BIOZENTRUM

The Center for
Molecular Life Sciences

Optimal Back-Projection for Tomography

3D Filter for Enhanced Reconstruction Quality

Zhiping Wang zhiping.wang@stud.unibas.ch

Master Student in *Physics of Life*

Biozentrum

11, March, 2025

Outline

1. Research Background
2. Methods
3. Results
   - i. On the SHREC dataset – good results
   - ii. On the real dataset – challenges remain
4. Discussion and Conclusion
5. Thanks

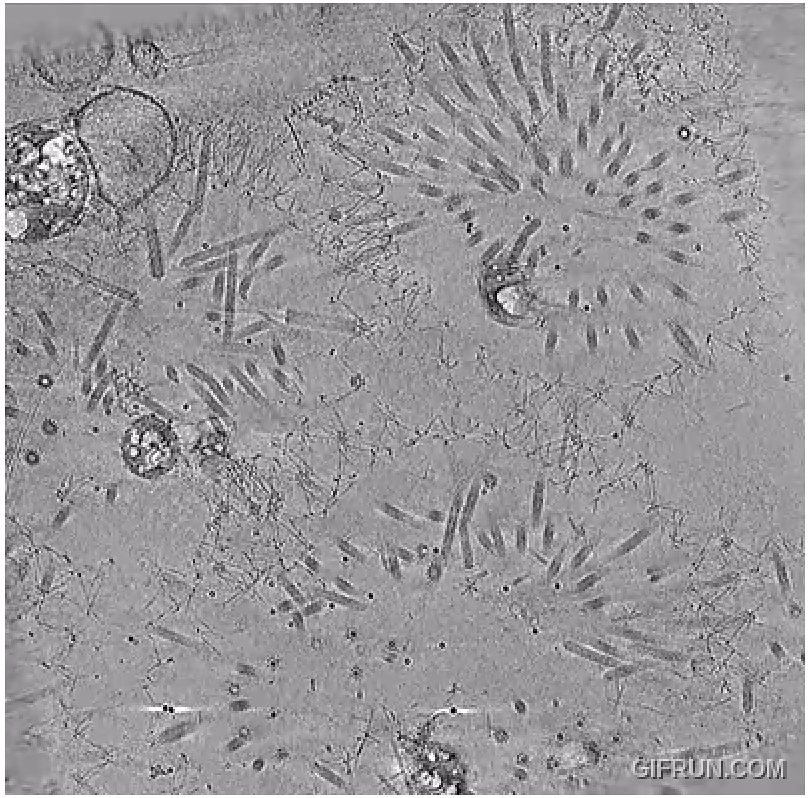
Research Background

Cryo-ET 3D reconstruction

raw data

(https://www.youtube.com/watch?v=SbEvuSgskWw)

3d result



(https://www.youtube.com/watch?v=mePLPzK15R8)

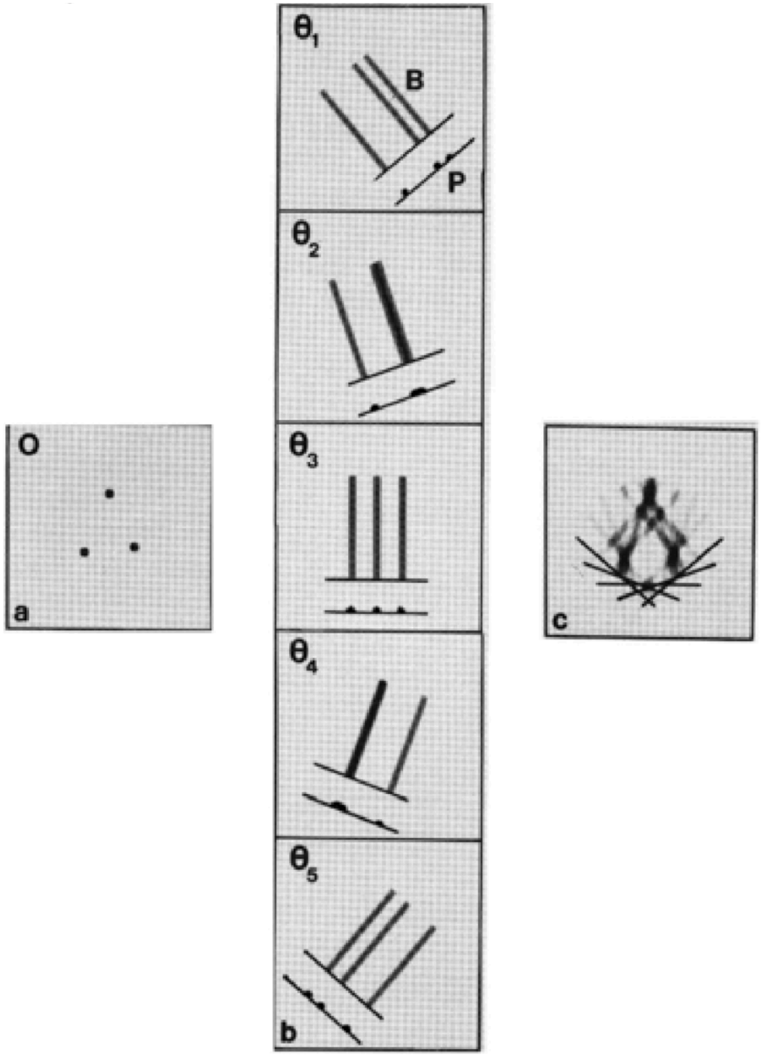
Back projection

- left: Projection of 3 points onto a line at various tilt angles.
- middle: Backprojection of projection data into reconstruction volume(Radon transform)
- right : reconstruction (inverse of the Radon transform)

  (From Rademacher, 1992)

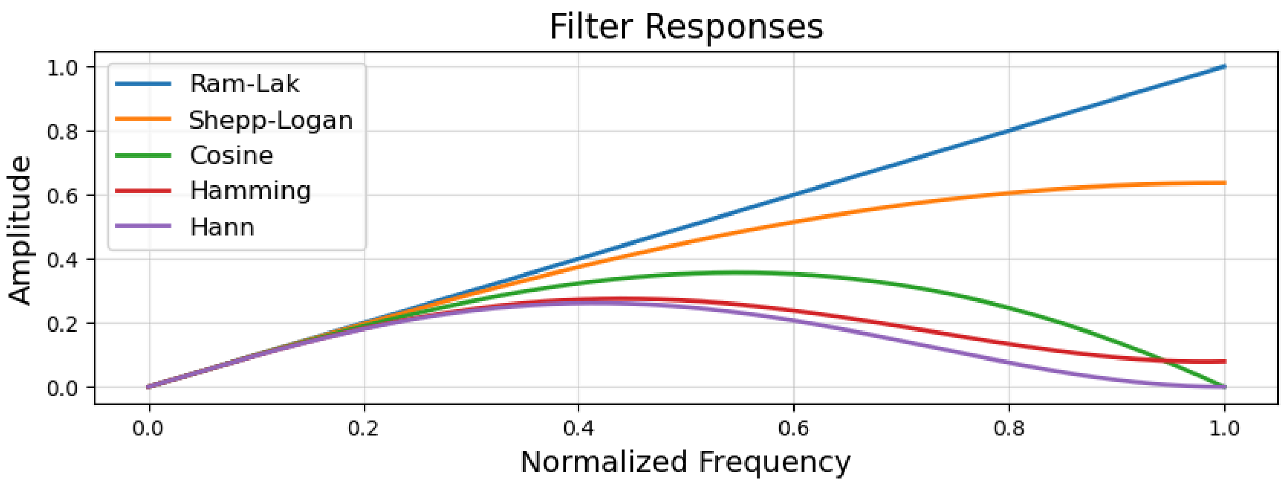
Cons

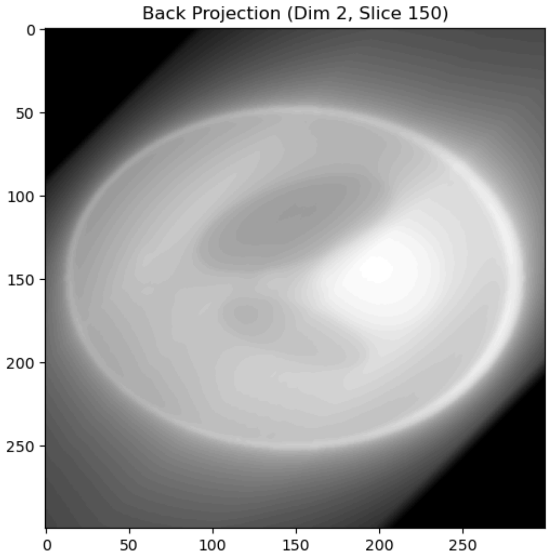
- Low Contrast
- Blur Artifacts
- ……

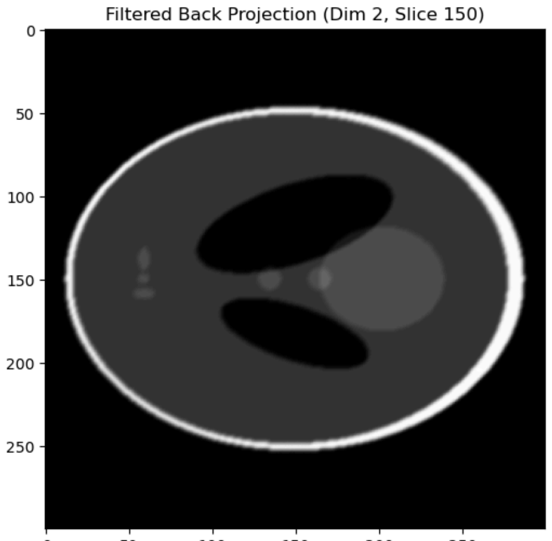


Filtered back projection

- Process:
  - Apply a filter to the projection data (e.g., Ram-Lak, Shepp-Logan).
  - Perform back projection to reconstruct the image.
- Advantage:
  - Significantly reduces artifacts such as blurring caused by plain back projection.
  - Improves spatial resolution and clarity in reconstructed images.

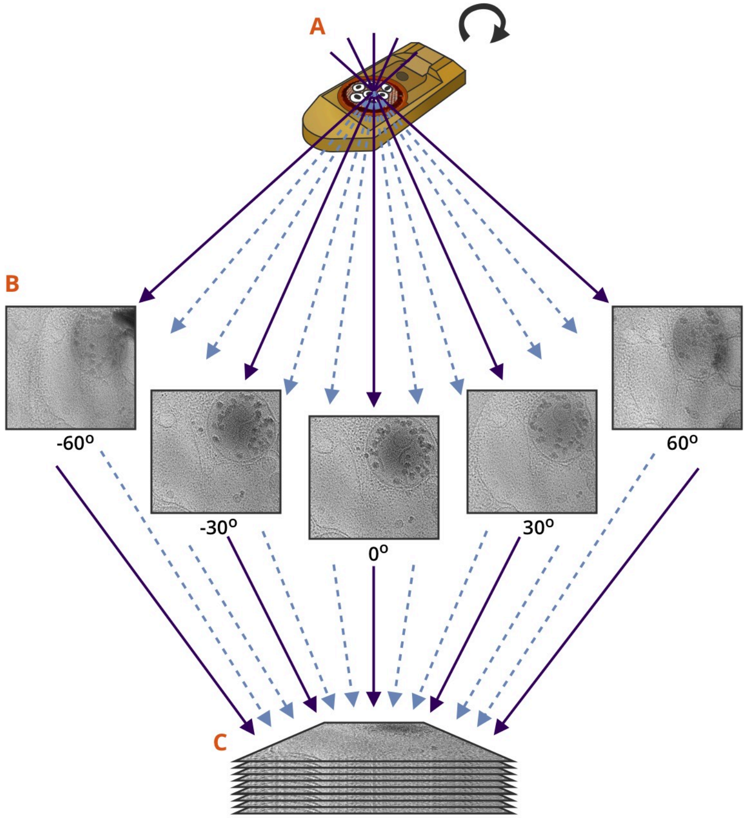






- Key Question:
  - **how can we get a great filter?**

Cryo-ET Data Acquisition and Radon Transform



The image on the left is from MyScopeCryoET2023, which illustrates the data acquisition process in Cryo-ET. The process can be mathematically described using the **Radon transform**

$$v(\vec{r}, \theta) = \int_{-\infty}^{\infty} u(z\cos(\theta) - \vec{r} \cdot \sin(\theta), z\sin(\theta) + \vec{r} \cdot \cos(\theta))\, dz.$$

Here, $v(\vec{r}, \theta)$ represents projections, $u(\vec{r}, z)$ is unknown 3D sample density.

We can also write it as:

$$v(\vec{r}, \theta) = \mathcal{R}(u(\vec{r}, z))$$

where $\mathcal{R}$ is Radon transform.

Mathematical Model of Filtered Back-Projection

- Filtering:
  - necessary due to the central slice theorem,
  - $$\hat{v}(\vec{r},\theta) = h(\vec{r},\theta) * v(\vec{r},\theta)$$
  - where h is the filter for projection
- Back-Projection:
  - inverse Radon transform
  - $$\tilde{u}(\vec{r},z) = \int_{-\pi}^{\pi}\int_{\mathbb{R}} v(\vec{s},\theta)h(n_\theta \cdot \vec{r} - \vec{s})\, d\vec{s}\, d\theta$$
  - $$\tilde{u}(\vec{r},z) = \mathcal{R}^{-1}(\hat{v}(\vec{r},\theta)) = \mathcal{BP}\left(h(\vec{r},\theta) * v(\vec{r},\theta)\right)$$
  - where $\mathcal{R}^{-1}$ is inversion of Randon transform, and $\mathcal{BP}$ is direct back projection.
- **Note**: In most cases, all filters are **identical along the tilt angle dimension**
  - This is precisely what my project aims to achieve.

My Project

Objective:

Design a filter that uses a **distinct filter for each tilt angle** to better capture information and achieve improved reconstruction results.

Implementation Method:

Develop an optimizable filter by iteratively refining it using publicly available models. This optimized filter is then generalized to more real experimental data.

Methods of Optimization

Optimizing Filter: Basic Idea and Dataset

Optimization Approach:

- Treat the filter as an optimizable parameter.
- Run filtered back-projection with a preliminary filter and compare the resulting with the **ground truth**.
- Iteratively fine-tune the filter: Employ an optimizer to adjust the filter parameters through iterative optimization, ultimately yielding an optimal filter.

Question: Where Do We Get the Ground Truth?

Obtaining true ground truth from real cryo-ET data is challenging. Therefore, we select the **SHREC** 2021 cryo-ET dataset (doi: 10.2312/3dor.20211307), which is a simulated cryo-ET dataset offering known ground truth information and is highly suitable for training. We first train on this dataset and then transfer the optimized filter to real data.

Data Representation and Dimensions

- **Ground Truth Volume**: $x \in \mathbb{R}^{z \times x \times y}$
- **Projection Data (Tilt-Series)**: $y \in \mathbb{R}^{m \times x \times y}$
- **Filter Parameterization**: $\beta \in \mathbb{R}^{m \times x \times y}$
- **Reconstructed Volume**: $\hat{x}(\beta) \in \mathbb{R}^{z \times x \times y}$

FBP Reconstruction:

$$\hat{x}(\beta) = \mathcal{R}^{-1}\big(\beta * y\big) = \mathrm{FBP}(y, \beta)$$

Optimization Process - Minimizing the Loss Funcation

1. **Loss Funcation**
   - L is the loss funcation, representing the difference between the predicted volume and the actual volume.
   - So we need to minimize it!

2. **Gradient Computation**

$$\nabla_{\beta} L = \mathrm{Backward}(\beta,\, y,\, \hat{x}(\beta))$$

   - Backward refers to PyTorch's automatic differentiation mechanism, which computes gradients via the chain rule automatically.
   - The computed gradient indicates how much and in which direction the filter parameters β should be adjusted during each iteration.

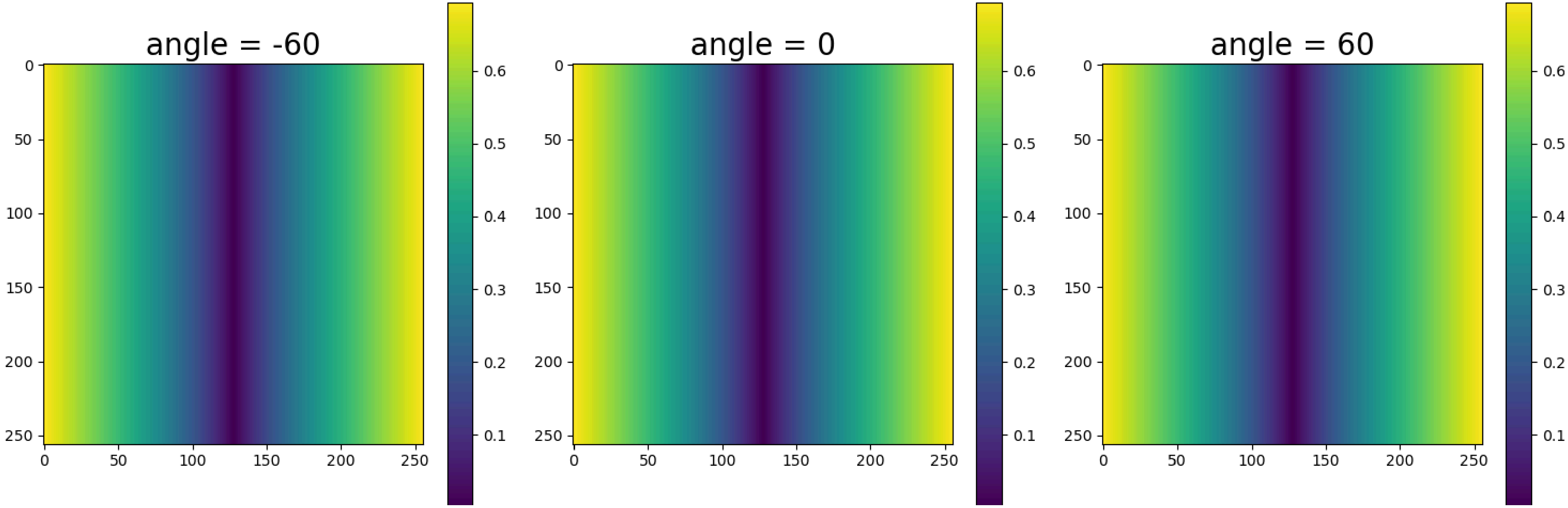
3. **Parameter Update**

$$\beta \leftarrow \beta - \alpha \nabla_{\beta} L$$

   - α is the learning rate controlling the step size.

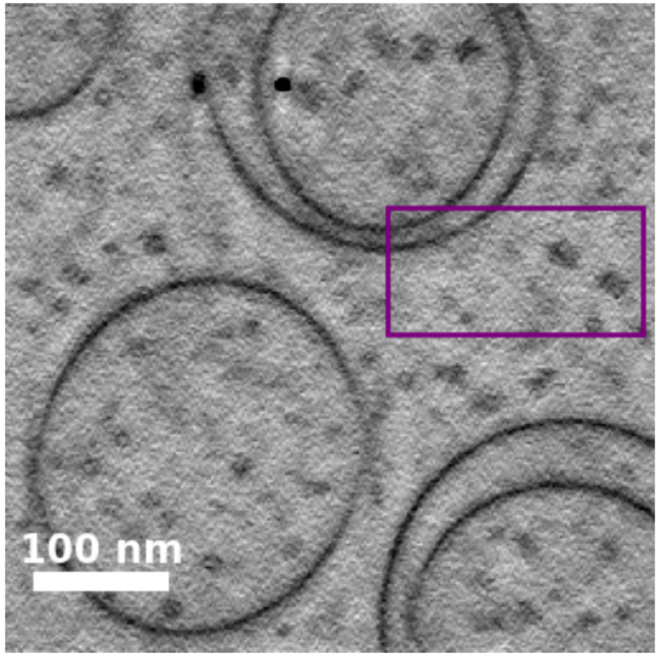
Result

Initial guess : Ramp Filter

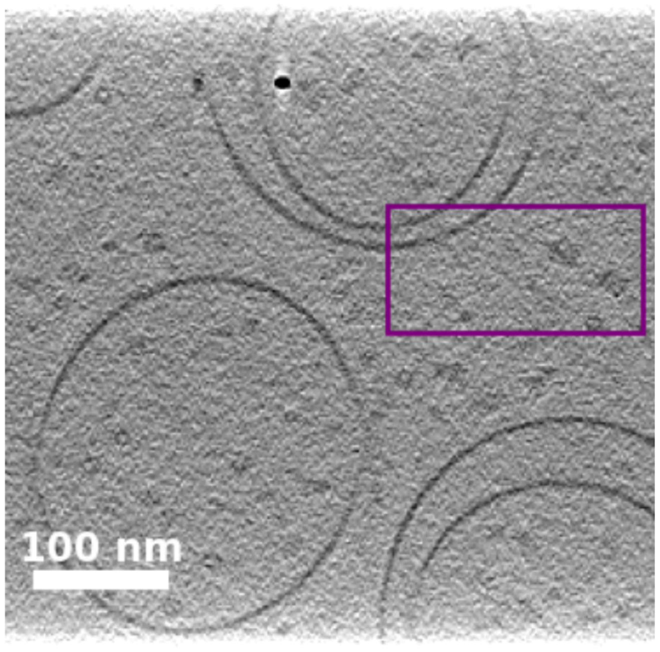


Enhances high frequencies and reduces blur naturally.

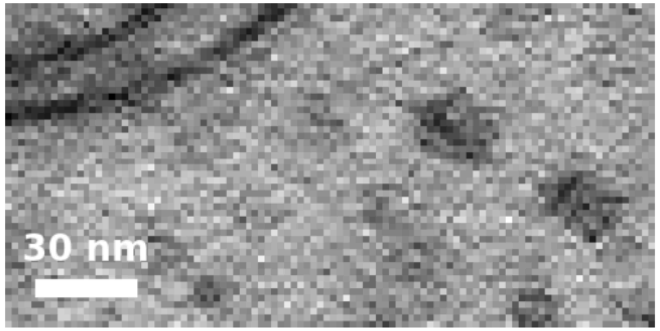
Result On SHREC 2021



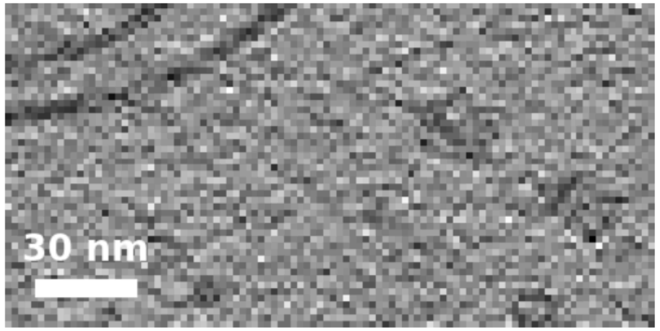
(a) Result of optimized filter



(b) Result of ODL's FBP

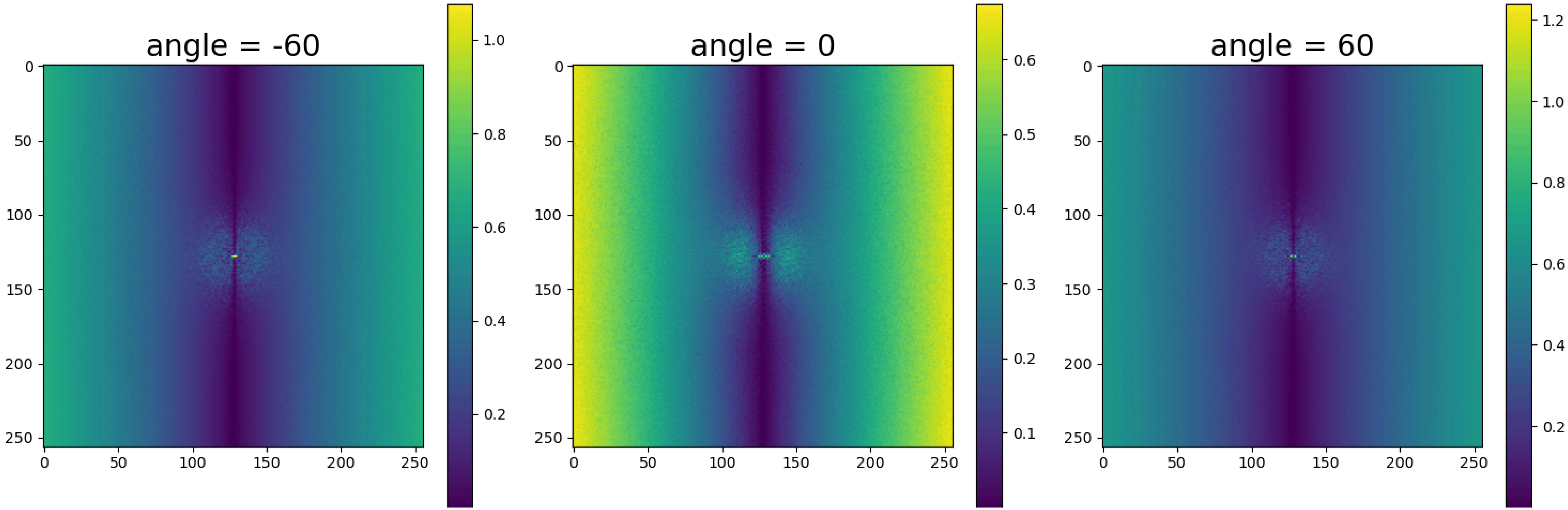


(c) Zoomed-in view of 2a



(d) Zoomed-in view of 2b

Result On SHREC 2021



We can observe significant changes in the low frequencies, which result in improved contrast in our reconstruction.

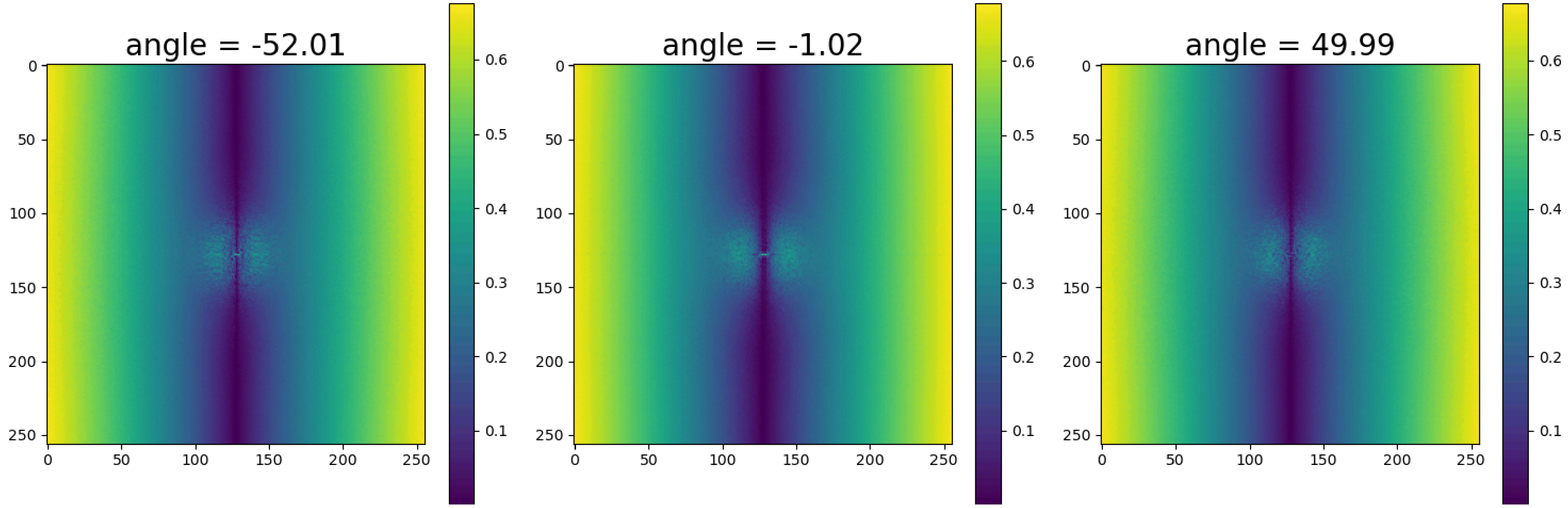
Result On Real Sample

So far, all training has been performed using the SHREC dataset, which significantly differs from real data. For real data, I employed a Chlamydomonas sample from EMPIAR-11830 (Kelley et al., 2024), which comprises 35 tilt angle projections for testing. In contrast, the SHREC dataset contains 61 angles and is based on a completely different numerical range, rendering the previously trained filter unsuitable for direct use.

Two Approaches to Address This:

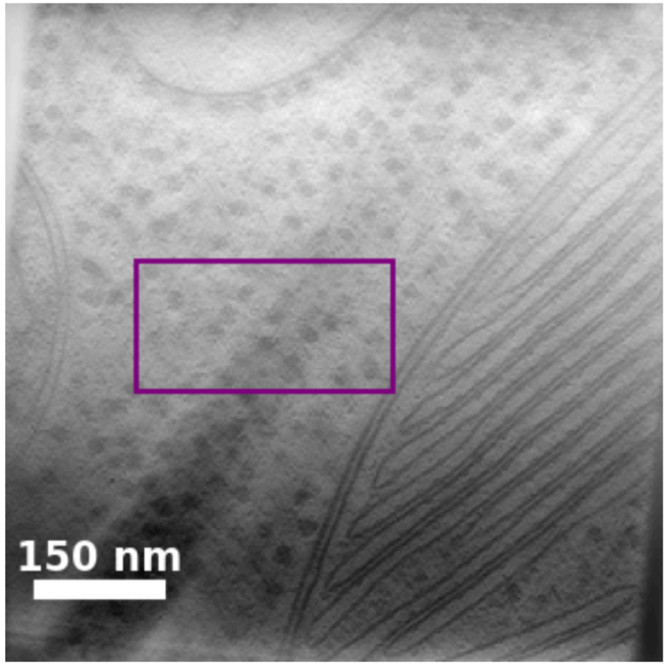
1. **Direct Interpolation:**
   Use the existing filter with interpolation to adjust for the size difference.
   - *Issue:* This method tends to introduce blur in the reconstructed real sample.
2. **Retraining with New Simulated Projections:**
   Generate synthetic projections that mimic the real sample's size and characteristics, then retrain the filter on these projections.
   - *Issue:* In comparison, no significant improvement has been observed.

Interpolated Filter

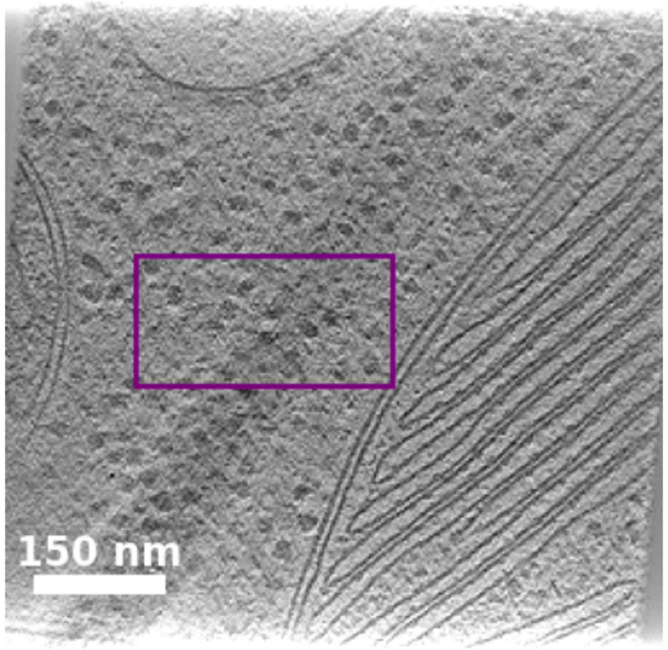


We can see that it is very **similar** to the filter of SHREC dataset since it is derived purely through linear interpolation. Although it does enhance contrast, the mismatch in the model results in some **blurring** in the image.

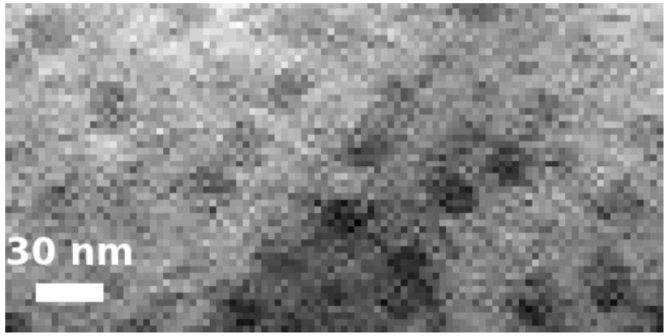
Interpolated Filter



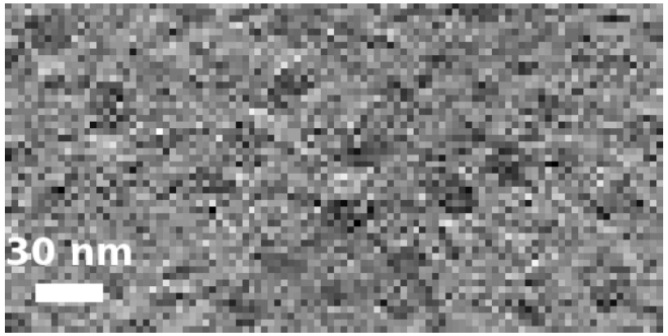
(a) Chlamydomonas Result of optimized filter



(b) Chlamydomonas Result of ODL's FBP



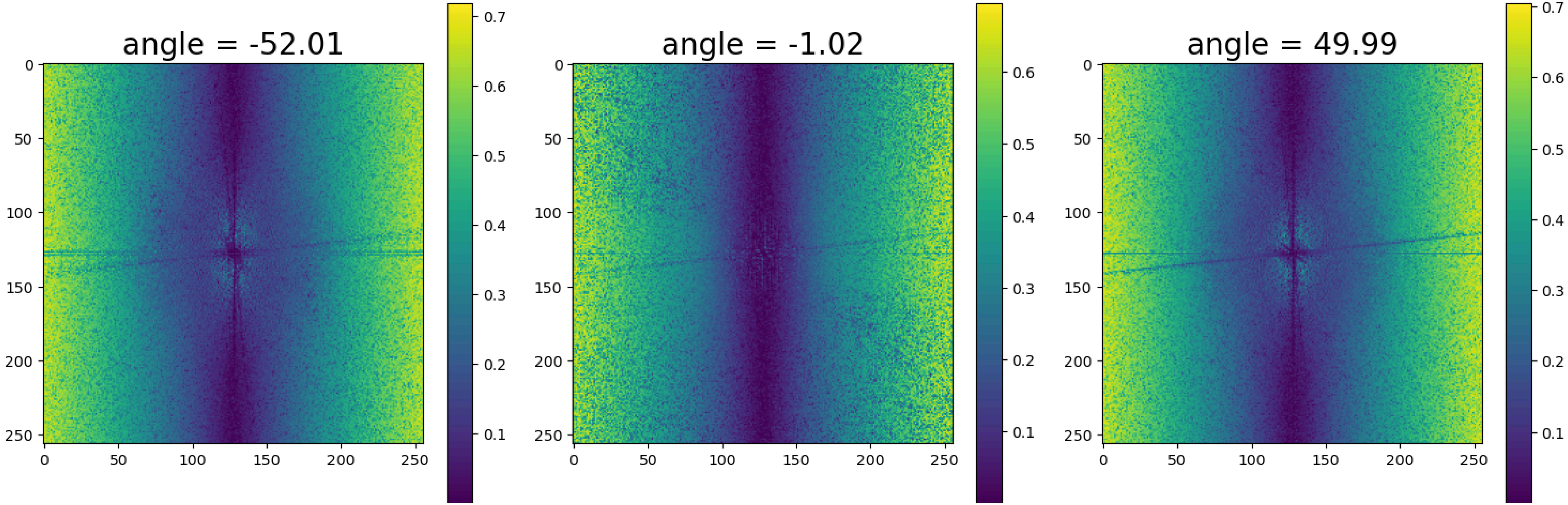
(c) re-normalizatied zoomed-in view of 5a



(d) Zoomed-in view of 5b

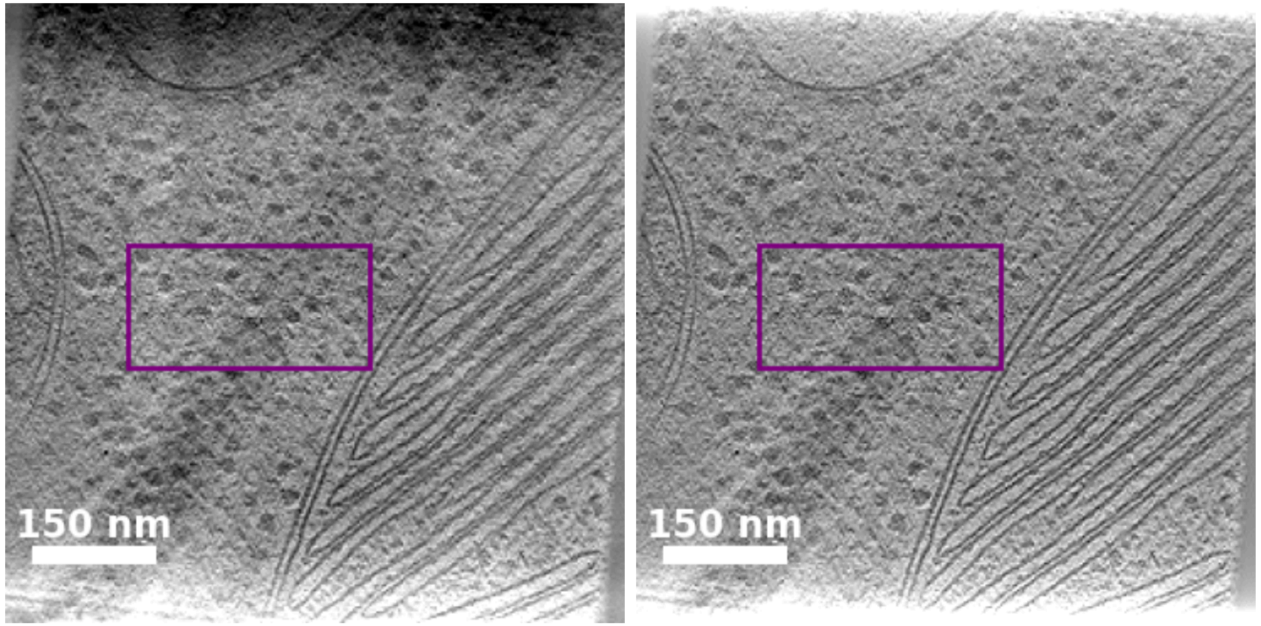
Based on the SHREC dataset with 60 tilt angles, the 4th-order least-squares fitting produced the best results. Although some blur and uneven background persist, the overall contrast is enhanced.

New Simulated Filter

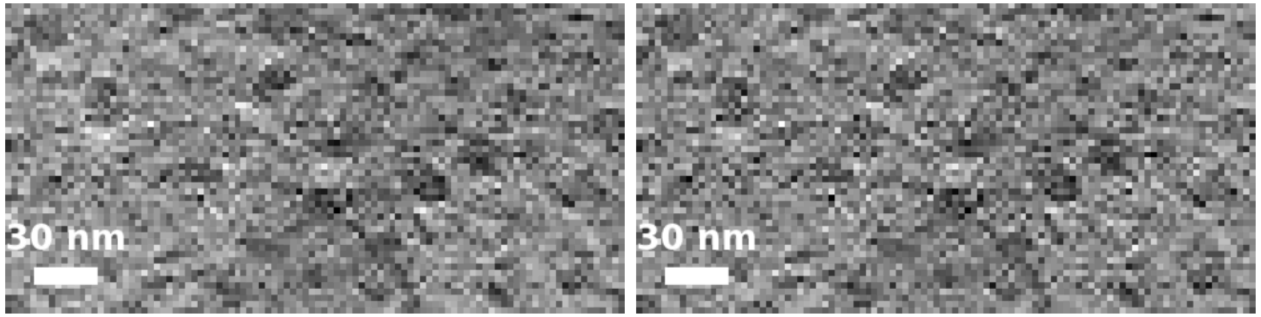


Training with simulated data is more challenging because there are significant mismatches between the two datasets, which result in severe artifacts. Ultimately, I added some regularization terms related to real data in the loss function, which successfully reduced these artifacts, although it also introduced some horizontal and vertical lines in the filter.

New Simulated Filter



(a) Chlamydomonas Result of optimized filter (b) Chlamydomonas Result of ODL's FBP



(c) Zoomed-in view of 4a (d) Zoomed-in view of 4b

At the same time, I also tried simulating projections that match the tilt angles of the Chlamydomonas sample using the SHREC dataset and retrained the model. Unfortunately, this approach did not yield a significant improvement, likely because the simulated projections differ too much from the real scenario.

Discussion and Conclusion

- Good results on SHREC data.
- Real data challenging due to limited training set and differences in data characteristics.

Future work may focus on:

- Acquiring more training data to enhance the network's generalization ability.
- Fine-tuning the model using a small amount of real data.
- Manually adding appropriate noise to reduce the distribution gap between synthetic and real data, thereby achieving better training and reconstruction results.

Thanks

During this three-month project, I have learned so much. I am grateful to everyone who has helped me, and I especially thank the following individuals for their attentive assistance and guidance:

- Dr. Ricardo Diogo Righetto
- Vinith Kishore
- Dr. Valentin Debarnot