CS230

Computer-Vision-Enabled Optical Characterization of Two-Dimensional Materials

David Saykin
Department of Physics
Stanford University
saykind@stanford.edu

Shivanand Ramesh Pujari
NVIDIA
pujari@stanford.edu

Abstract

Computer vision approach to analyze microscope photographs of two-dimensional materials is proposed. The application framework is explained, literature overview is given and possible approaches to tackle the problem are put forward.

1 Introduction

Nanofabrication is a crucial part of processor design, material studies, and cutting-edge research in many branches of science. But acquiring even such basic information about pieces of materials as its geometrical form is time-consuming: it takes hours to take AFM-image of a flake on a substrate like the one shown on the right part of Figure 1. It's much easier to take photo of the flake under the microscope like you could see on the left part of Figure 1. This would be a great for everyone working such microscopic images if information about the physical properties of the sample could be extracted from the optical image alone. In fact, people who work with flakes on a daily basis are able to roughly estimate how thick the flake is just by looking at its photo, which gives a natural desire to dedicate this problem to a deep learning algorithm.

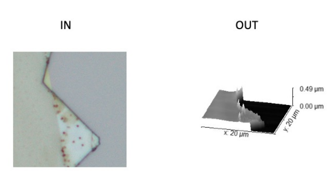

Figure 1: Model Concept

Input of our neural network would be an optical RGB–colored image taken by CCD camera connected to an optical microscope with 100x magnification with a resolution of as high as 256x256 pixels. And for the output, we would have a 2D-array of z-coordinates of each point of the sample. Each pixel of different color should be treated as a separate data sample, so one such pair of microscope and AFM image constitutes several thousand of labeled data points.

CS230: Deep Learning, Winter 2018, Stanford University, CA. (LateX template borrowed from NIPS 2017.)

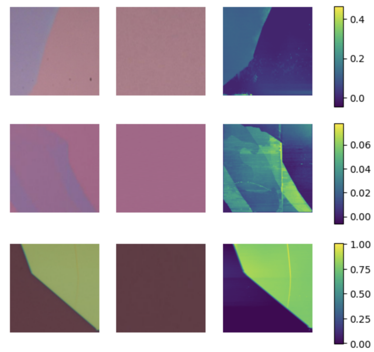

Figure 2: Example of the input data. First column is the optical image of the flake, second column is the image of the background extracted from the flake image. Third column together with the colorbar is the height profile.

2 Related work

In the Han, B.*et al*[1], authors have extensively worked on 2D material optical identification neural network (2DMOINet), where the model can predict, in real-time, various 2D materials physical properties in OM images regardless of variations in optical setups. But the paper limits to finding the material distribution the 3D rendering. So we are building model to enhance the more feature extraction along with density distribution projections. Adding to that [2] Long J, *et al* and [3]Ronneberger O, *et al* explore in building "fully convolutional" model to project inference and learning from a image so it coveres the groud truth extraction through Segmentation Architecture [6] J. Donahue **et al**. We will be using valuable insights from the Chen-Hsuan Lin **et al.**[5] work, which has fine details on the volumetric predictions using deep networks with 3D convolutional network. One of the limitation is paper delivers a volumetric rendering but does not provide much details on dimensions and physical properties specifics.

3 Dataset and Features

We have collected a dataset of 21 images of hexagonal boron nitride (hBN) with 256×256 resolution in RGB format paired with 21 topography profiles taken on atomic force microscope in application-specific format, which was then converted to 256×256 TIFF grayscale images, where the grayscale number represented the height of the flake.

Each input and output datapoint was rescaled to $[0, 1]$ insterval, which was achieved by choosing killo-angstrom as out height unit.

4 Experiments/Results/Discussion

4.1 vgg16- Regressor model

As part of the baseline model, we have used VGG16-unet model for feature extraction along with the linear-regressor as ouput height predictor are the floating numbers. We trained the VGG16 model and extracted features from the block1 cov2 layer and provided these inputs to the regressors as shown in 3. We have used MAE as a metric to define the performance of the models.Additionally we have done GRIDSEARCH parameter hyper tuning to get the best Ridge model parameters but it did not provide much improvement in the MAE metric.

4.2 vgg16- Regressor Model performance across different regressors

From the different regressors models, Ridge/RidgeCV gave the best performance with the MAE:0.04174nm4, which is quite a good performance for the 10nm microscopic images. From 5 we can see that model is predicting density distributions quite well but in 5th column (red) image we could see that model is not able to generalize the predictions.

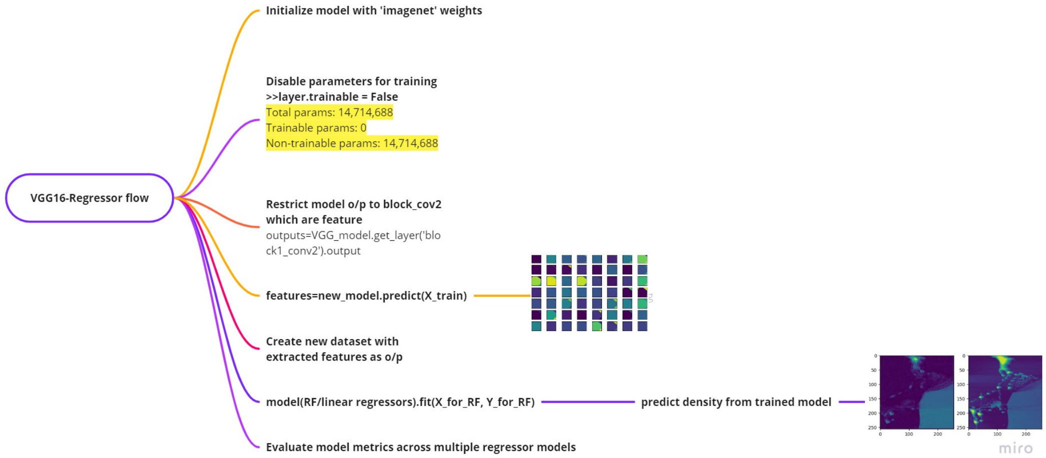

Figure 3: vgg16-Regressor model

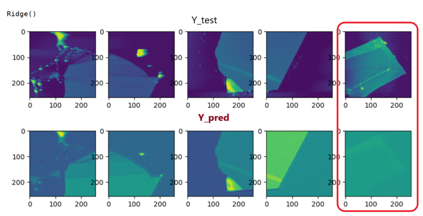

Figure 4: Baseline Ridge regressor best o/p

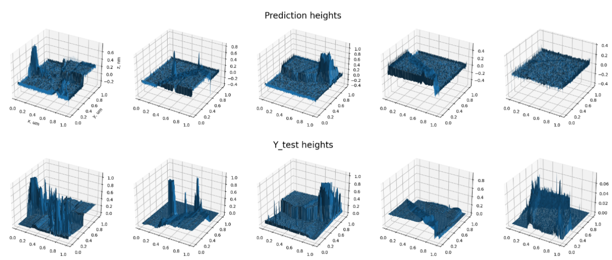

Figure 5: Ridge 3D height comparison o/p

	MAE(nm)	MSE	RMSE(nm)
Ridge	0.041740	0.008645	0.092980
RidgeCV	0.041734	0.008652	0.093014
LinearRegression	0.041738	0.008653	0.093022
ElasticNetCV	0.049363	0.008775	0.093675
LassoCV	0.048990	0.008800	0.093806
ElasticNet	0.130435	0.022245	0.149149

Figure 6: Baseline models performance (sorted descending)

4.3 Deep color-based model

As an alternative model we constructed a neural network which treats each pixel of the input image as a separate input point, and tries to predict the height of the flake at this pixel. The idea behind the model is to build a map from a color in the optical image reference by the background color to the height value. Since such map is expected to be nonlinear, deep NN structure is required. We have tried different value of the hyperparameters and concluded that the best result are achived by the NN with 3 layers, having 12, 6 and 1 nodes resceptively. First layer was chosen to be a 3×3 filter. SUmmary fo the model is provided in 7. We used mean-square loss function and mean-absolute-difference for our error. Our model has achives. Model training history is presented on 8. The visual example of the results is presented on 9. This model has improved the MAE to 0.03829nm (which was 0.04174nm in vgg16-Regressor).

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
input_5 (InputLayer)         [(None, 256, 256, 6)]     0
_________________________________________________________________
conv2d_8 (Conv2D)            (None, 256, 256, 12)      660
_________________________________________________________________
conv2d_9 (Conv2D)            (None, 256, 256, 6)       78
_________________________________________________________________
conv2d_10 (Conv2D)           (None, 256, 256, 1)       7
=================================================================
Total params: 745
Trainable params: 745
Non-trainable params: 0
_________________________________________________________________
```

Figure 7: Model summary. 3 RGB numbers from the optical image and 3 RGB numbers from the background are the input and a single number for the height of the flake in kÅ.

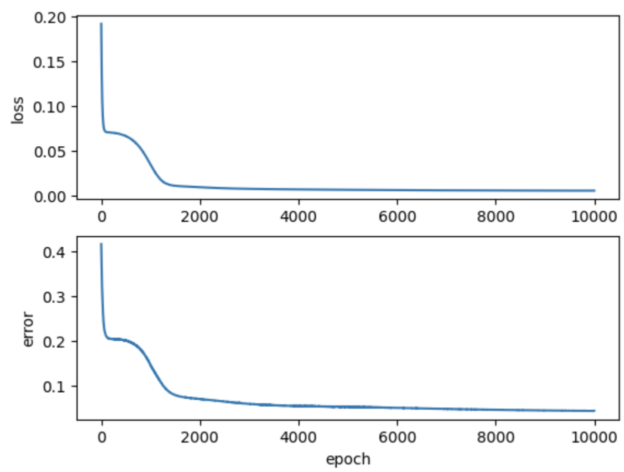

Figure 8: Model training history.

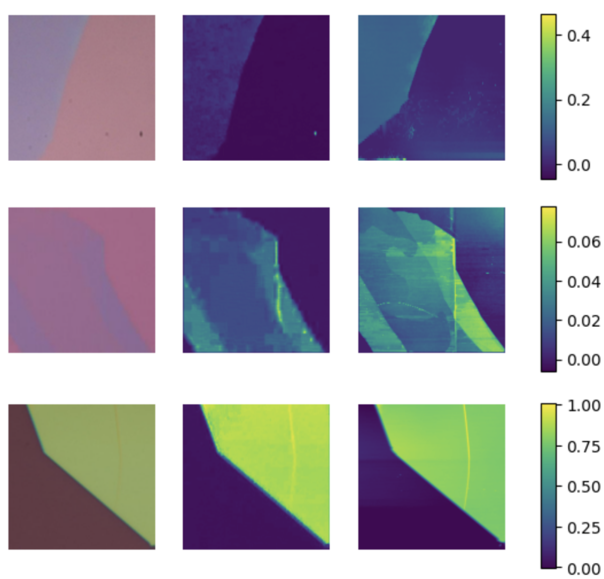

Figure 9: Example of the output data. The first column is the optical image of the flake, the second column is the prediction of the neural network. Third column is the true height profile. Colorbar is hared between the second and third columns.

5 Conclusion/Future Work

With our available data, Deep color-based model provided better performance than the vgg16-regressor model. However, vgg16-regressor runtime is quite low as it spends more time in the regression than the neural network training.

6 Contributions

This project was done in collaboration with David (Dept of Physics) and Shivanand (Nvidia). Shivanand has produced the baseline VGG16-Regrtessor and implemented and applied the ad-hoc feature extraction approaches. David has collected the dataset used and produced the original idea. David has developed a neural network-based model and fine-tuned it to the best performance model for use with feature subsets, and wrote the paper.

References

[1] Han, B., *et al* (2020). **Deep-Learning-Enabled Fast Optical Identification and Characterization of 2D Materials**. Advanced Materials, 32(29), 2000953.

[2] Long J, *et al* **Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)**; 2015 7-12 June 2015; 2015. p. 3431-3440.

[3]Ronneberger O, *et al* **Convolutional Networks for Biomedical Image Segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors.** Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015; 2015 2015//; Cham: Springer International Publishing; 2015. p. 234-241.

[4] Ferrari AC, **et al. Science and technology roadmap for graphene, related two-dimensional crystals, and hybrid systems.** Nanoscale 2015, 7(11): 4598-4810.

[5] Chen-Hsuan Lin **et al. Learning Efficient Point Cloud Generation for Dense 3D Object Reconstruction.** AAAI 2018: 7114-7121

[6] J. Donahue **et al. DeCAF: A deep convolutional activation feature for generic visual recognition**. In ICML, 2014. 1, 2

[7] Park, E. **et al.** 2017. **Transformation-grounded image generation network for novel 3d view synthesis. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition**