arXiv:2304.04297v1 [cs.CV] 9 Apr 2023

通过联邦资源实现实时 X 射线叠层成像数据分析的 AI辅助自动化工作流

Anakha V Babu、***Tekin Bicer***、***Saugat Kandel***、***Tao Zhou***、***Daniel J. Ching***、***Steven Henke***、***Sini˘sa Veseli***、***RyanChard***、***Antonino Miceli***、***Mathew Joseph Cherukara*** 阿贡国家实验室，美国伊利诺伊州莱蒙特

摘要

我们提出了一种端到端自动化工作流，利用大规模远程计算资源和边缘端嵌入式*GPU*平台，实现对所采集*X*射线叠层成像数据的人工智能/机器学习加速实时分析。叠层成像是一种无透镜方法，通过对来自相邻重叠扫描位置的大量衍射图样同时进行数值反演来对样品成像。这种采集方法能够利用*X*射线和电子实现纳米尺度成像，但通常需要非常大的实验数据集以及相应较长的周转时间，这可能会限制实时实验调控和低延迟监测等实验能力。在本研究中，我们介绍了一种能够自动执行叠层成像数据分析任务的软件系统。我们采用*PtychoNN*的改进版本来加速数据分析流水线。*PtychoNN*是一种基于机器学习的方法，用于解决相位恢复问题；与传统迭代方法相比，其速度提升了两个数量级。此外，我们的系统协调并重叠不同的数据分析任务，以尽量减少工作流不同阶段之间的同步开销。我们使用来自先进光子源*26ID*光束线的真实实验工作负载，以及来自阿贡领导力计算资源*ThetaGPU*集群的真实实验工作负载，对我们的工作流系统进行了评估。

引言

Neural Networks（NN）被认为是通用逼近器，并已成功解决计算机视觉、自然语言处理、自主控制以及许多其他领域中的问题 [1]。基于机器学习（ML）的逆成像问题替代方法近来受到了广泛关注，并已在医学成像、计算摄影、计算显微成像、地球物理成像及其他应用中取得了显著成功 [2]。具体而言，对于同步辐射设施，基于机器学习的重建方法已在布拉格相干衍射成像、高能衍射显微术和相干衍射成像模态中展现出在X射线成像方面的良好前景 [3–5]。这些基于机器学习的方法通常包含两个重要组成部分。首先，使用先前获取的（数据、解）对训练NN。其次，部署训练好的NN，以反演新的实验数据。这种基于NN的反演通常是单次处理，其速度可能比典型的迭代反演技术快几个数量级，从而为利用高分辨率成像数据进行实时实验反馈提供了可能性。

叠层成像是一种尤为重要的高分辨率相干衍射成像（CDI）技术，已使多种 X 射线和电子显微镜应用实现了纳米尺度成像这些应用包括集成电路成像 [6],藻类等生物样本成像 [7], 以及纳米线应变成像 [8, 9], ，甚至还包括纳米结构的原子尺度（亚埃级）成像 [10, 11]。在典型的 X 射线叠层成像实验中，聚焦的相干 X 射线束会在样品上扫描，探测器则采集大量远场强度衍射图样。CDI 实验的特性决定了相位信息会丢失，因此通常使用传统迭代方法，根据测得的强度恢复相位。为叠层成像开发的迭代重建技术通常需要大量数据和计算资源，并且需要对测量数据进行多次（数百至数千次）迭代 [12]。随着加速器技术的发展，例如阿贡光子源（APS）即将进行的升级，现代 X 射线叠层成像仪器的数据速率预计将提升数个数量级，这反过来又要求具备大规模计算资源，能够提供拍至艾 FLOPS 级的吞吐量，以执行传统的叠层成像分析任务 [13]。这使得使用传统叠层成像算法进行实时数据分析所需的计算资源变得难以承受。

在实时拼接成像重建中使用神经网络是一个颇具吸引力的方案，但这需要应对一些关键的后勤挑战。鉴于 APS-U 等下一代光源具有更高的相干性和亮度，神经网络训练必须处理海量数据，并将在远程高性能计算（HPC）资源上执行，以利用强大的计算吞吐能力实现快速训练。然而，由于数据传输所涉及的延迟，使用远程高性能计算集群进行神经网络推理无法向光束线提供实时反馈 [14]。因此，必须在边缘端或尽可能靠近数据生成源的位置执行推理，以实现实时数据分析。随着低成本嵌入式GPU 和专用人工智能/机器学习硬件加速器的普及，神经网络中的前馈计算能够以高吞吐量实现 [15]。训练好的模型可以定期部署到边缘设备上，用于低延迟操作，例如实验引导和/或监测。此外，一旦训练好的模型的重建结果达到用户质量指标，就可以释放宝贵的高性能计算资源并将其用于其他任务。这促使人们需要开发一种工作流系统，能够自动监控数据采集过程，并高效利用远程高性能计算资源，以实现更快的数据处理、机器学习训练和边缘端推理。

相关工作

多年来，人们已为叠层成像开发了许多重建算法 [16–19]。这些算法的并行化，以及在高端计算资源和加速器（如 GPU）上的高效执行，也受到了关注 [19–21]。三维叠层成像实验进一步提高了数据分析任务的计算需求；这类任务通过断层成像或层析成像技术扩展了二维叠层成像，使所采集实验数据集的规模增加了数个数量级 [12, 22]。尽管大多数三维重建任务采用单遍直接方法执行（由于这类方法的计算需求较低），但高质量的三维重建仍然依赖迭代技术。因此，人们开发了并行和分布式 [23, 24], 以内存为中心的 [25], 高性能三维重建算法 [26]和技术，这些算法和技术能够提供准实时结果 [27], [28, 29]。对于超大规模三维数据集，单遍滤波反投影方法的并行化也得到了研究 [30, 31]。然而，大规模实验数据集的分析（对于直接方法和迭代方法而言）需要使用集群和超级计算机，而这些资源通常难以交互。为了简化这些资源的使用，人们开发了多种运行时系统和工作流系统 [32–35], ; 它们的执行已在同步辐射图像分析任务中得到展示 [36–38], , 包括叠层成像工作流 [39]。尽管这些资源能够提供巨大的计算吞吐量，但由于作业调度、数据移动和资源分配所产生的开销，它们在反馈延迟方面仍然受到限制；因此，对于低延迟任务而言，最好使用靠近数据采集端的边缘设备。

人工智能/机器学习技术已被用于缓解叠层成像数据分析的计算需求并提升输出图像质量，涵盖2D [40–42] 和 3D 叠层成像 [43–45]。近年来，用于替代相位恢复算法的代理模型已展现出良好效果，且推理时间非常短 [3, 4, 46]。这些模型已部署到边缘端计算设备（例如NVIDIAJetson）上，使科学家能够获得实时反馈，并支持其他需要低延迟数据分析的下游任务 [5, 47]。然而，这些人工智能/机器学习模型的训练可能耗时较长，并且需要高端 GPU。此外，这些模型需要学习的样本和特征种类可能极其庞大，因此必须进行持续学习。

在本研究中，我们通过将人工智能/机器学习训练纳入我们的工作流系统 [39, 47], 补充了现有研究。具体而言，我们在高端计算资源上执行高延迟（连续）训练任务，而低延迟反馈操作则在光束线上使用边缘设备（NVIDIA Jetson）完成。训练任务与推理任务之间的松耦合执行，使我们能够在提供接近实时结果的同时，借助最新的人工智能/机器学习模型生成高质量输出。

目标

本研究聚焦于自动化机器学习加速的 X 射线叠层成像数据分析，以实现二维图像重建，并实时可视化由靠近数据采集设备的边缘设备生成的相位预测。工作流系统包含多个处理组件，例如，将数据从本地数据采集系统传输至 the Theta 超级计算机，启动重建流程，将结果传输至 Argonne 的 Petrel 数据分析系统，并将结果注册到 Globus 索引服务，以便进行可视化。工作流如图 1 所示。

Graphviz 远程高性能计算集群，然后启动重建流程以训练机器学习模型。训练完成后，训练好的模型会从远程高性能计算集群推送到边缘设备。我们的工作流系统协调不同的处理组件，例如数据采集系统与远程高性能计算集群之间的通信，利用可用资源在高性能计算资源中实现远程函数调用，并在不中断数据可视化的情况下执行边缘设备上的模型更新。在底层，我们的系统处理复杂任务，例如作业调度、任务同步、资源认证和管理。

方法与实验设计

Graphviz 主要包含四个不同阶段：(i) 数据采集，(ii) 相位恢复，(iii) 神经网络训练，以及 (iv) 实时 AI/ML 推理。我们将在以下小节中逐一说明这些阶段。

数据采集

实验在 APS 的硬X射线纳米探针光束线（26ID）上进行。光子能量为 10.4 keV。样品是一层厚度为 1.5 μm、蚀刻成随机图样的钨膜。使用像素尺寸为 55 μm、位于样品下游 1.55 m 处的 Medipix3 探测器采集 X 射线叠层成像图像和衍射图样。聚焦采用直径为 160 μm、最外层区带宽度为 30 nm 的菲涅耳波带片。光学元件被有意离焦 -150μm，以产生 FWHM 约为 600 nm 的过聚焦光束。每次叠层成像扫描均沿优化的螺旋路径进行，包含 963 个数据点，步长为 50 nm。

工作流系统组件

我们使用 Globus 工具实现数据分析流水线；即 Automate [48], funcX [49], 和 Transfer [50] 服务。清单 ***1***。*Globus Automate* 流程定义示例。

```
{"StartAt": "Init",          -> Beginning state
 "States": {                 -> State definitions
   "Init": {                 -> Init state
     "Type": "Action",
     "ActionUrl": "https://.../transfer", ......
     "Parameters": {
       "src.$": "...", "dest.$": "...",
       "items": [...]
     }, ...
     "Next": "Analyze"
   }, ......                 -> See Listing 3
```

Globus自动化服务帮助最终用户使用基于 JSON 的流定义语言实现工作流的步骤（或阶段）。这些阶段由操作组成，统称为流，包括数据传输和远程函数执行。这些操作通常通过与其他服务的应用程序编程接口（API）交互来触发或执行。清单 1给出了一个流定义示例。StartAt 的值为工作流的初始状态，指向 Init。后续的States 条目包含工作流状态的定义，其中 Init状态在开头定义。Init状态执行数据传输操作，源、目标和项目（文件夹或文件）作为参数传入。Next 键指向工作流中的下一个状态，

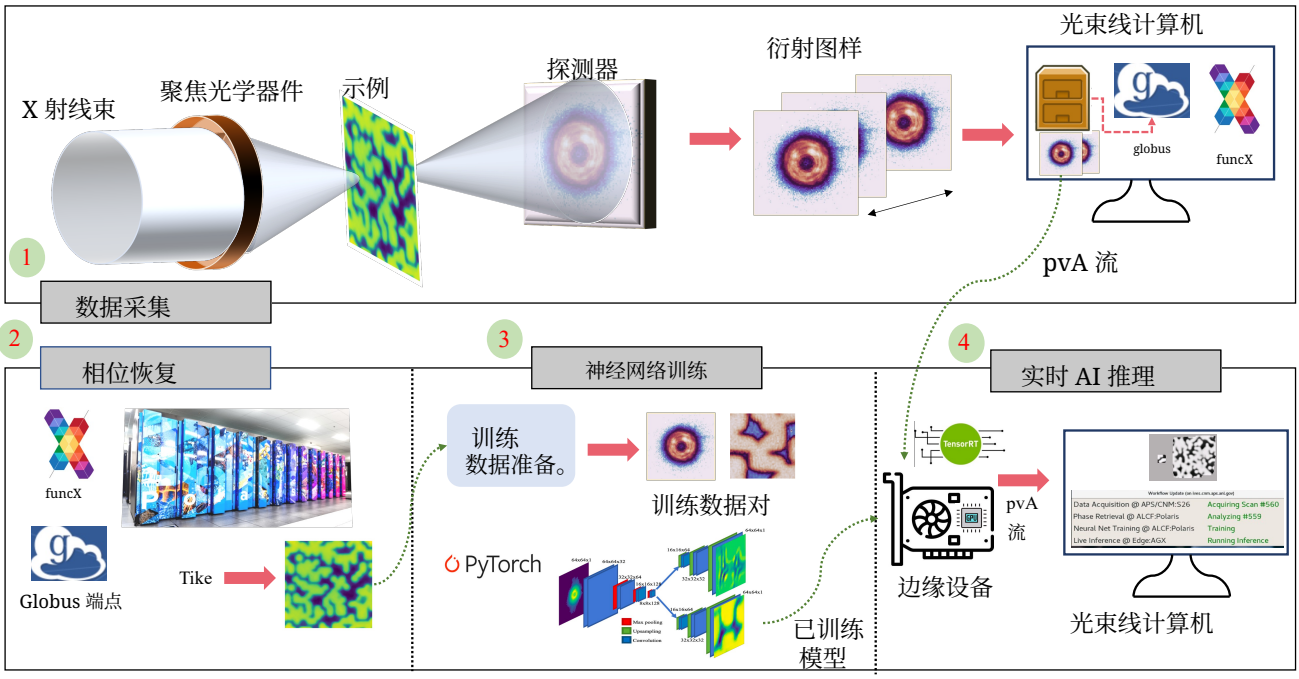

图 *1*。 一种利用联邦资源加速*X*射线叠层成像数据分析的自动化工作流。步骤*(1)* 表示叠层成像实验期间的数据采集过程。采集到的衍射图样被传输至高性能计算资源，并输入叠层成像重建任务，如步骤 *(2)* 所示。随后，重建得到的二维实空间图像用于训练 *PtychoNN*模型（步骤*(3)*）。模型训练是一个持续的过程，即随着测量数据的获取，工作流系统持续重建二维实空间图像，并将其用于训练。训练完成的最新模型会定期部署到光束线上，以进行实时推理，如步骤*(4)* 所示。我们使用 *Globus* 工具和服务执行数据传输、远程函数调用、身份验证以及任务和资源管理。

即“分析”（参见清单 3）。

Globus Automate 使用 Globus Auth 提供的身份验证基础设施，因此，与 Globus Auth 集成的服务可以受益于单点登录身份验证层 [51]。此功能显著简化了多种操作之间的交互，例如数据传输管理、与存储系统的交互以及对计算资源的使用。

清单 *2*
. 向 *funcX* 服务注册函数及示例用法。

```
from funcx.sdk.client import FuncXClient
# Python function to be run
def ptycho_func(**data):  ...

main(): ...
 fxc = FuncXClient() -> Get funcX client to
 ...                    interact with service
 # Serialize the local python function and
 # store it at funcX cloud service
 fxid = fxc.register_function(ptycho_func)
 fxep = getEndPointAddress()
 ...
 # Execute ptycho_func at fxep endpoint
 fxc.run(params,            -> Param. to pass
         endpoint_id=fxep, -> Remote resource
         function_id=fxid) -> Func. to run
```

我们使用 funcX 服务执行远程函数。使用 funcX 远程执行 Python 函数主要包含两个步骤：（1）将用户定义的 Python 函数注册到 funcX 云服务；（2）在目标计算资源上执行函数。步骤（2）需要 funcX 云服务与端点交互，因为用户定义的函数需要经过序列化、传输和调度，才能执行。这些步骤如清单 2 所示。具体而言，即 fxc.register function(...) 和 fxc.run(...)。清单 2 中的 funcX 函数用于注册和触发用户定义的 Python 函数（ptycho func(...)）。funcX 服务通过函数 ID 和端点 ID 参数，分别确定要运行哪个函数以及在哪里运行。注册的 funcX 函数可以在 Globus 自动化流程中调用，如清单 3 所示。funcX 端点通过系统的作业调度器管理大规模计算资源。

列表 *3.* 流定义中 *Globus Automate funcX* 操作的实现。

```
...                          -> See Listing 1
"Analyze": {                 -> Another state def.
  "Type": "Action", ...
  "ActionUrl": "https://automate.funcx.org",
  "Parameters": {
    "tasks": [{ "endpoint.$": "...",
      "function.$": "...",
      "payload.$": "..."
    }] },
  "Next": "Fin"              -> Next state
},
"Fin": { ...                 -> Final state
  "End": True
}, #Fin } #States, }         -> End of state defs.
```

工作流数据传输操作由 Globus Transfer 服务执行。与 funcX 服务类似，传输服务使用与端点交互的云服务来处理数据移动操作。数据传输操作具有容错能力，能够处理超大规模的数据移动。

我们的工作流系统会自动将采集的叠层成像数据从光束线传输到计算集群，启动重建任务，触发模型训练，并

使用 Globus 服务（在不同设施之间）、工作流脚本（在计算集群上）和 pvAccess（在光束线上）将最新模型部署到边缘端。Globus 服务与 Globus Auth 无缝集成，可安全访问大规模存储和计算资源。

相位恢复

图 1 展示了用于采集衍射图样的典型叠层成像实验装置。在实验过程中，会采集许多此类衍射图样，最终产生 TB 级测量数据。叠层成像重建过程旨在恢复相位信息，并根据衍射图样重建二维实空间图像。数据采集过程可用 $\psi^{i+1} = F(d, p\ \psi^i)$ 公式建模。这里，d 对应于衍射图样（探测器上的测量数据）及其对应的扫描位置，p是探针信息，而 ψ^i 是成像样品 [52] 的（初始）估计值。叠层成像重建过程尝试迭代收敛出一组与测量数据（d）一致的 ψ^{i+1} 值，同时求解相位恢复问题。

我们使用 Tike（版本 0.22） [53] 来实现相干衍射成像重建操作。Tike 提供多种迭代求解器，包括共轭梯度、差分映射、最小二乘法和正则化相干衍射成像迭代引擎（rPIE）；在工作流执行期间，我们使用了 rPIE 求解器。此外，这些求解器支持并行化，可在多个 GPU 和节点上高效执行 [52, 54]。我们通过 funcX 服务启动重建任务，该服务调用外部 Tike 脚本，并传入应用程序参数，例如迭代次数和探针数量、初始估计值，以及写入重建图像的频率。

神经网络训练

在本研究中，我们使用了略作修改的 PtychoNN（2.0 版），这是一种全卷积神经网络，[4] 中讨论了如何利用它在边缘端进行实时 AI 推理。我们采用持续训练方案，随着实验的推进不断增加训练数据（将新采集的衍射图样纳入训练数据集）。因此，每轮实验都无需从头训练模型，而是从先前保存的、验证损失最低的模型检查点继续训练。预留 10% 的训练数据集用于验证后，我们使用 PyTorch 框架训练神经网络 50 个 epoch，并采用循环学习率策略。我们使用自适应矩估计（ADAM）优化器更新模型的可学习参数，以最小化参考标签（通过相位恢复获得的重建结果）与神经网络输出之间的平均绝对误差（MAE）。50 个 epoch 结束时验证损失最低的训练好的模型，由工作流系统部署到光束线上（Jetson AGX 设备 [55]）。

实时 *AI/ML* 推理

实时 AI 推理由 NVIDIA 的 AGX Xavier [55]，嵌入式平台处理，如步骤 (4) 和图 1 所示。AGX Xavier（Jetson 系列）的计算能力为 32 TOPS，专为加速边缘端使用场景中的 ML 工作负载而设计。该设备在物理位置上靠近数据源或同步辐射光束线，并托管一个 EPICS [56]pvAccess 客户端，用于处理通过 pvAccess 协议 [57]从光束线计算机流式传输的衍射图样。如数据采集阶段所述，衍射图样在保存到文件系统的同时，还会作为 pvAccess 数据流从光束线计算机通过网络并发传输。边缘设备接收到的衍射数据由 pvAccess 客户端处理，以获得二维图像，然后馈送至推理引擎以预测相位。推理引擎使用来自高性能计算的最新训练好的模型，并将其转换为 ONNX 格式，这是一种与硬件无关的中间表示。我们使用 NVIDIA 提供的 TensorRT API [58], 库，在 NVIDIA 嵌入式 GPU平台上以更低的延迟执行推理。我们使用的批大小为 1，神经网络模型得到的相位预测结果会作为 pvAccess 数据流实时传输到光束线计算机。样品的实时单帧预测结果以及整体拼接相位图也会显示在用户界面中。使用 AGX Xavier 可实现最高 100 fps 的帧率，因为光束线计算机受限于 1 Gbps 的网络带宽，适用于 512×512 像素。

当探测器以 2 kHz 的峰值采集速率运行时，我们还扩展了边缘端实时流传输的能力。

结果与讨论

我们的工作流系统实现了同步辐射光束线与计算集群之间的无缝集成。具体而言，我们在 APS 的 26ID光束线上测试了该系统，并在 ALCF 的 ThetaGPU集群上执行重建和模型训练。训练好的模型会定期部署到位于光束线的 NVIDIA Jetson 边缘设备上。

我们的实验表明，使用 Jetson AGX Xavier 的边缘计算能力可实现高达 200 fps 的实时反馈，即每秒接收并推理 200 个衍射图样，同时在 ALCF 的 ThetaGPU集群上并行执行重建和机器学习模型训练。我们使用了 Tike 工具包（其中包含经过 GPU 优化的重建求解器），在 ThetaGPU 上高效执行相位恢复。Tike 求解器通过 funcX 调用；funcX 还负责 ThetaGPU集群上的任务和资源管理。

我们此前还在一次相干衍射成像实验中演示了该工作流，实验使用了 ALCF 的 ThetaGPU集群，最多动用了 64 块 A100 GPU（8 个节点），并同时运行了 168 个重建工作流。与在光束线上使用单块 A100 相比，该工作流可实现 29.6 倍的加速，计算扩展效率超过 81%（如 [39]所示）。新的工作流系统在我们此前工作的基础上扩展了人工智能/机器学习功能，进一步缩短了处理延迟。如果需要在边缘端实现更高的帧率，可以使用多个边缘设备或其他先进的 NVIDIA GPU 执行推理操作。我们的实验表明，使用 2 块 RTX A6000 GPU，推理速率最高可达 8000 fps。

我们的工作流系统通过 Globus 工具和服务提供跨设施集成的基本构建模块。这使得工作流可以轻松扩展，添加更多步骤（作为 Globus Automate 中的新状态）。同样，也可以利用这些构建模块轻松实现其他数据分析工作流 [32]。

致谢

在先进光子源、阿贡领导计算设施和纳米尺度材料中心开展的工作均获得了美国能源部科学办公室基础能源科学办公室的支持；这三者都是美国能源部科学办公室用户设施，资助依据为合同编号 DE-AC02-06CH11357。我们还感谢 ANL 实验室定向研究与开发项目 2023-0104 提供的支持。

参考文献

[1] K. Hornik, M. Stinchcombe, H. White 和 P. Auer。“多层前馈网络的逼近能力”。载于：Neural Networks 4.2（1991年），第251–257页。ISSN：0893-6080。[2] G. Ongie 等。“成像逆问题的深度学习技术”。载于：IEEE信息论选定领域期刊 1.1（2020年），第39–56页。[3] Y.Yao 等。“AutoPhaseNN：无监督的物理感知深度学习三维纳米尺度布拉格相干衍射成像”。载于：npjComputationalMaterials 8.1（2022），第1–8页。[4] M.J. Cherukara 等。“AI赋能的高分辨率扫描相干衍射成像”。载于：Applied PhysicsLetters 117.4（2020年），第044103页。[5] Z.Liu 等。“BraggNN：使用深度学习进行快速X射线布拉格峰分析”。载于：IUCrJ9.1（2022年1月）。[6] M. Holler 等。“集成电路的高分辨率无损三维成像”。载于：Nature 543.7645（2017），第402–406页。[7] J.Deng 等。“绿藻的同步冷冻X射线叠层成像与荧光显微术”。载于：美国国家科学院院刊 112.8（2015），第2314–2319页。[8] S. O.Hruszkewycz等。“通过单角度布拉格叠层成像实现高分辨率三维结构显微术”。载于：Nature Materials 16.2（2017年2月），第244–251页。ISSN：1476-4660。[9] M.O.Hill等。“利用相干X射线多角度布拉格投影叠层成像测量单根铟镓砷纳米线中的三维应变和结构缺陷”。载于：Nano Letters 18.2（2018年）。PMID：29345956，第811–819页。[10] Y.Jiang等。“二维材料的电子叠层成像，达到深亚埃¨分辨率”。载于：Nature559.7714（2018年7月），第343–349页。ISSN：1476-4687。[11] Z. Chen 等。“电子叠层成像达到由晶格振动设定的原子分辨率极限”。载于：Science 372.6544（2021年），第826–831页。[12] APS升级。 https://www.aps.anl.gov/APS-Upgrade。[访问日期：5月 2021]。[13] APS 科学计算战略。访问日期：08-5-2022。[14] A. Vasanthakumaribabu。“仿生网络的高效硬件实现”。博士论文。新泽西理工学院，2020年。[15] A. Reuther等。“AI加速器调查与趋势”。载于：CoRR abs/2109.08957（2021年）。[16] V.Nikitin 等。“通过贝叶斯重建实现三维对象的光子受限叠层成像”。载于：OSAContinuum 2.10（2019），第2948–2968页。[17] S.Marchesini 等。“SHARP：基于分布式GPU的叠层成像求解器”。载于：Journalof applied crystallography 49.4（2016年），第1245–1252页。[18] S. Aslan等。“采用可调学习先验的鲁棒叠层成像断层扫描分布式优化”。载于：arXiv 预印本 arXiv:2009.09498（2020年）。[19] D. A. Shapiro 等。“利用先进光源的nanosurveyor 仪器对纳米材料进行叠层成像”。载于：Journal ofPhysics: ConferenceSeries。第849卷。第1期。IOPPublishing，2017年，第012028页。[20] X. Wang 等。“用于并行且内存高效叠层成像重建的图像梯度分解”。载于：2022 SC22：高性能计算国际会议。计算、网络、存储与分析（SC）。IEEE 计算机学会。2022年，第 86–98 页。[21] Y. S. G. Nashed 等。“并行叠层成像重建”。载于：《Optics Express》22.26（2014年），第 32082–32097 页。[22] 情报高级研究计划局。《各种新兴纳米电子器件的快速分析》。 https://www.iarpa.gov/index.php/research-programs/raven [访问日期：5月 2021]。[23] T. Bicer 等。“Trace：面向大规模数据集的高吞吐量断层重建引擎”。载于：《Advanced Structural and Chemical Imaging》3.1（2017年），第 1–10 页。[24] T. Bicer 等。“通过大规模并行化实现快速断层图像重建”。载于：《欧洲并行处理会议》。Springer，2015年，第 289–302 页。[25] M. Hidayeto˘glu 等。“MemXCT：采用大规模并行化的以内存为中心的 X 射线 CT 重建”。载于：《国际高性能计算、网络、存储与分析会议》。2019年，第 1–56 页。[26] S. Aslan 等。“通过交替方向乘子法进行联合叠层成像断层扫描重建”。载于：《Optics Express》27.6（2019年），第 9128–9143 页。[27] T. Bicer 等。“利用流式滑动子集重建动态特征的断层图像”。载于：2020 IEEE/ACM 第二届极大规模实验在环计算年度研讨会（XLOOP）。IEEE，2020年，第 8–15 页。[28] X. Wang 等。“大规模并行三维图像重建”。载于：《国际高性能计算、网络、存储与分析会议论文集》。2017年，第 1–12 页。[29] M. Hidayetoglu 等。“Petascale XCT：多 GPU 节点上采用分层通信的三维图像重建”。载于：SC20：国际高性能计算、网络、存储与分析会议。IEEE 计算机学会，2020年，第 510–522 页。[30] P. Chen 等。“IFDK：可扩展的即时高分辨率图像重建框架”。载于：《国际高性能计算、网络、存储与分析会议论文集》。2019年，第 1–24 页。[31] P. Chen 等。“用于锥束 CT 重建的可扩展 FBP 分解”。载于：《国际高性能计算、网络、存储与分析会议论文集》。2021年，第 1–16 页。[32] R. Vescovi 等。“连接科学仪器与高性能计算：模式、技术与经验”。载于：《技术、经验》（）。[33] M. Salim 等。“Balsam：超级计算机上的近实时实验数据分析”。载于：2019 IEEE/ACM 第一届大规模实验在环计算年度研讨会（XLOOP）。2019年，第 26–31 页。[34] M. Basham 等。“数据分析工作台（DAWN）”。载于：《Journal of Synchrotron Radiation》22.3（2015年5月），第 853–858 页。[35] J. Donatelli 等。“Camera：能源研究应用高级数学中心”。载于：《Synchrotron Radiation News》28.2（2015年），第 4–9 页。[36] M. Klein 等。“面向实验数据驱动工作流的交互式超级计算”。载于：《高性能计算工具与技术》。Springer，2019年，第 164–178 页。[37] T. Peterka 等。“一种用于并行图像合成应用的可配置算法”。载于：《高性能计算、网络、存储与分析会议论文集》。IEEE，2009年，第 1–10 页。[38] T. Bicer 等。“同步辐射光源实验的实时数据分析与自主控制”。载于：IEEE 第 13 届 e-Science 国际会议（e-Science）。IEEE，2017年，第 59–68 页。[39] T. Bicer 等。“利用联邦设施进行高性能叠层成像重建”。载于：《通过实验、大数据以及建模与仿真的集成推动科学与工程发现》。Cham：Springer International Publishing，2022年，第 173–189 页。ISBN：978-3-030-96498-6。

[40] S. Aslan 等。“结合深度生成先验的叠层成像断层扫描”。载于：《Machine Learning: Science and Technology》2.4（2021年），第 045017 页。[41] S. V. Kalinin 等。《扫描探针显微镜中的大数据、深度数据与智能数据》。2016年。[42] S. Venkatakrishnan 等。“科学 CT 仪器的算法驱动进展：从基于模型的方法到基于深度学习的方法”。载于：arXiv 预印本 arXiv:2104.08228（2021年）。[43] Z. Liu 等。“深度学习加速光源实验”。载于：2019 IEEE/ACM 第三届超级计算机深度学习研讨会（DLS）。IEEE，2019年，第 20–28 页。[44] Z. Wu 等。“利用混合剂量测量进行基于深度学习的低剂量断层成像重建”。载于：arXiv 预印本 arXiv:2009.13589（2020年）。[45] Z. Liu 等。“TomoGAN：利用生成对抗网络进行低剂量同步辐射 X 射线断层成像：讨论 ”。载于：《 JOSA A 》 37.3 （ 2020 年），第 422–434 页。[46] M. J. Cherukara 等。“利用深度生成网络进行实时相干衍射反演”。载于：《Scientific Reports》8.1（2018年），第 16520 页。[47] A. V. Babu 等。《边缘端深度学习支持实时流式叠层成像》。2022年。[48] R. Chard 等。“Globus 自动化服务：跨时空连续体的研究流程自动化”。载于：《Future Generation Computer Systems》（2023年）。[49] R. Chard 等。“FuncX：面向科学的联邦函数服务架构”。载于：《第 29 届国际高性能并行与分布式计算研讨会论文集》。2020年，第 65–76 页。[50] K. Chard 等。“大数据的高效安全传输、同步与共享”。载于：《IEEE Cloud Computing》1.3（2014年），第 46–55 页。[51] S. Tuecke 等。“Globus Auth：研究身份与访问管理平台”。载于：2016 IEEE 第 12 届 e-Science 国际会议（e-Science）。IEEE，2016年，第 203–212 页。[52] X. Yu 等。“大规模叠层成像数据的可扩展、高精度多 GPU 图像重建”。载于：《Scientific Reports》12.1（2022年），第 5334 页。[53] 阿贡国家实验室。《 Tike：用于从叠层成像数据重建的断层成像工具箱》。在线访问日期：1月-。[54] X. Yu 等。“面向拓扑优化的多 GPU 叠层成像重建”。载于：《国际超级计算会议》。ICS ’21。美国，虚拟会议：ACM，2021年，第 354–366 页。ISBN：9781450383356。[55] NVIDIA Jetson AGX Xavier 开发套件。https://developer.nvidia.com/embedded/jetson-agx-xavier-developer-kit [16 2021]。在线；访问日期：-11月-。[56] 实验物理与工业控制系统（EPICS）。https://epics-controls.org。[57] pvAccess。https://epics-controls.org/resources-and-support/documents/pvaccess/。[58] I. J. utkebohle、BWorld 机器人控制软件。http://aiweb.techfak.uni-bielefeld.de/content/bworld-robot-control-software/ [19 2008] 2008。在线；访问日期：-7月-。

作者简介

Anakha V Babu 于*2010*年获得印度喀拉拉大学电子与通信工程技术学士学位，并于*2016*年获得印度孟买印度理工学院孟买分校微电子与超大规模集成电路设计技术硕士学位。她于*2020*年在美国新泽西州新泽西理工学院获得电气工程博士学位。她曾在阿贡国家实验室（*ANL*）*X* 射线科学部担任博士后研究人员，专注于在边缘端实现实时叠层成像数据分析并使其自动化。她目前担任硬件电子工程师，任职于阿贡国家实验室（*ANL*）的 *APS* 光束线控制与数据采集部门（*BCDA*）。

Tekin Bicer 是 *ANL* 数据科学与学习部的一名计算机科学家。他专长于高性能计算、大规模并行与分布式系统以及人工智能/机器学习方法，尤其专注于大规模*X*射线图像分析问题。*2014*年，他在俄亥俄州立大学计算机科学与工程系获得博士学位，期间研究了能够利用高端集群和云计算资源解决数据管理与分析问题的大规模计算系统。

Saugat Kandel 于*2008*年获得阿默斯特学院物理学文学学士学位，并于*2021*年获得西北大学应用物理学博士学位。此后，他一直在*ANL*的*APS*计算*X*射线科学组从事叠层成像的人工智能与优化方法以及自主实验方法研究。

TaoZhou 是 *ANL* 纳米材料中心的助理科学家，同时也是 *APS* 硬*X*射线纳米探针光束线的光束线科学家。他于 *2016*年 获得格勒诺布尔阿尔卑斯大学物理学博士学位。他的主要研究方向是开发由高级计算和机器学习方法辅助的新型 *X* 射线衍射成像技术。

Daniel J. Ching 于*2018*年获得俄勒冈州立大学材料科学博士学位。此后，他一直在*ANL*从事断层成像和叠层成像领域的计算成像研究。

Steven Henke 于*2013*年获得佛罗里达州立大学计算科学博士学位，并获得威斯康星大学物理学理学学士学位。自*2021*年起，他担任*APS*科学软件工程部门的数据科学家。

Siniˇsa Veseli 于*1992*年获得克罗地亚萨格勒布大学理学学士学位，并于*1996*年获得威斯康星大学麦迪逊分校物理学博士学位。他目前是*APS*科学软件工程与数据管理组的成员，负责*APS* 升级的控制数据采集软件，并参与*APS*数据管理项目。

Ryan Chard 获得了新西兰惠灵顿维多利亚大学计算机科学博士学位。在阿贡国家实验室担任玛丽亚·格佩特-梅耶研究员后，他加入*Globus*团队，致力于开发支持科学研究的网络基础设施。

AntoninoMiceli 是*ANL* 的*X*射线科学部探测器组组长，同时受聘于芝加哥大学和西北大学。他从事*X* 射线探测器研发，以推动科学进步，并负责运营 *APS* 探测器库，为*APS* 光束线和用户提供技术服务与支持。*Miceli* 博士获得了华盛顿大学的物理学博士学位。他于 *2005* 年加入 *ANL*。

*MathewCherukara Joseph*获得了普渡大学的计算材料科学与工程博士学位，并获得了印度马德拉斯理工学院的材料工程学士学位。他目前是 *APS* 计算 *X* 射线 *Science* 组的组长。他的研究利用*AI*开展自主实验、进行超越硬件限制的材料表征（计量学），并加速材料建模。

许可证

提交的稿件由UChicago阿贡有限责任公司（阿贡国家实验室运营方）（“阿-贡”）编制。阿贡是一家美国能源部科学办公室实验室，根据合同编号DE-AC02-06CH11357运营。美国政府为自身及代表其行事的其他方保留一项已付费、非独占、不可撤销的全球许可，以复制上述文章、制作衍生作品、向公众发行副本，并代表政府或由政府公开执行和公开展示。能源部将根据 DOE 公共访问计划向公众提供这些联邦资助研究成果的访问权限。
http://energy.gov/downloads/doe-public-access-plan。