# Ptychography SOM–CNN Reconstruction Pipeline

Technical Overview for Application Demo

Xinxin Sun

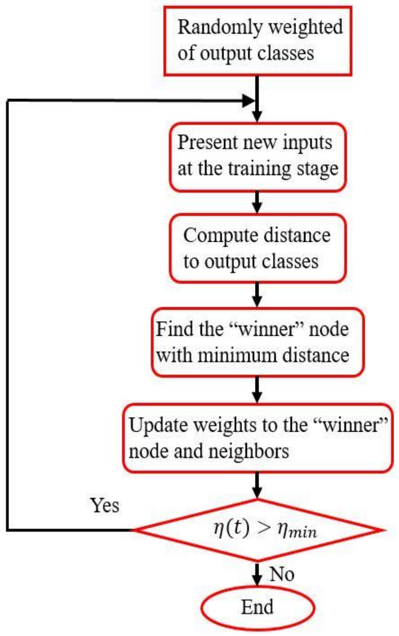
# Dataset Source

- CXIDB ID 65: Ptychography Gold Ball Example Dataset
- Depositor: Stefano Marchesini
- Method: Ptychography
- Sample: Nanometer gold balls
- Wavelength: 1.24 nm (1 keV)
- Data used: AuBalls_700ms_30nmStep_3_6SS_filter.cxi

# Method Overview

*Self-Organizing Map (SOM) for Label-Free, Physics-Aware Segmentation (Ptychographic Diffraction Frames)*

| Feature | Physical / Statistical Meaning in Coherent Diffraction |
|---|---|
| Intensity (I) | Bragg ring strength; speckle-rich coherent scattering |
| log(I) | Expands low-intensity region; enhances weak scattering |
| Local std | Speckle fluctuation magnitude; noise vs. structure contrast |
| Mean filter (3×3) | Local background estimator; smooths isolated speckles |
| Laplacian | Highlight ring boundary curvature and intensity roll-off |
| Gradient magnitude | Coherent edge sharpness; ring slope |
| Annular geometry (r) | Distance from beam center; ring radius signature |
| Angular pattern (θ) | Orientation-dependent speckle variation |
| Residual (I – Gaussian) | High-frequency structure; speckle–background separation |




**Workflow Summary**

**Extract 9 physics-aware features** from the 104×104 CXI diffraction frame.
**PCA reduction (9→6)** preserving >95% variance.
**Unsupervised SOM** forms physically interpretable clusters:
- Background noise
- Coherent ring body
- Central spot / defect-like core

**Automatic k-selection** via elbow method (best k=3 for all frames).
**Produce mask + cluster visualization** to quantify frame-to-frame variation (0/5/10/15/20 ms).

**Key Advantages**

**Label-free.** No ground-truth mask required; uses raw diffraction physics.
**Physics-aware.** 9-feature design captures speckle statistics, ring radius, symmetry, and artifacts.
**Interpretable.** Each SOM class corresponds to a distinct physical region of the Bragg pattern.
**Robust to noise.** Stable across time frames (0–20 ms) even with speckle variation.
**Transferable.** Works on any coherent diffraction dataset with minimal tuning.
**Efficient.** Runs in seconds per frame on CPU; ideal for fast screening or iterative pipelines.

# Method Overview (SOM → CNN Refinement)

*A Teacher–Student Framework for Robust Ptychographic Segmentation*

**Core Idea**

- SOM produces **stable, interpretable pseudo-labels** (class-3 ring region).
- CNN (U-Net) learns a **clean, high-resolution refinement** of SOM segmentation.
- The hybrid system removes boundary noise and stabilizes the ring mask.
- Optional **self-consistency loop** controlled by $\Delta$IoU threshold ($<0.005$).

**Key Advantages**

**Label-free pipeline.** No manual mask ever required.
**Noise-robust.** CNN trained on SOM's stable clusters avoids speckle noise.
**Explainable.** SOM provides physics interpretability; CNN adds geometric refinement.
**Transferable.** Same pipeline applicable to XCT, ghost imaging, quantum imaging, CDI, etc.
**Efficient.** Patch-based U-Net converges quickly on $104\times104$ diffraction patterns.

**Workflow**

**SOM Labeling (Teacher)**

- Provides initial cluster masks (background / ring / core).
- Class-3 → pseudo-label for CNN.

**CNN Training (Student)**

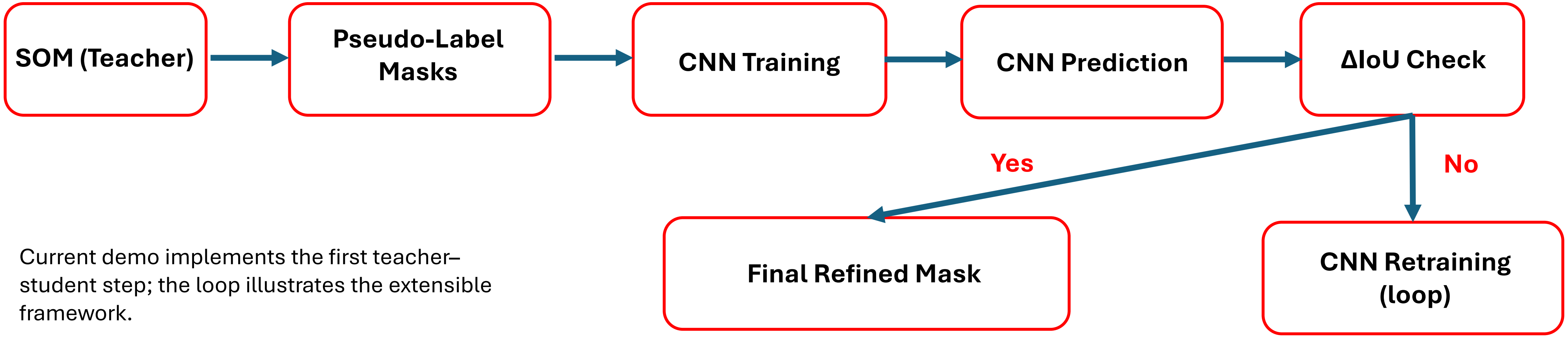
- Input: raw diffraction frame.
- Label: SOM pseudo-mask.
- CNN trains to denoise / refine / clarify the ring boundary.

**Pseudolabel Refinement Loop**

- CNN predicts improved mask.
- Compare IoU vs. previous iteration.
  - If $\Delta$IoU $< \varepsilon$ → convergence reached.
  - If $\Delta$IoU $\geq \varepsilon$ → update pseudo-label, retrain CNN.

**Final Output**

- High-fidelity, smooth ring segmentation for all selected frames.



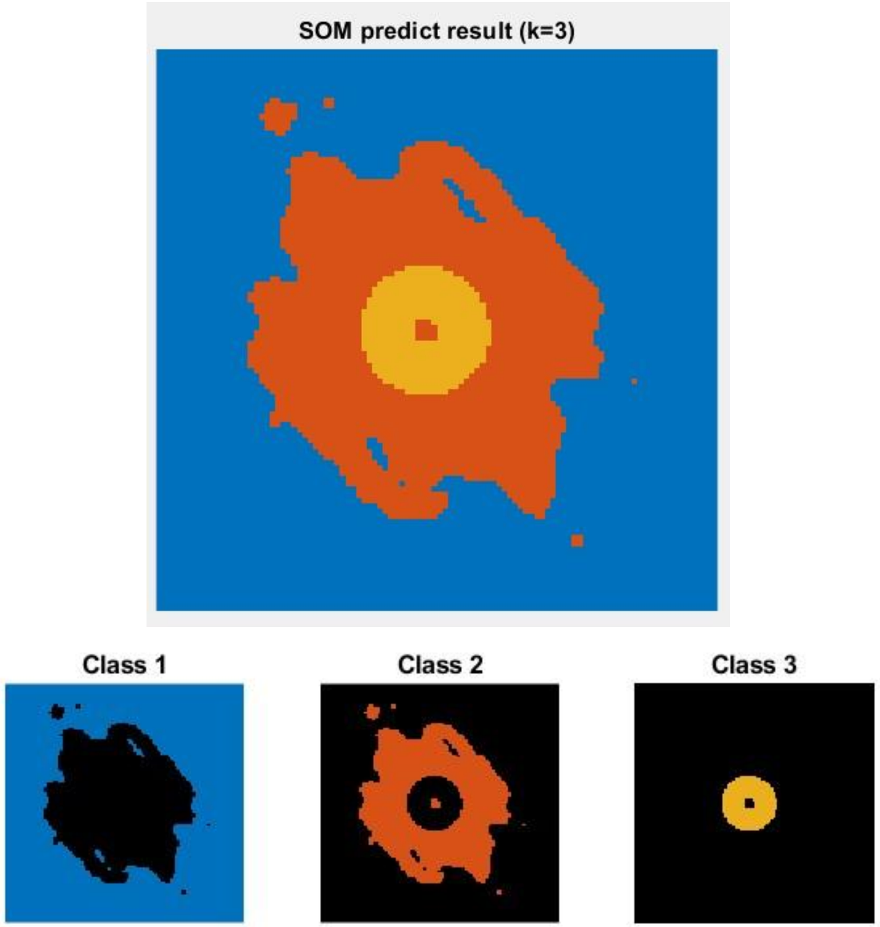

Current demo implements the first teacher–student step; the loop illustrates the extensible framework.

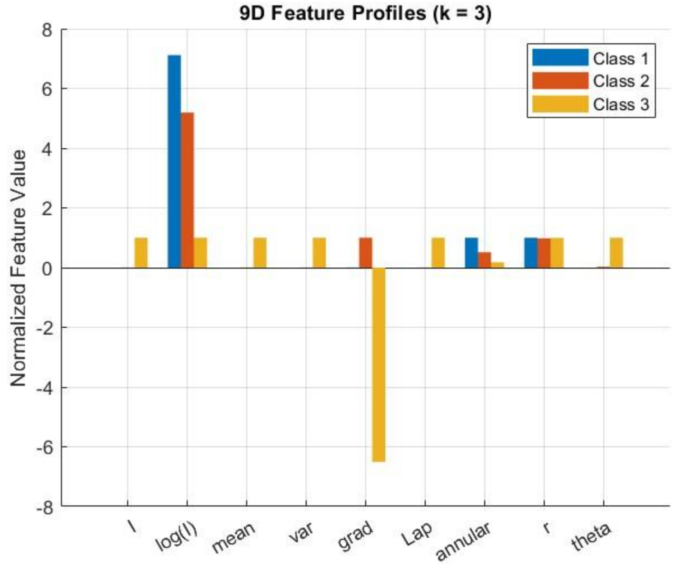
# Pipeline Architecture

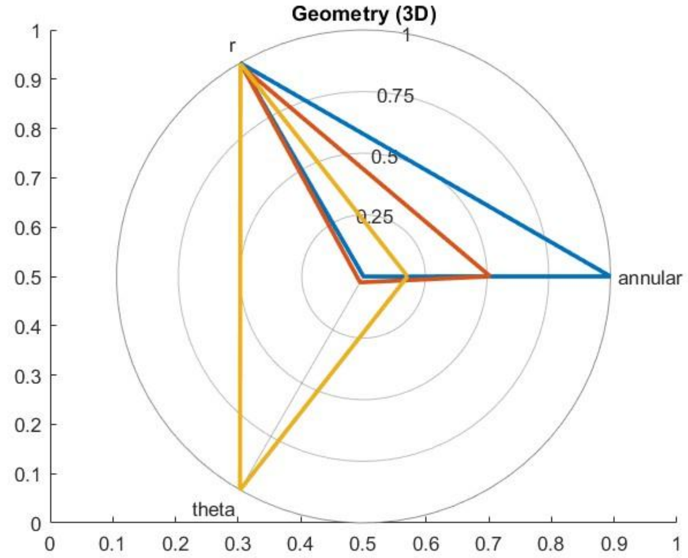
- 1. Extract diffraction frames (0,5,10,15,20)
- 2. Compute 9D handcrafted feature descriptor per pixel
- 3. PCA (6D) compression
- 4. SOM clustering (K selected via elbow)
- 5. Extract Class 3 (annular ring) mask
- 6. Iterative U-Net learning: 0→5→10→15→20

# SOM Feature Engineering

- **Feature set (9D):** intensity, log intensity, local mean, Laplacian, gradient, annular geometry, radial distance, angle
- **Core idea:** pixel-level SOM for physics-aware clustering
- **Transferable:** SOM centroids are reusable across frames / experiments



# Reusable SOM Cluster Centroids Across Diffraction Frames

**Physically-stable clustering**

- Unsupervised SOM learns **cluster centroids** in a 9-dimensional feature space reflecting diffraction physics (intensity, gradient, Laplacian, annular geometry).
- These centroids correspond to **statistically distinct scattering regimes** (core, annulus, background).

**Transferability across frames**

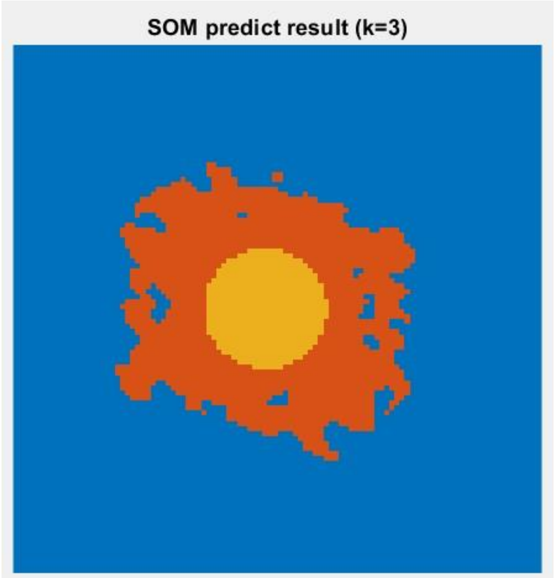
- Once the centroids are learned on one frame, they remain **stable across the entire scan** because scattering geometry and material composition are unchanged.
- SOM does not need to be retrained per frame.

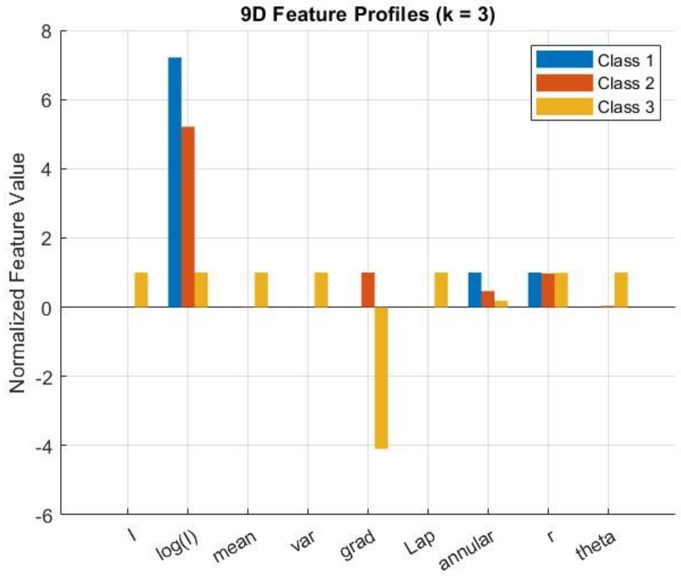
**Benefit for high-throughput experiments**

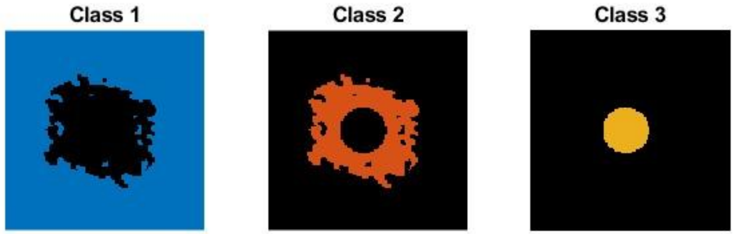
- “Train once, apply everywhere” segmentation.
- Consistent cluster identities across 0/5/10/15/20 frames.
- Enables scalable processing for large diffraction stacks or time-series data.

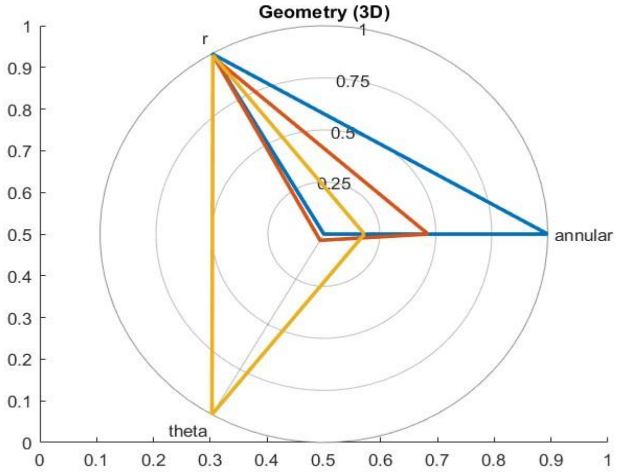
One set of learned feature prototypes generalizes across frames, providing physically-consistent segmentation without repeated optimization.

**Transfer of Learned Centroids (Train@0 → Apply@20)**



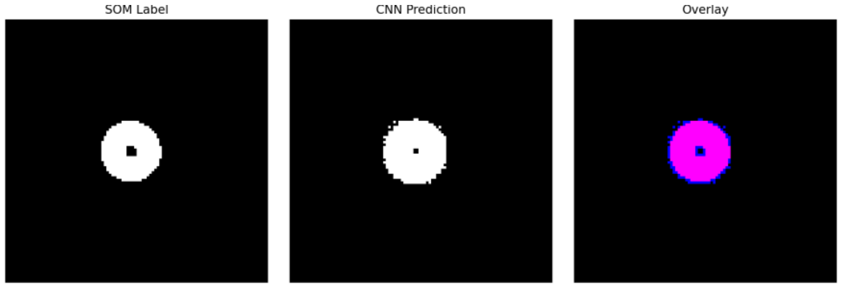



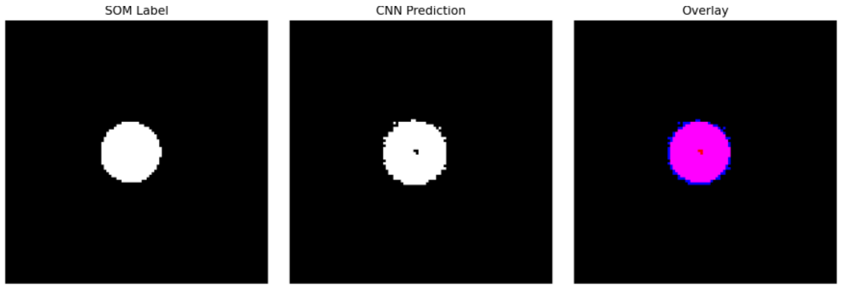
# SOM-to-CNN Self-Refinement: Frame-Consistent Mask Learning Across 0/5/10/15/20

- Stage 0: Train U-Net on Frame 0 (SOM Class 3)
- Stage 1: Predict Frame 5 → fine-tune on {0,5}
- Stage 2: Predict Frame 10 → fine-tune on {0,5,10}
- Stage 3: Predict Frame 15 → fine-tune on {0,5,10,15}
- Stage 4: Predict Frame 20 → final model
- CNN progressively regularizes the geometry and enforces cross-frame consistency

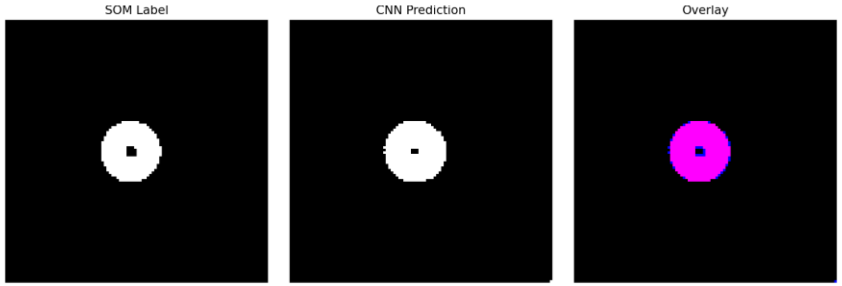
## Frame 0



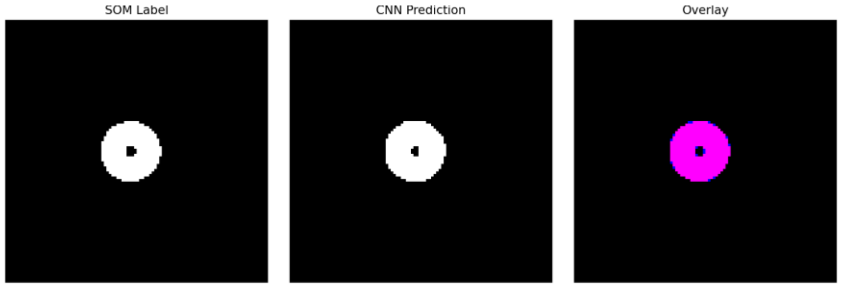

## Frame 5



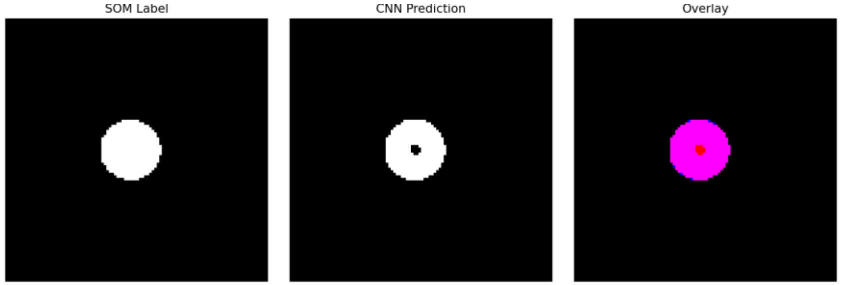

## Frame 10

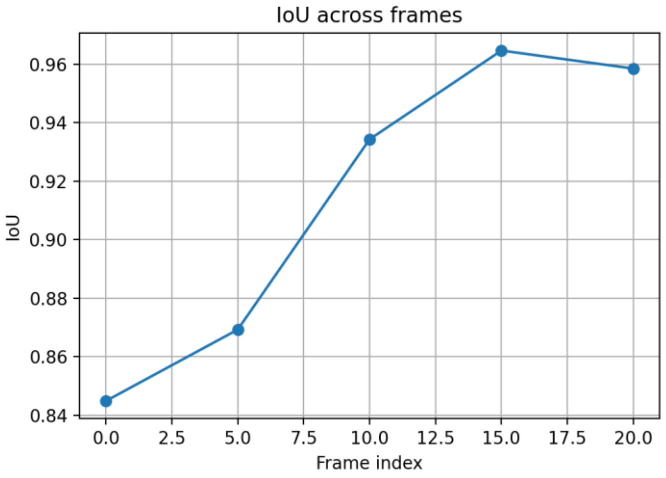



## Frame 15




## Frame 20




- SOM (teacher) provides stable, physics-informed pseudo-labels.
- CNN (student) learns high-fidelity masks that are smoother and more consistent.
- Overlays show tight alignment between SOM and CNN across all frames.
- IoU curve demonstrates convergence and robustness across scanning positions.

# Conclusion and Beamline Impact

## Conclusion & Key Takeaways

**1. SOM provides a physics-informed, noise-resilient supervisory signal**
The self-organizing map supplies stable and interpretable pseudo-labels derived directly from scanning geometry, enabling robust initialization without manual annotations.
**2. CNN learns high-fidelity, frame-consistent masks**
Once trained on SOM pseudo-labels, the CNN rapidly converges to smooth and accurate masks that remain consistent across scanning positions and experimental noise.
**3. SOM→CNN is a scalable teacher–student paradigm**
The workflow generalizes naturally to multi-frame refinement, sparse labeling scenarios, and dynamic scanning experiments.
**4. The hybrid approach reduces annotation cost to near-zero**
Only one SOM-labeled frame is required; all subsequent refinement is performed automatically via self-supervised fine-tuning.
**5. Convergence curve demonstrates model stability**
IoU evolution across 0/5/10/15/20 frames shows that the CNN stabilizes geometry and matches SOM-defined structures with high precision.

## Why This Matters for Your Beamline

- Enables **real-time or near-real-time segmentation** for coherent diffraction imaging.
- Reduces reliance on manual or heuristic mask design.
- Produces **consistent masks across long scanning sequences**, benefiting reconstruction stability.
- The framework is lightweight, interpretable, and easy to integrate with existing pipelines (CXI, ptychography, Bragg CDI).

## Next Steps

- Extend CNN to multi-class SOM labels (e.g., background, ROI, beam-stop).
- Integrate into automated reconstruction workflows.
- Validate on experimental datasets with drifting beam conditions or changing sample states.

*This SOM→CNN hybrid workflow extends naturally beyond gold-ball ptychography to Bragg CDI, ghost imaging, and other coherent-scattering datasets where stable segmentation boosts reconstruction quality and experiment throughput.*