Wafer Map Interpolation

Luis Castaneda (luismicas@stanford.edu)

https://youtu.be/SuEEfDXGWlc

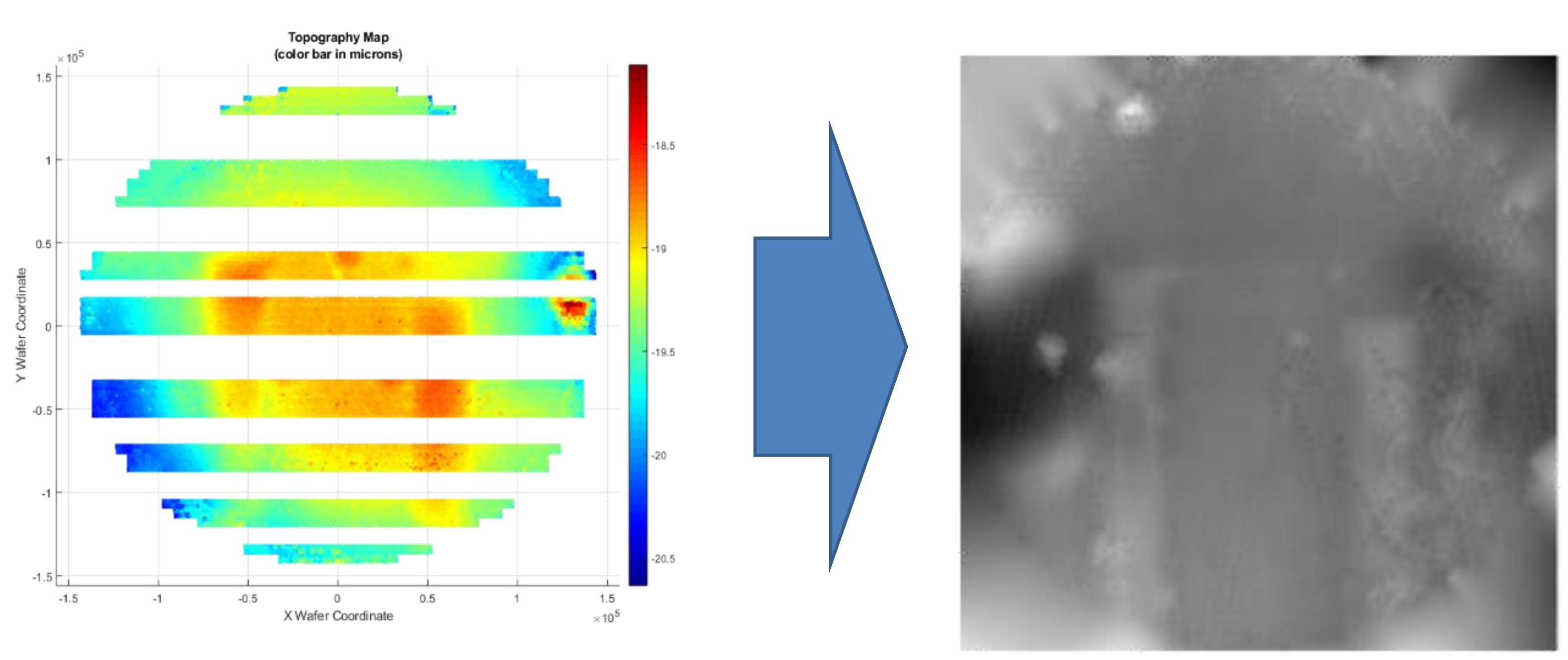
Stanford University

Purpose

Introduction

In the world of data storage, 3D NAND is at the forefront of memory technology, offering higher bit densities, faster write speeds, and lower power consumption. However, this architecture brings with it a myriad of new challenges. One of those main challenges is keeping wafers in focus. A work around for this problems involves creating a topographical map of the wafer which can be a very time consuming process to create through physical measurement.

I propose a novel application of deep partial convolutional networks to take topographical wafer maps with missing data and interpolate missing areas with meaningful context derived from surrounding data. By applying techniques from "Inpainting" algorithms, as well as enhancing the results with Single Image Super Resolution (SISR) algorithms, the goal of this project was to show that physical data collection could be reduced, while maintaining a high fidelity to the ground truth mapping.



(Input map and Final Image)

The Data

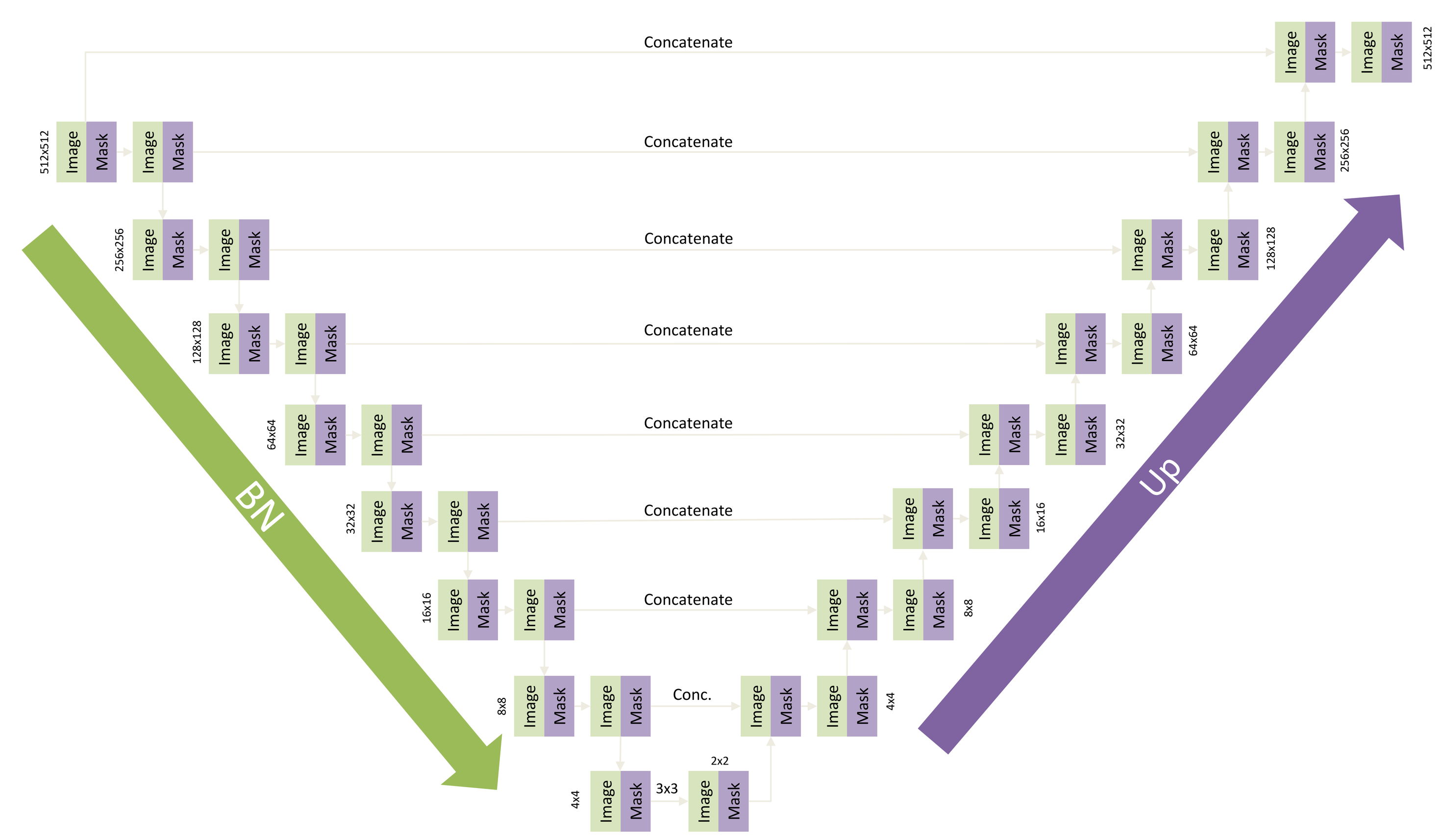
Data was collected on a few customer wafers which were available for study. Each wafer had a half-dozen or so mappings collected, these mappings were then converted from their original format into images which could be processed through existing CNN techniques. Each of these images was manipulated in various ways, including flip, blur, and translations to augment the total data set. In total there were 360 images used for training, 40 for dev, and 40 for test (all randomly sampled from the total data set).

Methods

InPainting

For this project I have elected to apply a technique developed by Guilin Liu, et. Al. from the NVIDIA corporation. This is a modification of the standard UNet architecture in Keras, but using partial convolutions and a mask updating step.

Total params:	32,876,128	Learning Rate: 0.0005
Trainable params:	32,865,248	Batch Size: 4



(Inpainting CNN Architecture)

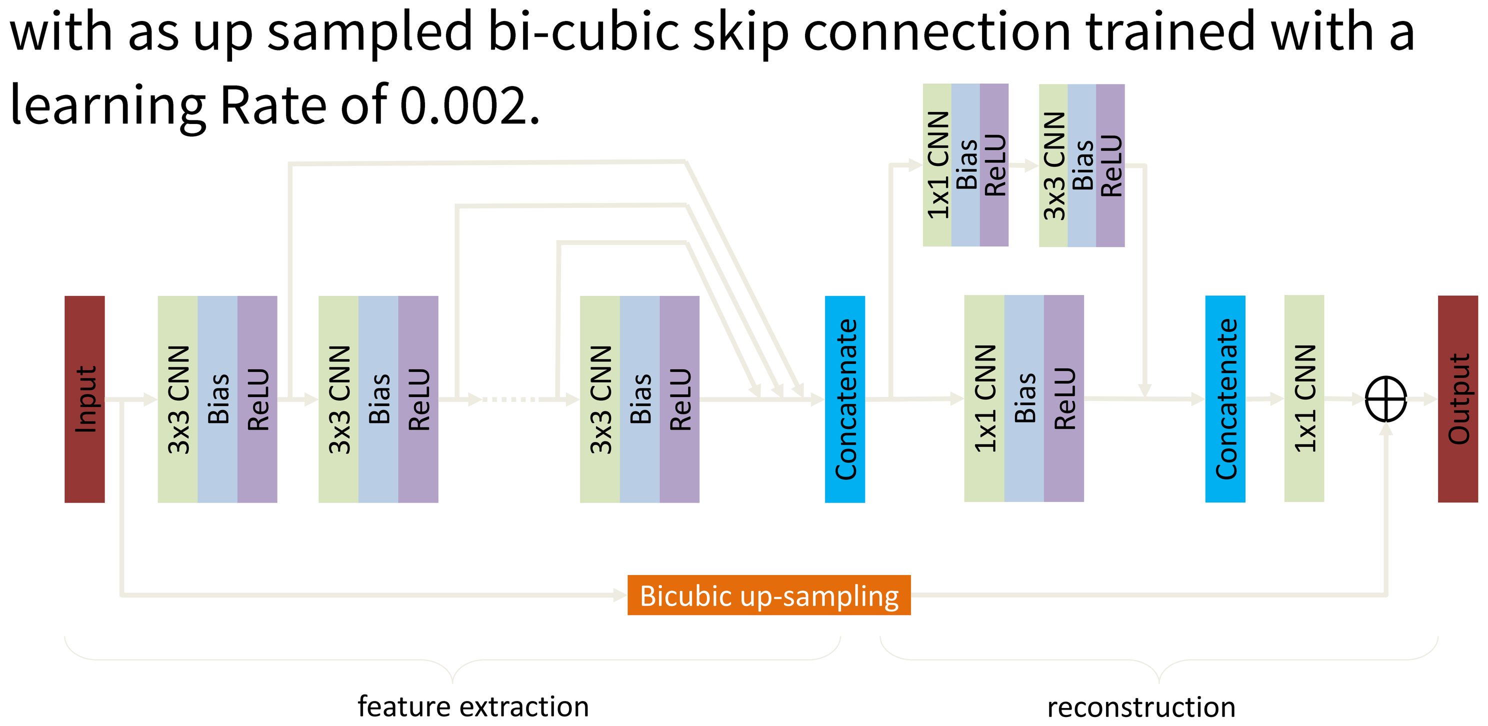
The loss function is a combination of 6 sub-loss functions.

$$L_{total} = L_{valid} + 6\, L_{hole} + 0.05\, L_{percept} + 120\left(L_{style_{out}} + L_{style_{comp}}\right) + 0.1\, L_{tv}$$

This architecture will recurrently update a mask as well as the image through each layer, reducing the size of the mask until its removed.

Upscaling

The goal of this step was to increase the contrast of the filled in context, as well as to upscale the image to provide adequate resolution for the final application. Standard CNN architecture with as up sampled bi-cubic skip connection trained with a learning Rate of 0.002.

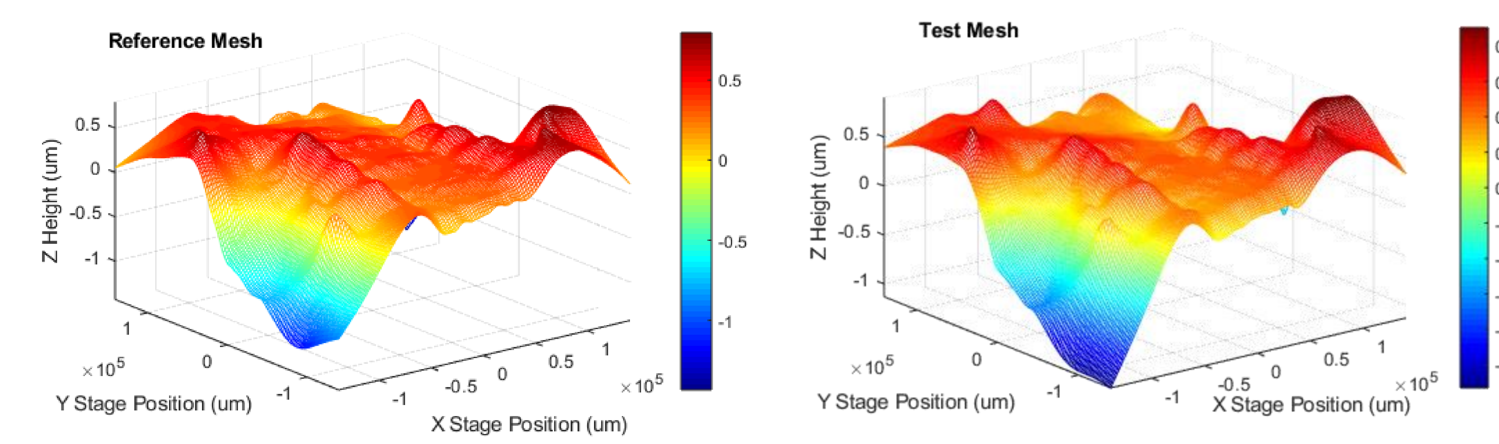


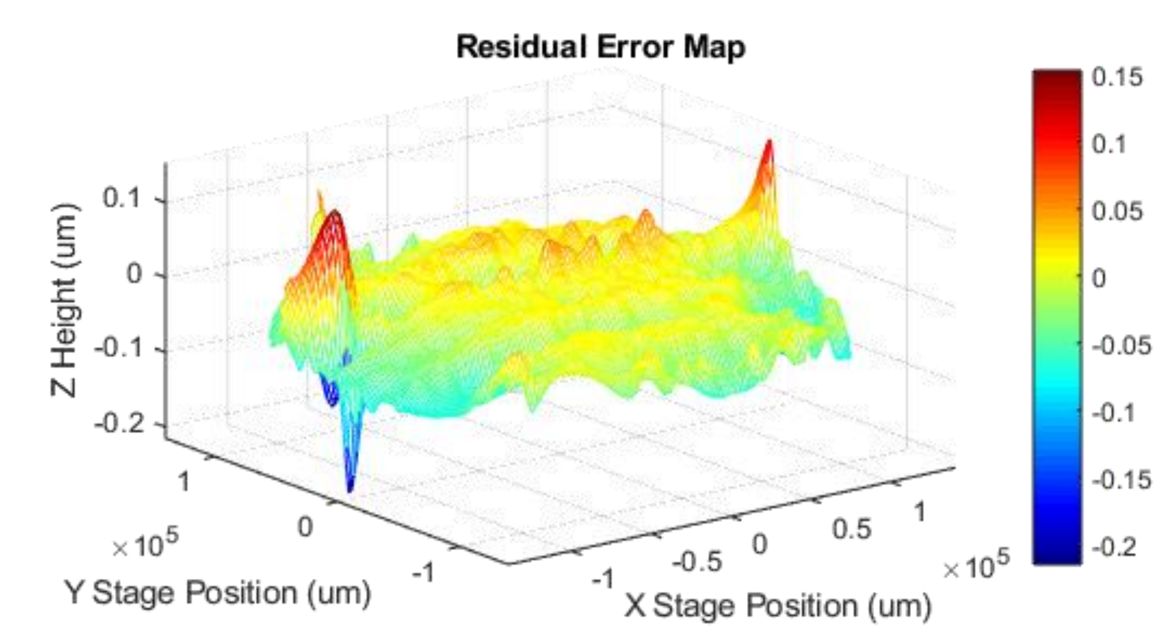
(Upscaling CNN Architecture)

Validation

Results

After interpolation, the upscaled images can be compared to the ground truth to calculate the residual errors. For this technique to be usable, the errors must be significantly less than the depth of focus of the inspection system (<90nm). The maps produced by this process were shown to be in tolerance.





Future Development

Discussion

This project demonstrated that the process is feasible, however there are a number of optimizations that could have been done. The UNet used for inpainting can be reduced significantly given that only the original data consists of 1 channel and that the context is simple and periodic. Additionally the upscaling could be integrated into the same network.

Acknowledgements

I'd like to thank the TA's of CS230 for their assistance and guidance, as well as the Coursera deeplearning.ai team.

References

[1]Guilin Liu, et. al. "Image Inpainting for Irregular Holes Using Partial Convolutions", https://arxiv.org/pdf/1804.07723.pdf

[2]Jin Yamanaka, et. al. "Fast and Accurate Image Super Resolution by Deep CNN with Skip Connection and Network in Network", https://arxiv.org/ftp/arxiv/papers/1707/1707.05425.pdf