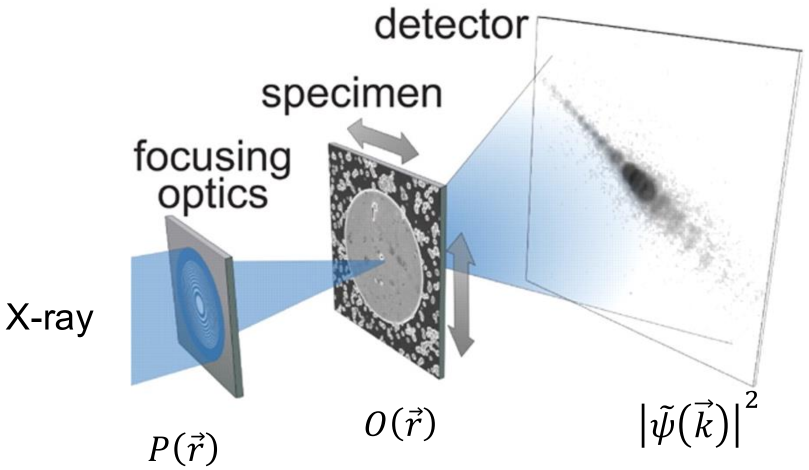
JUL 22, 2026

FROM BENCHMARKING TO AGENTIC WORKFLOWS

Foundation Models For X-ray Data Analysis

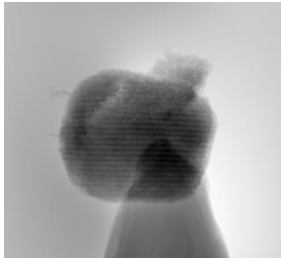
XIANGYU YIN
Assistant Computational scientist
Computational Science and AI Group
Advanced Photon Source
Argonne National Laboratory

Ptychography



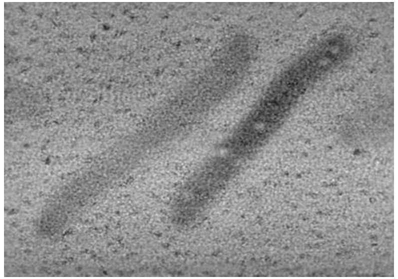
Retrieve object $O(\vec{r})$ (and phase $P(\vec{r})$) from a series of scanning diffraction patterns ($|\tilde{\psi}(\vec{r}_p,\vec{k})|^2$)

Sensitive to reconstruction parameters and prone to artifacts



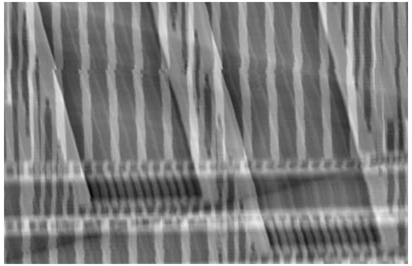
Sample: integrated circuit
Instrument: 2xfm
Diagnosis:
- Local scan drift

Actions:
- Turn on position correction
- Crop data to 64x64
- Increase the number of iterations



Sample: Ecoli
Instrument: Bionanoprobe
Diagnosis :
- Slowly-varying background
- Noisy artifacts

Actions:
- Use variable probe correction
- Use Poisson noise model



Sample: super crystal
Instrument: Ptycho-SAXS
Diagnosis :
- Non-unform background
- Singularity artifacts in the object

Actions:
- Use compact batch scheme with momentum acceleration
- Multislice ptychography

Benchmark LLMs/VLMs on Ptychography Tasks

Task 1: Visual artifact detection

Find artifacts (e.g., grid/aliasing/ring) in reconstructed ptychography images.

Input: reconstructed image
Output: diagnosis

Task 2: Parameter recommendation

Given a dataset's metadata, recommend new/better reconstruction parameters.

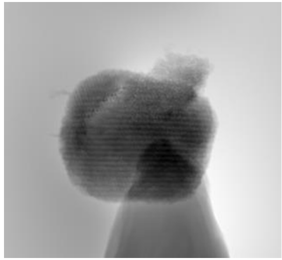
Input: sample metadata
Output: actions

Strategies compared

In-context learning (ICL): prompting w. examples

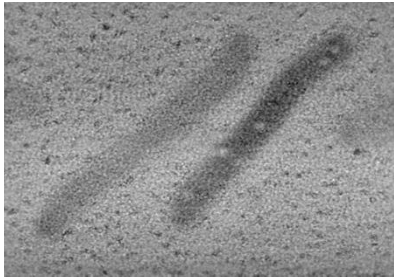
Sample-specific few-shot (SSFS): fetch relevant examples

Supervised fine-tuning (SFT): train adaptors w. few examples



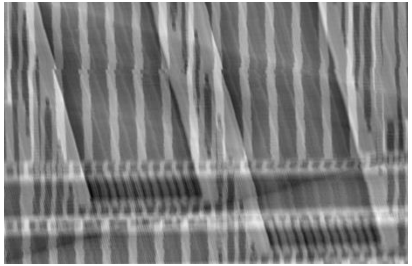
Sample: integrated circuit
Instrument: 2xfm
Diagnosis:
- Local scan drift

Actions:
- Turn on position correction
- Crop data to 64x64
- Increase the number of iterations



Sample: Ecoli
Instrument: Bionanoprobe
Diagnosis :
- Slowly-varying background
- Noisy artifacts

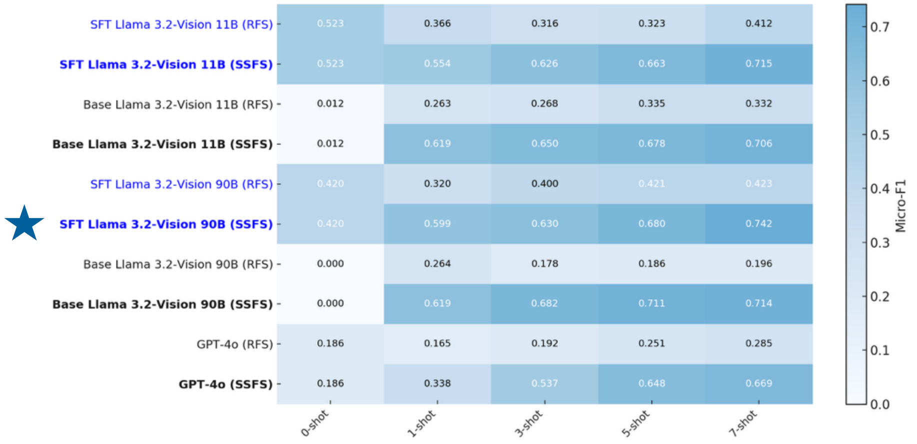
Actions:
- Use variable probe correction
- Use Poisson noise model



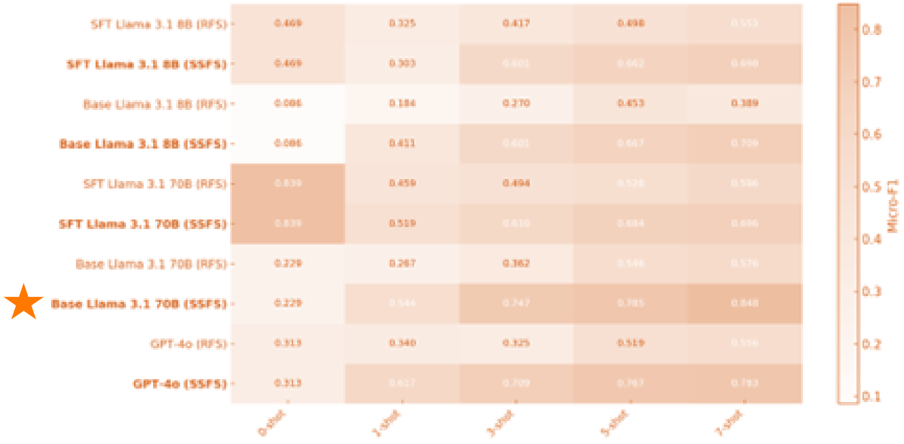
Sample: super crystal
Instrument: Ptycho-SAXS
Diagnosis :
- Non-unform background
- Singularity artifacts in the object

Actions:
- Use compact batch scheme with momentum acceleration
- Multislice ptychography

Benchmark LLMs/VLMs on Ptychography Tasks



Task 1: Visual artifact detection



Task 2: Parameter recommendation

Best recipe:
SFT + sample-specific few-shot examples

Example relevancy matters most, irrelevant context can hurt performance

Best recipe:
large base + sample-specific few-shots

SFT can produce “rigid expert”: great 0-shot but few shots degrades performance

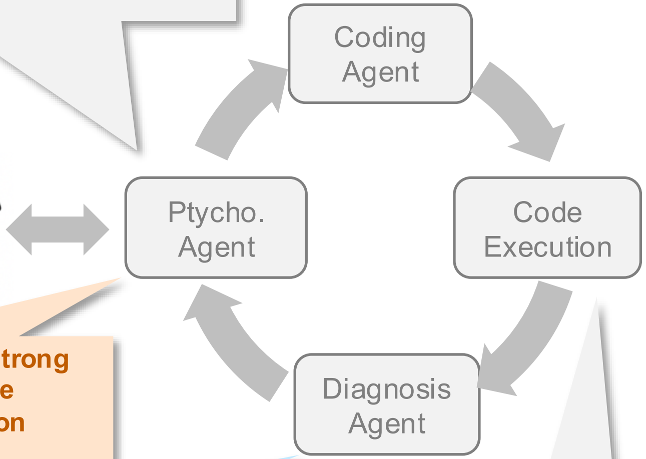
Build Workflows around Tested Models and Recipes

Can you help me process my ptychography dataset?

Certainly! Please answer some questions so that I can figure out your problem settings.

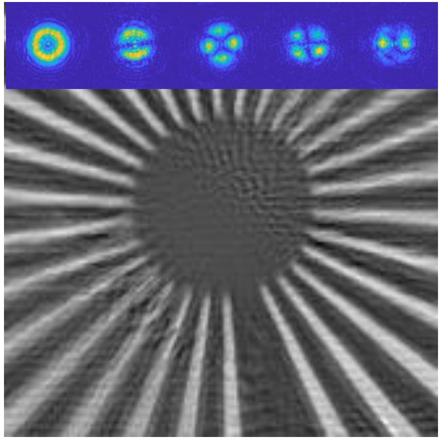
1. Which instrument did you use to acquire this data?
2. Was the data acquired in step-scan or fly-scan mode?
3. …

I suggest using 5 probe modes and single slice model.




```
% initialize parameters
par = {};
par.z = 1.92;
par.energy = 8.8;
par.Ndp = 256;
...
par.multislice_ptycho = false;
par.Nprobe = 5;
par.Niter = 1000;
...
% start reconstruction
ptycho_recon_exp_data(par);
```

Here are the reconstructed object and probe:



The reconstructed object seems to exhibit some grid or line artifacts. The final probe mode shows clear structures and a high-power percentage.

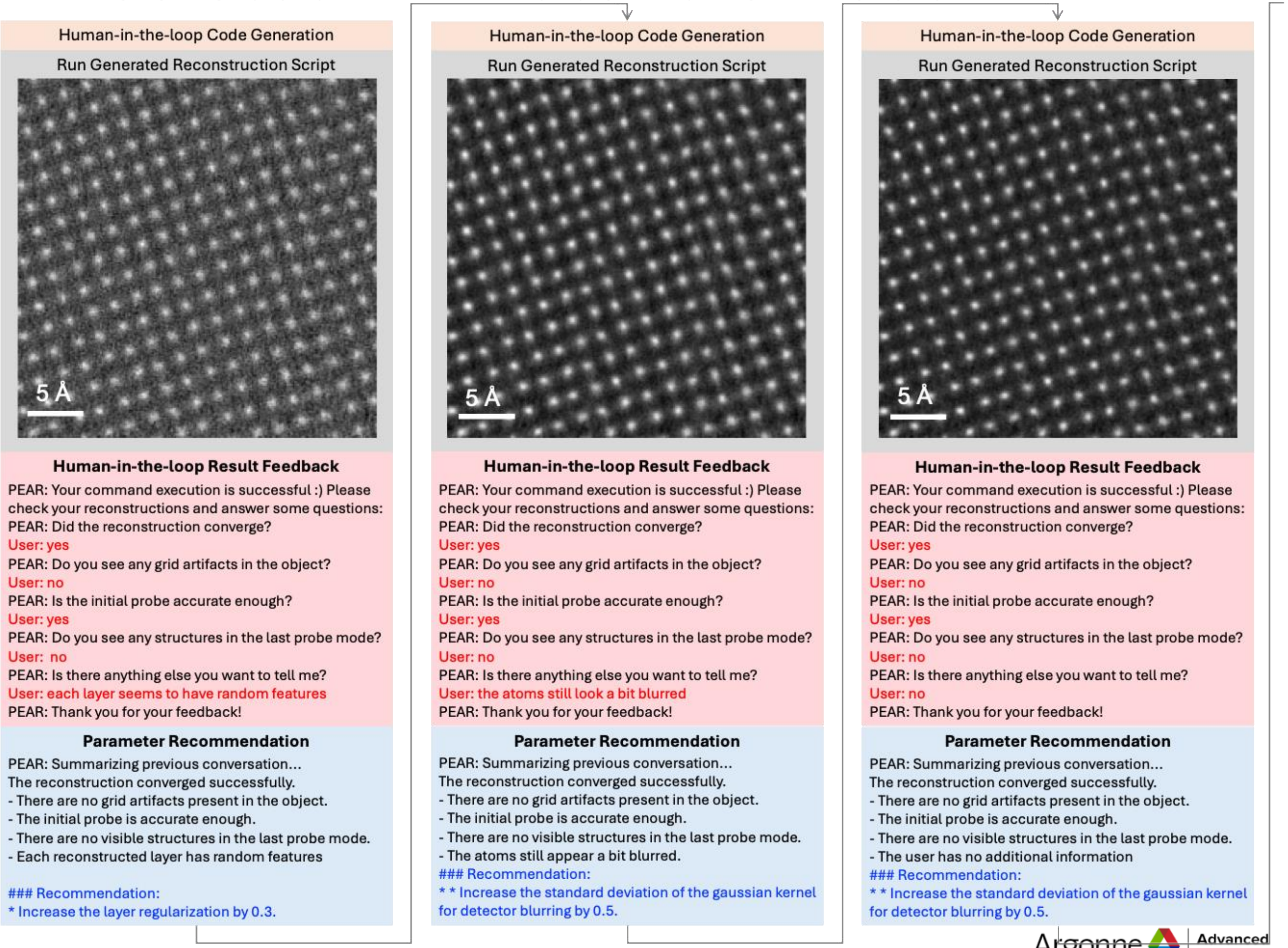
The artifacts could be caused by strong beam vibrations. Try increasing the number of probe modes and turn on variable probe correction.

Case Study: AI-Assisted 2D Materials Recon.

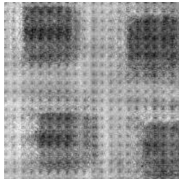
arXiv:2410.09034

AI workflow can improve reconstruction quality in a few iterations

Bayesian optimization can take tens/hundreds of iterations



From Parameter Tuning to Algorithm Discovery



Can you help me find a regularizer to suppress grid artifacts in IC dataset

Action: GENERATE

```
def regularize_llm(self):
"""Physics-aware
regulariser for hard-X-ray
IC ptycho."""
<function content>
self.set_data(updated_obj)
```

Here are the reconstructed object and probe:

Research Agent

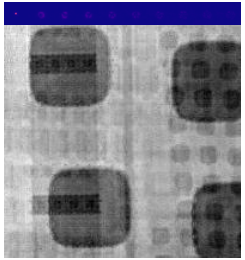
Coding Agent

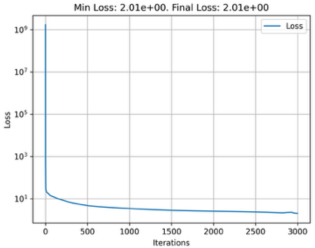
Code Execution

Evaluation Agent

Analysis Agent

ptychi





- **(Ground truth) metrics mode**
- **Human-in-the-loop mode**
- **Auto mode (tools & VLMs)**

Below is a compact "look-up sheet" you can keep next to your reconstruction script. It lists the regularizers, constraints and mitigation steps that have proved useful for hard-X-ray ptychography of planar integrated-circuit (IC) samples acquired in step-scan mode with (i) an inaccurate initial probe, (ii) very sparse/large scan steps and (iii) visible raster-grid artifacts.

Regularization techniques that work for this experiment

A. Spatial-domain sparsity & piece-wise constancy

- Isotropic Total-Variation (TV) or Total-Generalised-Variation (TGV) penalties suppress noise and fill in missing pixels when the overlap is small; TV was shown to recover 100 nm IC lines at only 15 % overlap (arxiv.org).
- Structure-Tensor or directional TV priors preserve the long, straight metal lines typical of IC routings (arxiv.org).

B. Frequency-domain / smoothness penalties

- L2/Tikhonov or gradient-penalising terms on the object and/or probe, implemented in modern autodiff engines (e.g. phaser/ptyRAD), keep high-frequency probe noise from leaking into the object (arxiv.org, ar5iv.labs.arxiv.org).

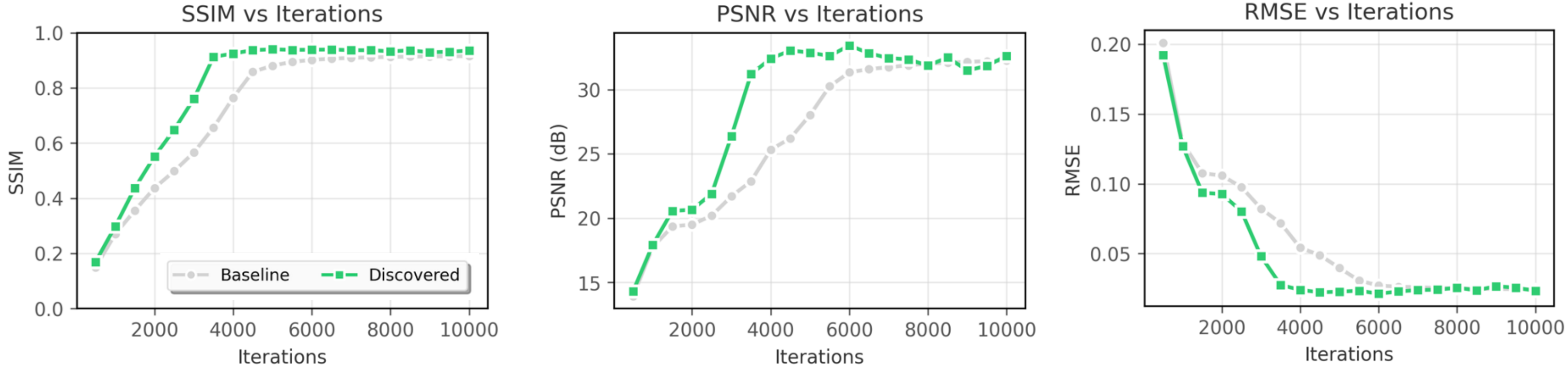
Action: EVOLVE

```
def regularize_llm(self):
"""Multi-scale anisotropic
DTGV-Huber regulariser with
adaptive notch filtering."""
<function content>
self.set_data(updated_obj)
```

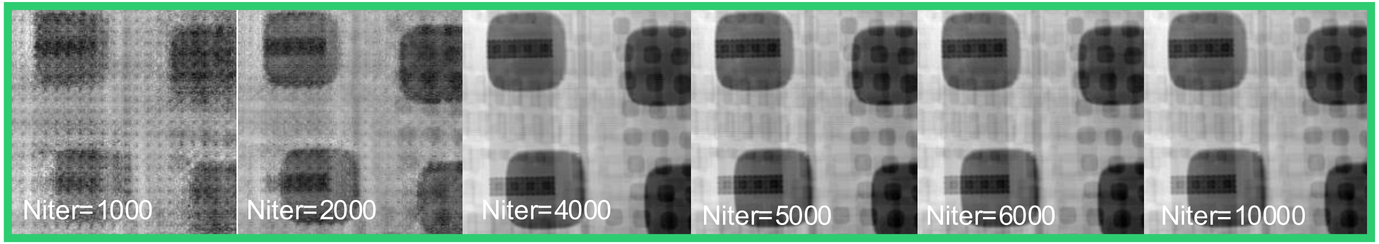
- Expose the main hyper-parameters (gamma, lam_init/min/max, huber_eps_factor, phase_tv_strength, softplus_beta_scale) through the external API so they can be tuned per dataset; grid/Bayesian search typically yields ≥2 dB PSNR
- Keep the Barzilai-Borwein step length as a tensor (do not call .item()) and compute it in full complex arithmetic; this preserves autograd graphs for unrolled or meta-learning pipelines.

…

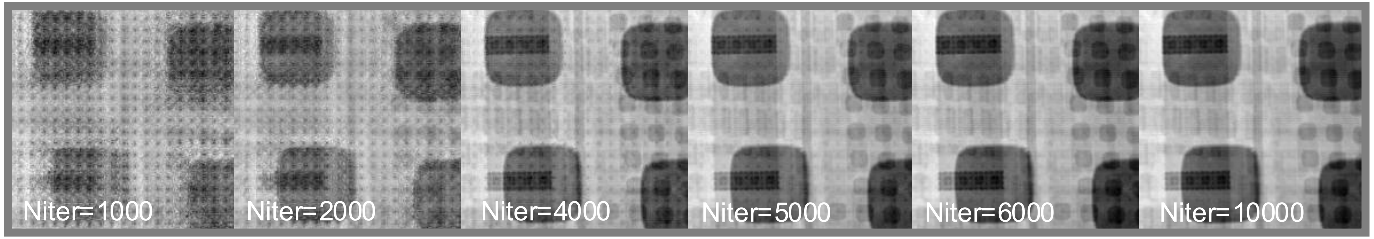
Discovered Regularizer v.s. the Hand-Crafted Baseline



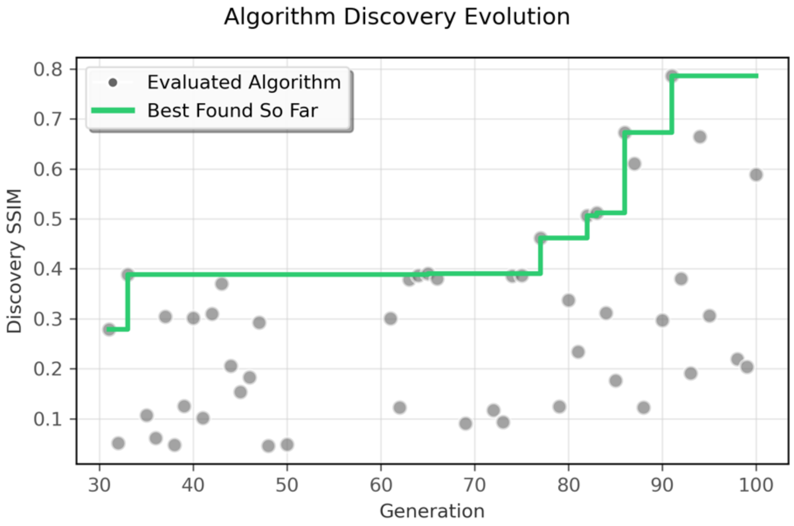
Discovered

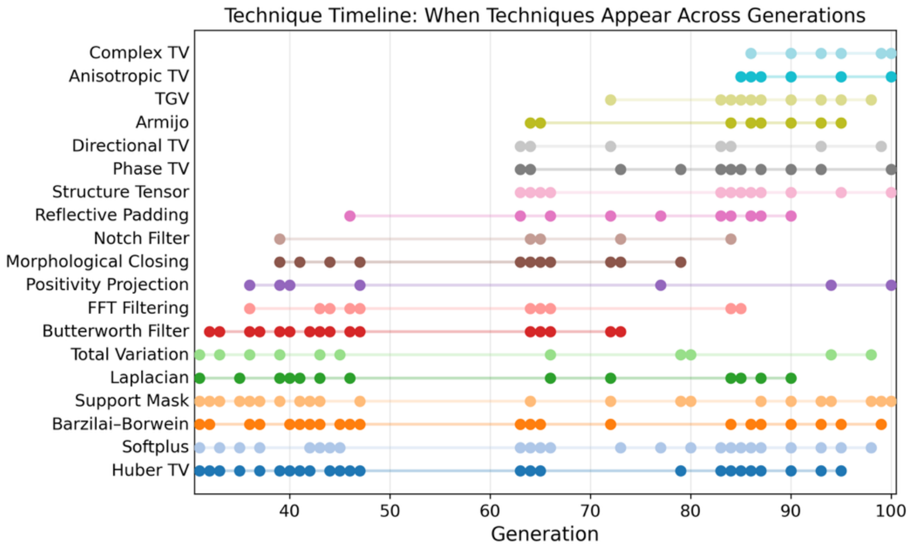


Baseline



Algo. Evolution History Recorded for Insights





Key Technical Insights:

Adaptive notch filtering:	FFT-based grid artifact suppression
Structure-tensor TV:	Edge-aware anisotropic smoothing
Barzilai-Borwein:	Adaptive step size acceleration

Most Data Analysis Problems are Unstructured

Ming Du

Structured problem

Optimizer

Loss

Model | Data & labels

Unstructured problem

Numerical optimization

A pool of TIFF images

Filtering

All sort of transforms

Morphological operations

Challenges for unstructured analysis

Development environment setup

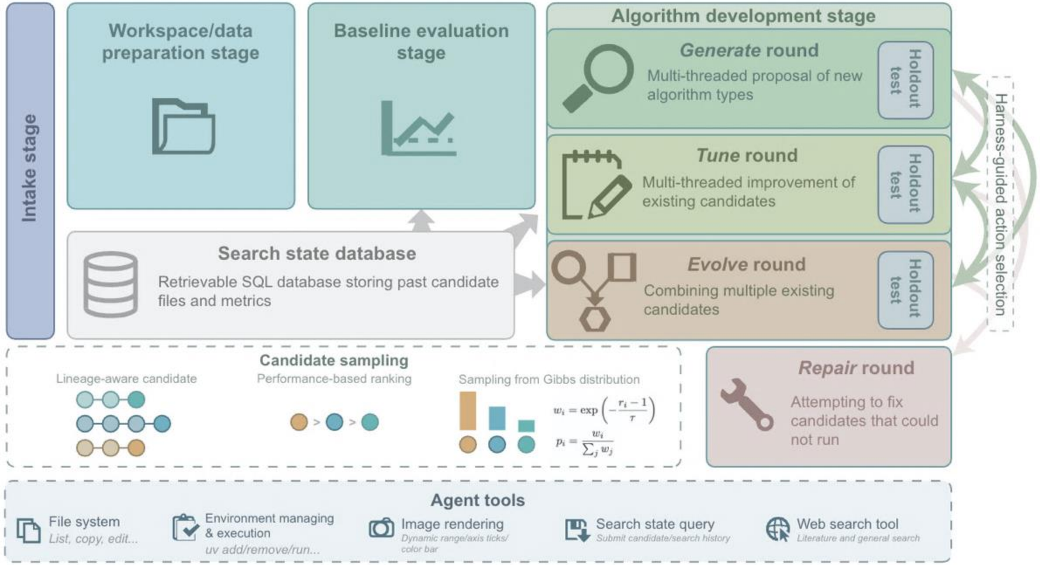
Evaluation script

Troubleshooting

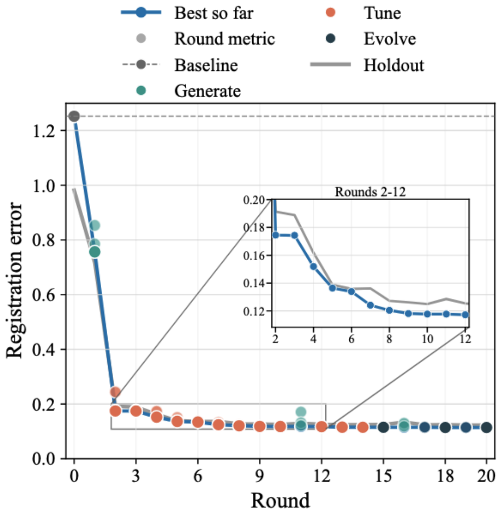
Exploration-exploitation tradeoff

Automated Agentic System for Unstructured Analysis

Ming Du



Example: finding a translational registration algorithm for small, noisy fluorescence images



Algorithm	Development set	Holdout set
Brute-force error minimization	1.25	0.98
Phase correlation	5.28	5.59
OpenEvolve	0.20	0.23
ERA	0.46	0.50
CVEvolve	**0.11**	**0.12**

Du, M., Yin, X., Luo, Y., Beniwal, D., Tang, S., Sharma, H., & Cherukara, M. J. (2026). CVEvolve: Autonomous Algorithm Discovery for Unstructured Scientific Data Processing. *arXiv preprint arXiv:2605.11359*.

https://github.com/AdvancedPhotonSource/CVEvolve

Skills + General Agent for Interactive Flexible Analysis

Yi Jiang

Ptycho Skills

If you need...	See...
Data formats & required HDF5 keys	references/data_formats.md
Parameter reference & suggestions (pear wrapper params)	references/ptychox_parameters.md
Troubleshooting artifacts & failures	references/troubleshooting.md
Quality assessment & "what to change next"	references/quality_assessment.md
Physics background & key equations	references/theoretical_foundations.md
Deep API customization (Options, Schedule, constraints, losses)	references/ptychox_api_guide.md
Warm-start optimization (multi-stage refinement)	references/warm_start_guide.md
Batch reconstruction (many scans)	references/batch_reconstruction.md
X-ray template (step-scan)	templates/ptychox_xray_standard.py
X-ray template (fly-scan)	templates/ptychox_flyscan.py
Near-field template	templates/ptychox_nearfield.py
Multislice template	templates/ptychox_multislice.py
Electron (4D-STEM) template	templates/ptychox_electron_stem.py
Custom regularizer template	templates/ptychox_custom_regularizer.py
JSON schema & sample configs	examples/
Running on ALCF (Polaris/Aurora)	See "Running on ALCF" section below
What makes PtychoX different (vs pty-chi/ptyrad/phaser/foldslice)	docs/unique_features.md
PtychoX vs pty-chi (algorithm-level differences, ported bug fixes)	docs/vs_ptychi.md

Self-updating of Ptycho Skills

"Can you process a new data?"

↓

"Explore more parameters/strategies to improve image quality"

↓

"Summarize the lessons we learned and update the Ptycho Skill"

The AI era in X-ray data science is not about replacing physics and scientists with black boxes.

It is about rebuilding data analysis systems so physics, AI, and human expertise work together.

Thanks! I look forward to the discussion.

Xiangyu Yin
Assistant Computational Scientist
Computational Science and AI Group
Advanced Photon Source
Argonne National Laboratory

Acknowledgement: Arthur Glowacki, Yi Jiang, Wendy Di, Grace Luo, Olga Antipova, Luxi Li, Si Chen, Junjing Deng, Barry Lai, Juanjuan Huang, Shelly Kelly, Peco Myint, Michael Prince, Sinisa Vesali, Nicholas Schwartz, Ming Du, Aileen Luo, Nina Andrejevic, Mathew Cherukara, Stefan Vogt, Alec Sandy, Jonathan Lang, Laurent Chapon

This research used resources of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357. I acknowledge funding support from Argonne **LDRD 2023-0049**