Distributed Holographic Image Processing with Neural Networks

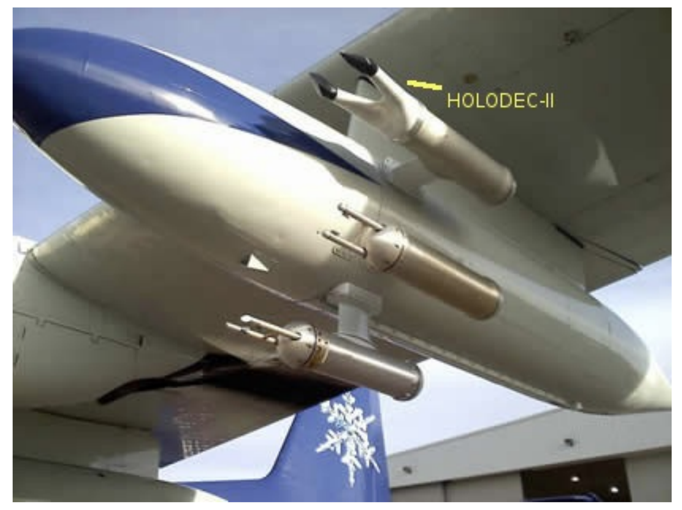
Jefferson Boothe[1,2,3], John Schreck[1], Matthew Hayman[1]

[1]NSF National Center for Atmospheric Research (NSF NCAR)

[2]University of Pittsburgh

[3]NSF Center for Space, High-Performance, and Resilient Computing (SHREC)

Introduction

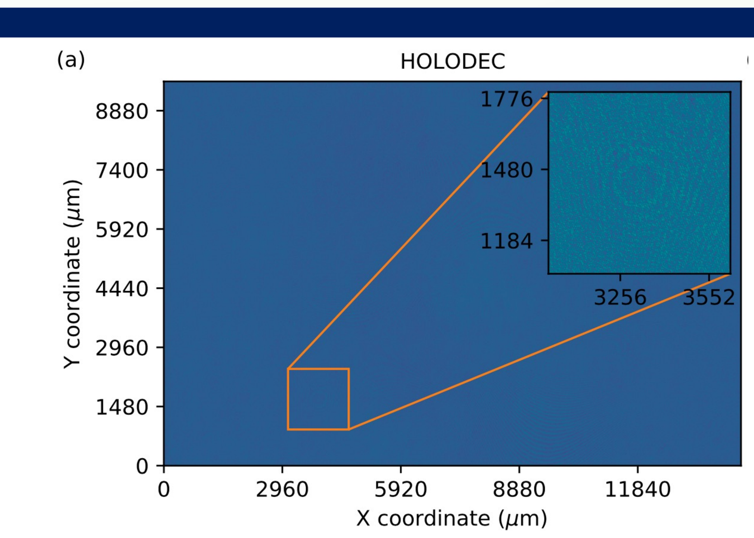


Goals:
- Improve performance and scalability of holographic image processing

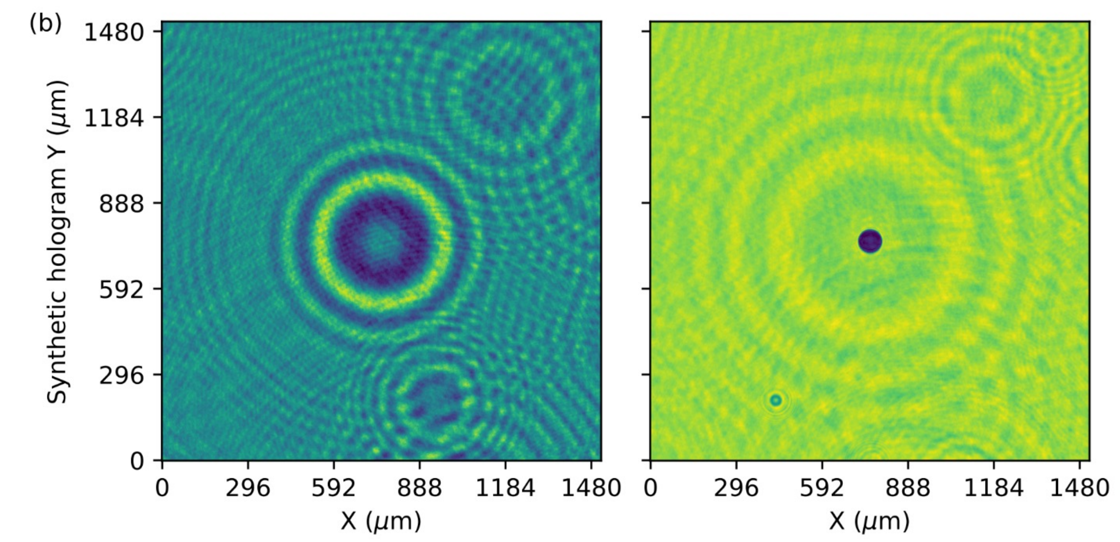
Motivation:
- ML has potential to process many types of holograms with minimal hand-tuning
- Standard method of hologram processing is computationally expensive

Challenges:
- Extremely large quantities of unlabeled data
- Each hologram represents 3D space to process

Holographic Data



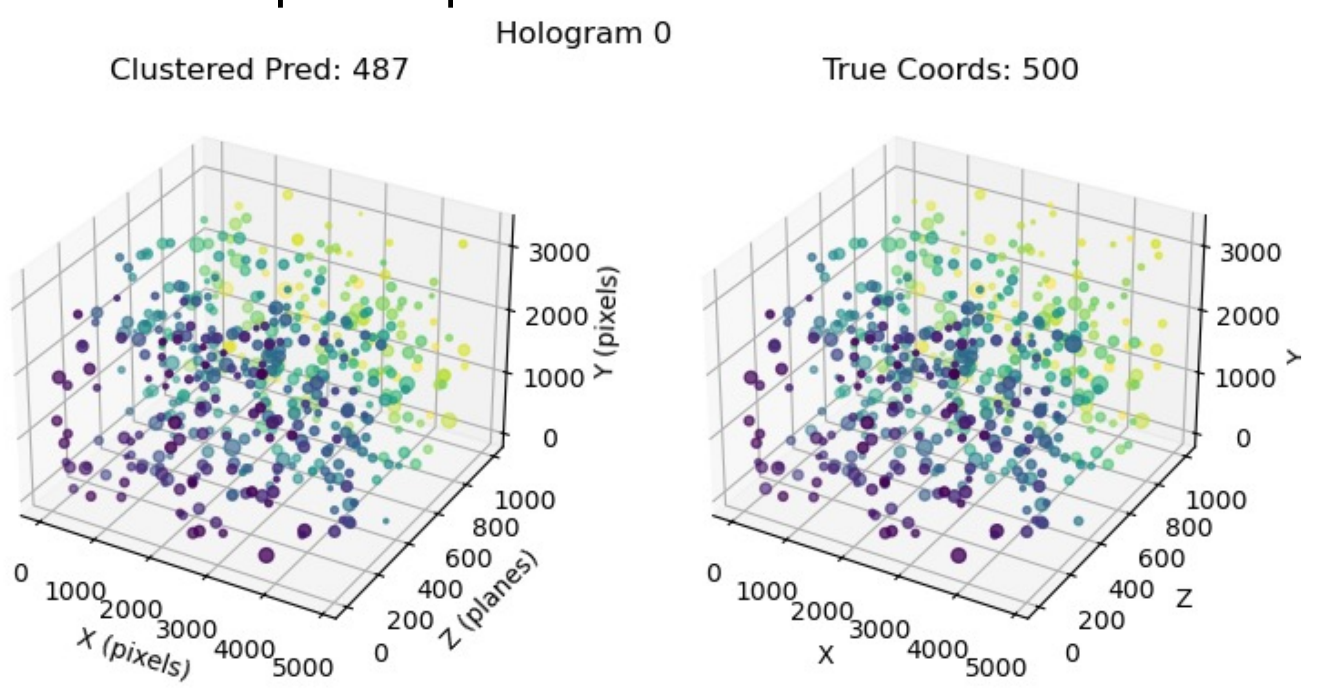
- Each hologram is megapixel in size
- 14.42 mm x 9.61 mm (2.96 μm/px)
- Holograms can be refocused to different depths
- Synthetic dataset matches HOLODEC properties and can be useful due to truth labels



Synthetic hologram at different z_j depths

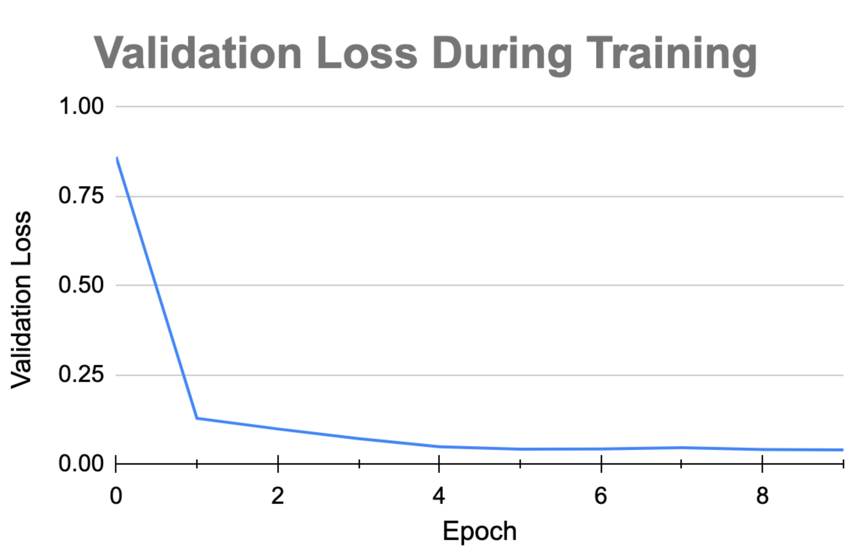
Inference

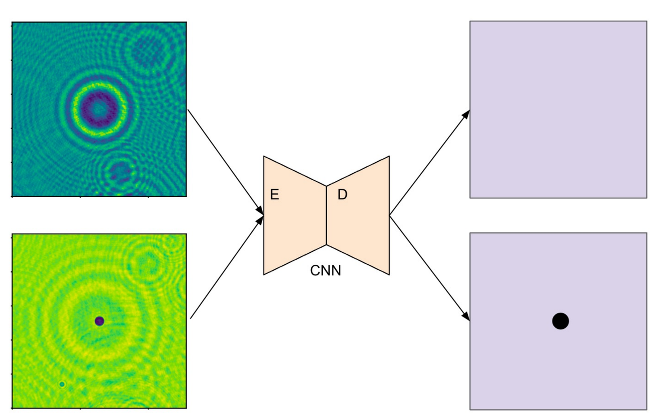
- Inference speed: 7 holograms/hr per NVIDIA A100
- Easily scalable across arbitrary number of GPUs
- Post-processing performed to extract particle coordinates and diameters from prediction masks
- Coordinates are clustered using OPTICS to remove duplicate predictions



Training

- Utilized Dice Loss to evaluate mask predictions against truth labels
- Hyperparameters tuned using ECHO over hundreds of trials
- Trained on 512x512 tiles of synthetic holograms
- Pytorch DDP utilized across 4 NVIDIA A100 GPUs on Derecho supercomputer





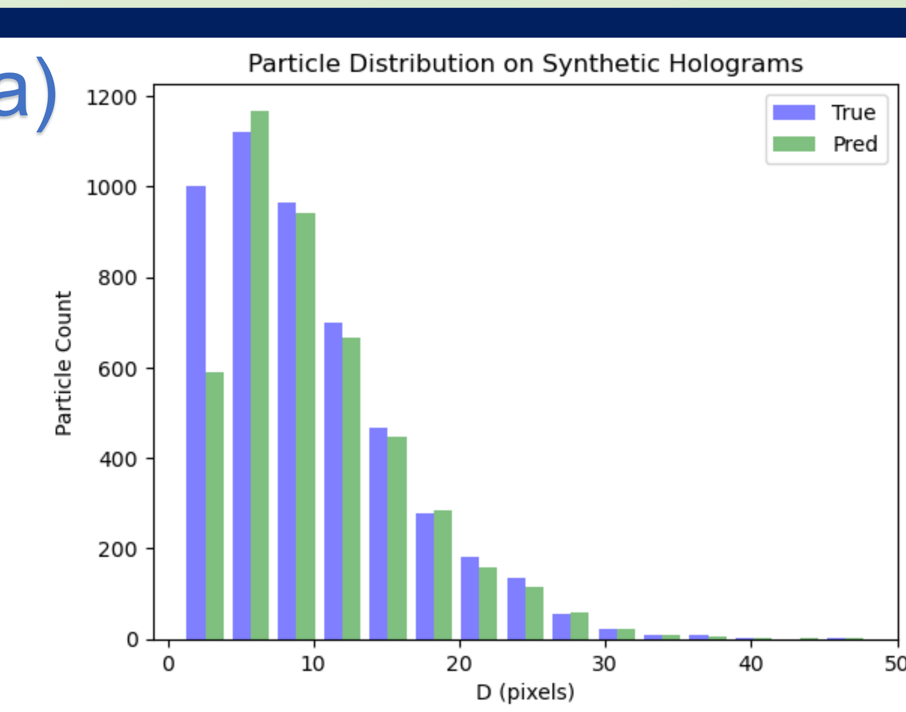
Semantic segmentation with CNN

Results

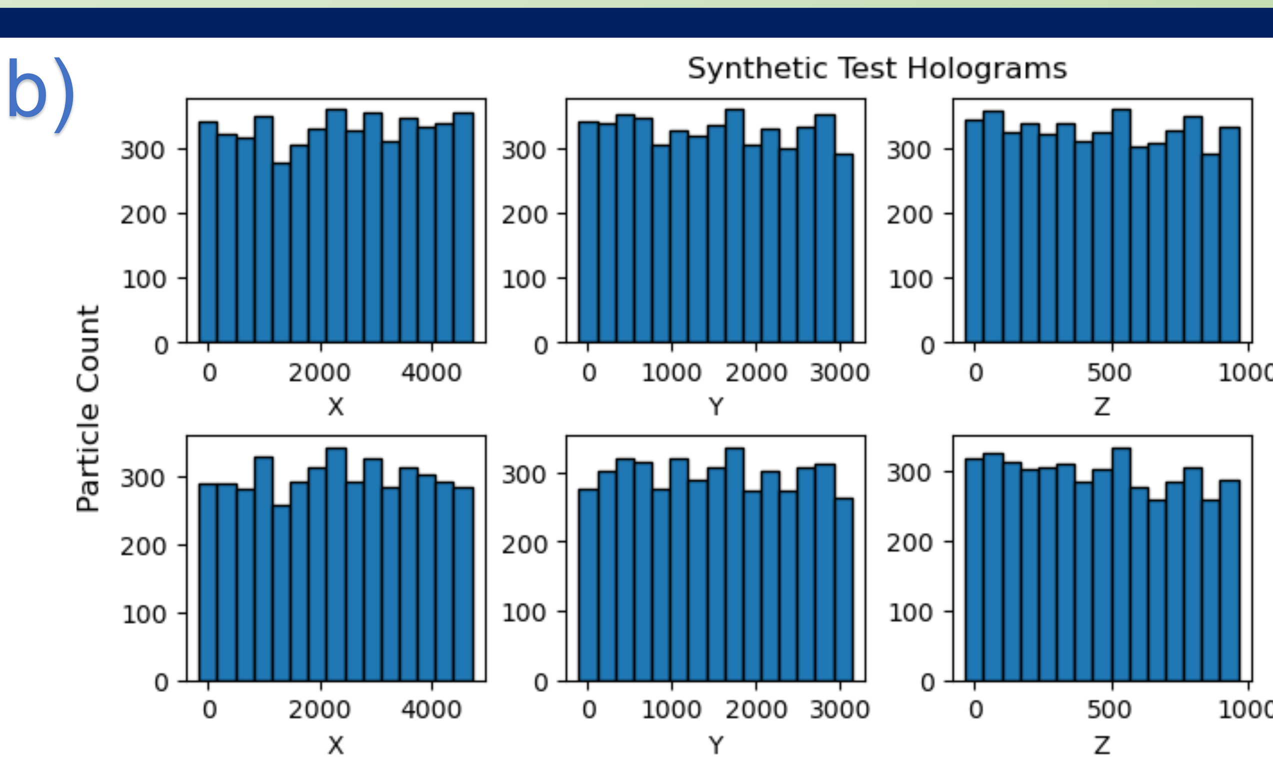
- Under-predicting particles with diameter below 3 pixels
- Larger errors visible at edges of image in X and Y
- Accurately predicting the number of particles does not correlate with accurate liquid density measurements

- Mean particle count difference: 10.0%
- Mean liquid density difference: 7.89%
- Mean effective radius difference: 6.52%

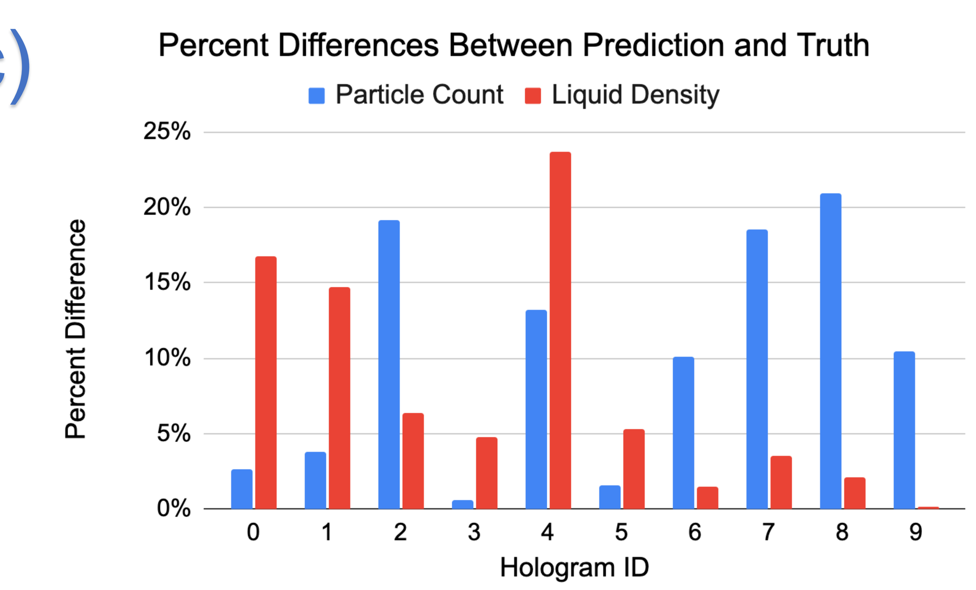
a)



b)



c)



a) A closer look comparing particles predicted and the truth organized by particle diameter

b) Particle count distributions across each variable; truth (top) and prediction (bottom)

c) Comparison of the percent difference between true and predicted particles for two properties

Future Work

- Utilize more complex model with depth lookahead and phase data built incorporated to improve particle prediction
- Create more diverse synthetic holograms to train on
- Compare performance against standard method on synthetic or CSET campaign data

Acknowledgements

This work was completed as part of the 2024 Summer Internships in Parallel Computational Science at NSF NCAR. I would like to thank my mentors and the MILES group for their continued support this summer. I also thank the SIParCS organizers, administrative team, and of course, my peers.