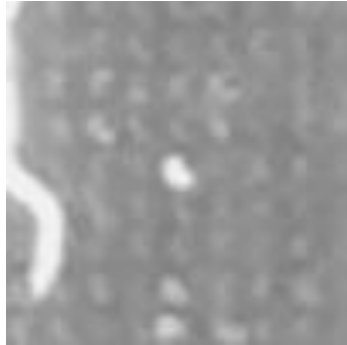


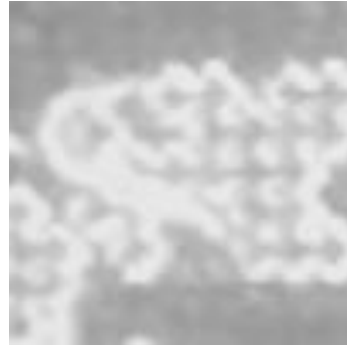
Pattern Recognition: Project 4

John Stroupe, David Hoefs, Alyssa Watson

Training Set



The above image is an example of what we decided should be considered a worm. We decided part of a worm still resembles a worm-like object. So, these images were moved from 0 to 1.



The above image is an example of what we decided should not be considered a worm. While it looks like there may be a worm, it is inconclusive to us. Images as bad as this moved from 1 to 0.

- Old Dataset Size:
 - o 0: 5500
 - o 1: 5500
- New Dataset Size:
 - o 0: 5454
 - o 1: 5546

Pre-Processing

- PCA
- Gamma Intensity Transform

Parameters of SVM

- Kernel Type: 'rbf'
- C: 20

Training/Validation/Test

- For model refinement, we chose 80% training, 0% validation, 20% test. The 0% validation is because we tried K-Fold against 20% unforeseen data and saw no obvious improvement over using 80% to train normally.
- We looped through various values of C for 'rbf', 'linear', 'polynomial' against the test set in multiple times each to build an average for each combination.

- The final turn-in model was trained with 90% of data and 10% test for a final verification. We decided 10% was sufficient for final testing because the dataset size was, in our experience, large enough to be well represented with 10%.

Confusion Matrix

		Actual	
		Worm	Not Worm
Predicted	Worm	502	47
	Not Worm	45	496

- Test Accuracy: 91.6%
- Sensitivity: .92
- Specificity: .91

Times

- SVM Training Time: 7.25 seconds
- SVM Testing Time: 0.75 seconds
- PCA Time: 5.56 minutes

Dependencies:

- opencv-python
- sklearn
- numpy