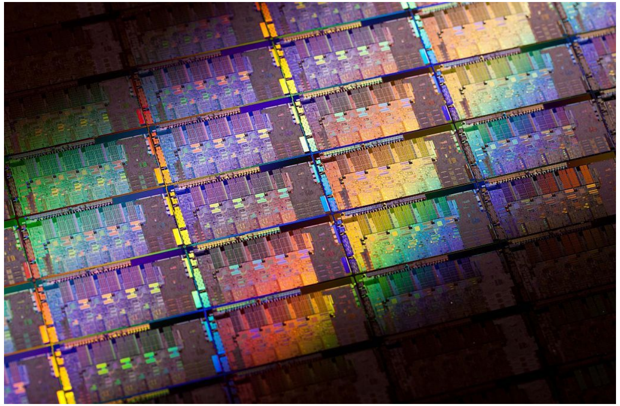
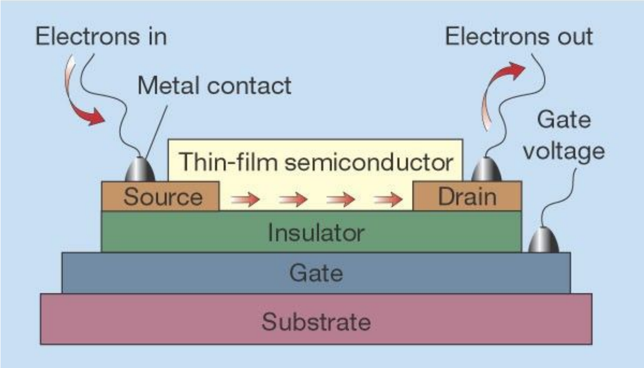
Determining the layer structure and thickness using model-based white light interferometry

Yian Cheng, Mingqian Ma, Changwen Xu
Nov. 26th 2024

Introduction :Thin Film Measurement





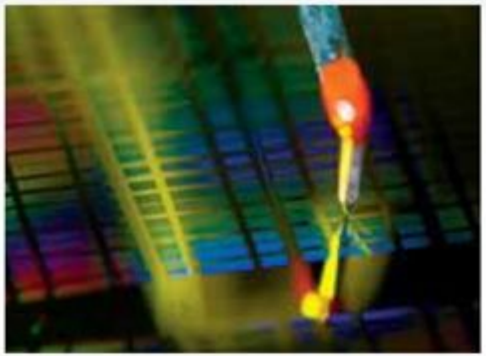
- Thin films are extremely thin (<1µm) layers of material applied to a substrate like metal or glass

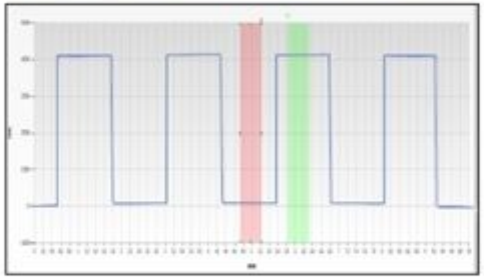
- Thin films play a key role in various industries like optics, electronics and material science, the thinness allow them have different properties than the bulk materials

- Thickness control is critical as it changes several properties of the film and the device

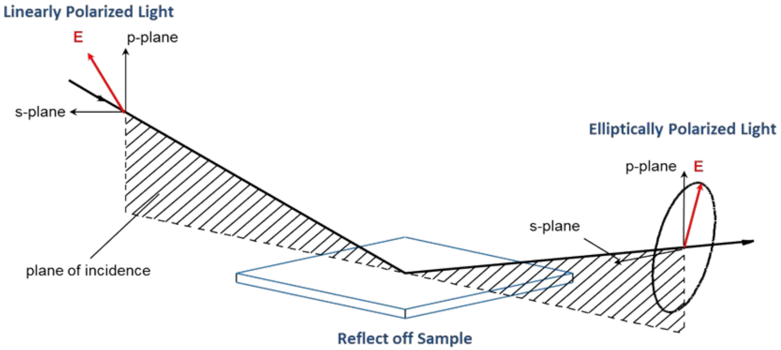
- We want to accurately measure the film thickness

Methods to Measure Thin Film Thickness

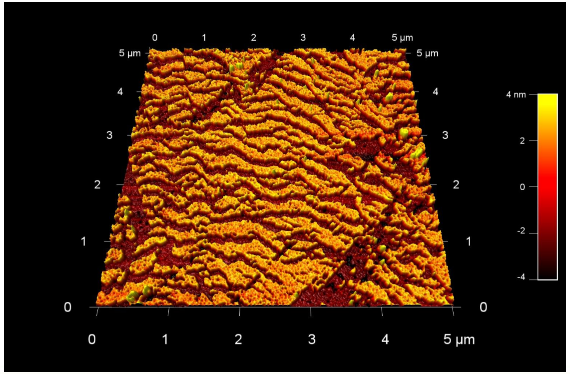




Stylus profilometry



Ellipsometry

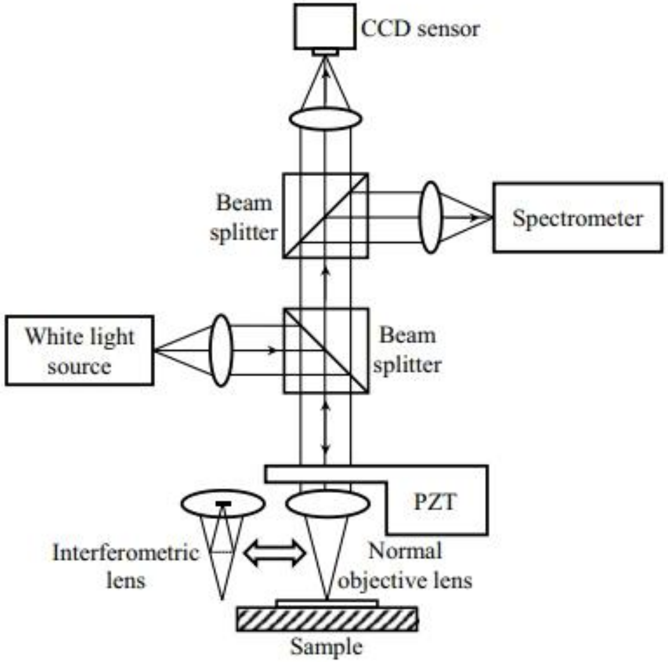


Atomic Force Microscopy (AFM)

White Light Interference

Advantages:

- Non-intrusive
- Cheap and fast
- Easy to set up



Inverse Problem: Create a fast and accurate model that can predict the structure and thicknesses of each layer of thin films based on the output of the signal

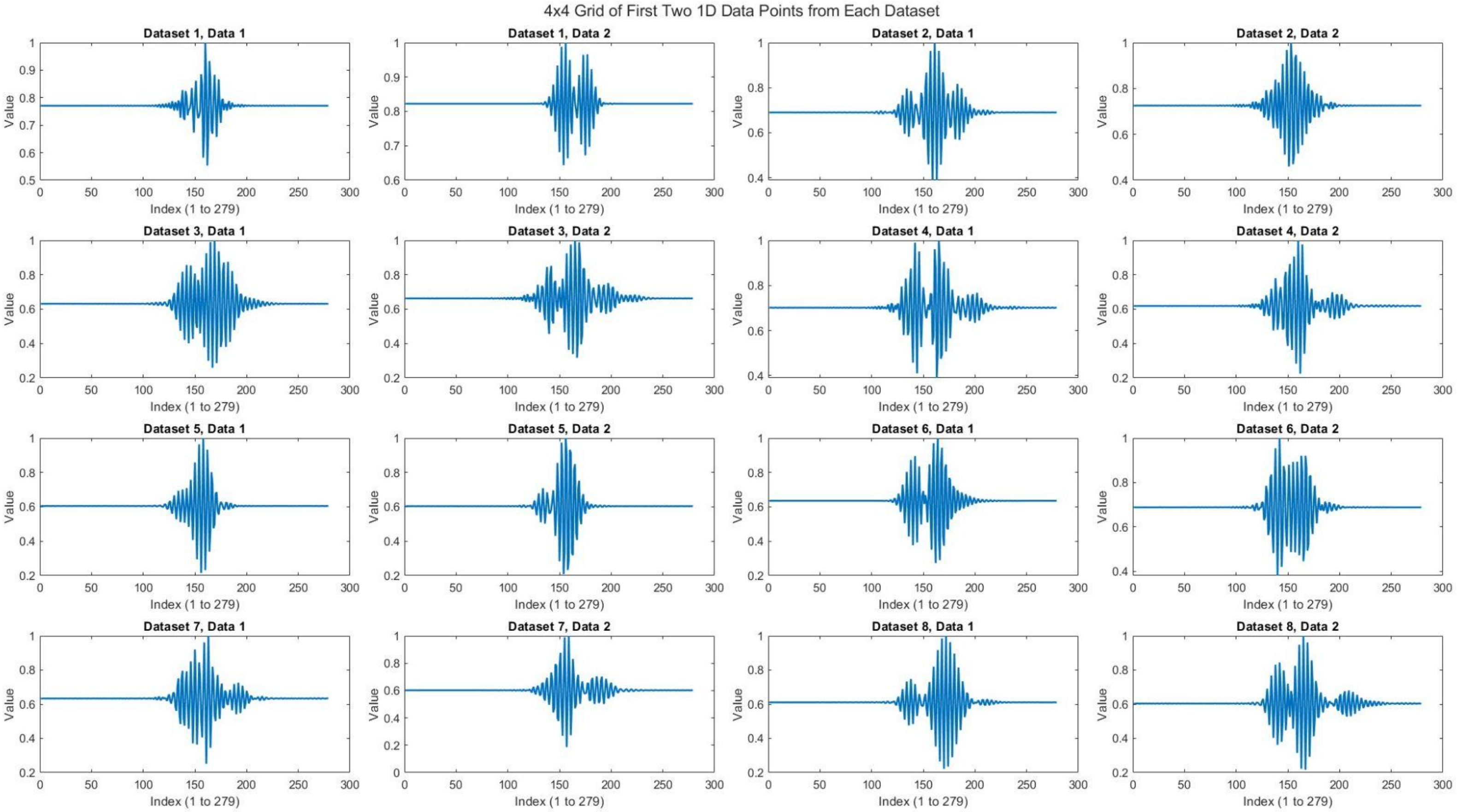
Dataset Generation

Table 1: Datasets generated with eight specific structures

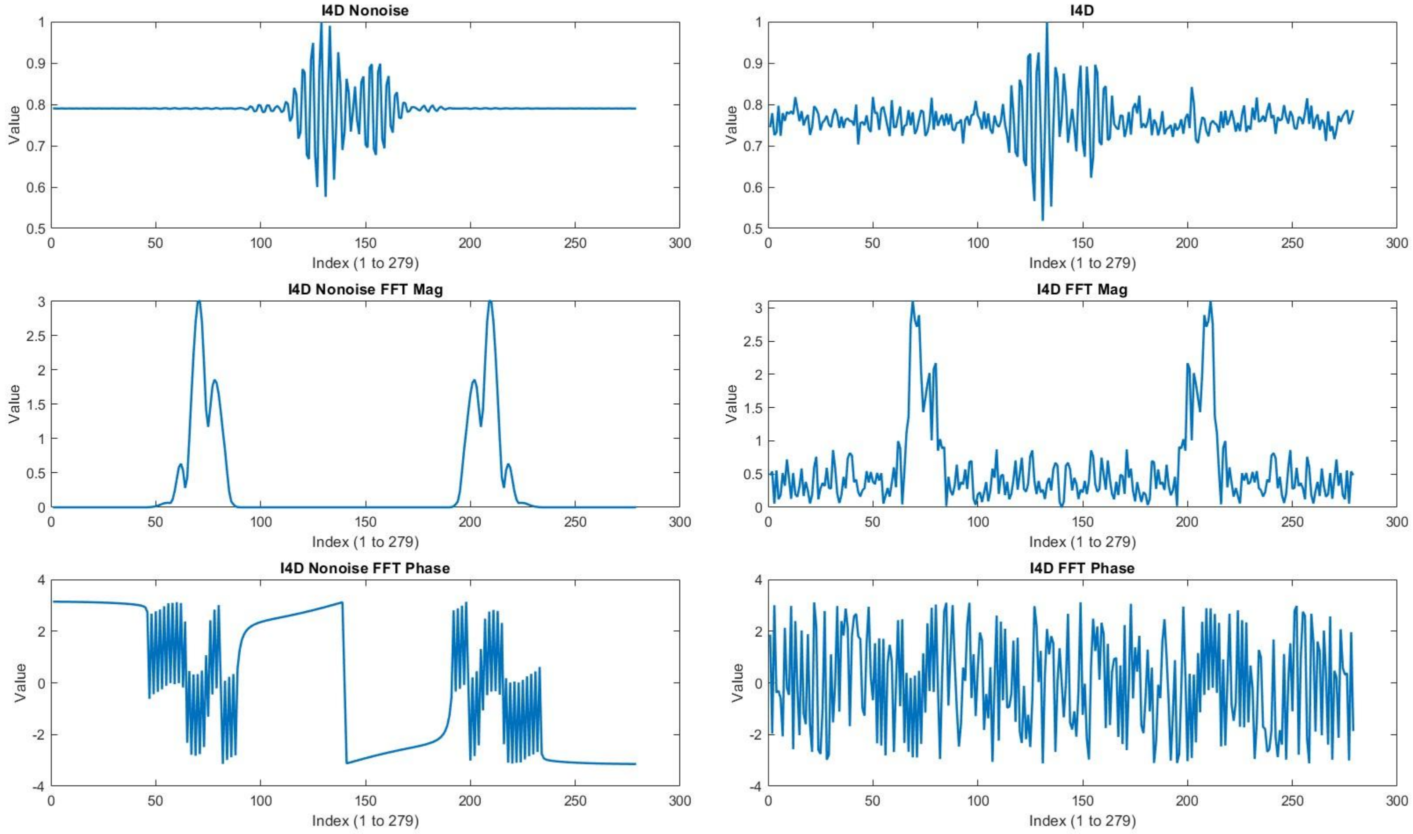
Dataset	Layers	Substrate
0	$SiO_2/Si_3N_4/SiO_2$	BK7
1	$SiO_2/Si_3N_4/SiO_2/Si_3N_4$	Glass
2	$SiO_2/Si_3N_4/SiO_2/p-Si/SiO_2$	BK7
3	$SiO_2/Si_3N_4/SiO_2/p-Si/Si_3N_4$	Glass
4	$SiO_2/Si_3N_4/SiO_2$	Si
5	$SiO_2/Si_3N_4/SiO_2/Si_3N_4$	Si
6	$SiO_2/Si_3N_4/SiO_2/p-Si/SiO_2$	Si
7	$SiO_2/Si_3N_4/SiO_2/p-Si/Si_3N_4$	Si

- Setup the parameter with different NA and wavelengths
- Generate layer structure and thicknesses and randomly select few combinations
- Simulation of the data with reflection coefficient matrix and TMM generation
- Random noises added

Dataset Visualization (1)



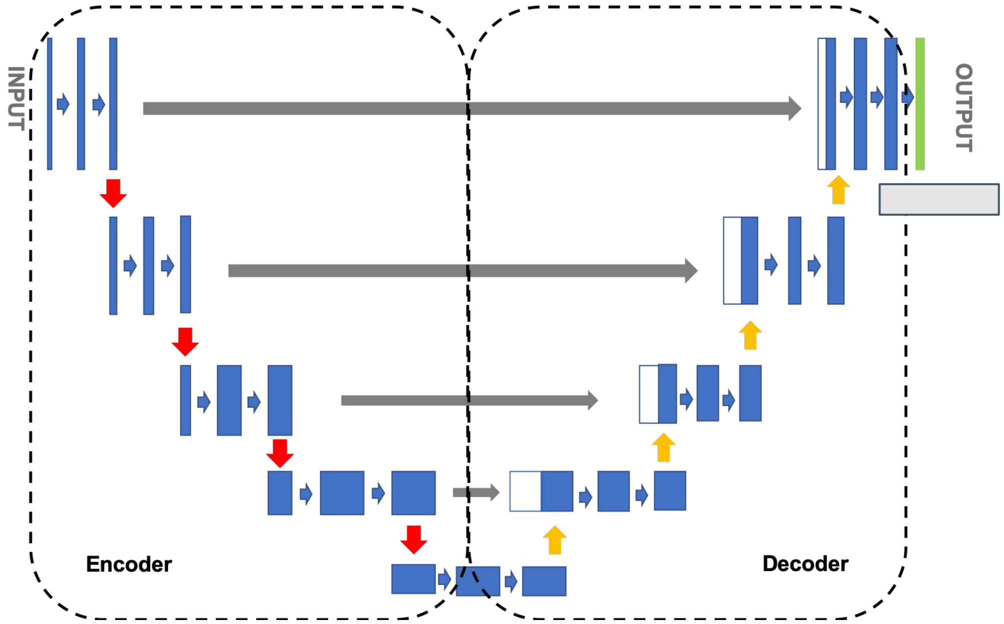
Dataset Visualization (2)



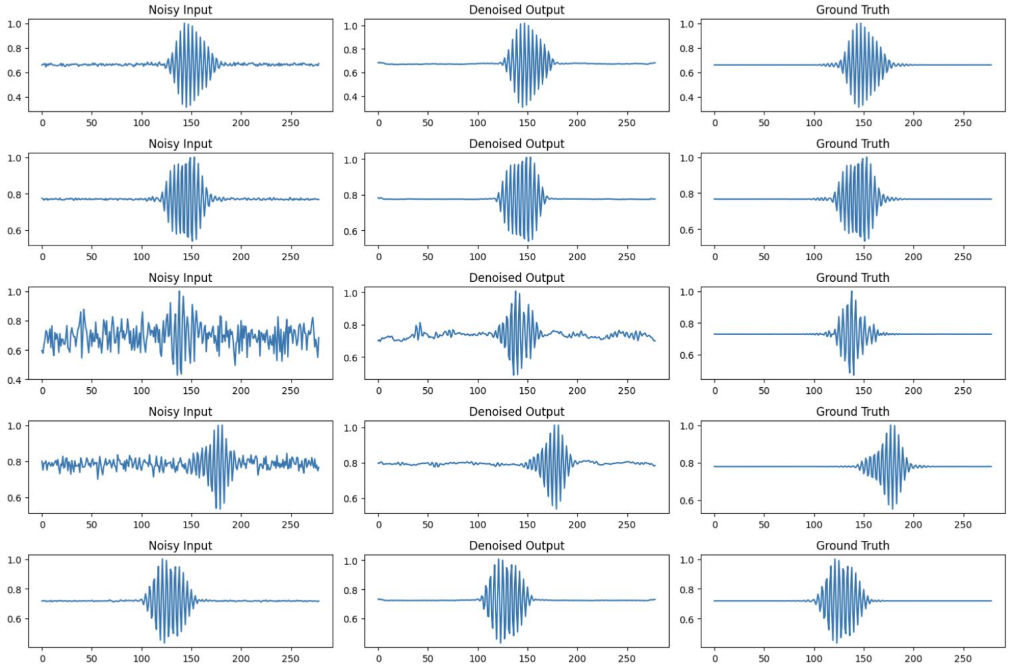
Signal Denoising - UNet

Applying U-Net to denoise the signal

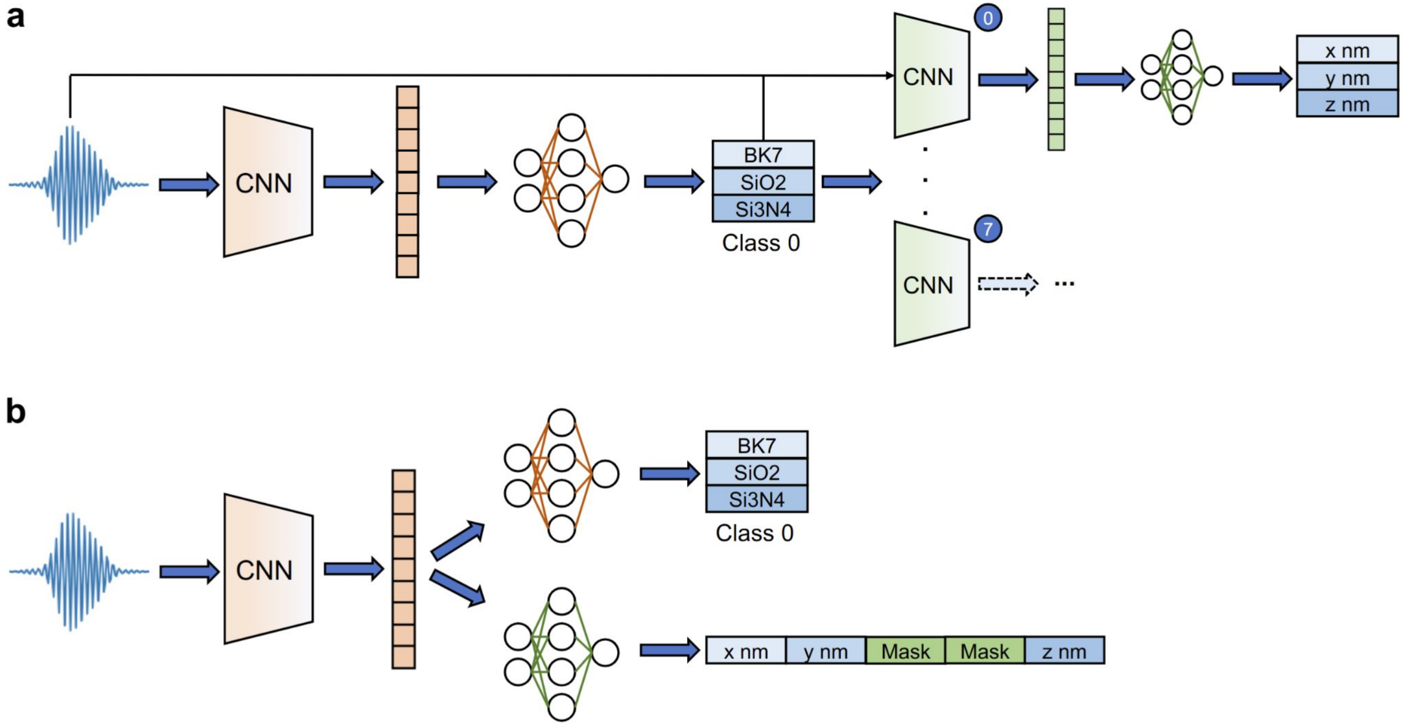
- Fully Convoluted Model
- Encode the signal then reconstruct
- Filtering various noise: Poisson, Gaussian noise



Signal Denoise Result - UNet

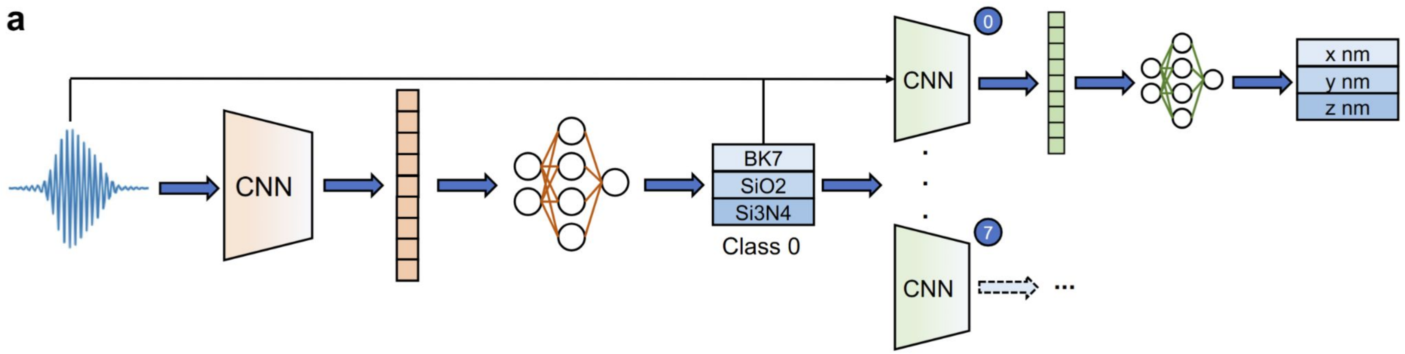


Model



a
CNN
CNN
BK7
SiO2
Si3N4
Class 0
0
CNN
x nm
y nm
z nm
7
CNN
...
b
CNN
BK7
SiO2
Si3N4
Class 0
x nm
y nm
Mask
Mask
z nm

Separate Model



Regression on a single architecture

Table 1: Mean Squared Error (MSE) of MLP and CNN on the test set of the four datasets we generated.

Dataset	Feature	MLP	CNN
I4D Small	No noise	0.0126	0.0070
	Noise	0.0134	0.0091
I4D Large	No noise	0.0128	0.0014
	Noise	0.0134	0.0077
I4D Cropped Small	No noise	0.0128	0.0125
	Noise	0.0129	0.0127
I4D Cropped Large	No noise	0.0124	0.0107
	Noise	0.0131	0.0126

Classification on eight architectures

CNN test accuracy:

- Trained on non-noisy data: 93.14%
- Trained on noisy data: 72.92%

MLP test accuracy:

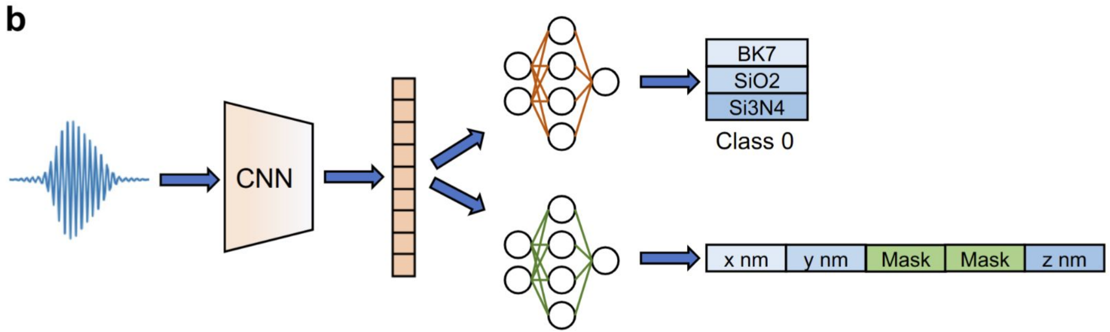
- Trained on non-noisy data: 33.86%

Regression on eight architectures

Table 2: Mean Squared Error (MSE) of CNN on the test set of the eight datasets we generated.

Dataset	Input feature	
	No noise	With Noise
0	0.0014	0.0077
1	0.0003	0.0051
2	0.0036	0.0056
3	0.0008	0.0057
4	0.0001	0.0035
5	0.0006	0.0054
6	0.0010	0.0055
7	0.0006	0.0049

Integrated Model



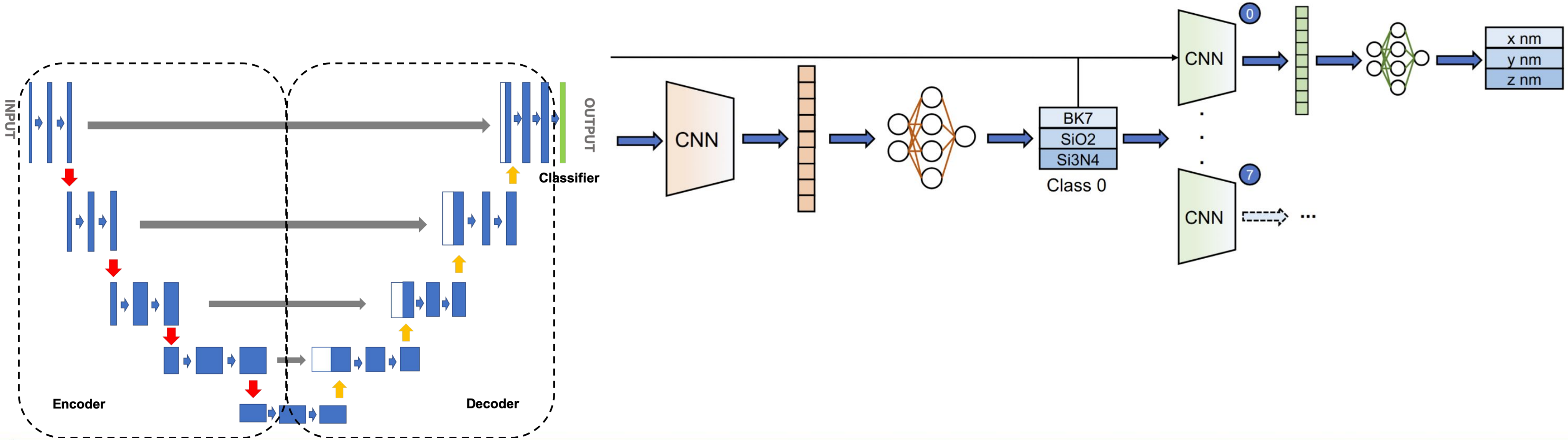
Integrated Model

- Test classification accuracy: 92.89%
- Test regression MSE: 0.2471

UNet + CNN

Table 3: Model performance of integrating UNet and CNN.

Input feature	Classification accuracy	Regression MSE
No noise	93.14%	0.0014
Noise	72.92%	0.0077
UNet denoised	71.72%	0.0075



Conclusion

We build a system which is a pipeline of models that can quickly identify the structure type and the thickness of each layer.

The model requires minimal data amount to achieve the purpose, even for data with noise injected.

Discussion and Future Work

- Q: Why hard to apply a universal model for different structures' regression?
  A: Inhomogeneity of different structures (#layers, optics property) make it hard to train an universal regression model.

- Q: Possible way to solve?
  A: Universal Architecture to capture the material properties and enable varying regression for different structures (predetermined).

- Q: Industrial standard of error tolerance?
  A: For single-layer film measurement, typical ellipsometry can achieve very accurate data of 0.1nm. However, the error goes to 1nm when the layers stack up. For multi-layer, it want to get several nm of error.