Image classification from the Single Particle Imaging (SPI) experiments with a fast object detection system based on a Convolutional Neural Network (CNN)

Alexandr Ignatenko

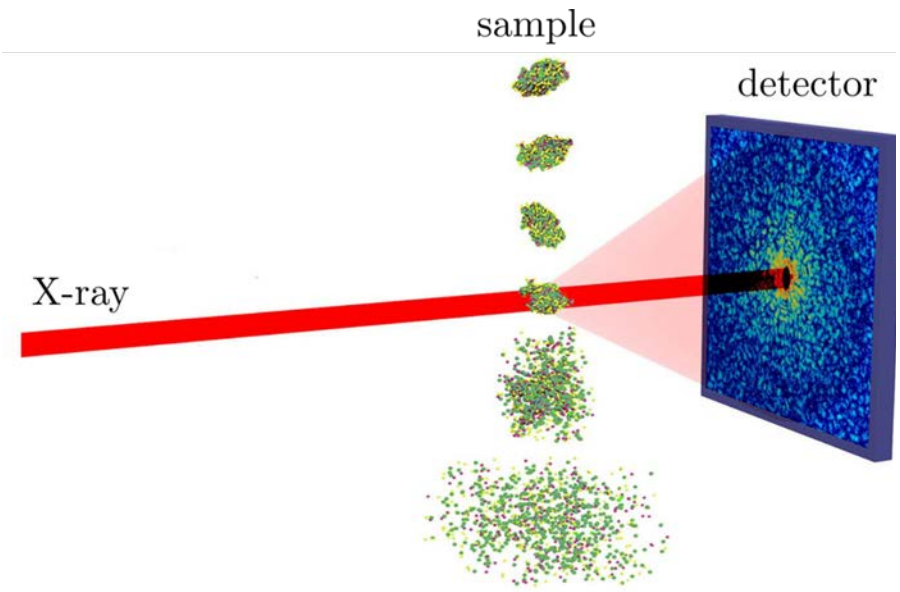
AMALEA Meeting
Hamburg, 02.12.2019

Overview

- Introduction
  - SPI experiments
  - YOLO
  - Metrics
- Training, validation details
- Results
- Summary & outlook

Introduction

SPI experiments



S. Bobkov et al., Journal of Synchrotron Radiation 22 (2015)

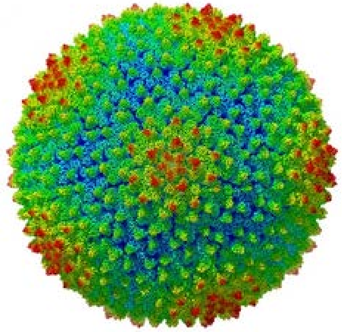
Large amount of data is collected: need for automated processing pipeline

Data pre-processing → Filtering → Classification → Object structure reconstruction

Classification task: select images of single particle under investigation for further analysis

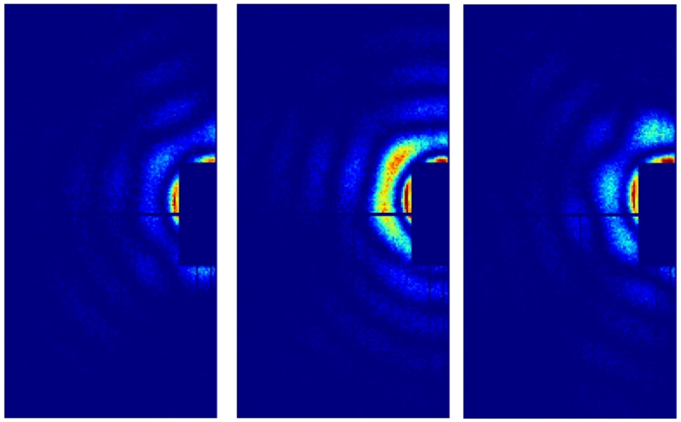
SPI experiment @ LCLS in 2018

- AMO beamline at LCLS
- Sample – bacteriophage PR772, expected size 60-75 nm
- E = 1.7 keV (λ = 7.29 Å)
- Detector distance = 125 mm
- Detector – pnCCD, a half of it was operational

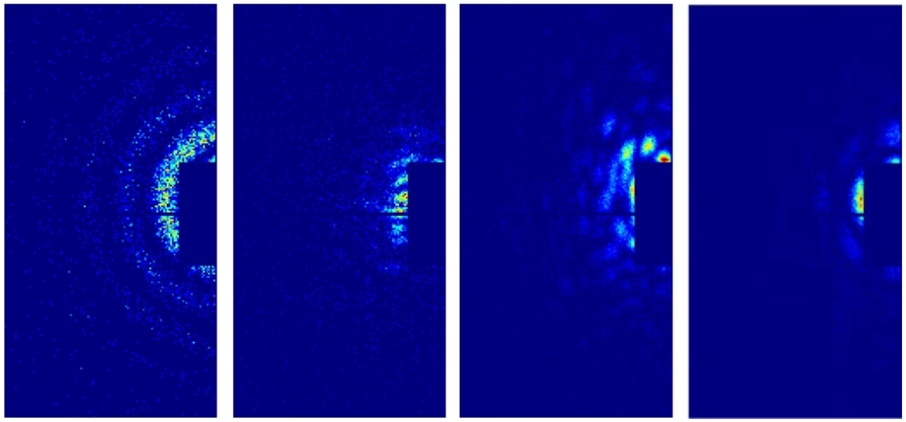


SPI experiment @ LCLS in 2018

Positive examples:



Negative examples:



Subtracted background, photon counts

YOLO

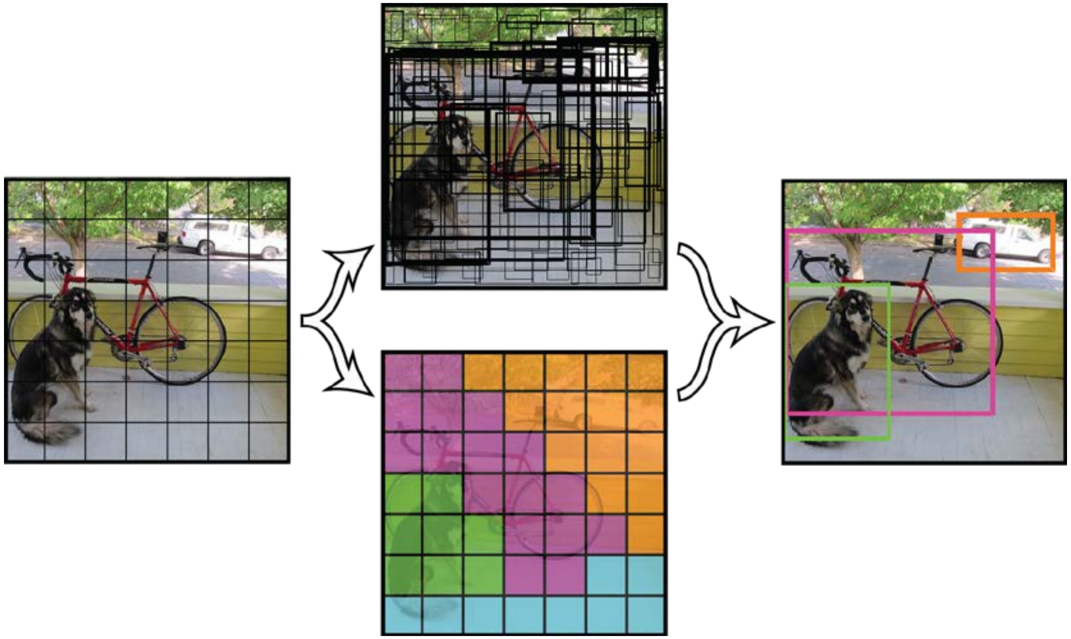
Different approaches in object detection:

- Region proposal networks (RPN), e.g. Region-based CNN (R-CNN) series

  Generating regions of interest (RoI) proposals → detecting object for each proposal

- Single shot detectors, SSD and YOLO (You Only Look Once)

  Object detection in a single pass



https://pjreddie.com/darknet/yolov2

Darknet YOLOv2

Layer	Filters	Size	Input	Output
Conv	32	3 x 3 / 1	416 x 416 x 3	416 x 416 x 32
Max		2 x 2 / 2	416 x 416 x 32	208 x 208 x 32
Conv	64	3 x 3 / 1	208 x 208 x 32	208 x 208 x 64
Max		2 x 2 / 2	208 x 208 x 64	104 x 104 x 64
Conv	128	3 x 3 / 1	104 x 104 x 64	104 x 104 x 128
Conv	64	1 x 1 / 1	104 x 104 x 128	104 x 104 x 64
Conv	128	3 x 3 / 1	104 x 104 x 64	104 x 104 x 128
Max		2 x 2 / 2	104 x 104 x 128	52 x 52 x 128
Conv	256	3 x 3 / 1	52 x 52 x 128	52 x 52 x 256
Conv	128	1 x 1 / 1	52 x 52 x 256	52 x 52 x 128
Conv	256	3 x 3 / 1	52 x 52 x 128	52 x 52 x 256
Max		2 x 2 / 2	52 x 52 x 256	26 x 26 x 256
Conv	512	3 x 3 / 1	26 x 26 x 256	26 x 26 x 512
Conv	256	1 x 1 / 1	26 x 26 x 512	26 x 26 x 256
Conv	512	3 x 3 / 1	26 x 26 x 256	26 x 26 x 512
Conv	256	1 x 1 / 1	26 x 26 x 512	26 x 26 x 256
Conv	512	3 x 3 / 1	26 x 26 x 256	26 x 26 x 512
Max		2 x 2 / 2	26 x 26 x 512	13 x 13 x 512
Conv	1024	3 x 3 / 1	13 x 13 x 512	13 x 13 x1024
Conv	512	1 x 1 / 1	13 x 13 x1024	13 x 13 x 512
Conv	1024	3 x 3 / 1	13 x 13 x 512	13 x 13 x1024
Conv	512	1 x 1 / 1	13 x 13 x1024	13 x 13 x 512
Conv	1024	3 x 3 / 1	13 x 13 x 512	13 x 13 x 1024
Conv	1024	3 x 3 / 1	13 x 13 x1024	13 x 13 x 1024
Conv	1024	3 x 3 / 1	13 x 13 x1024	13 x 13 x 1024
Route	16			
Reorg		/ 2	26 x 26 x 512	13 x 13 x2048
Route	26 24			
Conv	1024	3 x 3 / 1	13 x 13 x3072	13 x 13 x 1024
Conv	30	1 x 1 / 1	13 x 13 x1024	13 x 13 x 30

Detection

Feature learning

Classification

- Transfer learning - feature extractor pre-trained on ImageNet

- Main disadvantage – worse performance for smaller objects - is not very relevant:
  - ROI is almost the same → it is easy to adjust zoom

Total loss = Confidence loss + Localization loss + Classification loss

Metrics

$$\text{Recall} = \frac{TP}{TP + FN}$$

$$\text{Precision} = \frac{TP}{TP + FP}$$

TP – True positives

FP – False positives

FN – False negatives

False/True – with respect to the ground truth

$$F_\beta = (1 + \beta^2) \cdot \frac{\text{precision} \cdot \text{recall}}{(\beta^2 \cdot \text{precision}) + \text{recall}}$$

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

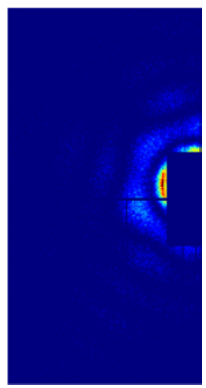
Training / validation details

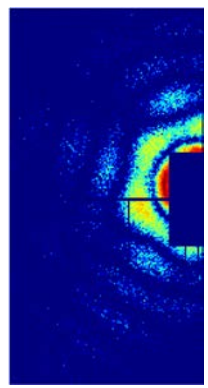
3 approaches

3 approaches:

1. Carefully select patterns for training across all runs, use linear scale for photon intensities
2. Select patterns for training from the first runs only, use linear scale for photon intensities
3. Select patterns for training from the first runs only, use log scale for photon intensities

Linear scale





Log scale

Questions to answer:

1. Can a reasonable selection be made?
2. Can we in general use the training data from the first runs only?
3. What scale is preferred, linear or log?

3 approaches

1 Representative(?) set of training examples, linear input

Training set

Positive set

- 165 patterns in the size range 60 -75 nm

Negative set

- 373 carefully selected patterns across all runs

Validation set

Positive set

- 53 patterns in the size range 60 -75 nm

Negative set

- 200 carefully selected patterns across all runs

3 approaches

2&3 Set of examples from the first runs, linear & log scale to the input&

Training set

Positive set

- 150 patterns in the size range 59 - 80 nm

Negative set

- 150 patterns in the size range 59 - 80 nm

Validation set

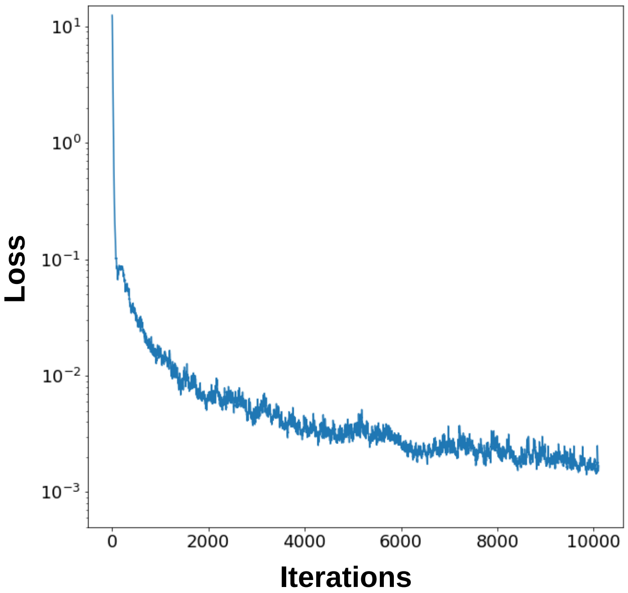
Positive set

- 100 patterns in the size range 59 - 80 nm

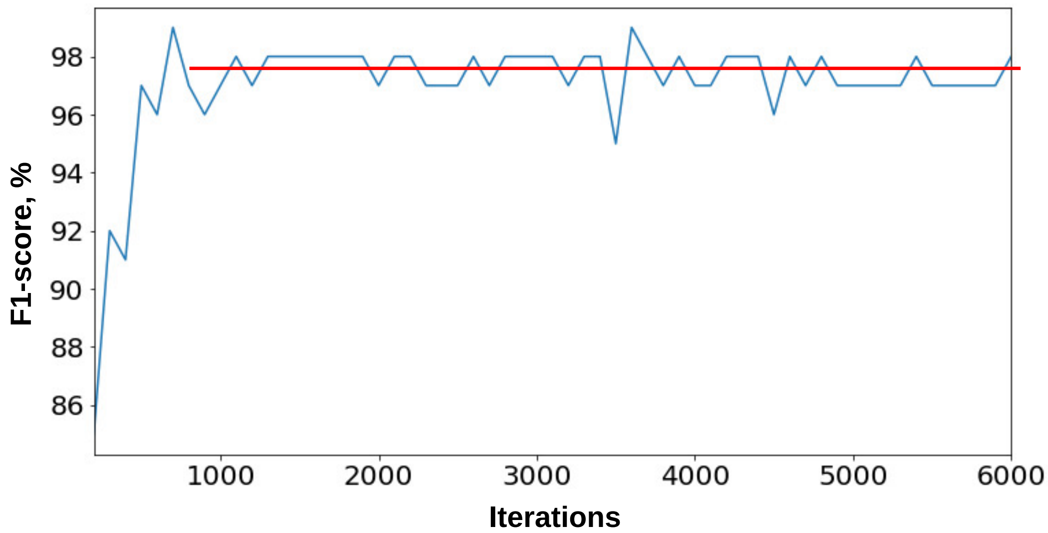
Negative set

- 100 patterns in the size range 59 - 80 nm

Training and model choice



Typical training loss

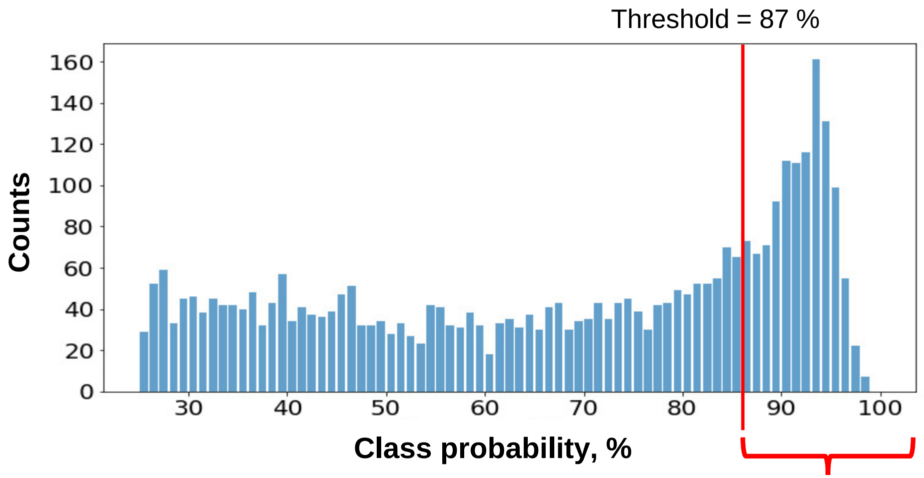


Apply model with the saturated F_1

Results

Representative negative set, linear input

Model run on ~18k patterns filtered by particle size

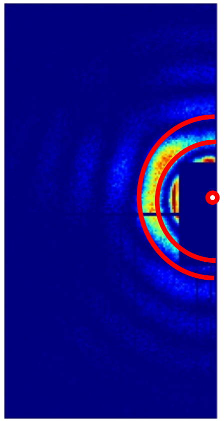


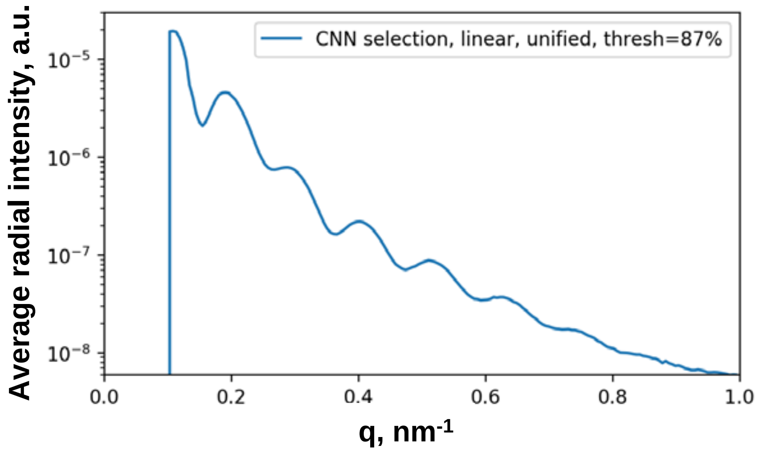
1036 patterns selected

For each pattern:

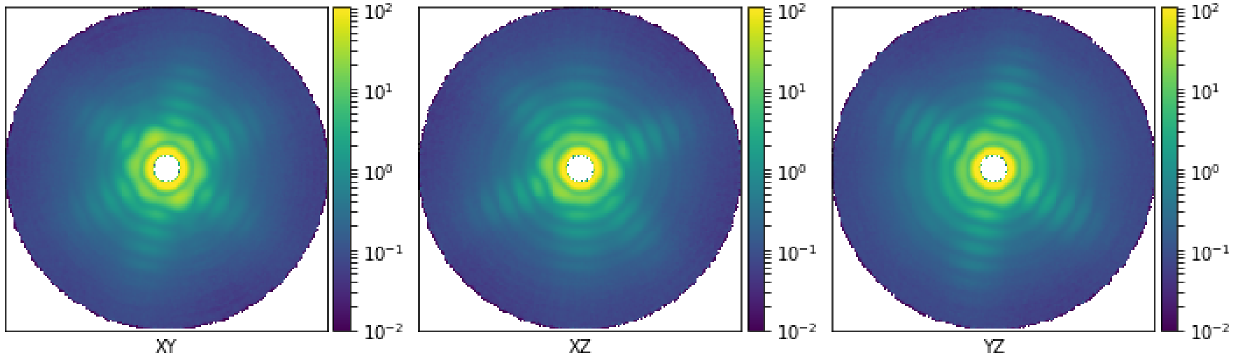
- calculate and average radial intensity

- translate distance to the beam center in pixels to distance in reciprocal space (q)





Representative negative set, linear input



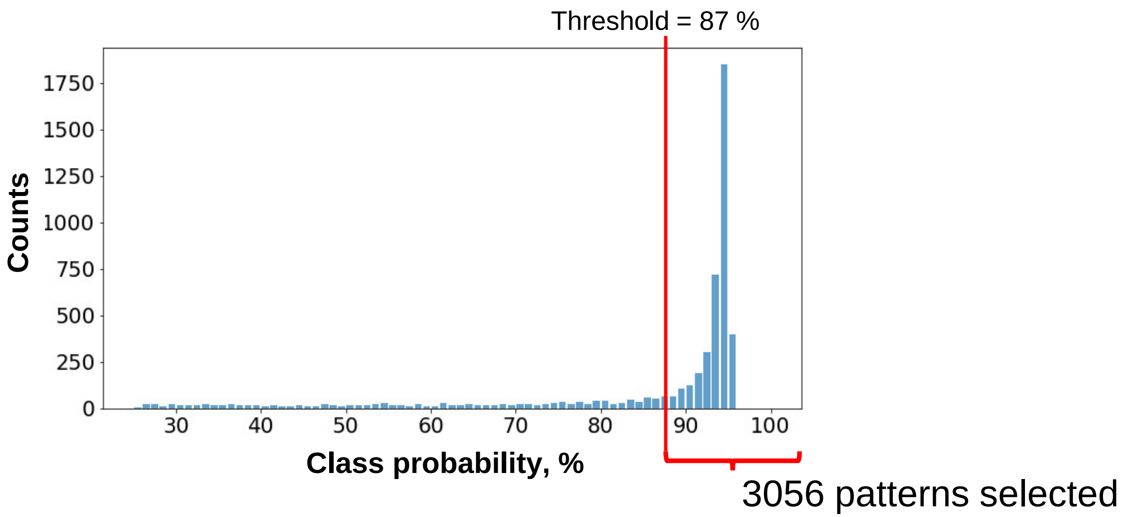
EMC reconstruction from 1036 patterns*

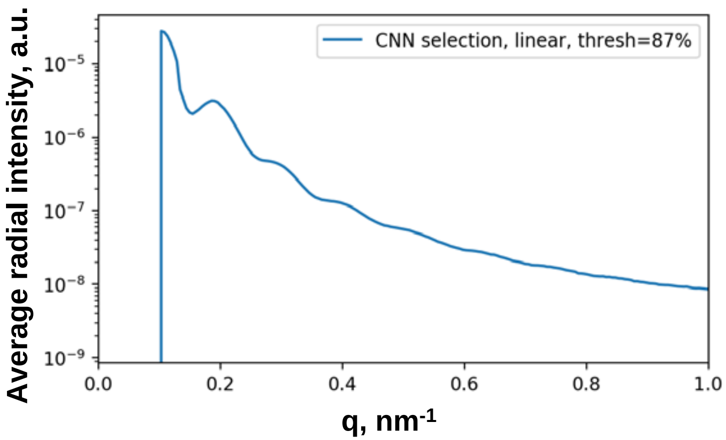
* here and thereafter – by S. Bobkov

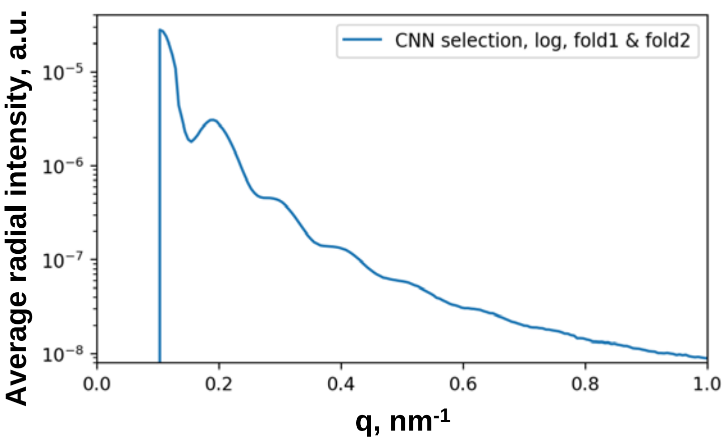
Set from the first runs

Linear & log input

Model run on ~18k patterns filtered by particle size

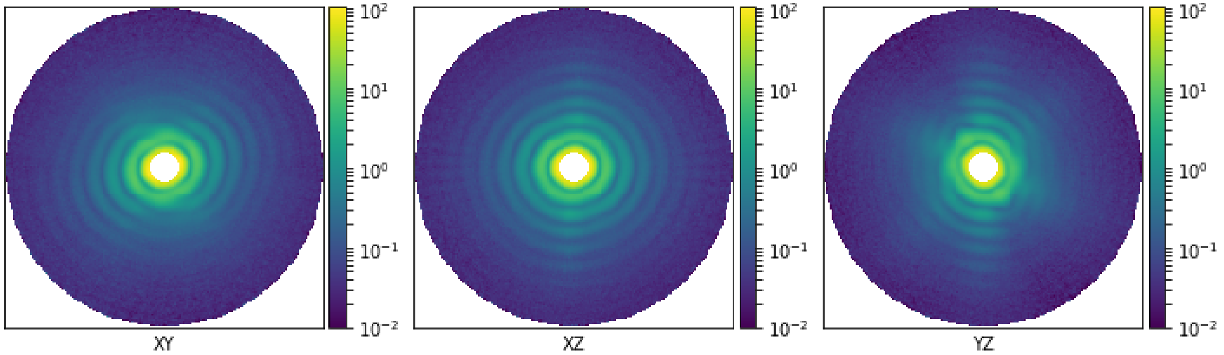






Set from the first runs

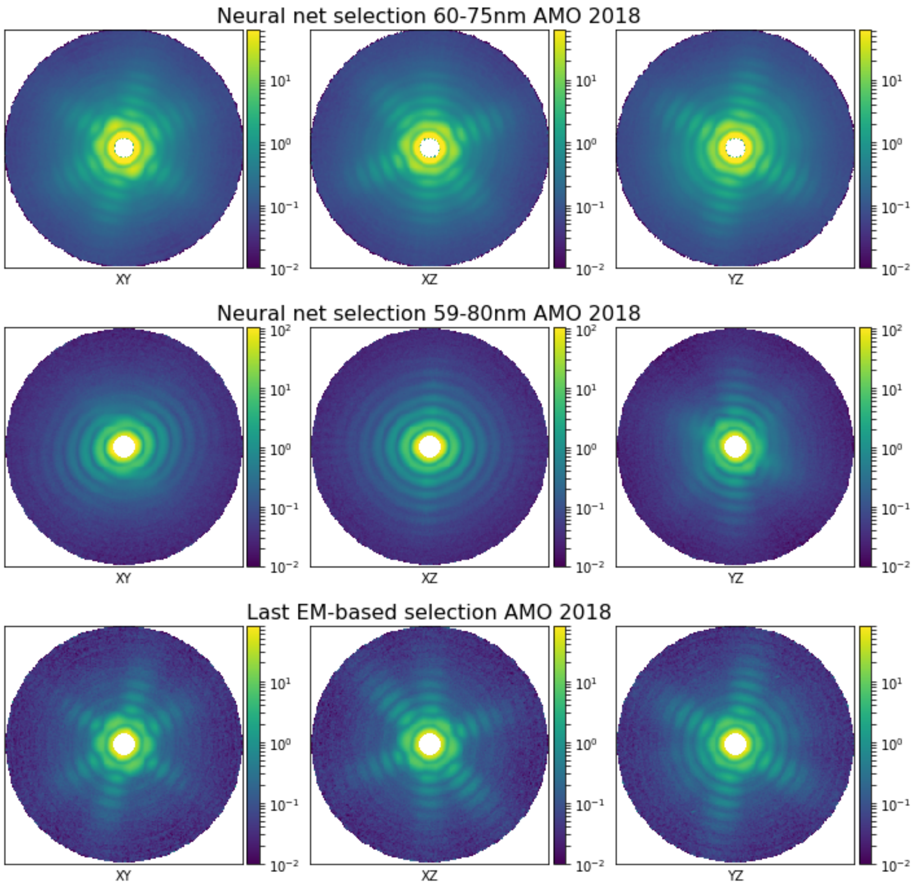
Log input



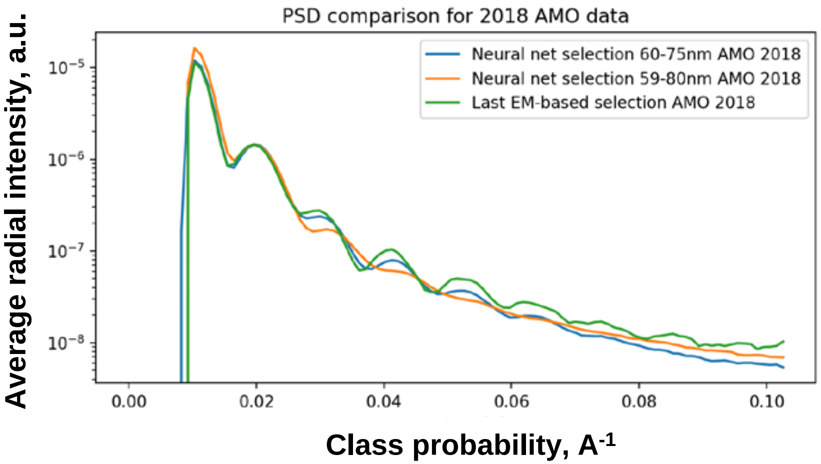
EMC reconstruction from 3056 patterns*

EMC reconstruction, log input*

Comparison



EMC reconstruction*



Average radial intensity for reconstruction (3D) *

Comparison

	# of selected patterns	Intersection with S. Bobkov's selection	Intersection with Haoyuan's selection
Representative negative set, linear input	1036	480	622
Set from first runs, linear input	3821	264	456
Set from first runs, log input	3056	283	513

1. Can a reasonable selection be made? Yes
2. Can we in general use the training data from the first runs only? Not a good idea
3. What scale is preferred, linear or log? Probably log, but it matters only if we are careful with the training set

Summary & Outlook

Summary & outlook

- CNN-based selection of single hits is promising – it was possible to make reasonable selection
- There is an indication that having representative (negative) examples can be vital and more important than the choice of either linear or log input
- Simulated data can give confidence in training / validation data (positive examples + partially negative examples)
- Customization of the CNN will be done (TensorFlow implementation, residual network, 1D input)
- Use of unsupervised learning

Acknowledgements

D. Assalauova (DESY)

L. Gelisio (DESY)

I. A. Vartanyants (DESY)

S. A. Bobkov (National research Center “Kurchatov Institute”)

SPI consortium @ LCLS

A. Aquila (LCLS)

Helmholtz Associations Initiative and Networking Fund and the Russian Science Foundation (ProjectNo.18-41-06001)

Thank you for your attention!