# Convolutional Neural Networks for DESY Photon Science

Philipp Heuser, DESY-IT 12/11/2019 Artificial Intelligence applied to Photon and Neutron Science - Grenoble





# ML/DL @ DESY (campus) photon science

#### **CFEL / European XFEL / DESY**

Number of projects for classification of diffraction data; Hit/No hit? Good hit? Double hit? ....

#### **Petra III**

Xiaogang Yang, Thu 10:15; Deep learning for Synchrotron X-ray Imaging

#### **EMBL**

**Daniel Franke**, Tue 15:15; Machine learning applications for Small Angle X-ray Scattering data collection and analysis at EMBL-Hamburg

#### **Machine Control Group**

#### Helmholtz Zentrum Geesthacht (HZG) / DESY-IT

Segmentation of 3D synchrotron radiation micro-computed tomography (SRµCT)

#### CSSB / DESY-IT

ML/DL for cryo electron microscopy/tomography

#### **DESY-IT**

Monthly seminars/lectures on ML Round table events

### **Projects**

#### Helmholtz Zentrum Geesthacht (HZG) / DESY-IT Segmentation of X-ray tomography data





200 400 100







#### **CSSB / DESY-IT Object Detection for Cryo Electron Tomography**



# **Automated Volumetric Interpretation**

Philipp Heuser, DESY-IT, Scientific Computing Julian Moosmann, Helmholtz Zentrum Geesthacht, HZG



Semantic segmentation of bone implants (HZG)



Certain bone implants will be absorbed over time. Investigating the physiological processes over time by Xray tomography requires an accurate segmentation of a significant number of comparably large volumetric datasets.



#### Segmentation U-Net



Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *arXiv:1505.04597* (2015).

### Weak ground truth



# **Segmentation**

#### Trained on ~ 100,000 images



DESY.

# Training with weak ground truth



retrain final model

### **Automated Volumetric Interpretation**

Trained on ~15,942 images from 11 3D datasets Each image 5 times **augmented**, yielding 95,652 images for training Training for 15 epochs using central 600 pixels takes 7 days using one V100 GPU (Keras)

#### Inference on Test-Sets (not used for training), yields very high accuracy



### **Segments**



DESY.

## **Final Segmentation (Mg5GD)**

Segmentation Mg5Gd

Segmentation Mg10Gd



Open Postdoc Position in this project in Hamburg @ HZG (julian.moosmann@hzg.de)

# **Object Detection for Cryo Electron Tomography**

Finding Type III Secretion System (T3SS) in minicells for subtomogram averaging





Hu et al., 2017

### **Object detection**

Philipp Heuser, DESY-IT, Scientific Computing Thomas Marlovits, Sean Miletic, CSSB



#### Finding T3SS in EM tomography for subtomogram averaging





### **Object detection**

#### How to detect a few tiny objects



- get 50x50 px subsamples from images, with stride 1
- Assign lable to each patch
- $\rightarrow$  Image Classification task



# LeNet-5 (1998) for classification

Yann LeCun, et al.

- pioneering 7-level convolutional network by LeCun et al in 1998,
- classifies digits, was applied by several banks to recognise hand-written numbers on checks digitized in 32x32 pixel greyscale images.



Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86**, 2278–2324 (1998).

# Challenge

#### Few true answers vs. lots of false answers





# **Segmentation**

#### **U-Net**



Manually corrected 12 particularly bad predictions



re-trained with 30 2D images



Applied to 4 3D volumes; corrected and added all falsely predicted and some randomly selected images to training data (509 2D images)



### **Final trained U-Net**





### **T3SS found!**

#### Searching only within the membrane



#### **T3SS candidates**



### **TP/FP**

#### **Ground truth?**

Currently we have about the same number of FP as TP.

FP TP



### **Conclusions/Outlook**

- Good results even with 'classic' networks
- Quality of ground truth crucial
- Even though, prediction by CNN can be better than training annotation
- Screws:
  - Generation of more training data from different materials and from other synchrotron
    → Final training
- Cryo ET for T3SS:
  - Better ground truth, new detector, full cells
  - 3D CNNs

### Acknowledgments Thanks to...



DESY-IT Frank Schlünzen



Thomas Marlovits Sean Miletic Wolfgang Lugmayr

#### Helmholtz-Zentrum Geesthacht

Zentrum für Material- und Küstenforschung

Julian Moosmann Diana Krüger

**HELMHOLTZ** Analytics Framework

Funding: Helmholtz Association Initiative and Networking Fund project number ZT-I-0003