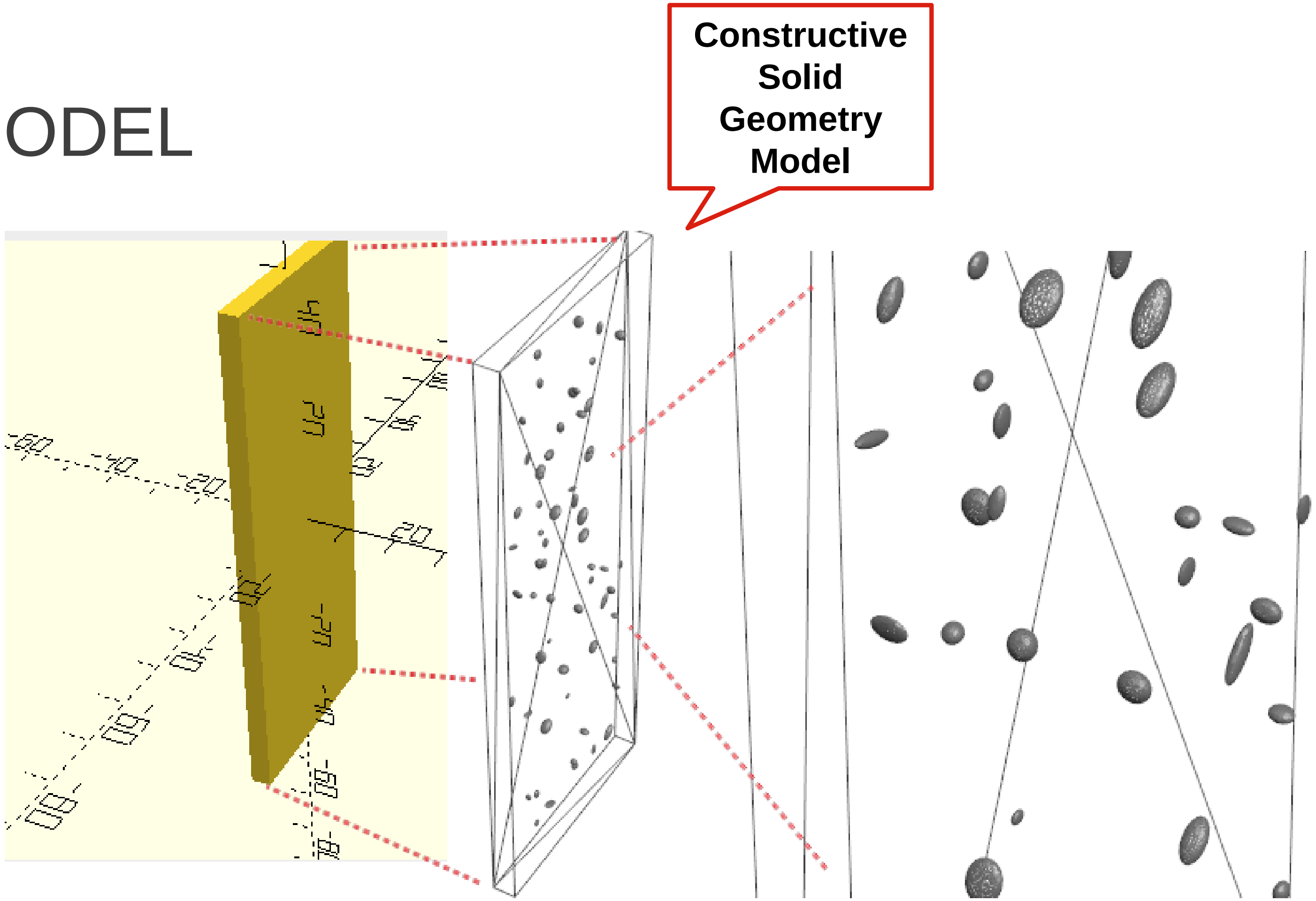
- 3D CAD modelling using automated model code generators, Monte Carlo simulation, and openSCAD
- X-ray simulation using own simulation software based on prooven and accurate gvxr/gVirtualXray library
- 3D CT reconstruction with Filtered Back Projection (using sine filters)
- Convolutional Neural Networks in different flavors
- Anomaly detectors applied to images and CT volume data

X-RAY SIMULATION

- **Input**: Polygon mesh grid (STL, Stereolithography file format) model
  - An STL file describes a raw, unstructured triangulated surface
  - Decomposition of multi-material structures in single density parts (finally merged in simulator)
  - 3D Model design: Constructive Solid Geometry (CSG)
- **Output**: X-ray intensity image with a specific detector resolution (number of pixels) and pixel size, floating point or integer data format (at least 16 Bits)
- Spatial source, object, and detector geometries can be fully parametrized including rotatated planes
- **Core software library**: gvxr / gVirtualXray using GPU computations and the OpenGL Shading Language (faster than 1ms / image)
  - https://gvirtualxray.fpvidal.net/
  - Based on the **Beer-Lambert law to compute the absorption of light** (i.e. photons) by 3D objects (here polygon meshes).

X-RAY SIMULATION: CAD MODEL

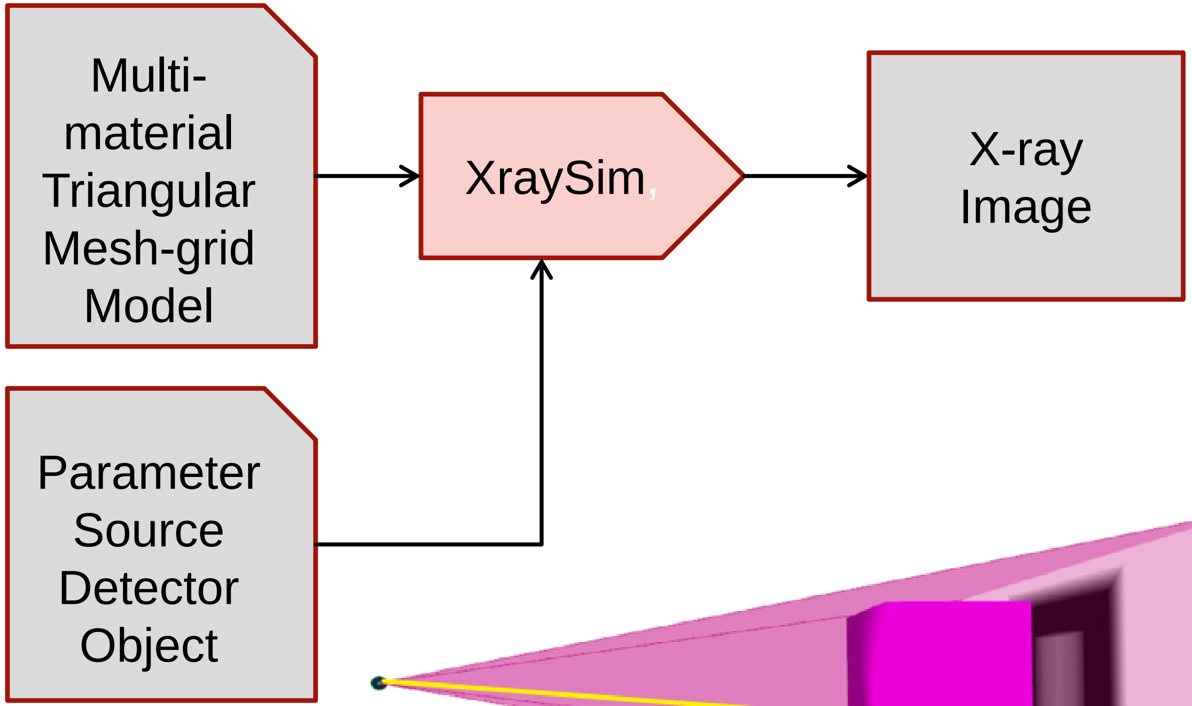
```
rotate ([90,90,90])
difference () {
  rotate ([90,0,0]) cube([100,4,40],true);
  union () {
    translate([3.17,6.14,0.67])
      rotate ([0,0,-1.43])
      scale([1.15,1.12,0.31])
      sphere(r=0.5,$fn=20);
    translate([-16.66,-4.05,0.39])
       rotate ([0,0,40.14])
       scale([0.89,2.21,1.46])
       sphere(r=0.5,$fn=20);
…
  }
}
```

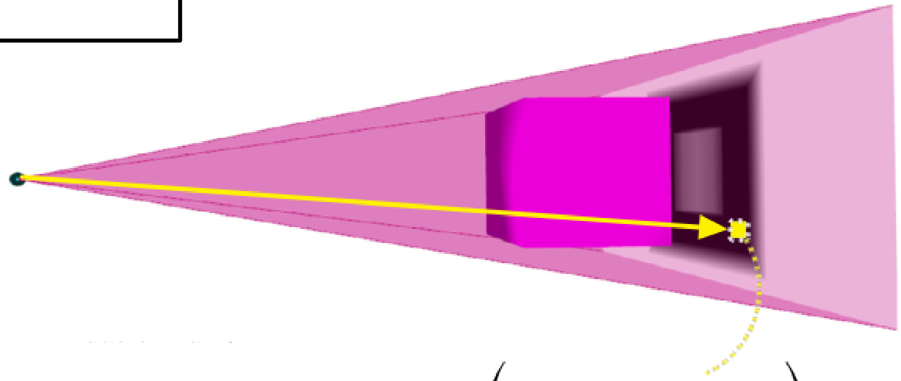


X-RAY SIMULATION

- C++ simulation library gvxr/gVirtualXray[1]
- Integrated in own simulator program XraySim: https://github.com/bslab/xraysim
- GPU/OpenGL Ray tracing using Beer-Lambert law
- Attentuation along direct transmission path from source to detector – no scattering and reflection

> “I(x,y) is the integrated energy in eV received by pixel (x,y). In the polychromatic case, the beam spectrum is discretised in several energy channels. E_i corresponds to the energy in eV of the i-th energy channel. $D(E_i)$ is the number of photons emitted by the source at that energy E_i. The detector response $R(E_i)$ mimics the use of a scintillator by replacing the incident energy E_i with a smaller value, i.e. $R(E_i) < E_I$. $\mu_j(E_i)$ is the linear attenuation coefficient of the j-th material at energy E_i. $d_j(x,y)$ is the path length.”



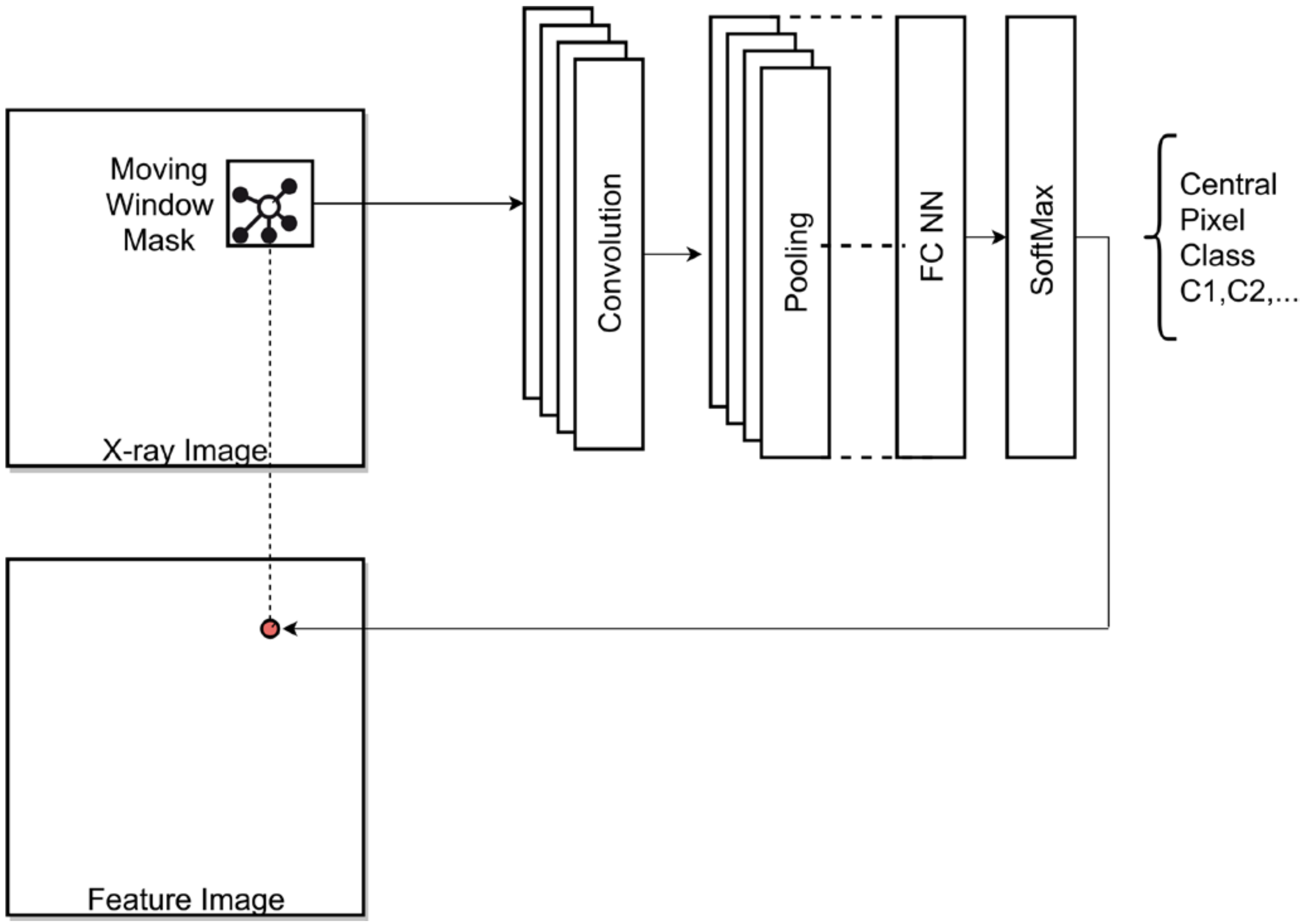


$$\mathbf{I}(x,y) = \sum_i \mathbf{R}(E_i)\,\mathbf{D}(E_i)\,\exp\left(-\sum_j \mu_j(E_i)\,\mathbf{d}_j(x,y)\right)$$

[1] Simulation of X-ray projections on GPU: Benchmarking gVirtualXray with clinically realistic phantoms, Jamie Lea Pointon, Tianci Wen, Jenna Tugwell-Allsup, Aaron Sújar, Jean Michel Létang, and Franck Patrick Vidal Computer Methods and Programs in Biomedicine, 2023.….

SEMANTIC CNN PIXEL CLASSIFIER

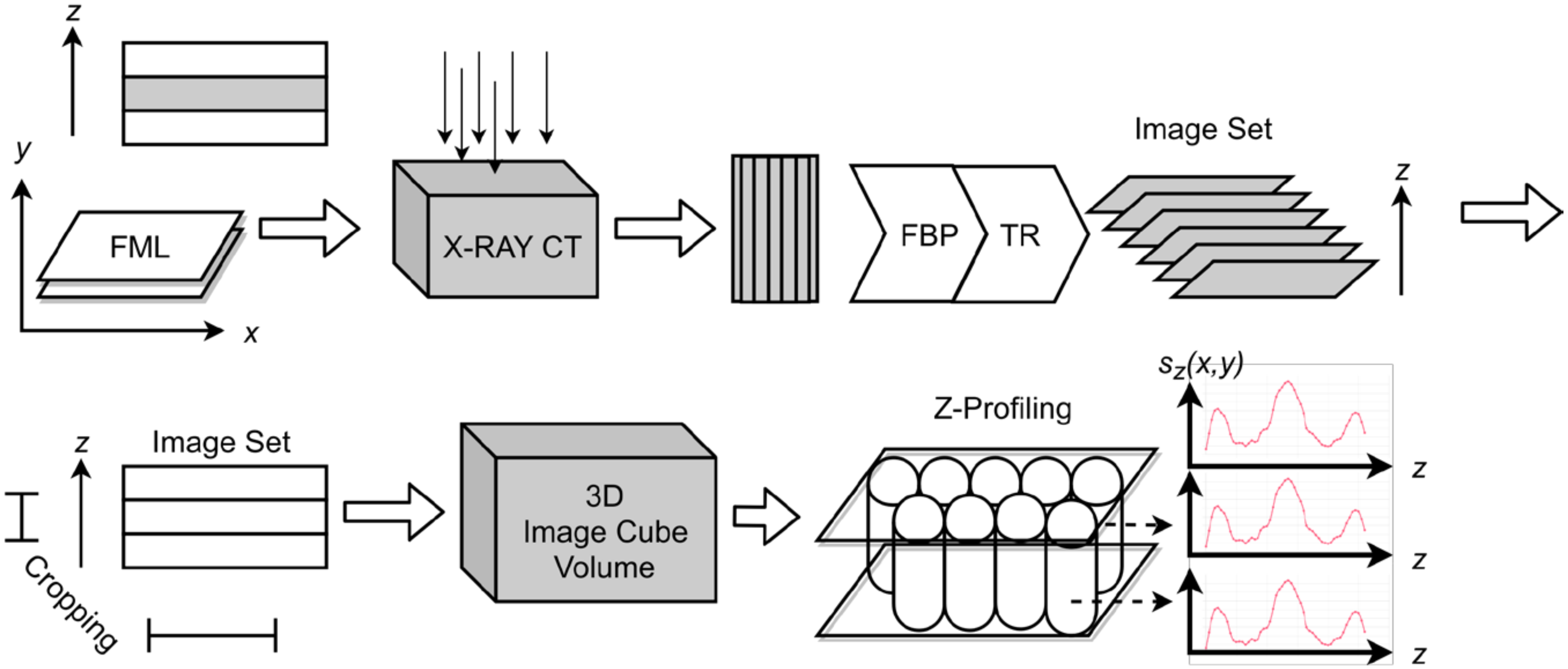
- **Input**: A sub-window of an X-ray image
- **Output**: The object class to which the central pixel of the window belongs
- The CNN classifier is applied to all pixels of an input images and produces an equally sized feature marking output image
- Point clustering (e.g., using DBSCAN) can be used to extract list of geometric objects (pores, damages, …)
- Supervised positive training (classification of known features classes)



Universität Bremen

DAMAGE DETECTION IN FML CT DATA

- Goal: Find (or mark) damages (deformations, cracks, delaminations) in 3D CT volumes
- Method: Z-Slicing of 3D CT volumes and application of a data-driven feature detector to z-profiled slices



DAMAGE DETECTION: POSITIVE VS. NEGATIVE TRAINING

Negative Training

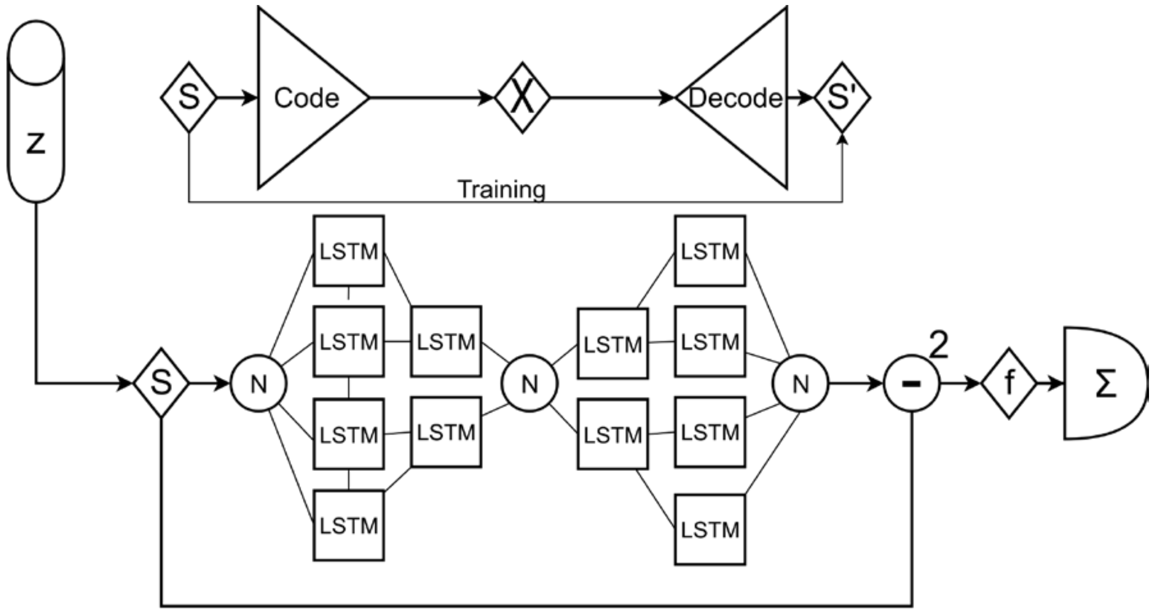
The predictor model is trained with well known defects and damages (classifier). Suitable if there is a solid reference data set and an already existing knowledge base.

Positive Training

The predictor model is trained with the base-line reference without defects and damages (anomaly detector). Suitable to cover defects and damages with known and unknown characteristics.

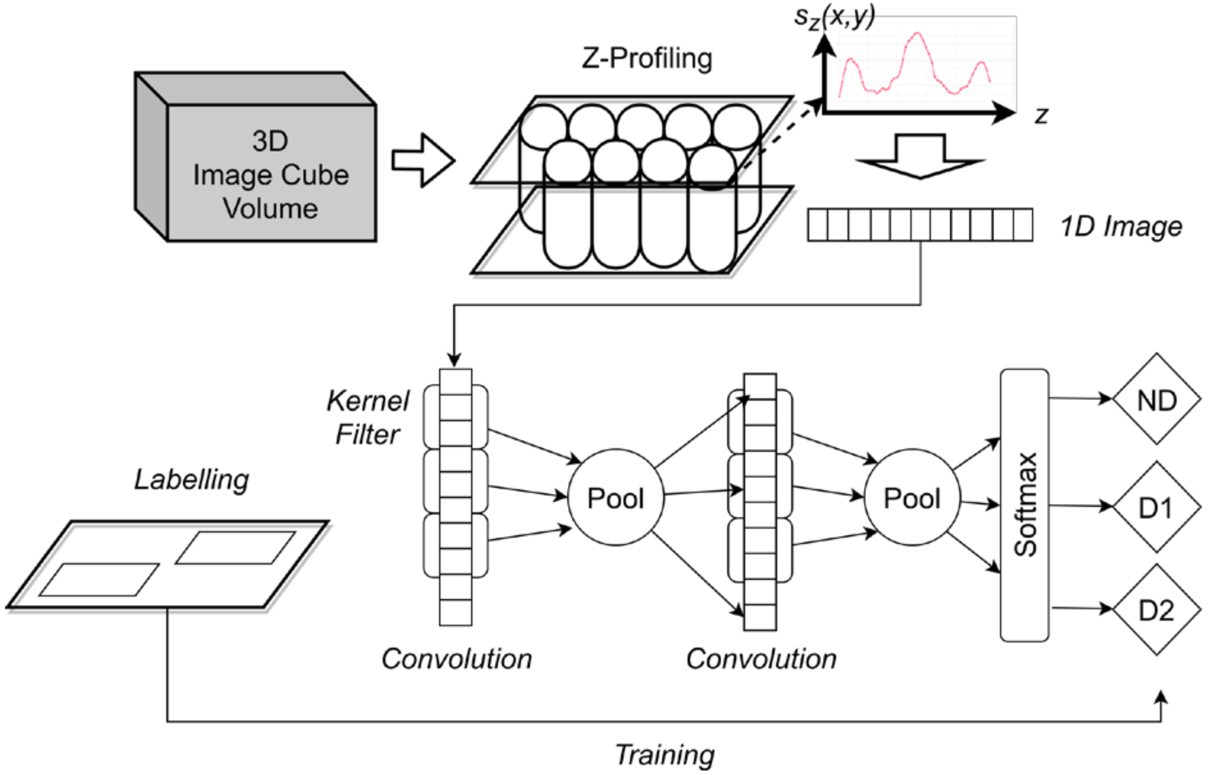
ANOMALY DETECTION IN FML CT DATA: NEGATIVE TRAIN.

- An anomaly detector is build with a Autoencoder, either using a CNN or a LSTM-ANN
- The AE is trained with z-profile slices without defects or damages (base-line, ground truth data)
- The AE „learns“ the z-profile structure of the FML plates and outputs a simplified representation (neg. Train.)
- If there is a damage/defect, the AE is not able to reconstruct the base-line structure, and an error occurs



ANOMALY DETECTION IN FML CT DATA: NEGATIVE TRAIN.

- A CNN is trained with damaged z-profiles to classify damaged and undamged z-profile slices

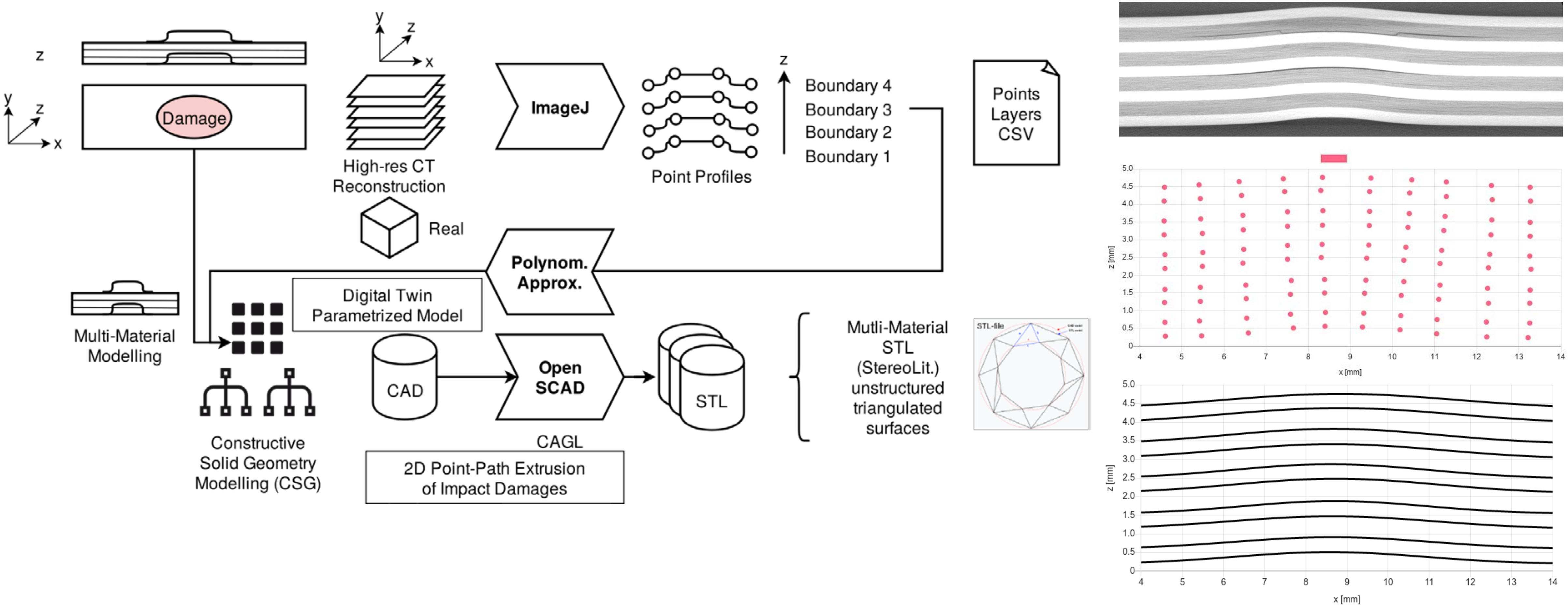


Universität Bremen

ANOMALY DETECTION IN FML CT DATA: SIMULATION

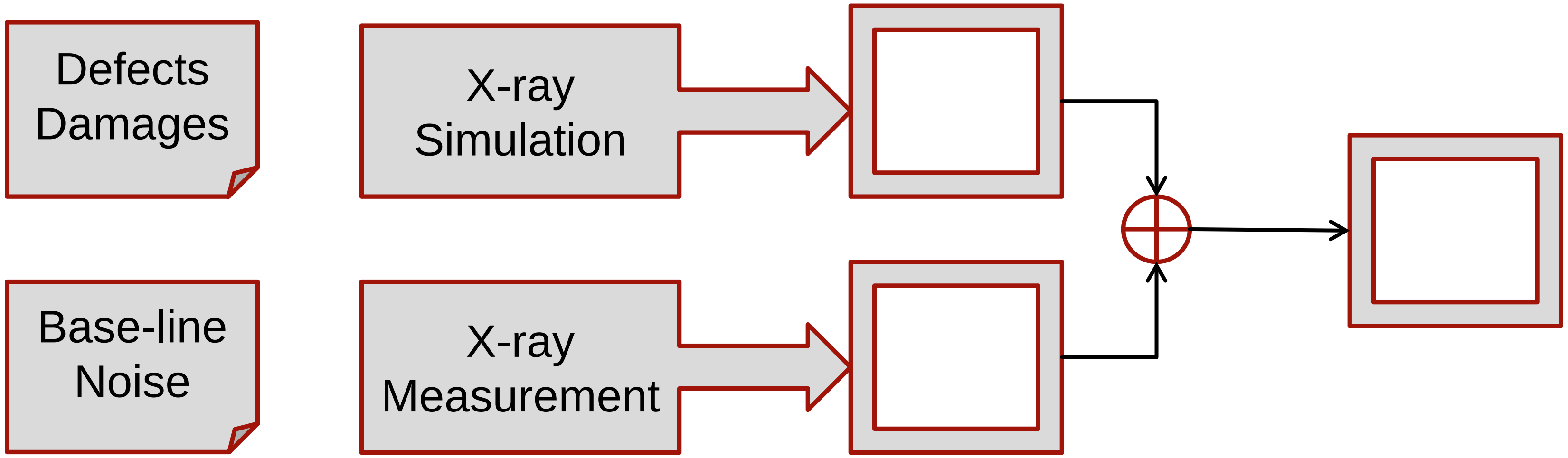
- A typical sample set contains less than 10 different specimens, each with a distint and unique impact damage (and base-line = no damge)
- **Data augmentation by simulation is required to increase feature and data variance!**
- But in contrast to mechanical pore modelling in homogeneous materials, modelling of impact damages in FML is much more complicated reaching high accuracy (wrt. real structures and images)
- Hand-made layer boundary point-marking using image tools
- Functional approximation $\rightarrow$ 3D CAD model $\rightarrow$ X-ray simulation

ANOMALY DETECTION IN FML CT DATA: SIMULATION



FUSION OF REAL AND SYNTHETIC X-RAY IMAGES

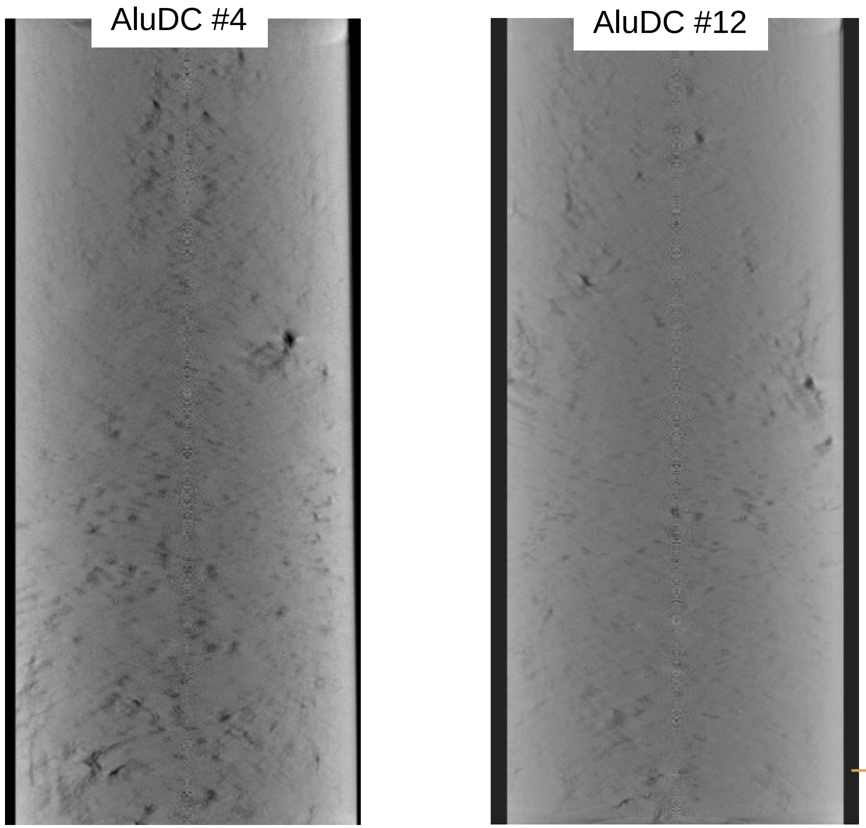
- Modeling of specific material structures and image patterns can be a challenge:
  - Fibre Materials (Irregular, unknown geometric placement, clusters)
  - Non-gaussian X-ray Noise (e.g., depends on X-ray tube)
- Solution: Modeling of target features (damages) and homogeneous materials + Overlay of real measured images (without defects, base-line)



RESULTS

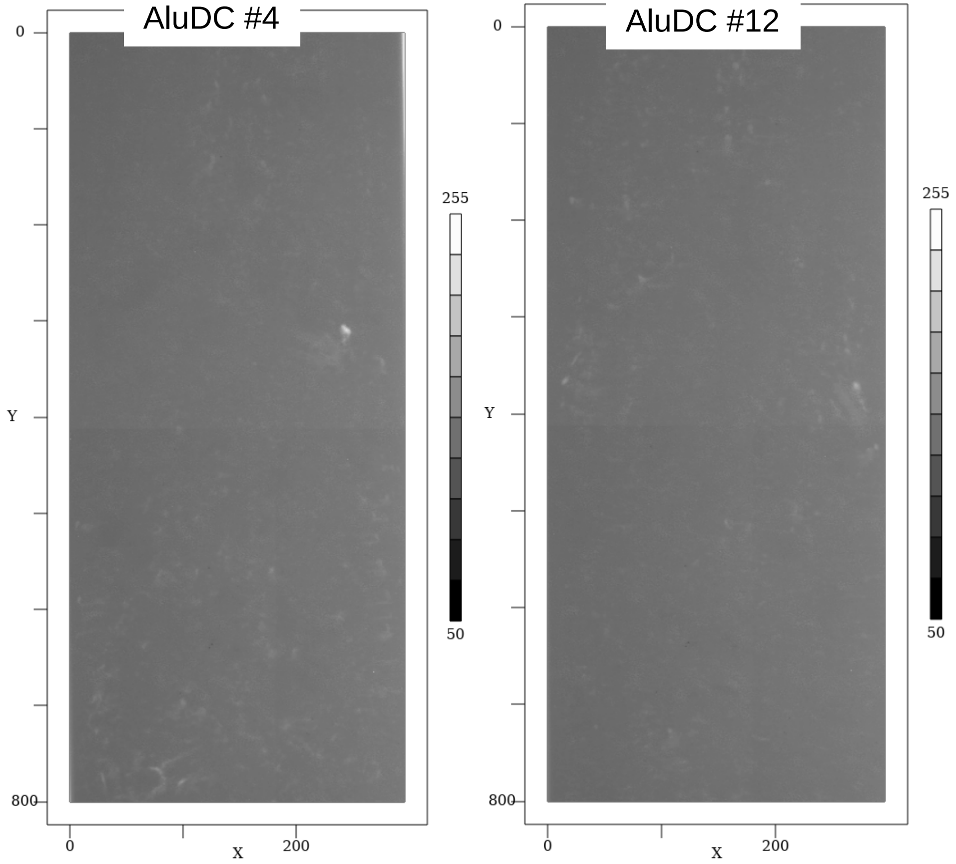
- Convolutional Neural Networks applied to X-ray radiography images of Aluminum die casting plates
- Anomaly detectors (CNN and LSTM-AE) applied to X-ray CT volume data

COMPARISON RECONSTRUCTED 3D CT (MID-Q) AND SINGLE PROJECTION X-RAY IMAGES (ALUDC)



Feature Marking: Second Challenge!

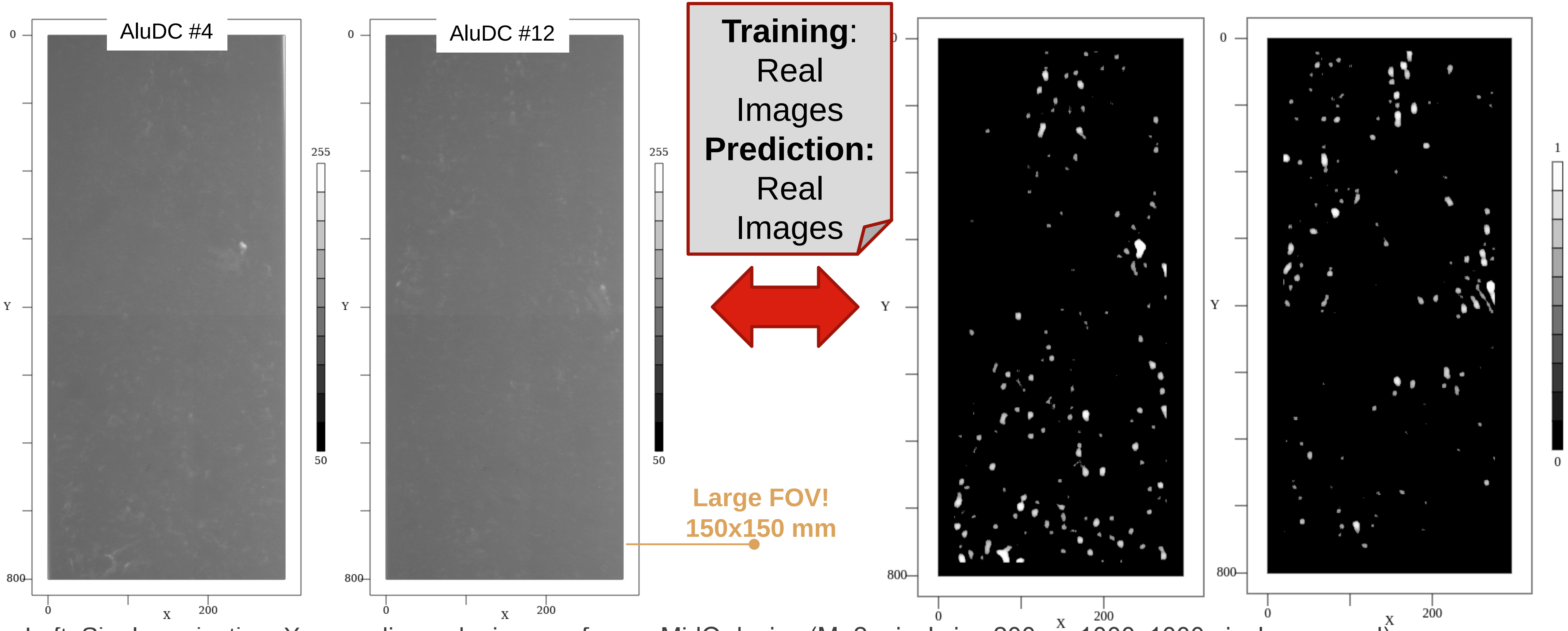
Large FOV!
150x150 mm



- Left: Volume projection of reconstructed CT images with data from a Mid-Q device (400/800 projections, rec. with classical fbp alg.)
- Right: Single projection X-ray radiography images from same Mid-Q device

Universität Bremen

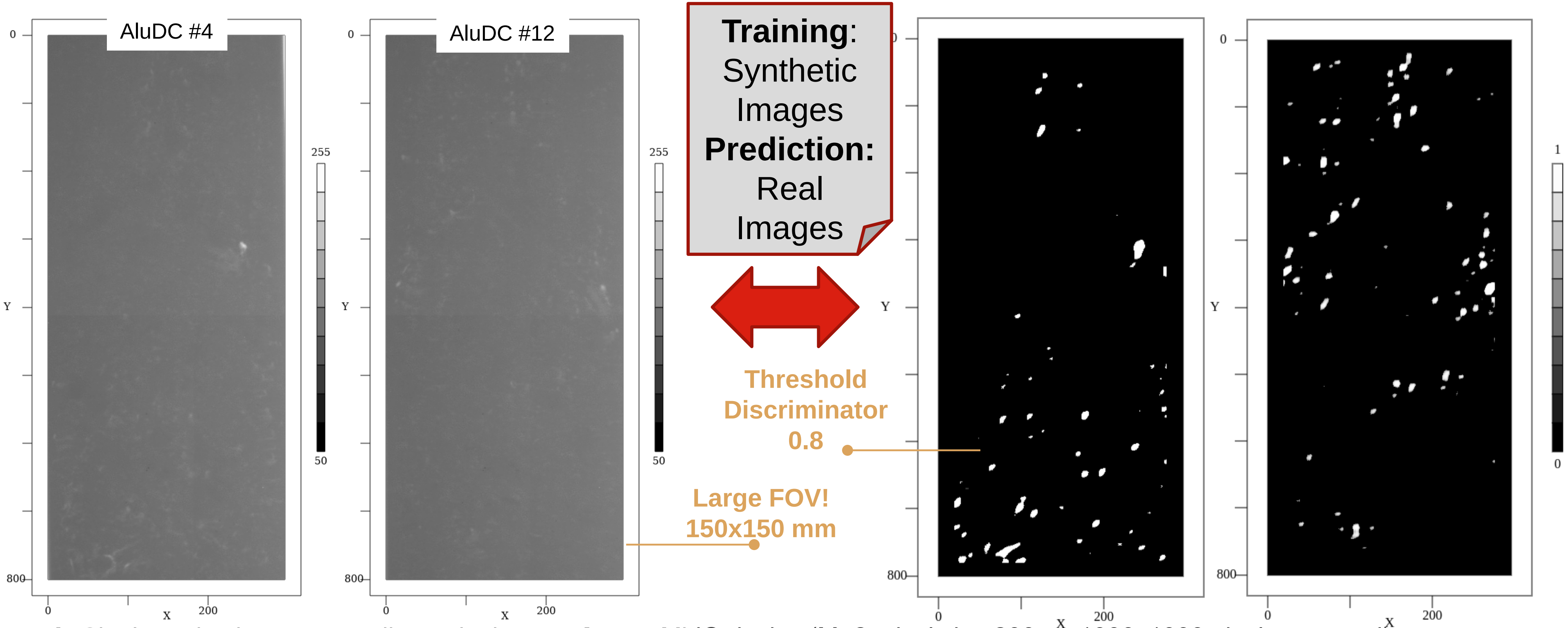
MID-Q RADIOGRAPHY AND CNN PORE FEATURE MARKING (R)



- Left: Single projection X-ray radiography images from a MidQ device (M=2, pixel size 200µm 1000x1000 pixels, cropped)
- Right: CNN Pixel Classifier Feature Marking predicted from single projection image (MidQ), trained with real images [8-4]

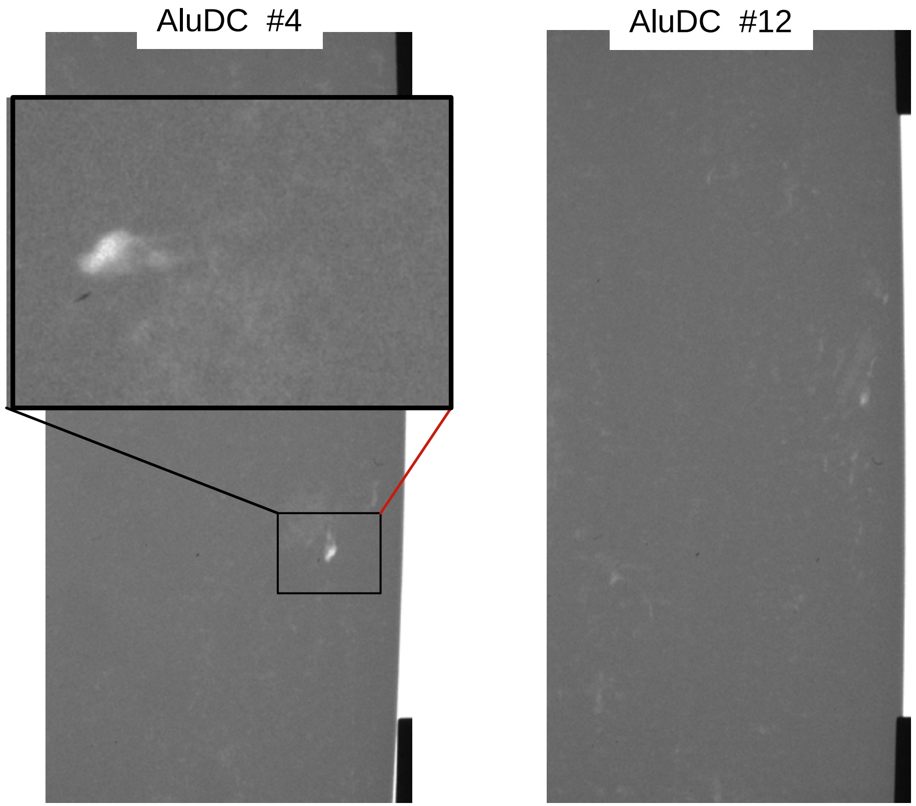
Universität Bremen

MID-Q RADIOGRAPHY AND CNN PORE FEATURE MARKING (S)



- Left: Single projection X-ray radiography images from a MidQ device (M=2, pixel size 200µm 1000x1000 pixels, cropped)
- Right: CNN Pixel Classifier Feature Marking predicted from single projection image (MidQ), trained with synthetic images [8-8-4]

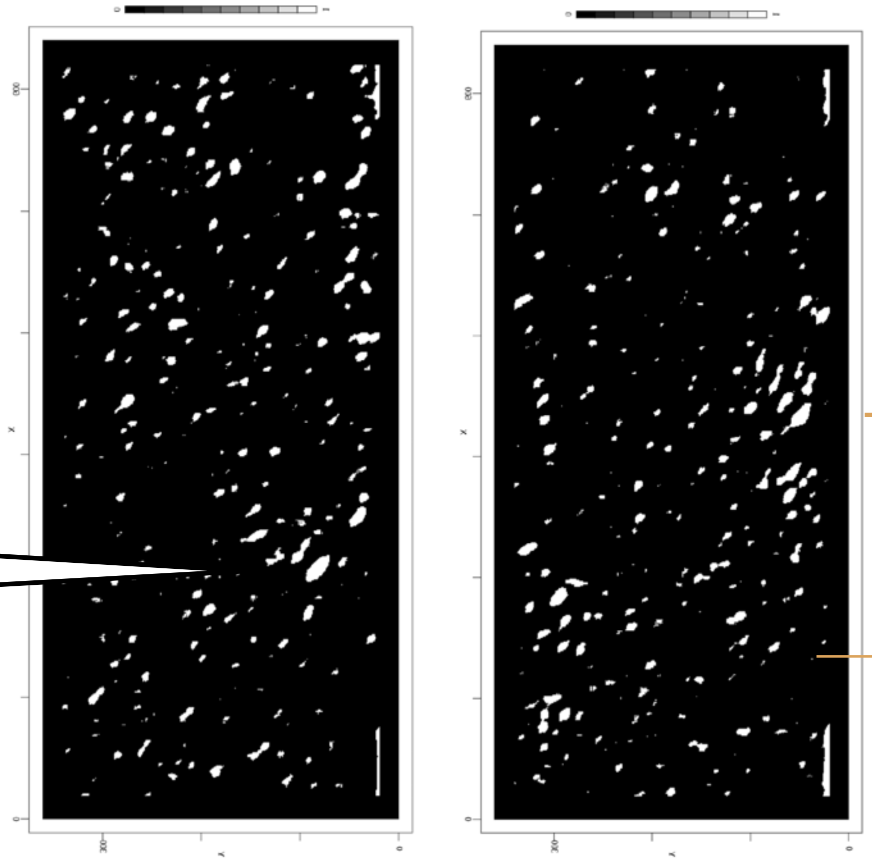
LOW-Q RADIOGRAPHY AND CNN PORE FEATURE MARKING (S)



Training: Synthetic Images **Prediction:** Real Images

Noise or Pores?

Smaller FOV! 70x40 mm



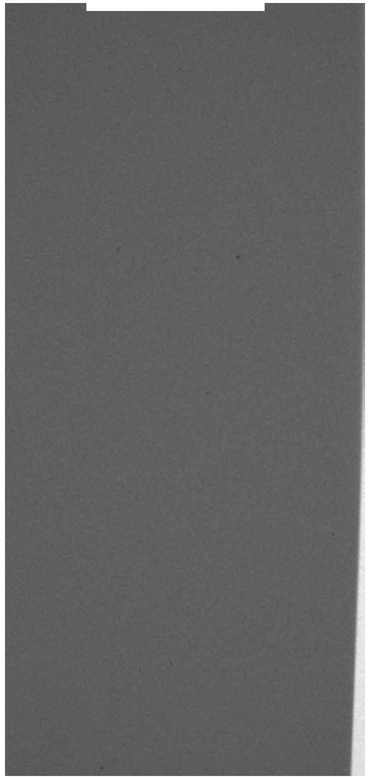
Threshold Discri-minator 0.95

- Left: Single projection X-ray radiography images from an Imaging LowQ device (M=1, eff. pixel size 40µm 1920x1080 pixels)
- Right: CNN Pixel Classifier Feature Marking predicted from single projection image (LowQ), trained with synthetic images [8-8-4]

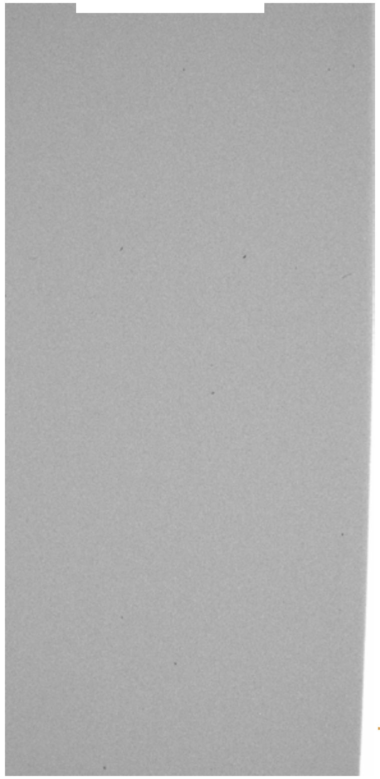
Universität Bremen

LOW-Q RADIOGRAPHY AND CNN PORE FEATURE MARKING (S)

AluRO #1



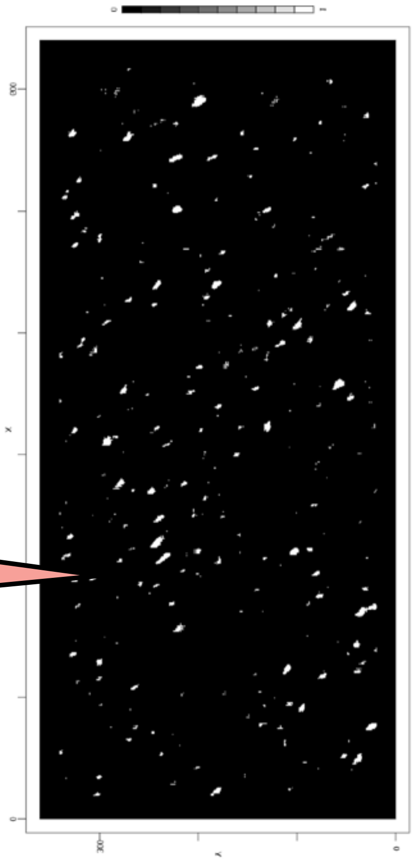
AluRO #1



Polished Alu Rolled Plate: No Pores & Defects!

Noise!

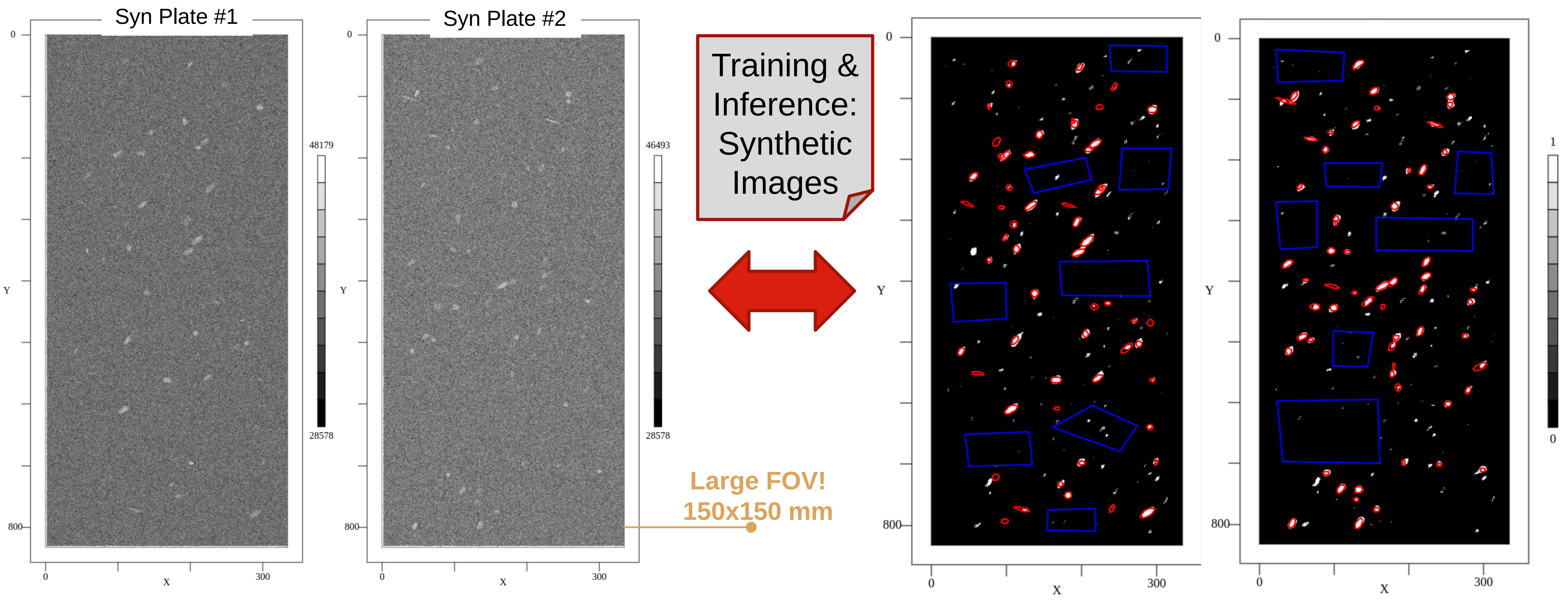
Smaller FOV! 70x40 mm



Threshold Discri-minator 0.95

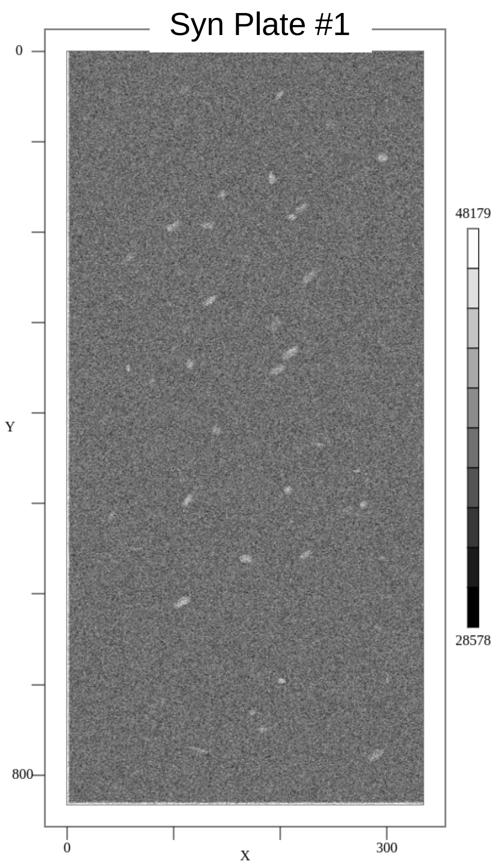
- Left: Single projection X-ray radiography images from an Imaging LowQ device // **Extruded aluminum plates** (d = 2 mm)
- Right: CNN Pixel Classifier Feature Marking predicted from single projection image (LowQ), trained with synthetic images [8-8-4]

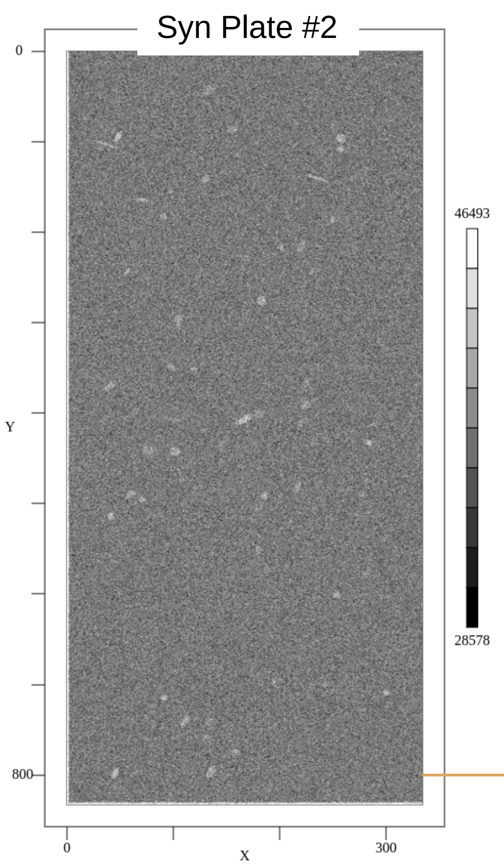
SIMULATED RADIOGRAPHY AND CNN PORE FEATURE MARKING (GROUND TRUTH)



- Left: Single projection X-ray radiography images from XraySim (M=2, pixel size 150µm 1000x1000 pixels, cropped) // **Synthetic Plate**
- Right: CNN Pixel Classifier Feature Marking predicted from single projection image [8-8-4]

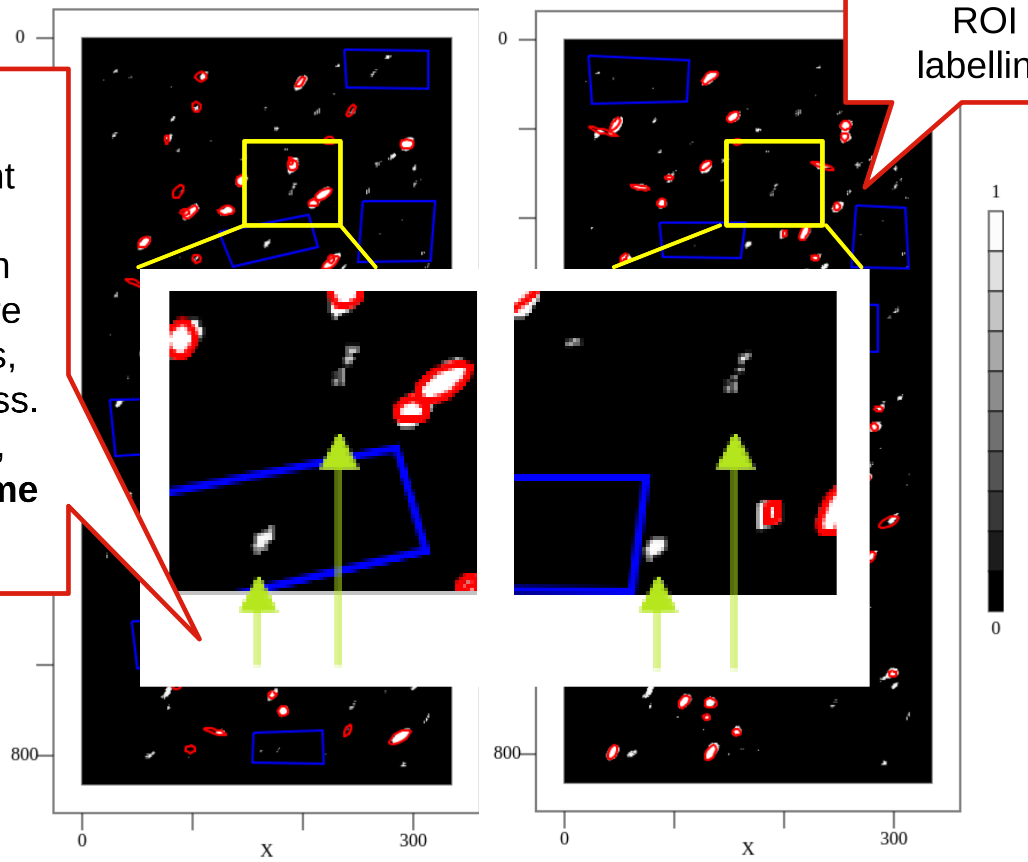
SIMULATED RADIOGRAPHY AND CNN PORE FEATURE MARKING (GROUND TRUTH)





Large FOV!
150x150 mm

Two independent plates and images with different pore distributions, different gauss. pixel noise, **produce same artifacts**!



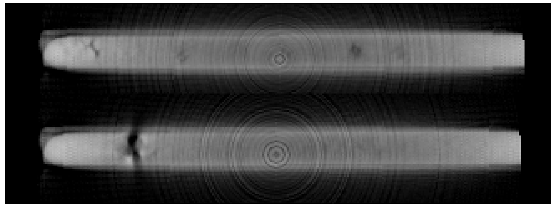
Trained w **Synthetic** Images and CAD model-based automated ROI labelling

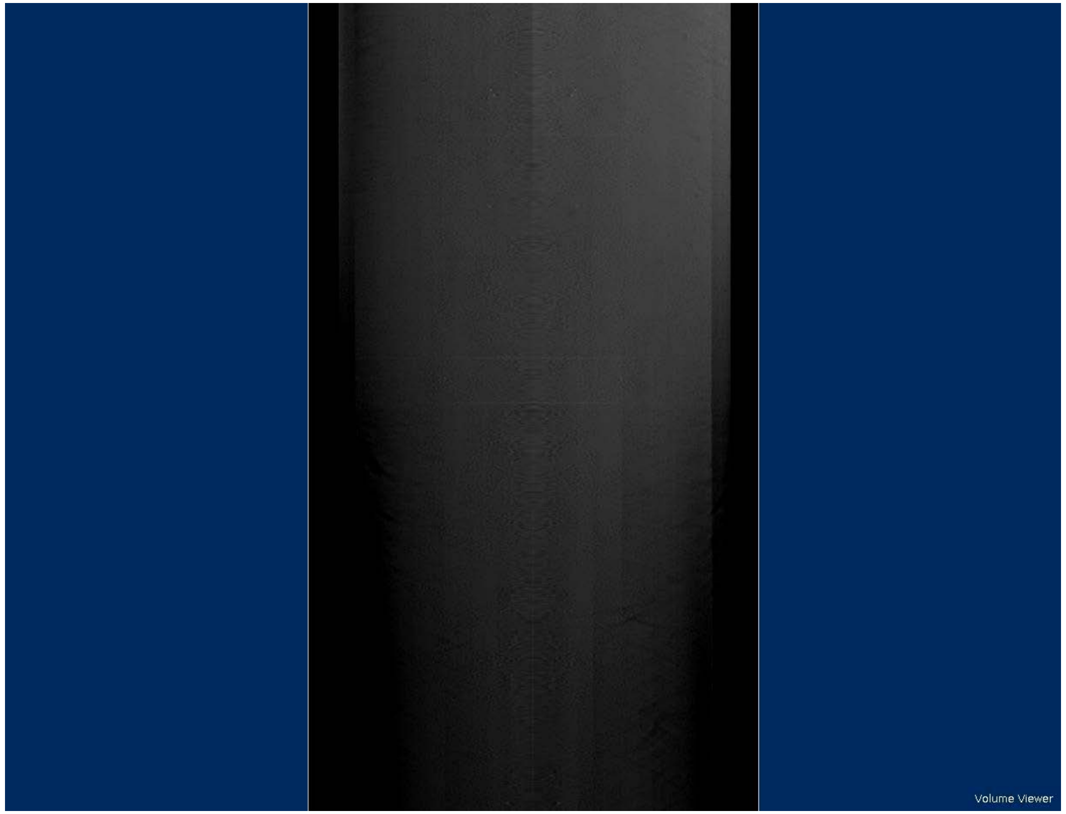
- Left: Single projection X-ray radiography images from XraySim (M=2, pixel size 150µm 1000x1000 pixels, cropped) // **Synthetic Plate**
- Right: CNN Pixel Classifier Feature Marking predicted from single projection image [8-8-4]

PORE INSPECTION AND CHARACTERISATION BY CT

It is a challenge to estimate pore shapes (geometry, size), density, spatial distribution, and to distinguish reconstructed pores from image artifcats and noise!

- Manual measuring of shape parameters of selected pores (e.g., using ImageJ analysis software) with ellipse approximation
- Automated pore analysis by point clustering methods and ellipsoid approximation

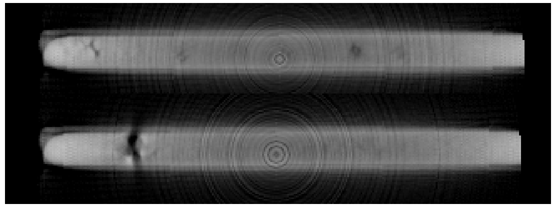


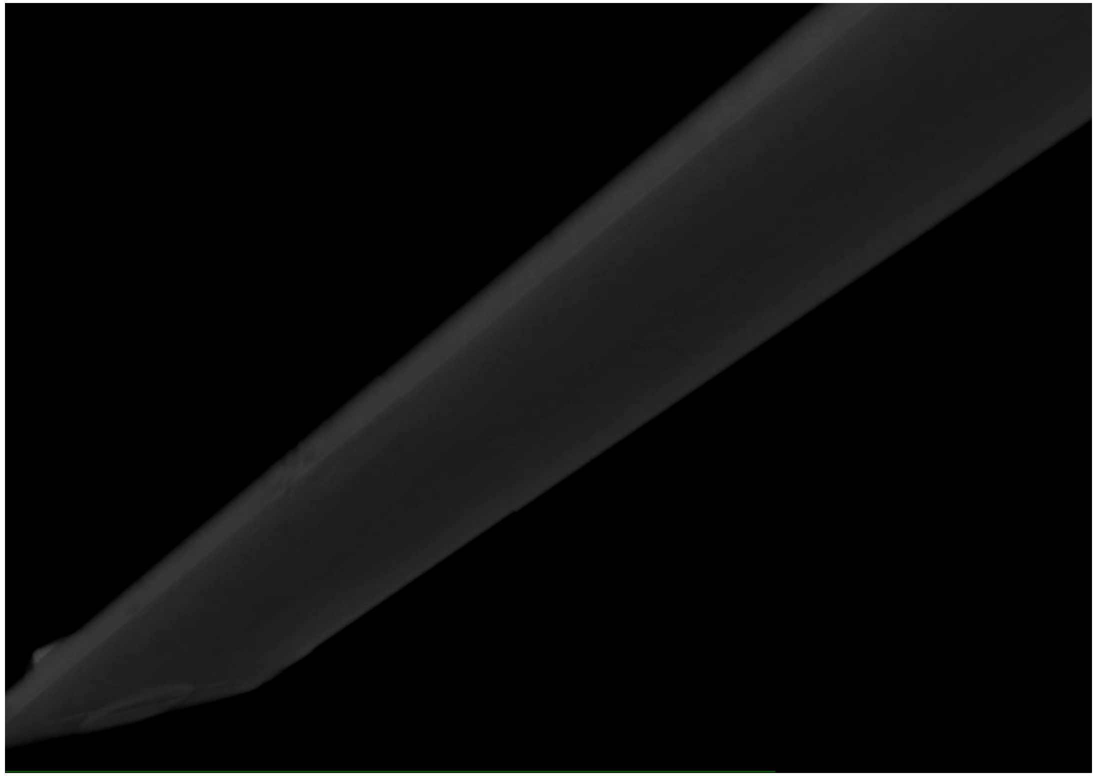


PORE INSPECTION AND CHARACTERISATION BY CT

It is a challenge to estimate pore shapes (geometry, size), density, spatial distribution, and to distinguish reconstructed pores from image artifcats and noise!

- Manual measuring of shape parameters of selected pores (e.g., using ImageJ analysis software) with ellipse approximation
- Automated pore analysis by point clustering methods and ellipsoid approximation

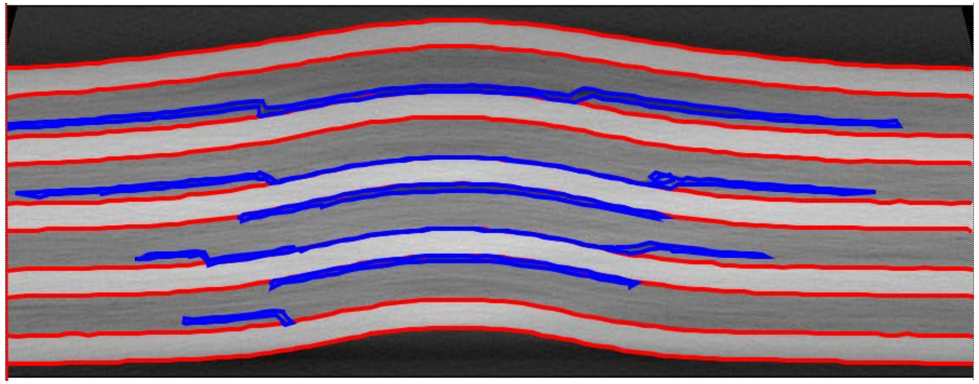


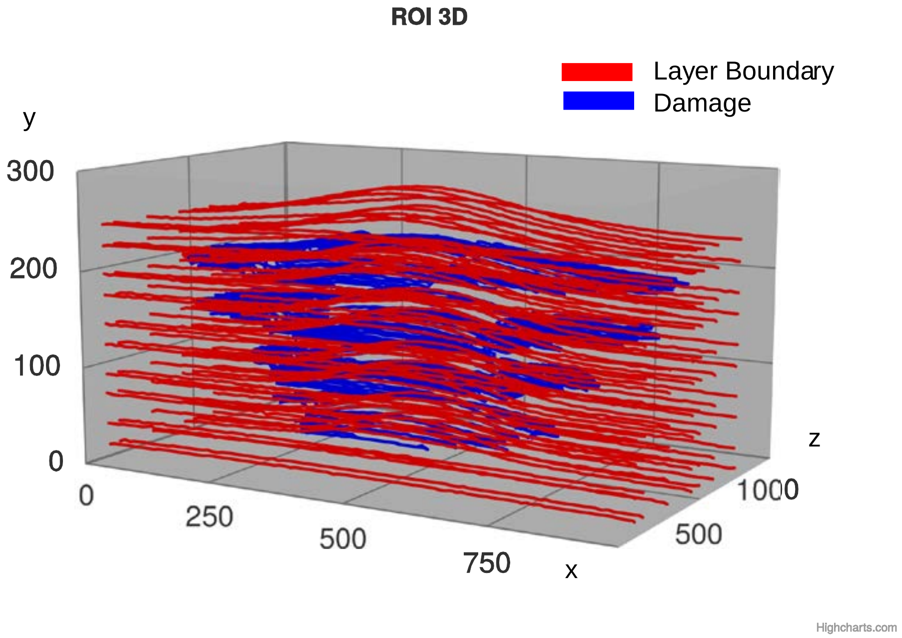


Universität Bremen

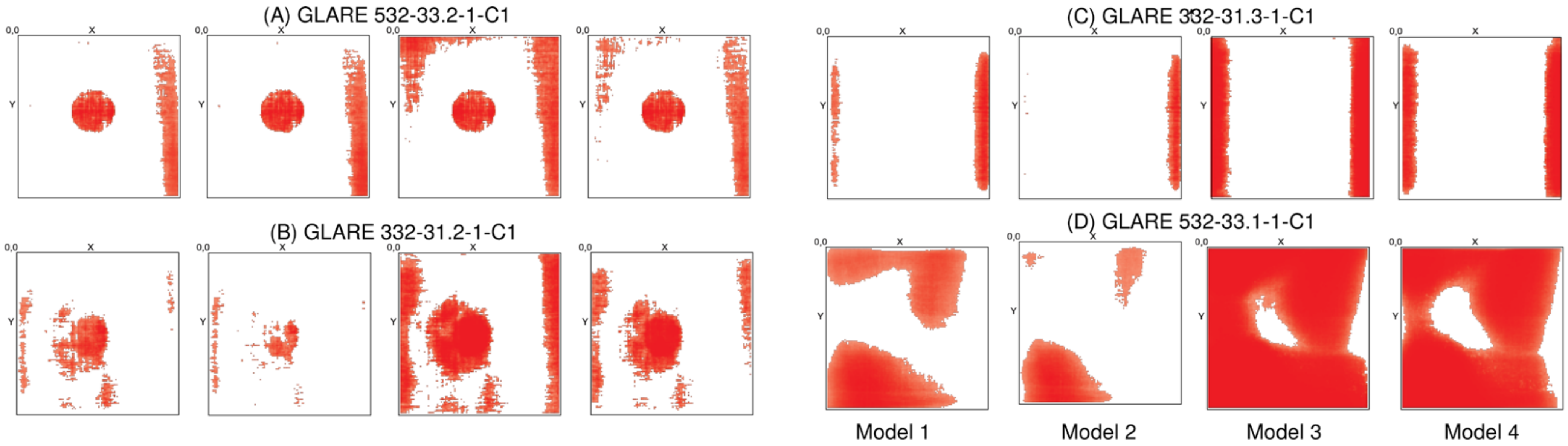
FML HOST MATERIAL AND DAMAGE MODELING

- Still work under progress
- Even the host material composite structure is a challenge
- Damages and Deformations must preserve material mass and volume!
- **ROI and composite layer boundary marking with semi-automated tracker (and Canny edge detection)**





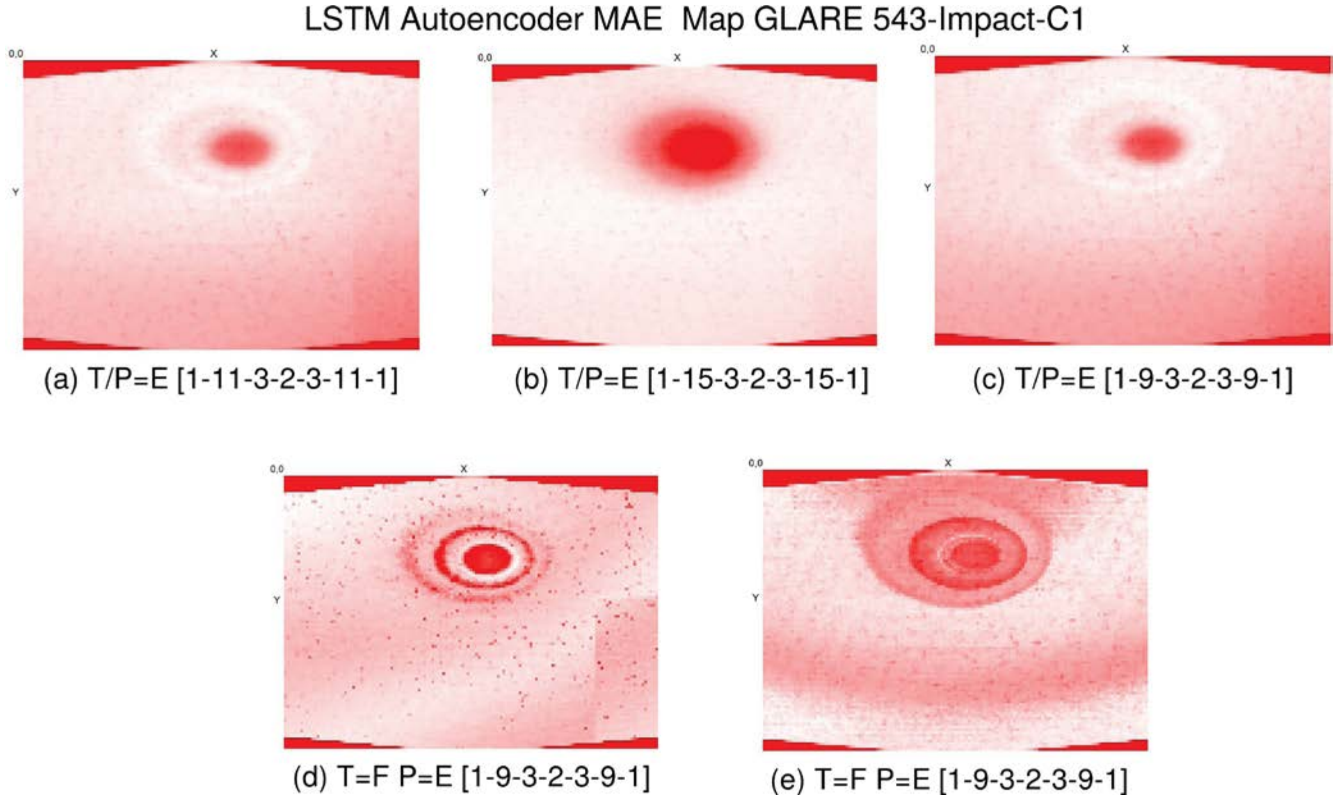
ANOMALY DETECTION IN FML CT DATA (POSITIVE TRAIN.)

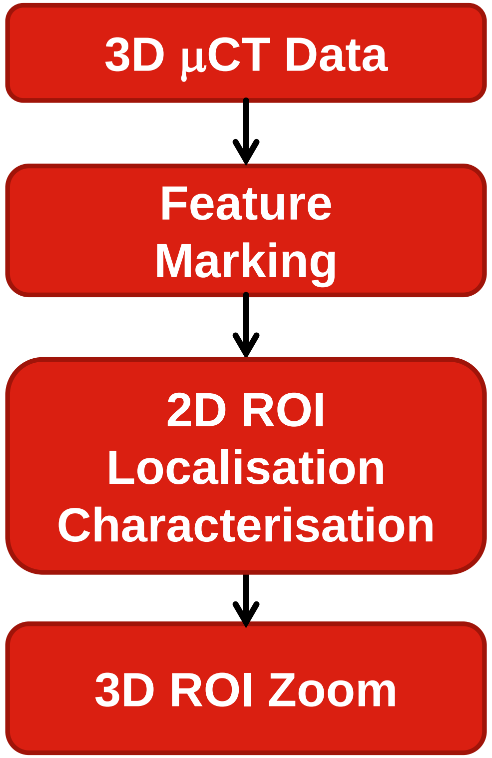


- A **CNN** is used to detect anomalies in a CT volume (feature marking of damage candidates) // Data from **HighQ** device
- Specimen: FML plate with different damages: A: foil pseudo defect,, B: Resin washout B, C: Baseline, D: Layer delamination:

1 Chirag Shah, Stefan Bosse, and Axel von Hehl. Taxonomy of Damage Patterns in Composite Materials, Measuring Signals, and Methods for Automated Damage Diagnostics, Materials 15 (MDPI), no. 13 (2022): 4645 ....

ANOMALY DETECTION IN FML CT DATA (NEGATIVE TRAIN.)





- A **LSTM Autoencoder** is used as an anomaly detector. Shown is the feature marking of the AE (top view of the X.ray CT volume)
- Specimen: FML plate with impact damage. A-E: Different AE model configurations and trainings // Data from **HighQ** device

CONCLUSIONS

Data

- **Single- and Multi-Proj. X-ray Images**
- **Data and feature variance is always limited**!
  - CT scans require high measuring time and produce big data volumes
  - Noise (LowQ)
- Supervised Learning: Hand-made **labelling is a challenge** and inaccurate
  - Relation between image and target features can be very low (contrast)
- CT data can not be used directly for **labelling** due to geometrical distortions (wrt. single projection input data)

Methods

- 3D CT reconstruction using **Filtered Back Projection** (sine wave filters)
- **Convolutional Neural Networks** for pore and damage feature marling (data-driven negative training) and **LSTM** anomaly detectors (positive training)
- **X-ray simulation** based on Beer-Lambert law and multi-material polygon mesh models
- **Monte Carlo simulation** of materials with defects and damages (openSCAD, **Constructive Solid Geometry**)
- **Measuring devices**: LowQ, MidQ, HighQ

Results

- **A pure data-driven feature marking model (semantic image pixel classifier) trained with synthetic images only can be applied to real images**
- **The semantic pixel feature marling model is capable to highlight low-contrast features (e.g., hidden pores)**
- **X-ray noise has significant impact on feature prediction results**
- **Accurate and representative training examples (labelling, simulation models) are a pre-requisite for robust data-driven models and a challenge!**

THANK YOU

Stefan Bosse

sbosse@uni-bremen.de

www.edu-9.de