Image Classification of Dewetting Microscopy Using Artificial Neural Networks

[1] Kim C, [2] Sutrisno R, [3] Chua K, [4] Karim A, [2] Toti G

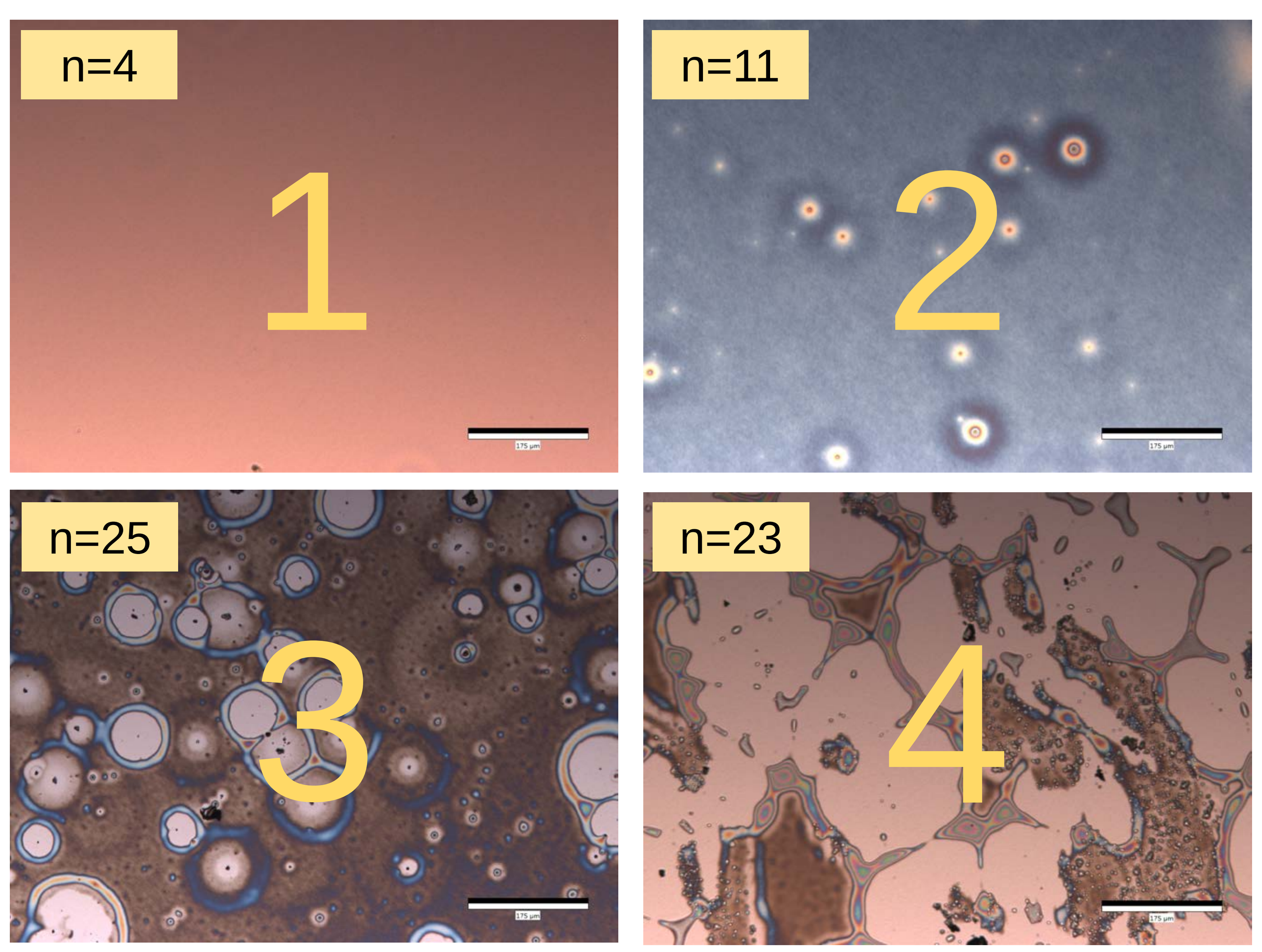
[1] Brown University, Providence, RI; [2] Department of Computer Science, University of Houston, Houston, TX;
[3]Cornell University, Ithaca, NY; [4] Department of Chemical and Biomolecular Engineering, University of Houston

ABSTRACT

Contemporary methods to analyze dewetting stages from optical microscopy are limited to manual classification. The project seeks to automate this process by using image processing techniques and machine learning. Magnitude independent features, such as pixel skew, variance, and entropy, along with their local deviations, were used to train a simple feed forward neural network. From a dataset of 64 images, tuning was achieved by selecting the neural network hyperparameter configuration with the highest peak cross validation score. The selected model accurately classified approximately 80% of the testing set.

BACKGROUND

- Dewetting is a process that leads thin polymer coatings to separate from their substrate. It occurs when conditions are more favorable for molecules of a liquid to coalesce.
- Chemical engineers are interested in determining what conditions bring to different degrees of dewetting.
- Machine learning can be used for classification of dewetting stages.



Figs. 1-4: Dewetting Stages (None, Early, Intermediate, Late)

FEATURE EXTRACTION

1. Convert to **grayscale** and **saturation** images
2. Apply Mean Filter Normalization (MFN) to reduce microscopy artifacts

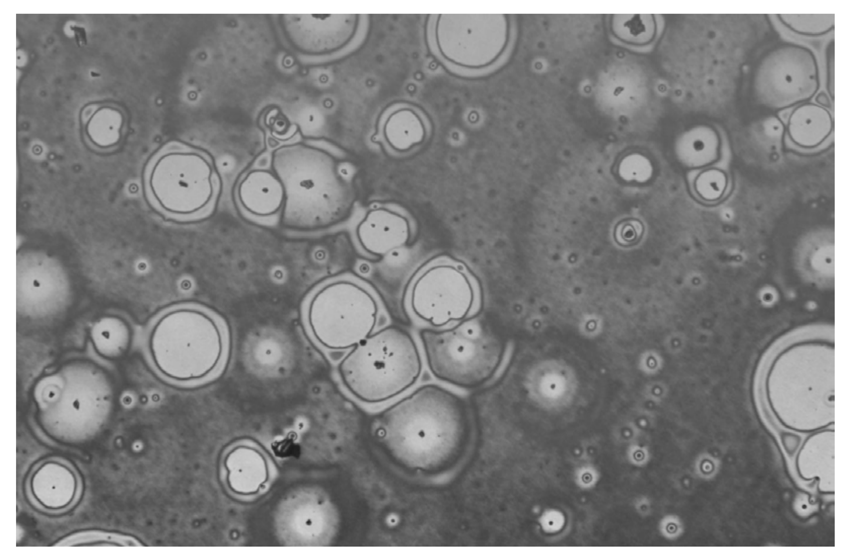


Fig. 5: Grayscale Image After MFN

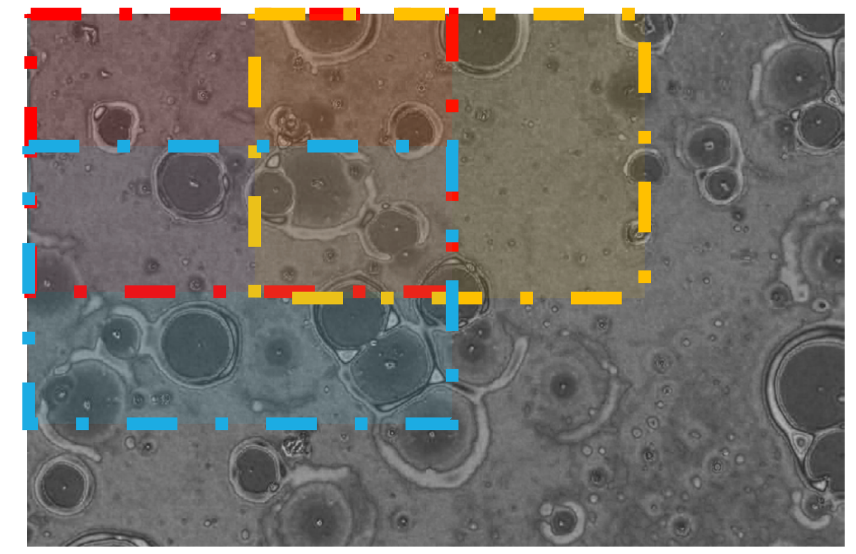


Fig. 6: Saturation Image after MFN with Striding Grid Representation

3. Calculate magnitude independent features from entire image

intensity variance	**entropy**	**skewness**
σ^2	$-\sum_{i=0}^{255} p(i)\log_2 p(i)$	$\sum\left(\frac{X-\mu}{\sigma}\right)^3$

4. Quantify homogeneity by calculating **local** deviations of features with respect to **global**

= 12 features

NEURAL NETWORK (NN)

Grid Search Hyperparameter Space:

Architecture

12 → Hidden Layers → 4

- 1st Hidden Layer Nodes: **10**, **8**, **5**
- 2nd Hidden Layer Nodes: **8**, **5**

Regularization

- **None**
- **L1**
- **L2**

→ Penalties @ **10^{-3},10^{-4}**

= 30 models

5. Measure training and cross-validation (*k*=10) scores every 25 epochs up to 800 epochs
6. Compare optimal cross-validation scores before overfitting across models

RESULTS

Highest CV Model

Architecture: (12,8,8,4)

Regularization: L1(10^{-3})

Epoch: 500

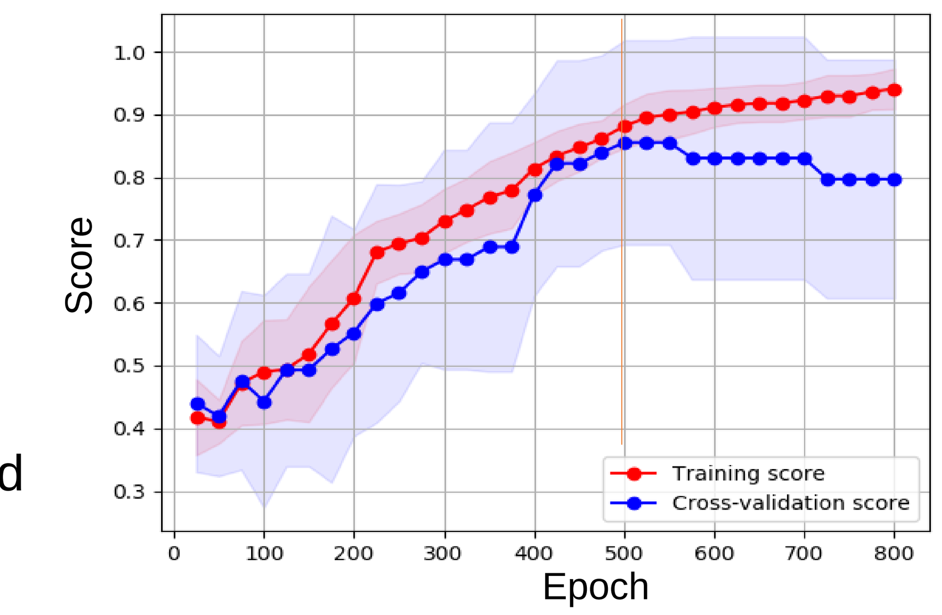
Accuracy Metrics

*20% of dataset was reserved for testing

Training: 0.882 ± 0.035

Validation: 0.855 ± 0.163

Testing: 0.786



		Predicted			
		1	2	3	4
Actual	1	1	0	0	0
	2	0	1	2	0
	3	0	0	4	0
	4	0	0	1	5

Fig. 7 (top): Epoch Curve for Selected NN; shaded regions represent $\mu \pm \sigma$
Fig. 8 (right): Confusion Matrix

CONCLUSIONS

1. Meaningful supervised learning is evident.
2. Decision boundary between class 2 and 3 needs improvement.
3. Proximity of training and cross-validation accuracies suggest model is not overfitting to the training sample.

FUTURE WORK

1. Increase training set size to enhance performance
2. Test robustness of model across different polymer and image acquisition methods.

ACKNOWLEDGEMENTS

This research was made possible by a NSF grant to the University of Houston Department of Computer Science.

REFERENCES

- Meredith, J. C., Smith, A. P., Karim, A., & Amis, E. J. (2000). Combinatorial materials science for polymer thin-film dewetting. *Macromolecules*, *33*(26), 9747-9756.
- Raschka, S., & Mirjalili, V. (2017). *Python machine learning*. Packt Publishing Ltd.