

基于深度学习下电子束成像设备的像差值预测

摘要

电子束成像设备（如透射电子显微镜）在材料科学、生物学和纳米技术领域具有重要应用。然而，成像过程中常因像差（如球差和色差）导致图像失真和分辨率下降。传统像差校正方法通常依赖复杂的物理模型和精密的硬件设备，难以适应动态实验条件。为解决上述问题，本文提出了一种基于深度学习的像差值预测方法，通过模拟生成包含 35,050 张干涉图像的高质量数据集，填补了当前公开数据集中缺乏专注于光学像差预测任务的空白。本文通过深度学习技术建模干涉图像与球差参数之间的非线性关系，实现了从图像到球差参数的直接预测，为球差校正任务提供了新思路和技术路径。实验结果表明，基于 ResNet18 模型更改的回归模型在验证集上 C_1 的最低平均绝对误差 (MAE) 达到了 4.76 而验证集上 A_1 的最低平均绝对误差 (MAE) 达到了 6.66，研究充分验证了深度学习技术在像差预测中的潜力。

关键词：深度学习；像差；图像处理；电子束成像

1 引言

1.1 研究背景与电子波长计算

电子束成像设备（如透射电子显微镜 Transmission Electron Microscope, TEM 和扫描电子显微镜 Scanning Electron Microscope, SEM）广泛应用于材料科学、生物学和纳米技术等领域。这些设备通过极短的电子波长实现高分辨率成像，为微观结构研究提供了重要支撑。然而，成像过程中由于球差（Spherical Aberration, C_1 ）和像散（Astigmatism, A_1 ）的存在，导致图像分辨率下降和伪影的产生，严重影响了实验数据的质量和精度。因此，准确预测并校正像差 C_1 和 A_1 对于提升电子束成像质量具有重要意义。

电子波长是影响成像质量的关键因素之一，其计算公式为：

$$\lambda = \frac{h \cdot c}{\sqrt{E_0 \cdot q \cdot m_0 \cdot c^2 \cdot 2 + (E_0 \cdot q)^2}} \quad (1)$$

其中， h 是普朗克常数 ($6.626 \times 10^{-34} \text{ J} \cdot \text{s}$)， c 是光速 ($3.0 \times 10^8 \text{ m/s}$)， q 是电子电荷 ($1.602 \times 10^{-19} \text{ C}$)， m_0 是电子静止质量 ($9.109 \times 10^{-31} \text{ kg}$)， E_0 为电子加速电压。以 $E_0 = 200 \text{ kV}$ 为例，计算得到的电子波长约为 2.51 pm。这一极短波长虽然赋予了电子束极高的分辨率，但也使得球差和像散对成像质量的影响更加显著。

1.2 球差与像散效应

在电子显微成像中，球差 C_1 是一种轴对称的低阶像差，其主要效应表现为光束聚焦的球面误差，导致非线性相位误差的累积。球差引起的相位误差 $\chi_{C_1}(\omega)$ 可用以下公式描述：

$$\chi_{C_1}(\omega) = \frac{\pi C_1}{\lambda} |\omega|^2 \quad (2)$$

其中， $|\omega|$ 为频率坐标的模长， λ 为电子波长， C_1 为球差系数。

此外，像散 A_1 是另一种常见的低阶非对称像差，导致电子束在不同方向上的聚焦不一致。像散的相位误差 $\chi_{A_1}(\omega)$ 可以描述为：

$$\chi_{A_1}(\omega) = \frac{\pi}{\lambda} \text{Re}(A_1 \cdot \omega^2) \quad (3)$$

其中， A_1 为像散系数的复数形式，用于表示不同方向的聚焦误差， ω 为频率坐标的复数形式。像散会引起光束在某一方向的拉伸或模糊，从而导致图像的对称性和分辨率下降。

1.3 像差模拟与干涉图像数据生成

为了研究深度学习方法在球差 C_1 和像散 A_1 预测任务中的应用，本实验采用模拟方法生成了具有不同球差系数和像散系数的干涉图像数据集。模拟生成基于电子显微成像的相位对比传递函数 (CTF)，其表达式为：

$$T(\omega) = E(\omega) \cdot \exp(i \cdot (\chi_{C_1}(\omega) + \chi_{A_1}(\omega))) \quad (4)$$

其中， $E(\omega)$ 为电子波的初始振幅函数， $\chi_{C_1}(\omega)$ 和 $\chi_{A_1}(\omega)$ 分别表示球差和像散引起的相位误差。

为了模拟真实实验条件中的复杂性，干涉图像生成过程还引入了以下随机扰动因素：

- **电子束倾斜：**使用高斯分布模拟倾斜不稳定性，使倾斜参数 k_0 在不同图像中随机波动。
- **光束发散角：**使用复数形式表示光束发散角，模拟其随机扰动对成像的影响。
- **探测器噪声：**模拟实际 CCD 探测器中的泊松噪声和高斯暗电流噪声。

最终的图像强度由以下公式计算：

$$I(k) = |\mathcal{F}[T(\omega) \cdot E(\omega)]|^2 + N_{\text{CCD}} \quad (5)$$

其中， \mathcal{F} 表示傅里叶变换， N_{CCD} 为探测器噪声。

1.4 数据集特点与样例

实验生成了包含 35,050 张干涉图像的模拟数据集，其中 25,050 张用于球差 C_1 的预测任务，另外 10,000 张用于像散 A_1 的预测任务。这些数据样本覆盖了多种球差和像散参数的组合，为深度学习模型的训练和评估提供了充足的支持。

球差 C_1 数据集：在球差 C_1 数据集中，每 50 张图像对应不同的球差系数 C_1 ，模拟了从小到大的球差变化，同时引入了随机光束倾斜和发散角的扰动。所有图像经过归一化处理，分辨率调整为 512×512 ，以适配深度学习模型的输入需求。

像散 A_1 数据集：像散 A_1 数据集包含 10,000 张图像，每 100 张图像对应不同的像散系数 A_1 。这些图像在模拟过程中，通过调整像散系数的复数形式，生成了不同方向上的像散失真效应。此外，数据集同样引入了随机光束倾斜和探测器噪声等扰动因素，以更真实地模拟实际成像条件。

最终图像强度计算：数据集中每张图像的强度 $I(k)$ 通过以下公式计算得到：

$$I(k) = |\mathcal{F}[T(\omega) \cdot E(\omega)]|^2 + N_{\text{CCD}} \quad (6)$$

其中， $T(\omega)$ 为包含球差 C_1 和像散 A_1 信息的传递函数， $E(\omega)$ 为初始振幅分布， \mathcal{F} 为傅里叶变换， N_{CCD} 为模拟的 CCD 噪声。

数据集总结：球差 C_1 数据集和像散 A_1 数据集分别针对不同的像差参数进行了细致的模拟，具有以下特点：

- 球差 C_1 ：每 50 张图像对应不同的球差系数，共计 25,050 张图像；
- 像散 A_1 ：每 100 张图像对应不同的像散系数，共计 10,000 张图像；
- 模拟过程中引入了随机扰动因素（光束倾斜、发散角、探测器噪声）以增加数据的多样性和真实性。

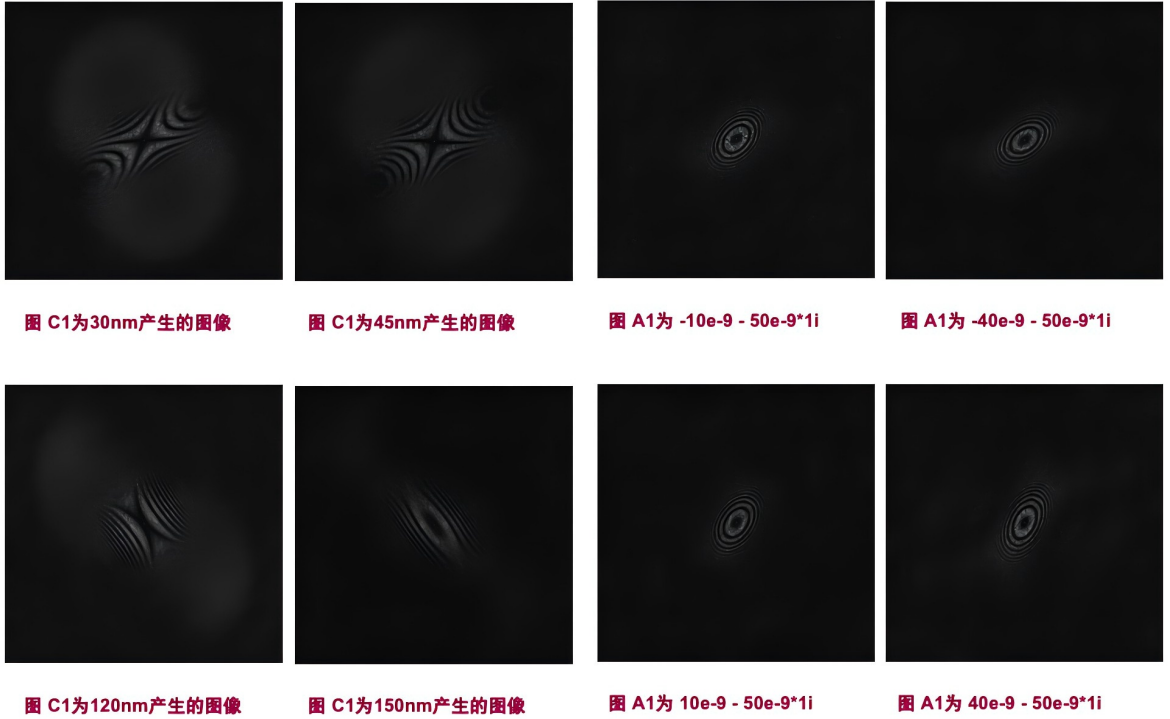


图 1. 数据集样例：(左) 球差 C_1 干涉图像样例；(右) 像散 A_1 干涉图像样例

图 1 展示了数据集中球差和像散干涉图像的典型样例。左图为不同 C_1 参数下的干涉图像，右图为不同 A_1 参数下的干涉图像，可见 A_1 图像具有更显著的方向性失真特征。

2 相关工作

球差 (C_1) 作为光学系统中常见且重要的像差之一，对高分辨率成像系统的性能有显著影响。传统的球差校正方法主要依赖于硬件设备，例如多极校正磁透镜和球差校正器，通过调

节电子光学系统中的场强分布来补偿球差 [9]。尽管这些方法在一定程度上改善了成像质量，但其高昂的硬件成本、复杂的操作流程以及对实验条件的高依赖性限制了其应用范围。

传统的光场传输理论虽然可以模拟光束在含像差的光学系统中的传输过程，但从已知光强分布中提取像差参数却十分困难，尤其是在存在多种像差的情况下。随着计算能力和数据量的提升，深度学习技术迅速发展，为这一问题提供了潜在的解决方案。深度卷积神经网络 (DCNN) 被广泛应用于各类任务中，如图像分割、面部对齐 [10]、目标识别、故障诊断和目标检测等 [3]。然而，随着网络层数的增加，DCNN 会面临退化问题。为了解决这一问题，何凯明团队在 2015 年提出了深度残差网络 (ResNet)，通过引入残差学习框架，增强了网络的鲁棒性并解决了退化问题 [2]。ResNet 在 ImageNet 和 CIFAR-10 等分类任务中取得了优异的性能，扩展了深度学习在学术和应用研究中的使用范围，并被应用于疾病检测、垃圾图像分类及医学影像分类等领域。

在像差估计研究中，Möckl 等人 [6] 使用矢量衍射理论模拟了具有不同焦点和 Zernike 像差的点扩散函数 (PSF) 图像，并通过在包含 10 万张图像的数据集上训练神经网络，实现了对像差系数的高精度预测。这一方法成功验证了神经网络在复杂 PSF 图像中提取像差信息的能力。Vanberg 等人 [7] 使用深度卷积神经网络 (CNN) 预测焦平面图像的 Zernike 模态系数，证明了深度学习在波前感知任务中的可行性。Ma 等人 [5] 通过使用两组不同焦点的强度图像作为 AlexNet 的输入，成功从强度图像中提取了相应的 Zernike 系数。Wang 等人 [8] 提出了基于 U-Net 架构的相位恢复深度卷积神经网络 (PRDCNN)，该方法能够直接测量光学系统中的像差结构，展现出更直接、更准确的波动性和鲁棒性。这些研究证明了深度卷积神经网络 (DCNN) 能够从光强图像中提取像差信息并进行估计。

3 本文方法

本文提出了一种基于深度学习的光学像差预测方法，分别针对电子束成像设备中的球差 (C_1) 和像散 (A_1) 进行独立建模与预测。通过生成包含 25,050 张干涉图像的高质量模拟数据集，并结合三种深度学习模型 (VAE、AlexNet 和 ResNet18) 更改的回归模型对 C_1 进行实验，验证其预测性能。鉴于基于 ResNet18 更改的回归模型在 C_1 任务中的出色表现，进一步单独使用该回归模型进行 A_1 的预测任务，评估其在不同像差参数上的适用性。具体模型及优化方法如下：

3.1 变分自动编码器 (VAE)

变分自动编码器 (Variational Autoencoder, VAE) 是一种生成模型，其核心思想是通过潜在空间中的概率分布建模输入数据，并生成近似输入数据的新样本。VAE 由编码器和解码器两部分组成，其中编码器将输入映射到潜在分布，解码器则根据潜在分布生成新的数据样本。

VAE 的损失函数由重构误差 ($\mathcal{L}_{\text{reconstruction}}$) 和 KL 散度 (\mathcal{L}_{KL}) 两部分组成，其总损失函数为：

$$\mathcal{L} = \mathcal{L}_{\text{reconstruction}} + \beta \mathcal{L}_{\text{KL}}, \quad (7)$$

其中 β 为权重超参数， $\mathcal{L}_{\text{reconstruction}}$ 用于衡量重构图像 x' 与原始图像 x 的差异：

$$\mathcal{L}_{\text{reconstruction}} = \|x - x'\|^2, \quad (8)$$

\mathcal{L}_{KL} 用于衡量潜在分布 $q(z|x)$ 与标准正态分布 $p(z)$ 的差异：

$$\mathcal{L}_{\text{KL}} = \frac{1}{2} \sum (1 + \log(\sigma^2) - \mu^2 - \sigma^2), \quad (9)$$

其中 μ 和 σ^2 为编码器输出的均值和方差。

尽管 VAE 在高维特征建模中具有潜力，但其在球差预测任务中的性能较其他方法略有不足，尤其在复杂非线性特征提取方面，重构误差可能限制模型精度。

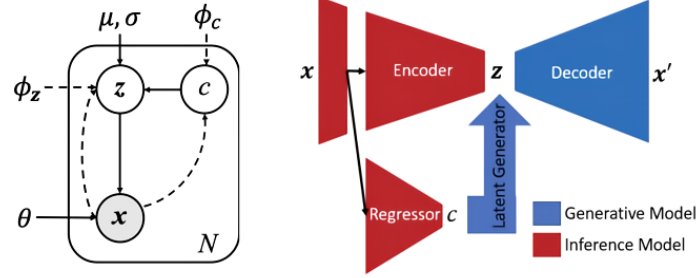


图 2. 基于 VAE 的回归模型的概率图 (左) 和图形图 (右)。假设每个图像 x 由其表示 z 生成，而 z 又取决于像差 c (蓝色方块)。推断模型 (红色方块) 构建一个概率编码器来确定潜在表示，以及一个概率回归器来预测像差 [11]。

3.2 AlexNet

AlexNet 是一种经典卷积神经网络 (CNN)，通过多层卷积和池化操作提取图像的特征，最终通过全连接层进行分类或回归任务 [4]。AlexNet 的主要结构包括 5 个卷积层和 3 个全连接层，并引入了 ReLU 激活函数和 Dropout 技术以加速训练和减少过拟合。

为了适配球差预测任务，本研究对 AlexNet 的全连接层进行了如下修改：

$$\text{FC}_{\text{modified}} = \text{Linear}(\text{Input Size}, 1), \quad (10)$$

其中 Input Size 为上一层的输出特征维度，最终输出一个回归值以预测 C_1 。

AlexNet 在本任务中展现了一定的预测能力，但由于其结构较为简单，在处理复杂干涉图像时泛化能力有限，验证集误差波动较大。

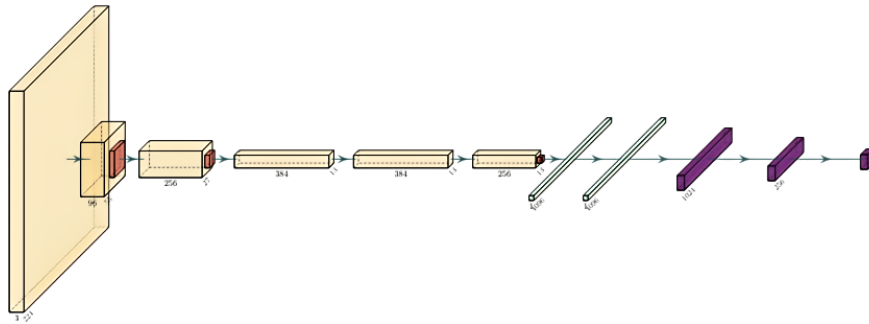


图 3. 基于 AlexNet 的回归模型结构图

3.3 ResNet18

ResNet (Residual Network) 是深度学习中的一种重要模型，通过引入残差连接 (Residual Connection)，有效缓解了深层网络中的梯度消失问题 [1]。ResNet18 是 ResNet 系列中较浅的一种版本，包含 18 层网络，适合中小型数据集的特征提取任务。

残差块的核心是通过跳跃连接，将输入直接传递到输出：

$$\mathbf{y} = \mathcal{F}(\mathbf{x}, \{W_i\}) + \mathbf{x}, \quad (11)$$

其中 $\mathcal{F}(\mathbf{x}, \{W_i\})$ 表示卷积、Batch Normalization 和 ReLU 操作的组合， \mathbf{x} 为输入， \mathbf{y} 为输出。

本研究对 ResNet18 的最后一层进行了修改，将全连接层替换为适用于回归任务的线性层：

$$\text{Output} = \text{Linear}(\text{Input Features}, 1), \quad (12)$$

其中 Input Features 表示最后一层特征的维度，输出为预测的球差值 C_1 或 A_1 。

相比于 VAE 和 AlexNet，ResNet18 凭借其深层网络结构和残差连接的优势，在球差预测任务中表现最为优异。残差连接有效缓解了梯度消失问题，使其能够高效提取复杂非线性特征，并显著降低验证集误差。

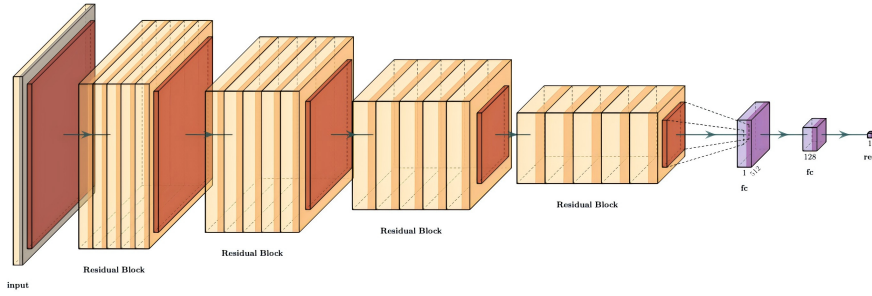


图 4. 基于 ResNet 的回归模型结构图

3.4 实验设置

本研究设置如下：1. **全连接层的修改**：针对深度学习模型的架构设计，特别是全连接层部分，进行了针对性优化，以更好地适应像差值预测任务。2. **学习率调度**：采用 StepLR 学习率调度器，每 10 轮训练将学习率衰减为原来的 0.1，以稳定训练过程。3. **损失函数**：选用平均绝对误差 (Mean Absolute Error, MAE) 作为损失函数，定义为：

$$\mathcal{L}_{\text{MAE}} = \frac{1}{N} \sum_{i=1}^N |P_{i,\text{true}} - P_{i,\text{pred}}|, \quad (13)$$

其中 $P_{i,\text{true}}$ 为第 i 个样本的真实 C_1 值， $P_{i,\text{pred}}$ 为预测值。

通过上述模型设计，本研究验证了深度学习模型在球差值预测任务中的有效性与潜力。

4 复现细节

4.1 与已有开源代码对比

没有参考任何相关源代码。

4.2 创新点

4.2.1 数据集构建与应用

针对这一重要且复杂的光学像差问题，本文通过模拟生成了包含 35,050 张干涉图像的高质量数据集。该数据集充分涵盖了多种不同的球差参数和噪声条件，为深度学习模型的训练和评估提供了丰富的样本支持。这一工作填补了当前公开数据集中缺乏专注于光学像差预测任务的空白，为相关研究领域的数据需求提供了有力保障。

4.2.2 任务驱动的研究创新

与传统的基于物理模型的球差校正方法不同，本文通过深度学习技术分别建模干涉图像与球差 (C_1) 和像散 (A_1) 参数之间的非线性关系，实现了从图像到球差及像散参数的直接预测。该方法不仅避免了传统方法中复杂的物理建模和优化过程，还展现了优越的实时性和高效性，为球差和像散校正任务提供了新思路和技术路径。

5 实验结果分析

通过实验结果可以看出，AlexNet 回归模型在球差预测任务中的表现优于 VAE 回归模型。具体而言，VAE 回归模型的训练损失在早期快速下降，但验证损失波动较大，且训练后期验证集损失与训练集损失之间存在较明显的差距，这表明 VAE 回归模型在泛化能力方面存在不足，可能出现过拟合问题。同时，验证损失的波动也可能反映出 VAE 回归模型在处理高维非线性特征时的能力有限。AlexNet 回归模型在训练过程中的损失收敛更快，验证损失与训练损失的差距较小，且两者在训练后期均趋于稳定。

相比之下，ResNet18 回归模型展现了更强的特征提取能力和泛化性能。在初始训练 10 轮时，其 C_1 的验证集的最低 MAE 达到 9.45；进一步延长训练至 20 轮和 50 轮后， C_1 的验证集 MAE 进一步下降至 6.03 和 4.76， A_1 的验证集在 50 轮时 MAE 达到了 6.66。ResNet18 回归模型的损失曲线表现平稳，验证集与训练集误差之间的差距较小，表明其对复杂非线性特征的建模能力更强。此外，ResNet18 的残差结构有效缓解了梯度消失问题，使其能够在更深的网络中保持高效训练。这些结果表明，深度学习模型的结构深度和优化策略对像差值预测的性能有重要影响，其中 ResNet18 回归模型为实现高效、准确的像差预测提供了最佳方案。

表 1. 不同模型的实验结果

模型	验证集 MAE(A_1 的 MAE)
VAE(10epoch)	19.25
AlexNet(10epoch)	16.15
ResNet18(10epoch)	9.45
ResNet18(20epoch)	6.03
ResNet18(50epoch)	4.76 (6.66)

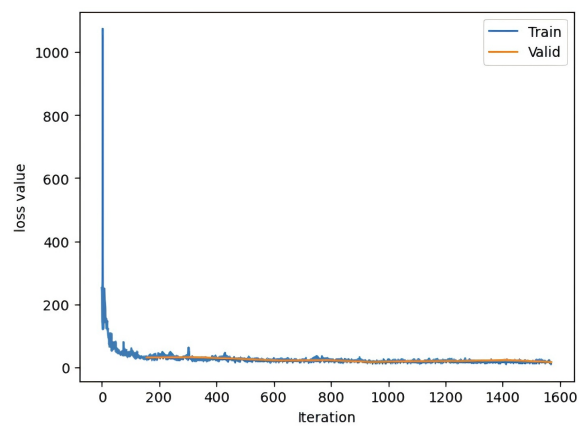
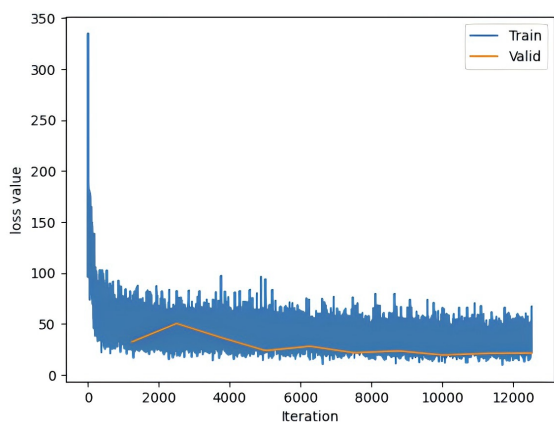


图 5. (左) 基于 VAE 回归模型进行 10 轮 C1 的结果图。(右) 基于 AlexNet 回归模型进行 10 轮 C1 的结果图。

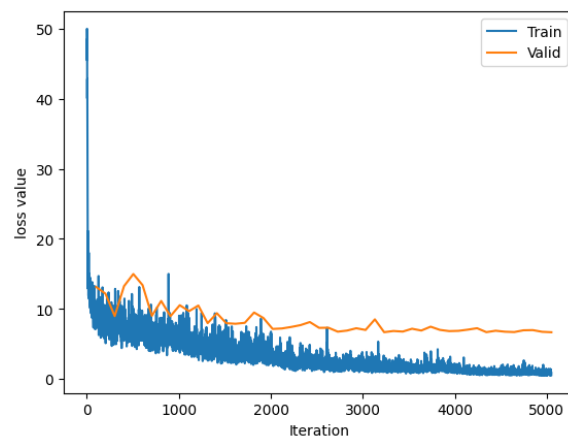
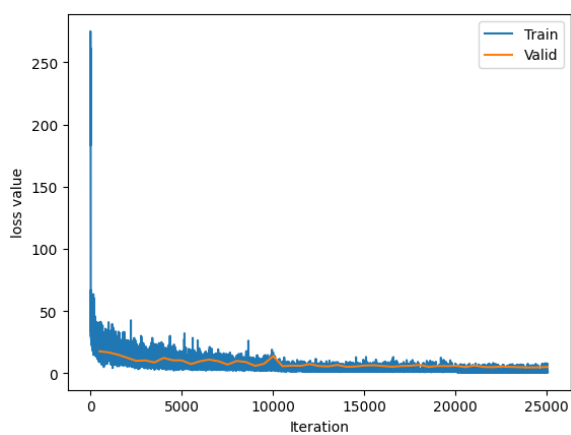


图 6. (左) 基于 ResNet 回归模型进行 50 轮 C1 的结果图。(右) 基于 ResNet 回归模型进行 50 轮 A1 的结果图。

6 总结与展望

本文研究了基于深度学习的电子束成像设备像差值预测方法，分别针对球差 (C_1) 和像散 (A_1) 的非线性特性，设计并评估了三种深度学习模型 (VAE、AlexNet 和 ResNet18) 修改的回归模型。实验结果表明，ResNet18 回归模型凭借其残差结构和深层网络的强大特征提取能力，在球差预测任务中取得了最低 MAE 为 4.76 的优异表现，并在像散预测任务中取得了 MAE 为 6.66 的优异表现，优于其他两种回归模型。这表明深度学习技术不仅能够有效应对像差校正任务中的复杂非线性问题，还在实时性和高效性方面具有显著优势。

未来的研究将着眼于更复杂的网络结构（如 Vision Transformer 和混合模型架构）的探索，并结合更大规模和更多样化的干涉图像数据集进一步提升模型性能。此外，将物理先验知识与深度学习模型相结合，以优化网络设计和特征提取能力，为电子束成像设备的实时像差校正和智能化操作提供更强大的技术支持。

参考文献

- [1] Dongwei Chen, Fei Hu, Guokui Nian, and Tiantian Yang. Deep residual learning for nonlinear regression. *Entropy*, 22(2):193, 2020.
- [2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [3] Yunbo Li and Xiaofeng Li. Automatic segmentation using deep convolutional neural networks for tumor ct images. *International Journal of Pattern Recognition and Artificial Intelligence*, 37(03):2352003, 2023.
- [4] Lan Lin, Ge Zhang, Jingxuan Wang, Miao Tian, and Shuicai Wu. Utilizing transfer learning of pre-trained alexnet and relevance vector machine for regression for predicting healthy older adult’ s brain age from structural mri. *Multimedia Tools and Applications*, 80(16):24719–24735, 2021.
- [5] Huimin Ma, Haiqiu Liu, Yan Qiao, Xiaohong Li, and Wu Zhang. Numerical study of adaptive optics compensation based on convolutional neural networks. *Optics Communications*, 433:283–289, 2019.
- [6] Leonhard Möckl, Petar N Petrov, and WE Moerner. Accurate phase retrieval of complex 3d point spread functions with deep residual neural networks. *Applied physics letters*, 115(25), 2019.
- [7] Pierre-Olivier Vanberg et al. Machine learning for image-based wavefront sensing. 2019.
- [8] Yangyundou Wang, Hao Wang, Yiming Li, Chuanfei Hu, Hui Yang, and Min Gu. High-accuracy, direct aberration determination using self-attention-armed deep convolutional neural networks. *Journal of microscopy*, 286(1):13–21, 2022.
- [9] Kenta Yoshida, Alexander Bright, and Nobuo Tanaka. Direct observation of the initial process of ostwald ripening using spherical aberration-corrected transmission electron microscopy. *Journal of electron microscopy*, 61(2):99–103, 2012.
- [10] Shaohua Zhang, Hua Yang, and Zhouping Yin. Multiple deep convolutional neural networks averaging for face alignment. *Journal of electronic imaging*, 24(3):033013–033013, 2015.
- [11] Qingyu Zhao, Ehsan Adeli, Nicolas Honnorat, Tuo Leng, and Kilian M Pohl. Variational autoencoder for regression: Application to brain aging analysis. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part II 22*, pages 823–831. Springer, 2019.