HPC+AI-ENABLED X-RAY SCIENCE

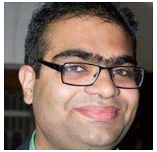
MATHEW J. CHERUKARA
Group Leader, Computational X-ray Science
Advanced Photon Source

MJCherukara
mcherukara

Argonne NATIONAL LABORATORY | 75 1946–2021

ACKNOWLEDGEMENTS: CXS GROUP

HPC/Workflows



Hemant Sharma

Tekin Bicer

Daniel Ching

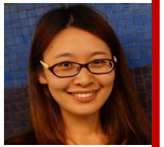
Applied Math Image/Signal processing

Ash Tripathy

Zichao (Wendy) Di

Doga Gursoy

Deep Learning Automatic Differentiation



Yudong Yao

Saugat Kandel

Nina Andrejevic

Jay Horwath

Anakha Babu

Theory + Computational Physics/Materials

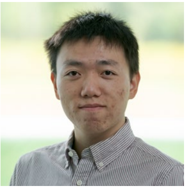
M. Van Veenendaal

Bob Von Dreele

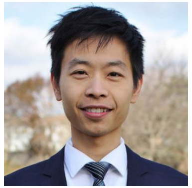
Brian Toby

Xiaoxu (Shirley) Guo

ACKNOWLEDGEMENTS



Tao Zhou, CNM, Argonne



Henry Chan, CNM, Argonne

Ross Harder, APS, Argonne

APS@ANL: Antonino Miceli, Barbara Frosik, Yi Jiang, Steven Henke, Sinisa Veseli

MCS@ANL: Prasanna Balaprakash, Zichao Di (also CXS)

DSL@ANL: Tekin Bicer (also CXS), Zhengchun Liu, Ryan Chard

MSD@ANL: Stephan Hruszkewycz, Charudatta Phatak

CNM@ANL: Subbu Sankaranarayanan, Martin Holt

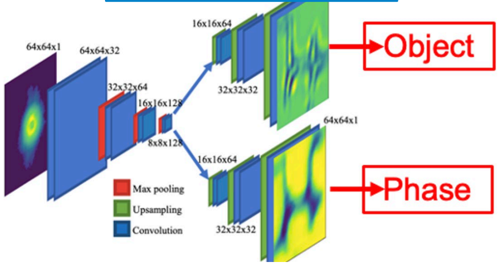
LBNL: Pablo Enfedaque, Alex Hexemer

NVIDIA: Ekaterina Sirazitdinova, Geetika Gupta

FUNDING:
Argonne LDRDs: AICDI, AutoPtycho
AI SUF: Digital Twin for In-Silico Experiments

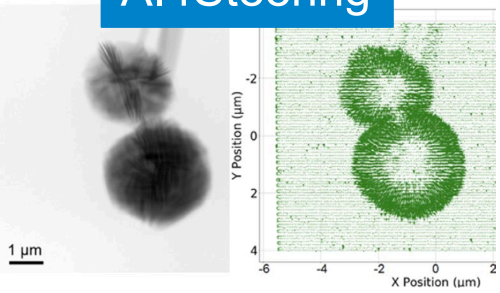
OUTLINE: AI4SCIENCE

AI4Analysis



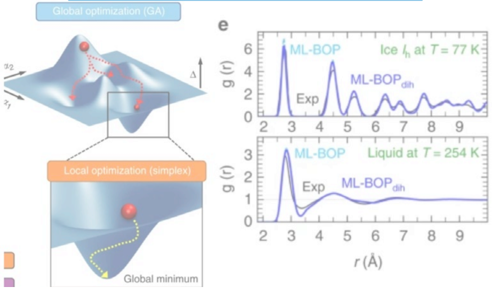
- AI@Edge: >100X faster and (sometimes) more accurate analysis
- Enables real-time analysis on Gb/s data streams

AI4Steering



- AI@Edge: Self-driving experiments & instruments:
  - maximize info gain in minimal time

AI4Knowledge



- Learn physics directly from scattering data

Outlook

- Even more HPC
- MLOps
- AI on sensors

AI is a *necessity* to fully exploit instruments' capabilities

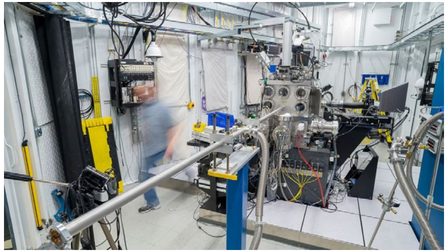
X-RAY LIGHT SOURCES OF THE WORLD

- >50 across the world
- Current and future upgrades to increase brightness and coherence

Enable scale-bridging, multi-modal view of materials *operando*

Source: Xu, W., et al. "The complexity of thermoelectric materials: why we need powerful and brilliant synchrotron radiation sources?." Materials Today Physics 6 (2018): 68-82.

THE ADVANCED PHOTON SOURCE @ ARGONNE



~5,700 researchers per year from academia, industry, and government

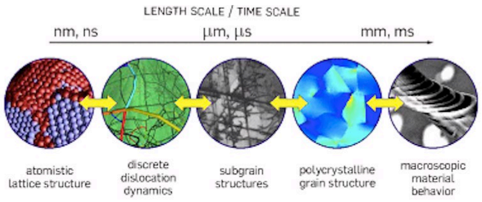
~70 beamlines (instruments) capable of independent operation; all unique; <u>many multimodal</u>

- Imaging (tomography); Scanned probe microscopy (strain + fluorescence mapping); Coherent scattering (XPCS, Ptychography); Diffraction (MX, powder, PDF, HEDM, stress/strain, SAXS, GISAXS); Spectroscopy (IXS, nuclear resonant scattering, XMCD, XAFS

Need unique data workflows and compute solutions

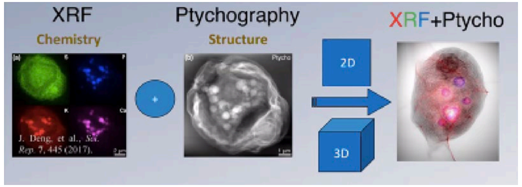
X-RAY MICROSCOPY IN A NUTSHELL

Scale



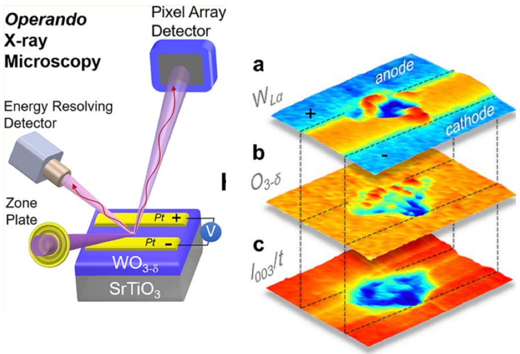
Scale-bridging imaging
- 5-6 orders of magnitude in a SINGLE instrument

Multi-modal



Simultaneous imaging
- Composition
- Structure
- Defects
- Oxidation state
- Strain
- Photovoltaic response

Operando



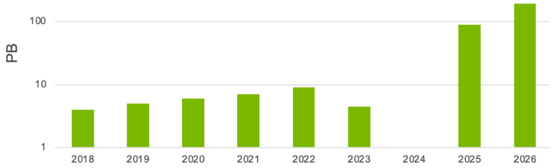
Environmental imaging
- Electrochemcial
- Material synthesis
- Cryogenic
- High pressure
- Magnetic

Slide adapted from: Barry Lai, APS, ANL

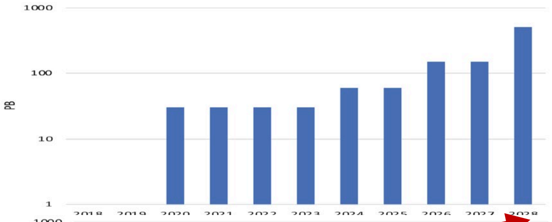
MOTIVATION: DATA RATES AND COMPUTE NEEDS

Data volumes

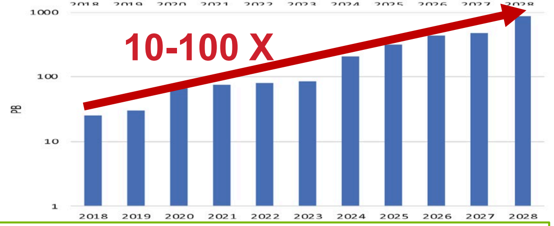
APS



LCLS

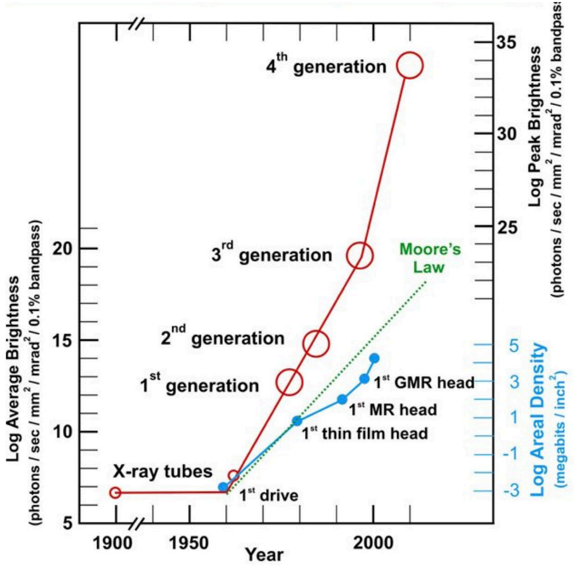


DOE Total



Schwarz, N., Campbell, S., Hexemer, A., Mehta, A., & Thayer, J. (2020, August). In Smoky Mountains Computational Sciences and Engineering Conference (pp. 145-156).

Compute limits



http://archive.synchrotron.org.au/images/AOF2017/Boland---AOF---Future-light-sources-2017-05-29.pdf

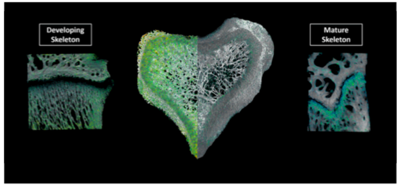
Compute needs outpace Moore's law

AI is a *necessity* to fully exploit instruments' capabilities

MOTIVATION 2: INVERSE PROBLEMS IN MATERIALS CHARACTERIZATION

① IMAGING TAKING A SNAPSHOT

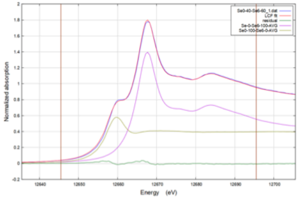
Synchrotron X-rays allow us to take an image of a sample. By studying the interaction of light with an object, we are able to get information about the structure or the function of whatever we are imaging. Our beamlines can take a picture of the tiny airways in a lung or get a three-dimensional image of materials like steel pipelines.



E.g.: Projections -> 3D image

② SPECTROSCOPY ANALYZING THE CHEMISTRY

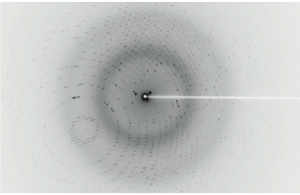
We can see how different wavelengths of light interact with matter, allowing us to analyze what the sample is made of. With spectroscopy we can look at the matter inside of a lentil or model the molecules that exist in space.



Spectra -> chemical composition

③ DIFFRACTION AND SCATTERING UNDERSTANDING THE STRUCTURE

Sometimes light can bounce off a sample and create a unique pattern. This pattern allows us to gain insight into the structure of the object. With diffraction and scattering we are able to understand the shapes of proteins inside of living things or visualize the structure of crystalized materials.



Diffraction -> atomic structure

Inverse problems are computationally expensive!

Source: https://www.lightsource.ca/public/images-pdfs-tour-posters/2020.light.pdf

INVERSE PROBLEMS IN MATERIALS CHARACTERIZATION

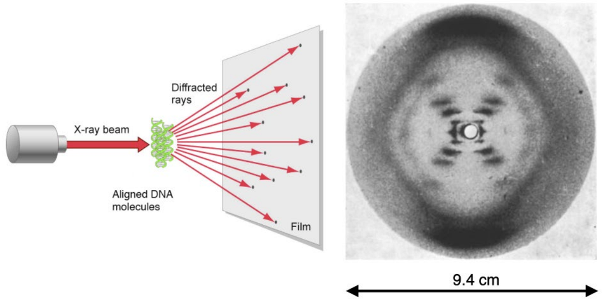
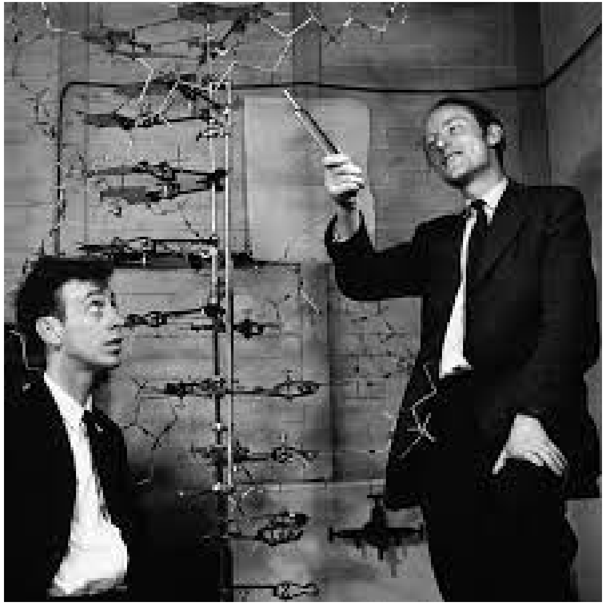


Figure 1: *The experimental geometry used by Franklin and Gosling for their 1952 x-ray diffraction experiments on aligned DNA fibers (left). The famous "Photo 51" taken by Franklin and Gosling that allowed Watson and Crick to figure out the structure of DNA and was published in [1]. The actual width of this image is 9.4 cm (as indicated below the image).*

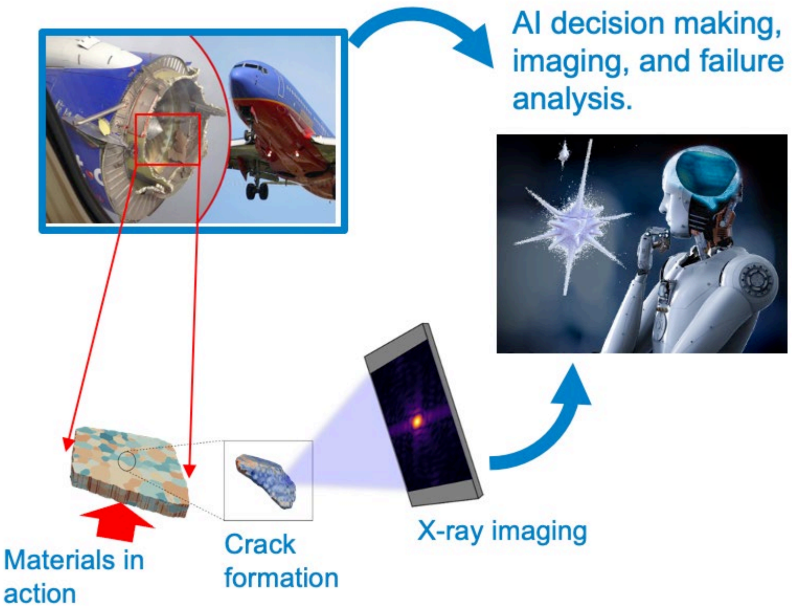


Nobel prize in Medicine : 1962

AI4Analysis

MOTIVATION 3: REAL-TIME FEEDBACK

Experimental steering



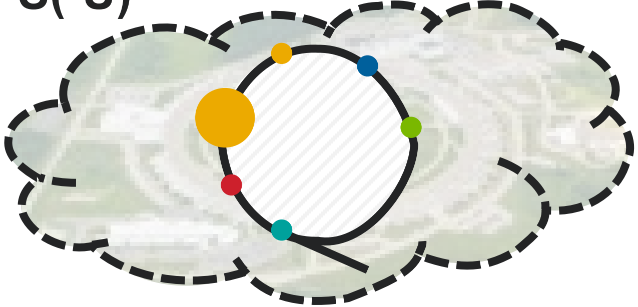
Autonomous experiments need real-time data inversion
- Need to invert data on order of seconds or less

- Autonomously identify and track the most participatory volumes in bulk material.
- Sample only as much as necessary
  - Minimize radiation damage

POLARIS – INSTRUMENT 2 EDGE (I2E)

Tightly coupling APS instruments with ALCF supercomputers

APS(-U)



ALCF

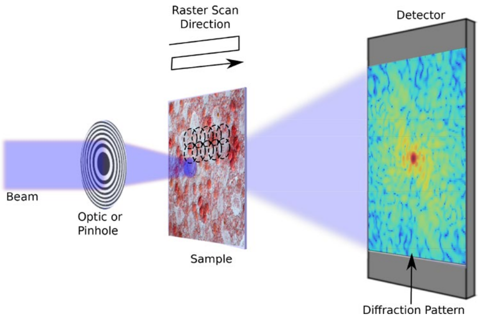
Workflows:
- Scalable software solutions for inverse problems
- Online and offline DL training at scale, deploy at edge

Polaris:
- 560 nodes:
  - 128-core AMD Milan CPU
  - 4X NVIDIA A100
- ~44 PFLOP/s peak performance (double precision)
- ~4 PFLOP/s on-demand use by experimental facilities including APS

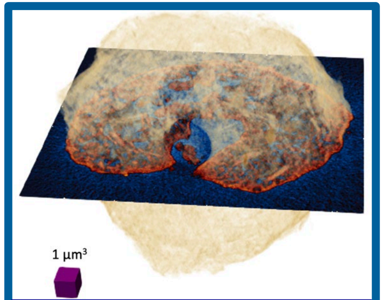
AI4ANALYSIS: COHERENT IMAGING

COHERENT IMAGING IN APSU

Ptychography



Scan:
- ~500 nm beam
- ~10nm resolution



Catalytic particle: Y.Jiang, J.Klug, J.Deng, Y.Yao et. al. @APS

Today: mm^2 @ 10 nm in hours

APSU: cm^2 @ 10 nm in 1 day

- **equivalent to mapping USA @ few meter resolution**
- **data rate = streaming 100,000s of HD movies simultaneously**

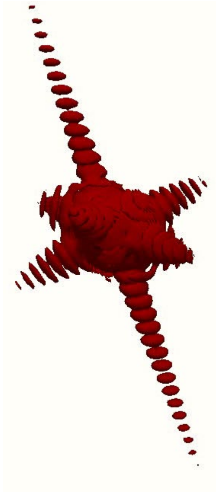
CURRENT STATE: ITERATIVE PHASE RETRIEVAL

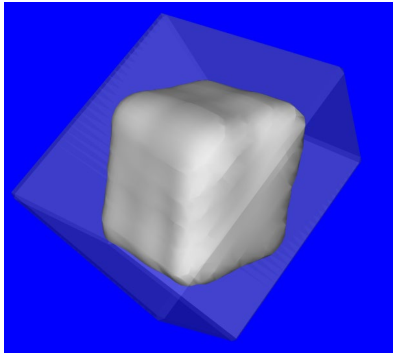
FFT

$\tau^{(n)}$

[Reciprocal Space Constraints]

Experimental Amplitudes
Oversampled, Thresholded, ZeroPadded





[Direct Space Constraints]

Evolving Support
Mixing old and new iterates.

$u^{(n)}$

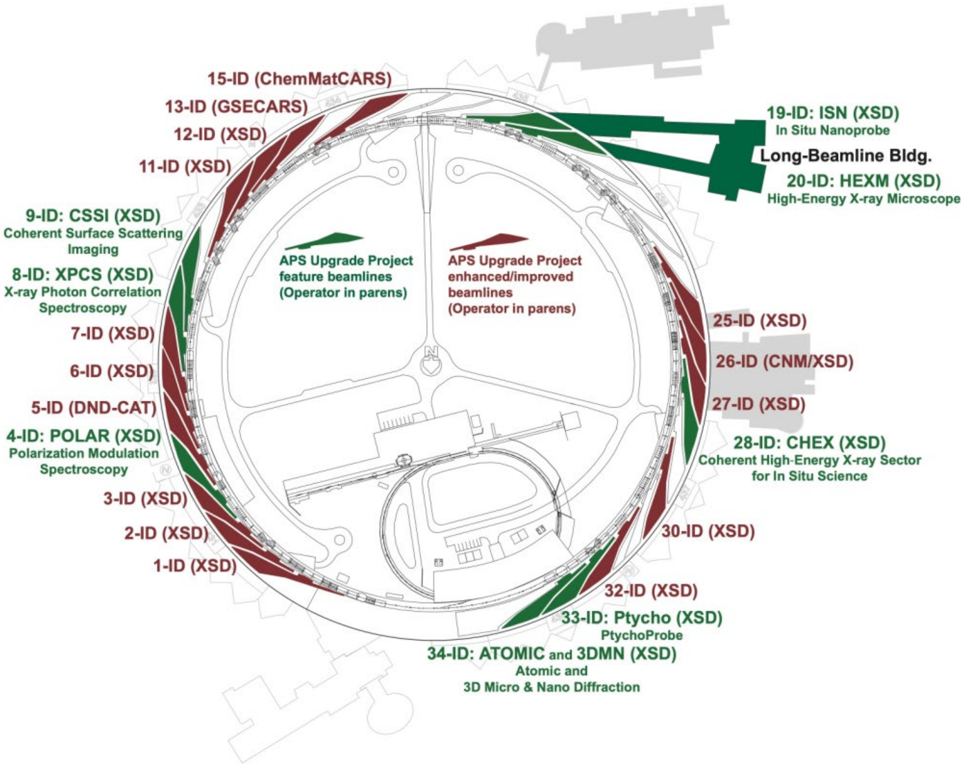
FFT^{-1}

100-1000's of iterations and multiple random starts required for high fidelity data!

- Sensitive to choice of parameters: need multiple tries and expert input

COHERENT IMAGING IN APSU

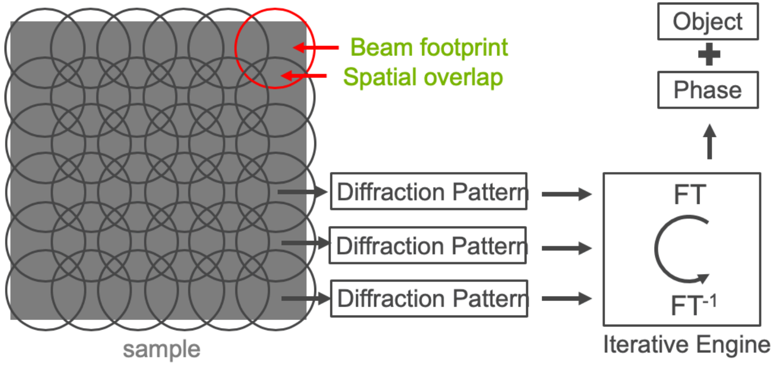
- Ptychography:
  - 5-10 beamlines
- CSSI: Coherent Surface Scattering
- ATOMIC: 3D Bragg CDI
- POLAR: Polarization Modulation Spectroscopy
- CHEX beamlines: Bragg CDI, XPCS, etc.
- HEXM: High energy 3D Bragg CDI
- S26 Nanoprobe: 3D Bragg Ptychography
- …. Many more



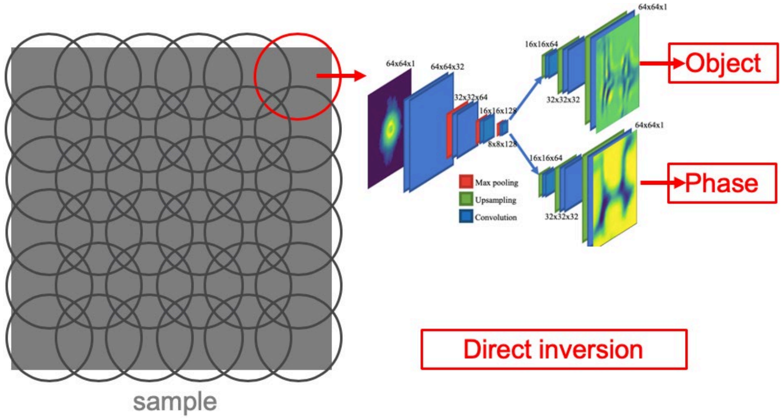
100X-1000X increase in data rates post APSU!

REINVENTING COHERENT IMAGING DATA INVERSION

Phase retrieval



AI driven



Requires HPC to keep up with experiments

PtychoNN is >100X faster
Needs 25X less data

Cherukara, Mathew J., Tao Zhou, Youssef Nashed, Pablo Enfedaque, Alex Hexemer, Ross J. Harder, and Martin V. Holt. "AI-enabled high-resolution scanning coherent diffraction imaging." *Applied Physics Letters* 117, no. 4 (2020): 044103.

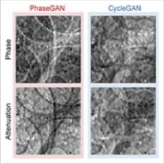
DL FOR PTYCHOGRAPHY

PtychoNet: Fast and High Quality Phase Retrieval for Ptychography

Ziqiao Guan[1]
ziguan@cs.stonybrook.edu
Esther H. R. Tsai[2]
etsai@bnl.gov
Xiaojing Huang[3]
xjhuang@bnl.gov
Kevin G. Yager[2]
kyager@bnl.gov
Hong Qin[1]
qin@cs.stonybrook.edu

[1] Department of Computer Science
Stony Brook University
Stony Brook, NY 11794, USA
[2] Center for Functional Nanomaterials
Brookhaven National Laboratory
Upton, NY 11973, USA
[3] National Synchrotron Light Source II
Brookhaven National Laboratory
Upton, NY 11973, USA

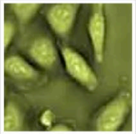
Optics Express Vol. 29, Issue 13, pp. 19593-19604 (2021) • https://doi.org/10.1364/OE.423222



PhaseGAN: a deep-learning phase-retrieval approach for unpaired datasets

Yuhe Zhang, Mike Andreas Noack, Patrik Vagovic, Kamel Fezzaa, Francisco Garcia-Moreno, Tobias Ritschel, and Pablo Villanueva-Perez

Optics Express Vol. 26, Issue 20, pp. 26470-26484 (2018) • https://doi.org/10.1364/OE.26.026470

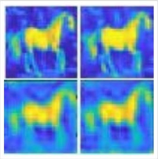


Deep learning approach for Fourier ptychography microscopy

Thanh Nguyen, Yujia Xue, Yunzhe Li, Lei Tian, and George Nehmetallah

Author Information ▾ Find other works by these authors ▾

Optics Express Vol. 28, Issue 12, pp. 17511-17520 (2020) • https://doi.org/10.1364/OE.393961



Deep neural networks in single-shot ptychography

Omri Wengrowicz, Or Peleg, Tom Zahavy, Barry Loevsky, and Oren Cohen

Author Information ▾ Find other works by these authors ▾

Deep Learning Coherent Diffractive Imaging

Dillan J. Chang[1,†], Colum M. O'Leary[1,†], Cong Su[2,3,4], Salman Kahn[2,3,4], Alex Zettl[2,3,4], Jim Ciston[5], Peter Ercius[5] and Jianwei Miao[1*]

DL MODELS CAN REPLACE ITERATIVE PHASE RETRIEVAL

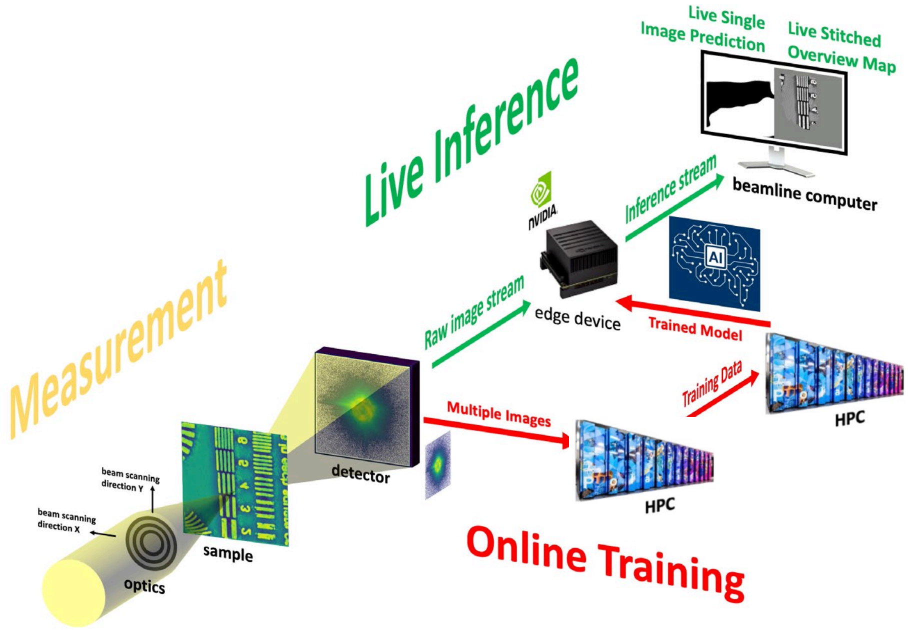
--

HOW TO IMPLEMENT THEM ON HIGH-RATE (>GB/S) INSTRUMENTS?

AI@EDGE ENABLES REAL-TIME PTYCHOGRAPHY

Train AI @ ALCF, deploy AI @ beamline

SCAN ME

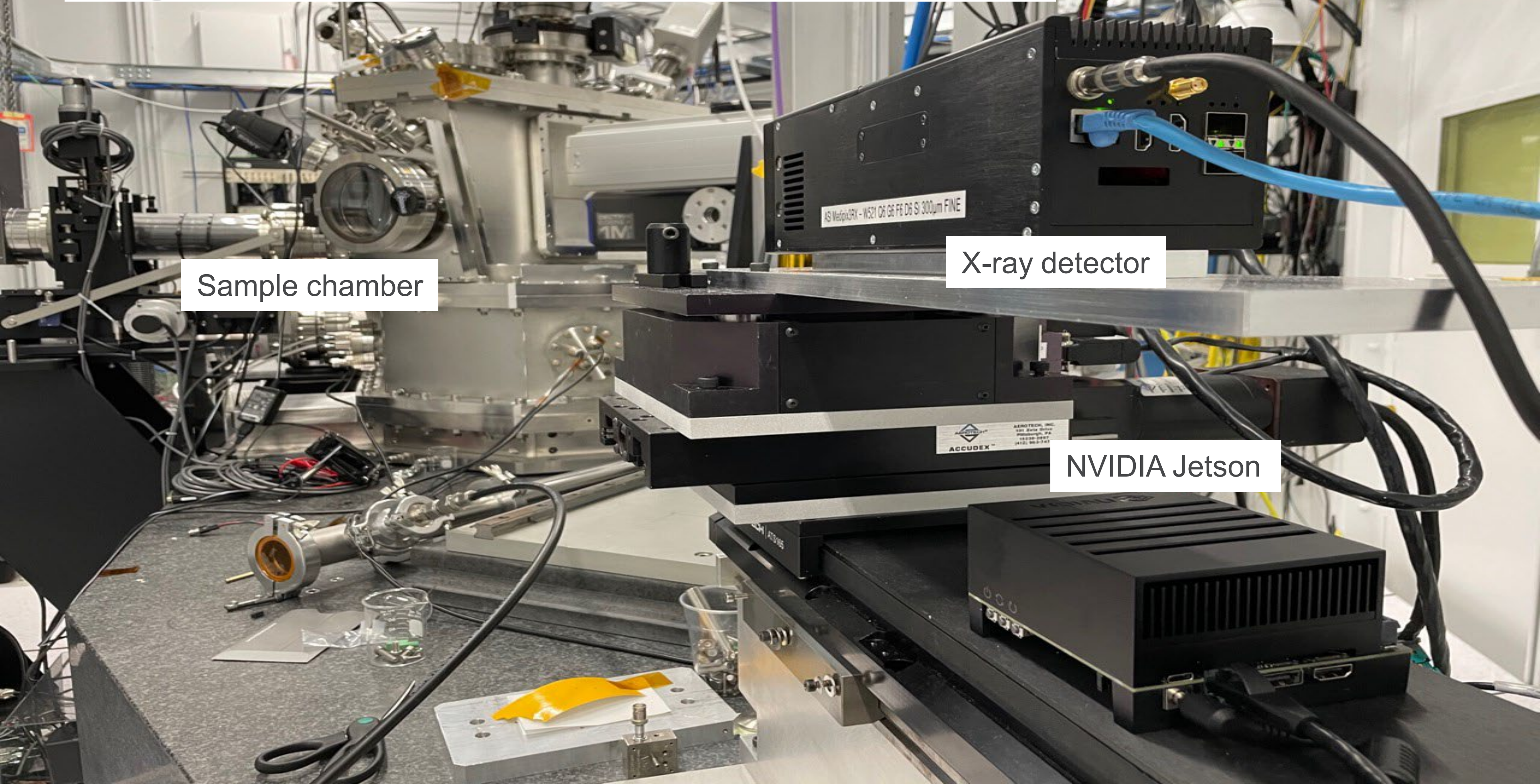


- Real-time imaging: >100X faster than phase retrieval
  - Live inference at **2 KHz** on 128x128 detector images (1 Gb/s)

Cherukara, Mathew J., Tao Zhou, Youssef Nashed, Pablo Enfedaque, Alex Hexemer, Ross J. Harder, and Martin V. Holt. "AI-enabled high-resolution scanning coherent diffraction imaging." *Applied Physics Letters* 117, no. 4 (2020): 044103.

A. V. Babu, T. Zhou, S. Kandel, T. Bicer, Z. Liu, W. Judge, D. Ching, Y. Jiang, S. Veseli, S. Henke, R. Chard, Y. Yao, E. Sirazitdinova, G. Gupta, M. V. Holt, I.T. Foster, A. Miceli and M. J. Cherukara, "Deep learning at the edge enables real-time, streaming ptychography", arXiv:2209.09408

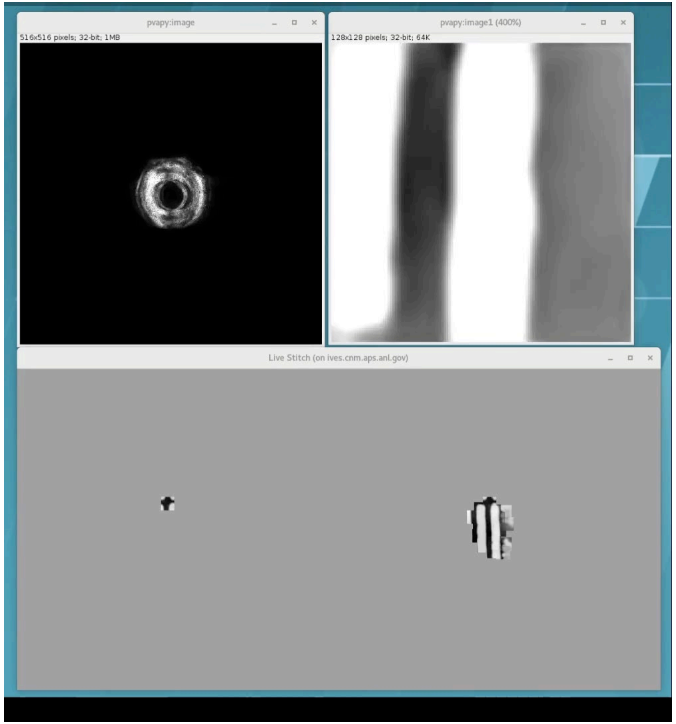
AI@EDGE FOR PTYCHOGRAPHY



AI@EDGE ENABLES REAL-TIME PTYCHOGRAPHY

Train AI @ ALCF, deploy AI @ beamline

SCAN ME

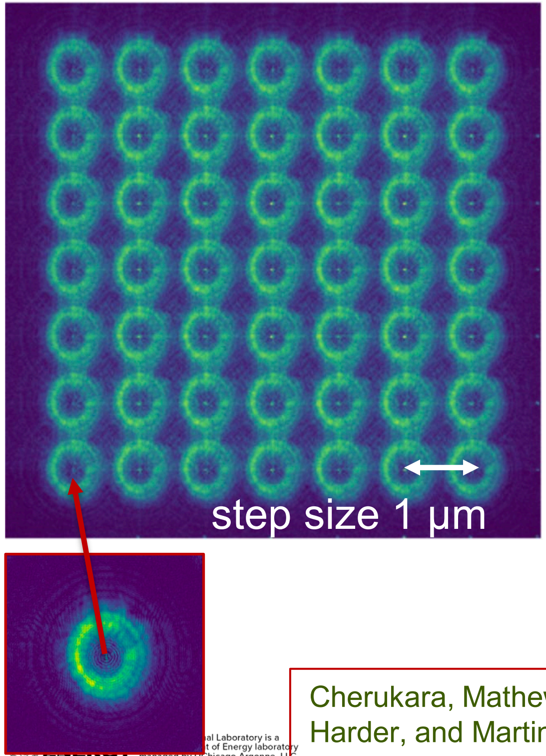


A. V. Babu, T. Zhou, S. Kandel, T. Bicer, Z. Liu, W. Judge, D. Ching, Y. Jiang, S. Veseli, S. Henke, R. Chard, Y. Yao, E. Sirazitdinova, G. Gupta, M. V. Holt, I.T. Foster, A. Miceli and M. J. Cherukara, “Deep learning at the edge enables real-time, streaming ptychography”, arXiv:2209.09408

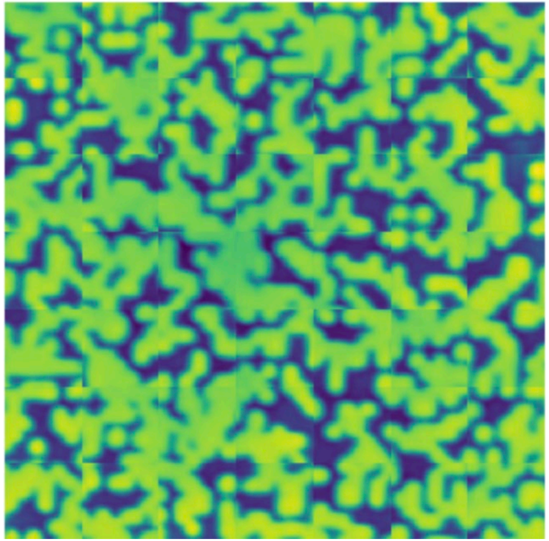
Cherukara, Mathew J., Tao Zhou, Youssef Nashed, Pablo Enfedaque, Alex Hexemer, Ross J. Harder, and Martin V. Holt. "AI-enabled high-resolution scanning coherent diffraction imaging." *Applied Physics Letters* 117, no. 4 (2020): 044103.

SPARSE DATASET : REDUCED OVERLAP

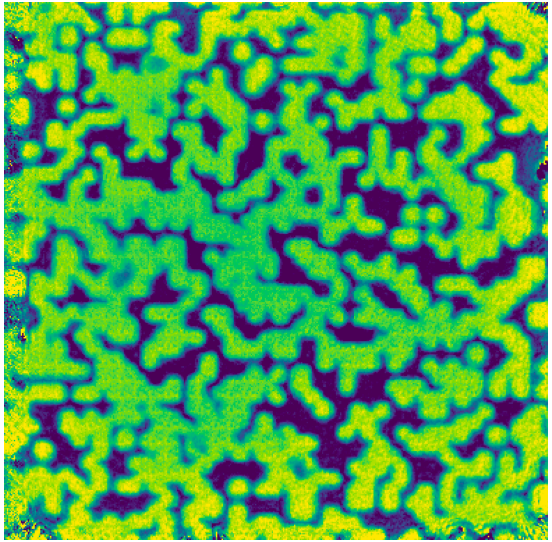
1000 nm step



NN prediction (1000 nm step)



ePIE (200 nm step)



AI ENHANCED Imaging:
- **25X less beam damage**
- **25X faster data acquisition**

Cherukara, Mathew J., Tao Zhou, Youssef Nashed, Pablo Enfedaque, Alex Hexemer, Ross J. Harder, and Martin V. Holt. "AI-enabled high-resolution scanning coherent diffraction imaging." *Applied Physics Letters* 117, no. 4 (2020): 044103.

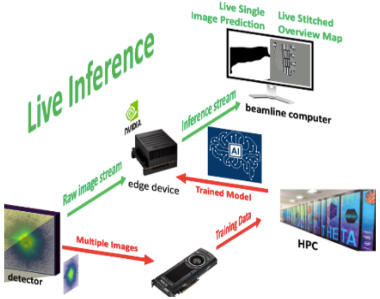
HPC+AI@EDGE TRANSFORMS EXPERIMENTAL SCIENCE

No HPC

- Reconstruction time: Days-Weeks
- Data needed: full

HPC

- Reconstruction time: Minutes-Hours
- Data needed: full



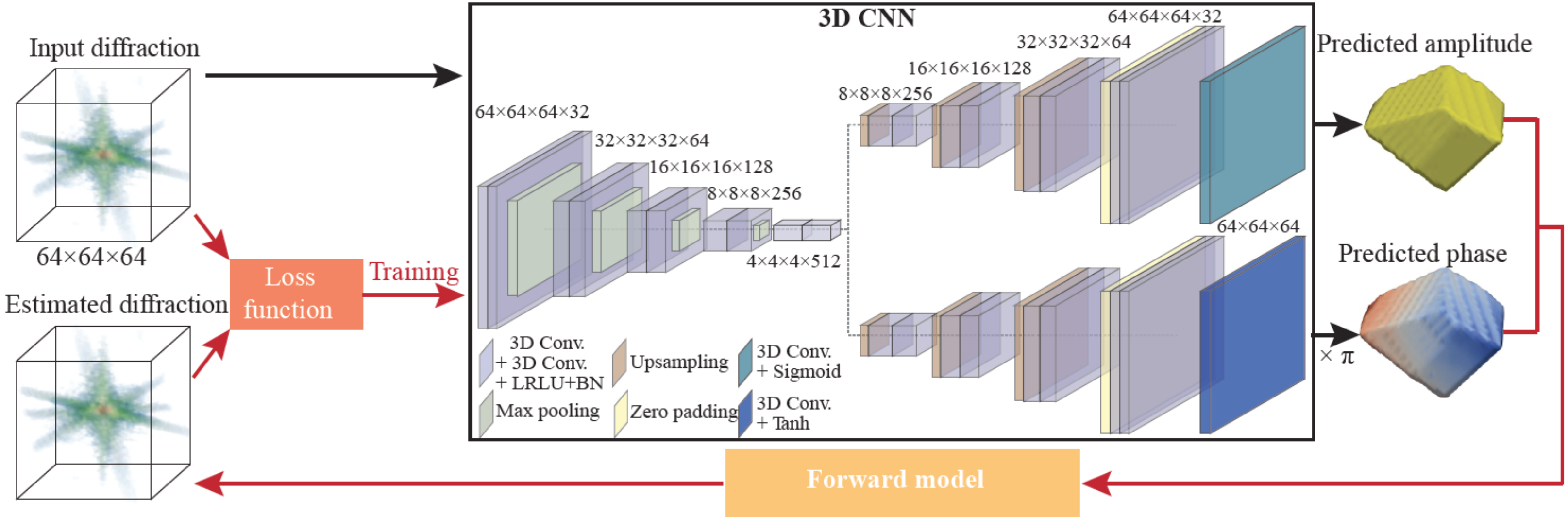
HPC+AI@Edge

- Reconstruction time: Miliseconds
- Data needed: >25X less

PHYSICS-AWARE ML IN PRODUCTION FOR BCDI

AUTOPHASENN

Unsupervised NN for 3D BCDI phase retrieval



3D convolutional neural network:

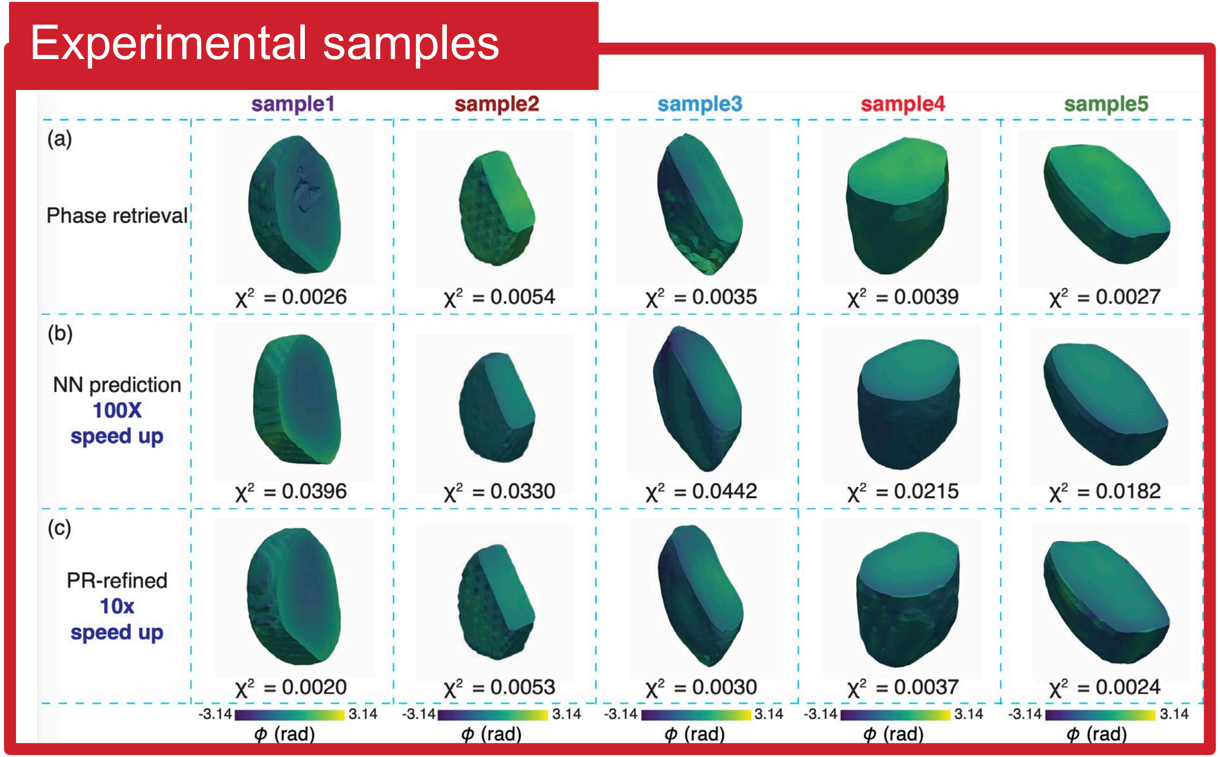
Learn the inversion from input intensity to images of object

Forward model:

Eliminate the need for ground truth image in training

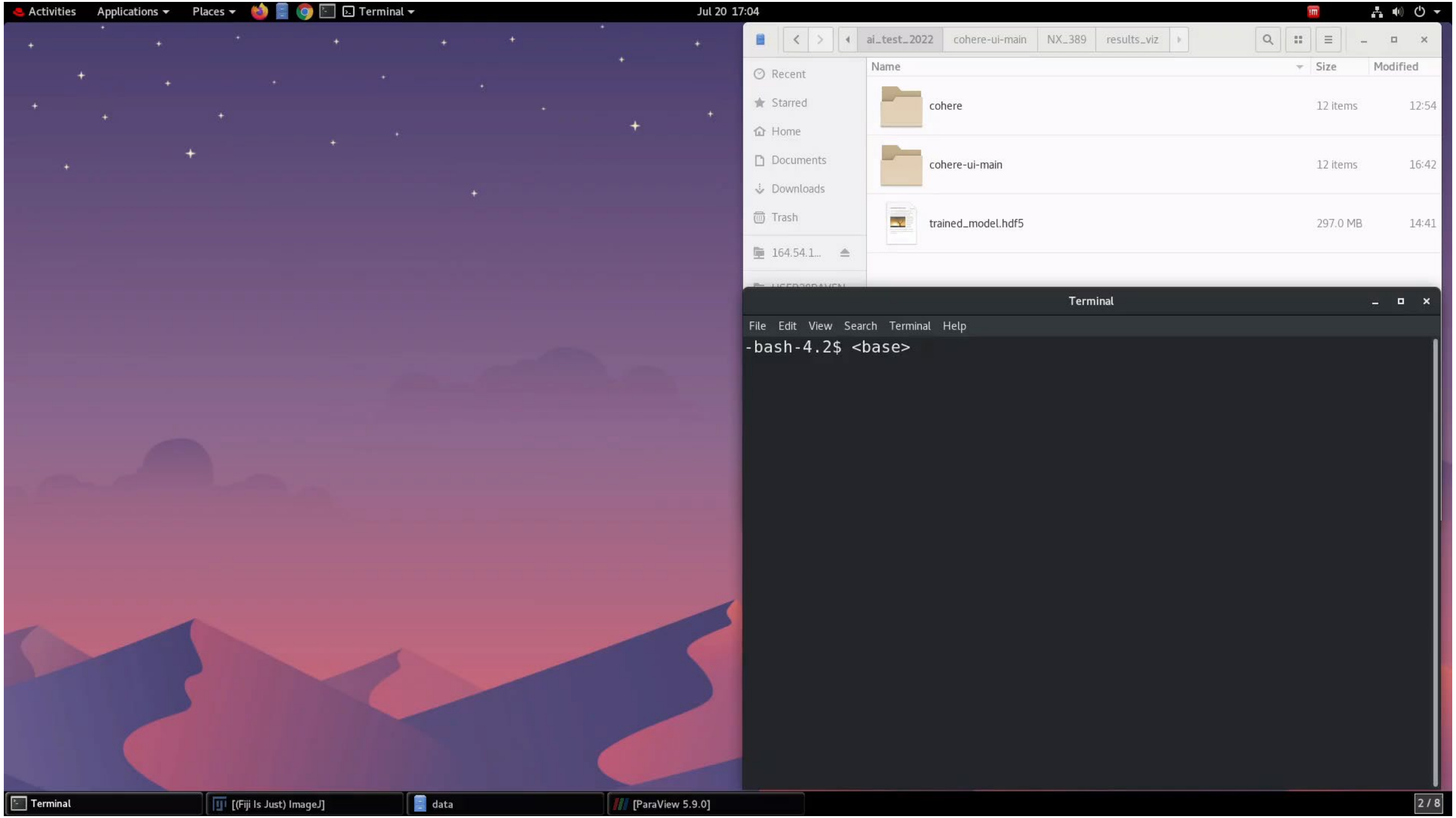
AI4Analysis

3D BCDI NN: FASTER, MORE ACCURATE



Yao, Y., Chan, H., Sankaranarayanan, S., Balaprakash, P., Harder, R. J., & Cherukara, M. J. (2022). AutoPhaseNN: unsupervised physics-aware deep learning of 3D nanoscale Bragg coherent diffraction imaging. npj Computational Materials, 8(1), 1-8.

AUTOPHASENN IN PRODUCTION



Frosik, B. and Harder, R. https://github.com/AdvancedPhotonSource/cohere

TAILORED DL SOLUTIONS FOR DIFFERENT MODALITIES

	PtychoNN	AutoPhaseNN
Training	Supervised Online	Unsupervised Offline
Size	< 1 M params	> 10 M params
Inference time	< 1 ms	< 1 s
Generalizability	None	All convex objects, weak phase

AI4STEERING

SMART DATA ACQUISITION

AI4Steering

Experiment:
- Scanning Bragg diffraction imaging (008 peak) of layered material (WSe_2)

Problem:
- Given an unknown sample, how should we acquire data to maximize information gain in minimal time?

Approach:
- Sample a few (~1%) points randomly
- Use a pre-trained NN to predict the most important points to acquire.
  - Decision is made in ~ 1s

Result:
- AI approach reconstructs image with far fewer points

TRAINING SMART ACQUISITION NETWORK

AI4Steering

- Trained on 1 image
  - 100 different masks at different sampling

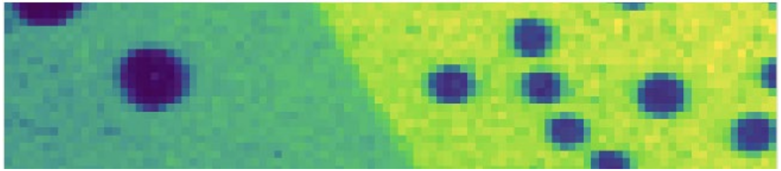
S. Kandel, T. Zhou, A. Babu, Z. Di, X. Ma, M. Holt, A. Miceli, C. Phatak, M. Cherukara, “AI-driven steering of high-resolution scanning microscopes”

Zhang, Y., Godaliyadda, G. M., Ferrier, N., Gulsoy, E. B., Bouman, C. A., & Phatak, C. (2018). Slads-net: supervised learning approach for dynamic sampling using deep neural networks. Electronic Imaging, 2018(15), 131-1.

SMART DATA ACQUISITION

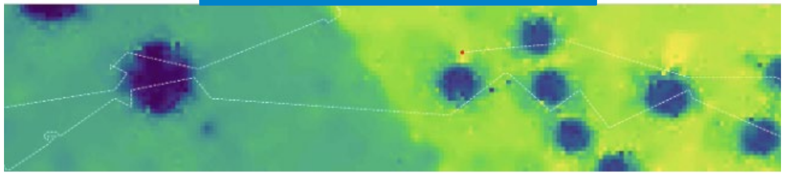
AI4Steering

Full-res image

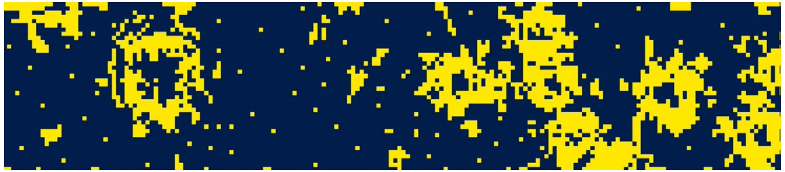


'Ground truth' : 100 nm steps

AI-guided acquisition



4.3X less points



Locations chosen by AI to scan
- Each yellow dot is a scan point

S. Kandel, T. Zhou, A. Babu, Z. Di, X. Ma, M. Holt, A. Miceli, C. Phatak, M. Cherukara, "AI-driven steering of high-resolution scanning microscopes"

Zhang, Y., Godaliyadda, G. M., Ferrier, N., Gulsoy, E. B., Bouman, C. A., & Phatak, C. (2018). Slads-net: supervised learning approach for dynamic sampling using deep neural networks. Electronic Imaging, 2018(15), 131-1.

AI@EDGE DRIVES THE EXPERIMENT

SCAN ME

AI4Steering

S. Kandel, T. Zhou, A. Babu, Z. Di, X. Ma, M. Holt, A. Miceli, C. Phatak, M. Cherukara, “AI-driven steering of high-resolution scanning microscopes”

OPEN SOURCE CODE + DATA

PtychoNN: https://github.com/mcherukara/PtychoNN

AutoPhaseNN: https://github.com/YudongYao/AutoPhaseNN

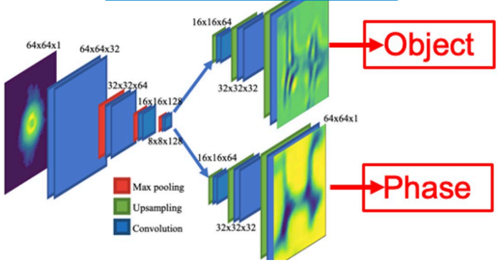
BraggNN: https://github.com/lzhengchun/BraggNN

Smart scanning: https://github.com/saugatkandel/sladsnet_new

RECAP: AI4SCIENCE

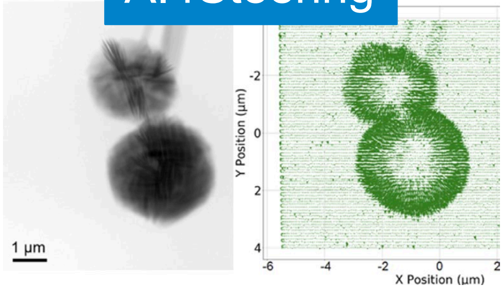
AI will be an integral part of APSU beamlines

AI4Analysis



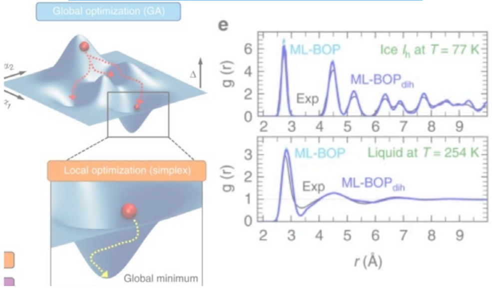
- AI@Edge: >100X faster and (sometimes) more accurate analysis
- Enables real-time analysis on Gb/s data streams

AI4Steering



- AI@Edge: Self-driving experiments & instruments:
  - maximize info gain in minimal time

AI4Knowledge



- Learn physics directly from scattering data

Now — **1-2 years** — **5-10 years**

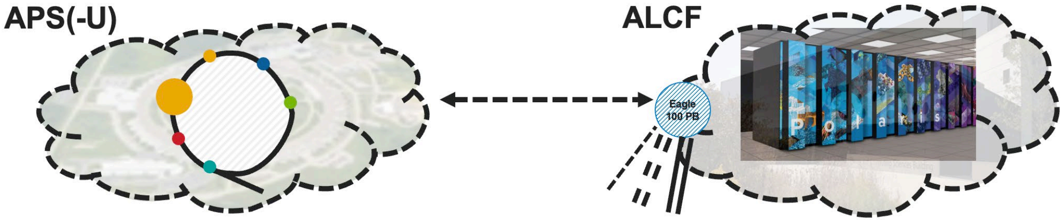
Production timeline

OUTLOOK

Outlook : 1 year

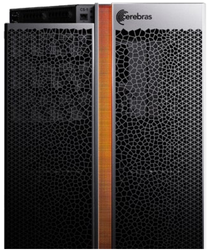
EVEN CLOSER COUPLING TO ALCF

Polaris



- Scalable software solutions for inverse problems
- Online and offline DL training at scale, deploy at edge
- Preemptable queues: On-demand HPC

AI-accelerators



	WSE-2	A100	Cerebras Advantage
Chip Size	46,225 mm2	826 mm2	56 X
Cores	850,000	6912 + 432	123X
On-chip memory	40 Gigabytes	40 Megabytes	1,000 X
Memory bandwidth	20 Petabytes/sec	155 Gigabytes/sec	12,733 X
Fabric bandwidth	220 Petabits/sec	4.8 Terabits/sec	45,833 X

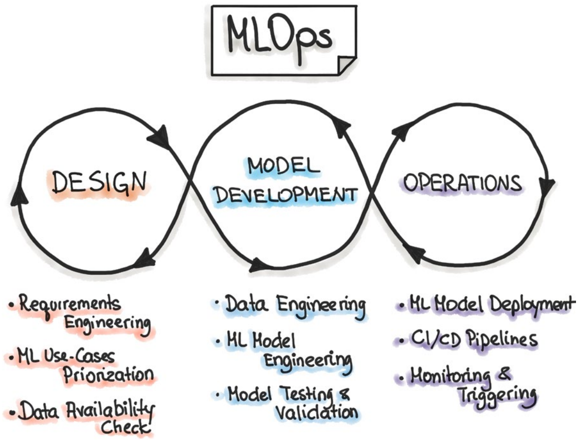
Source: https://cerebras.net/chip/

- E.g. Need to invert 3Kx3Kx3K imaging data for 3D nanoscale imaging
- ~ 1 TB of memory

More details: https://www.alcf.anl.gov/alcf-ai-testbed

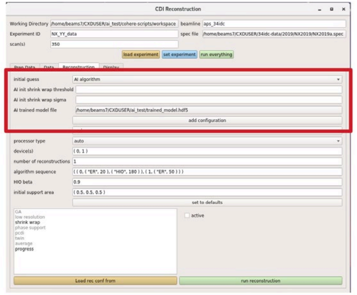
Outlook: 2-3 years

MLOPS



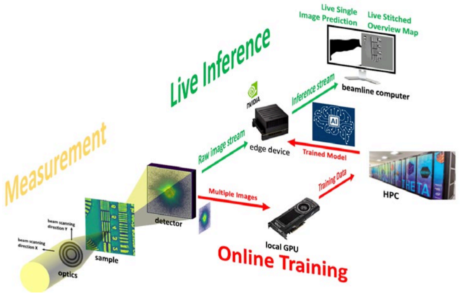
https://ml-ops.org/content/mlops-principles

User model serving



- Update cycle : ~months
- Model size: > 10 million

Edge model serving



- Update cycle : ~minutes
- Model size: <million parameters

SAGE
A Software-Defined Sensor Network
Cyberinfrastructure for AI at the Edge
www.sagecontinuum.org

Pete Beckman, Nicola Ferrier, Raj Sankaran et al @ ANL

Outlook: 5-10 years

ON-SENSOR (DETECTOR) COMPUTE

Edge (on-detector)

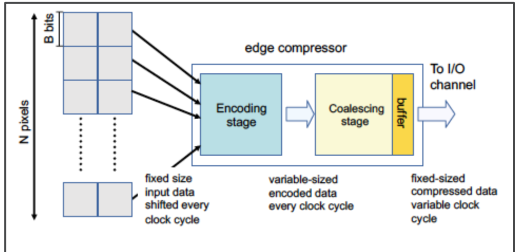


Figure 11: Basic architecture of the streaming edge compressor

How do you get ~ 1 Terabit/s off a 4 cm^2 piece of detector silicon ASIC chip?

- 256 x 256 x 16-bits x 1 MHz = 1 Tbps
- Realistic off-chip bandwidth in 65-nm is ~20 Gbps.

Future Needs

- On-chip lossless 10X data compression for ptycho, HEDM etc.
- Edge analysis and ML inference
- Both on-chip and on custom hardware very near

Hammer, M., Yoshii, K., & Miceli, A. (2021). Strategies for on-chip digital data compression for X-ray pixel detectors. *Journal of Instrumentation*, *16*(01), P01025.

OPEN QUESTIONS

(Selected) Things that we need help with

1. How do we incorporate UQ without slowing infererence?
   1. Beyond MC-dropout, ensembles or Bayesian NNs

2. Can we build foundational models for various phase retrieval applications?
   1. Across x-ray modalities?
   2. Across optical, electron, x-ray?

3. How do we abstract useful information and/or intelligently acquire only relevant information from massive datasets?
   - E.g.: mm^2 at 10 nm resolution

4. How do we effectively collate information from complementary instruments to span spatiotemporal scales and many modalities?

WE WOULD LOVE TO WORK WITH YOU!

WE HAVE DATA ☺

mcherukara@anl.gov

MJCherukara

mcherukara

POSTDOC POSITION: PHYSICS-AWARE DL

Got a PhD in optical, electron or x-ray characterization?

Apply now!

SCAN ME

1. Build physics-aware DL models
2. Train at scale on world's biggest computers
3. Deploy at cutting-edge multimillion dollar microscopes

https://www.linkedin.com/jobs/view/postdoctoral-appointee-sciml-physics-aware-deep-learning-for-x-ray-science-at-argonne-national-laboratory-3292186972