# Microcrystal segmentation for SSX

# Fixed Target Serial Synchrotron Crystallography (FT- SSX)

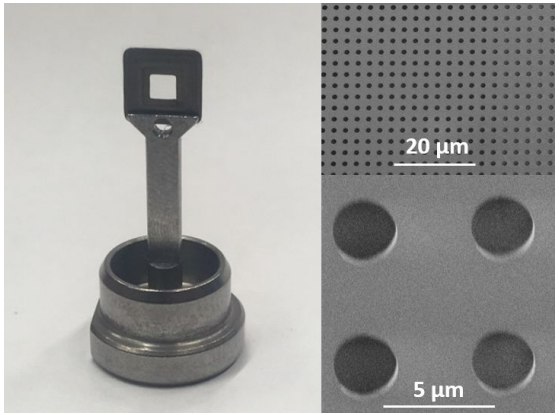
- **FT-SSX**:

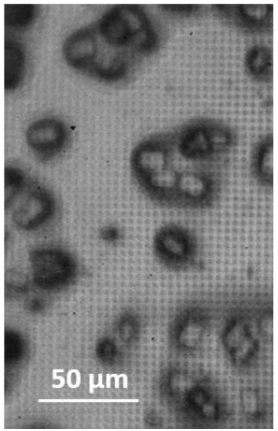
  FT: crystals are mounted in a sample holder and maintain their position relative to the holder during all the experiment.

  SSX: collection of diffraction data from multiple microcrystals contained in one or more experimental supports until obtaining a complete dataset.

- **Emerging technique** in X-ray Free Electron Lasers and Synchrotrons (specially at micro-focus beamlines) facilities:

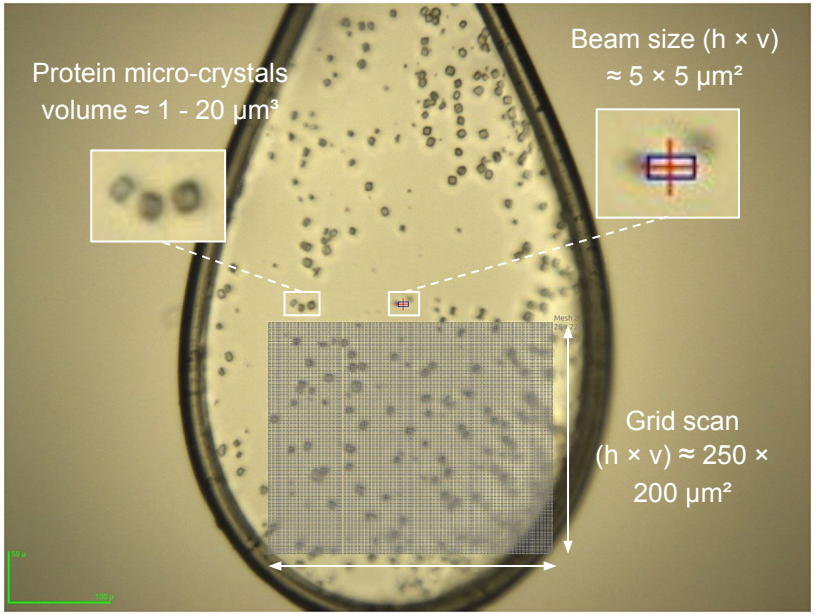
  Regarding FT-SSX experiments in ALBA: new micro-focus BL06-XAIRA beamline (beam size at sample position = 1-2 × 1-2 $\mu m^2$)

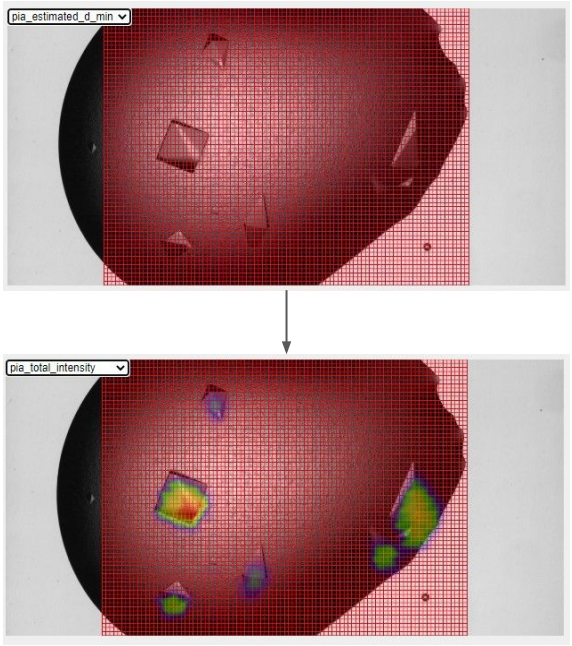



*https://www.psi.ch/en/lnb/supports-for-serial-protein-crystallography*

# Current beam-crystal alignment




Specially when dealing with micro-crystals: crystals are severely damaged by the X-ray beam exposure (**radiation damage**)

*https://www.embl.org/groups/macromolecular-crystallography/user-information/crystal-mounts/*
*https://www.diamond.ac.uk/Instruments/Mx/VMXi/VMXi-Manual/Using-ISPyB0/Data-Management/Grid-Scans.html*

# Alternative solution: crystal segmentation with a deep learning approach

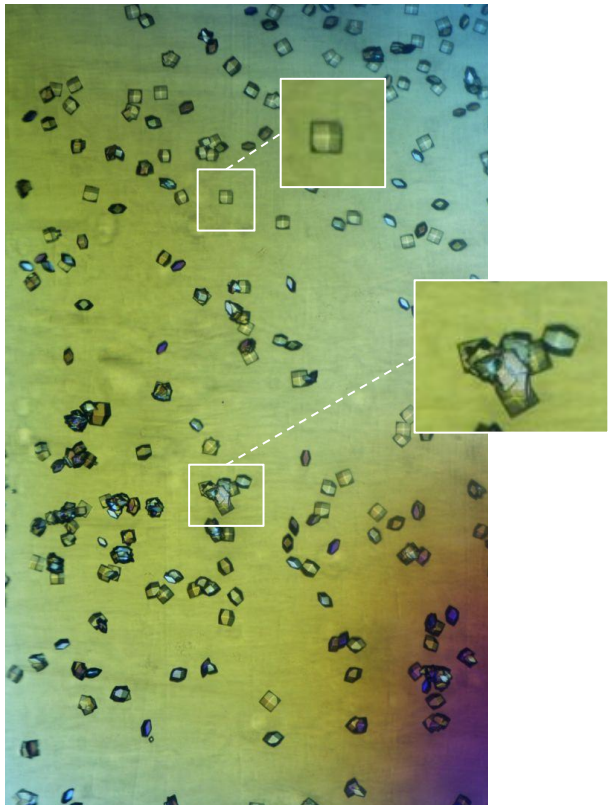
**Strengths**:

- Radiation damage-free methods
- Faster than grid scans

but, **problems**:

- Absence of
  - structured
  - curated
  - open
  - experimental (real)

  data of segmented crystals
- and, inherent complexity to segment own/third datasets:
  - needs expertise
  - time-consuming
  - different escenarios

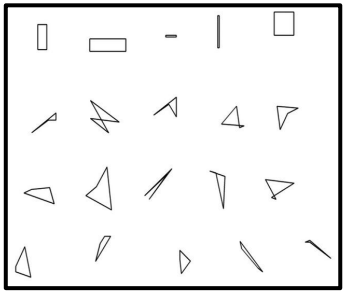


# Current DL crystal detection/segmentation approaches in the literature

Some available open data (and models):

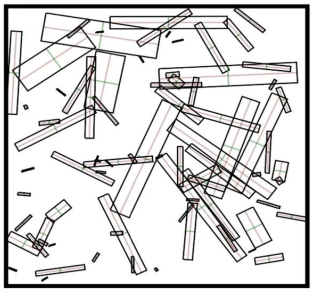
- based on synthetic data:

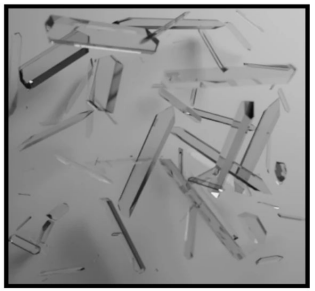
*DeepCentering: Ito et al., (2019)*



*Model trained on basic geometrical figures*

*Bischoff et al, (2022)*

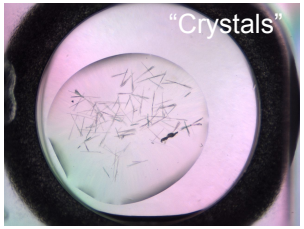




*Renderized hyper-realistic crystals with automatic extraction of bounding boxes*

- based on experimental data:

MARCO

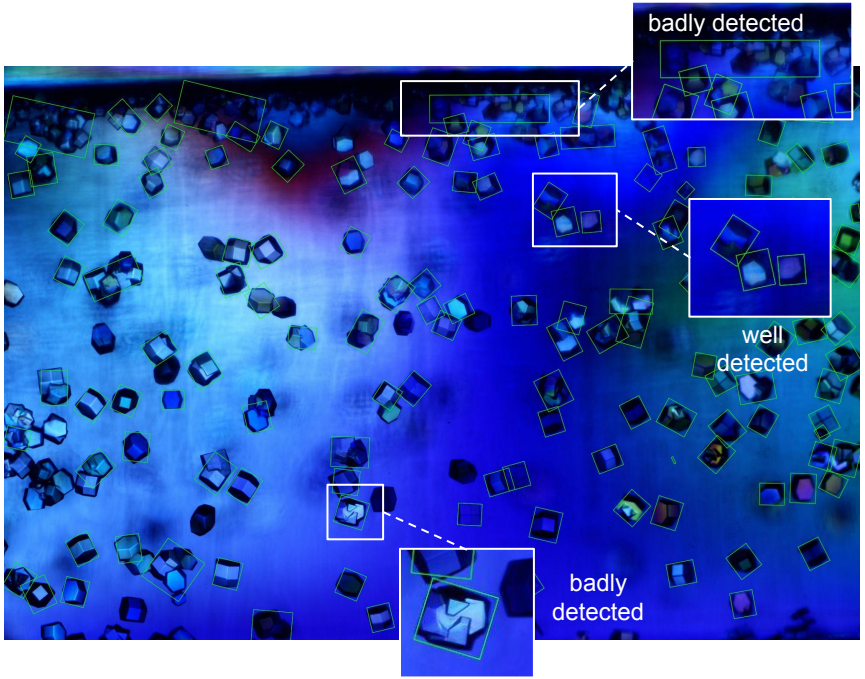


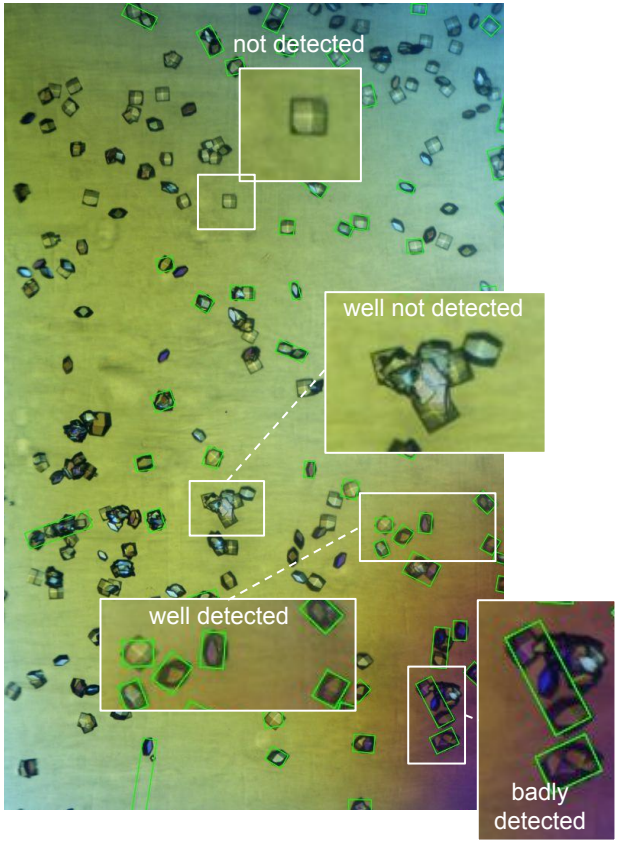


*Curated dataset for classification: not useful for our goal*

# Current approaches: Bischoff et al. model vs. experimental data

Due to the high complexity of context synthetic data-based, DL models are not perfect: we need to add new context through the training datasets using experimental data



# Conclusions

- Microcrystal segmentation is desirable for SSX experiment: avoid radiation damage, automatization, sample characterization before data collection.
- Synthetic datasets do not cover all experimental escenarios.
- Curated and open experimental datasets are mandatory to improve DL-based crystal segmentation.
- How to segment microcrystals from experimental data in a semi-automatic manner is still not clear.