MACHINE LEARNING FROM MULTIMODAL DATA

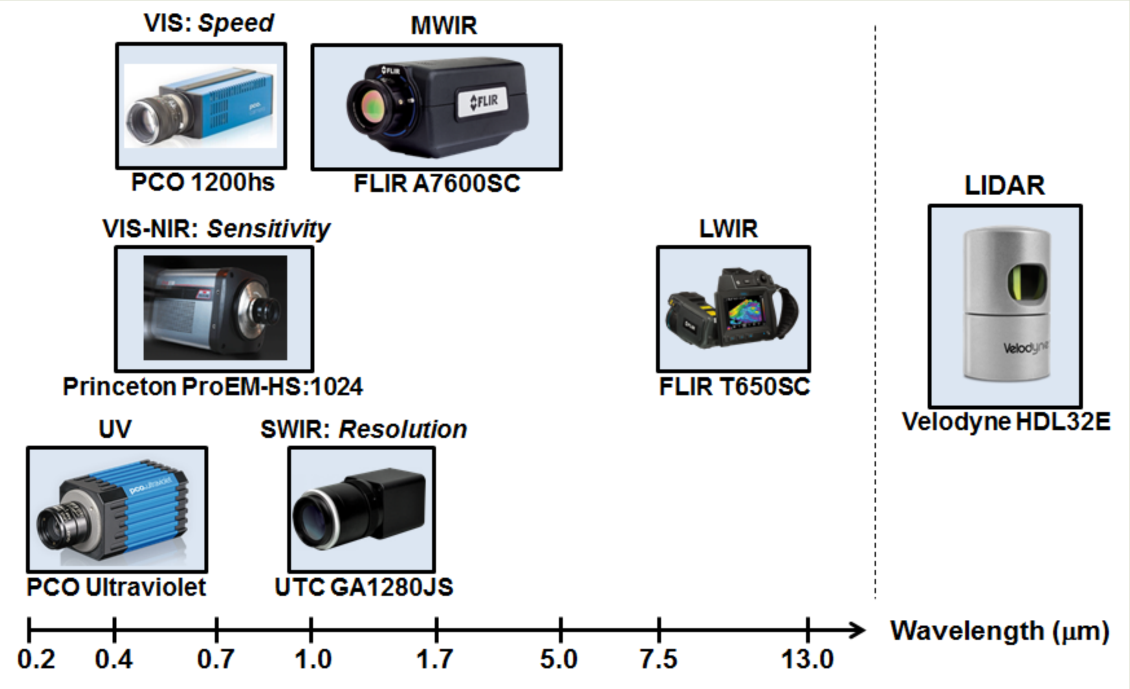
Andrea Burns

Dr.
Waheed Bajwa

Dr.
Mark Pierce

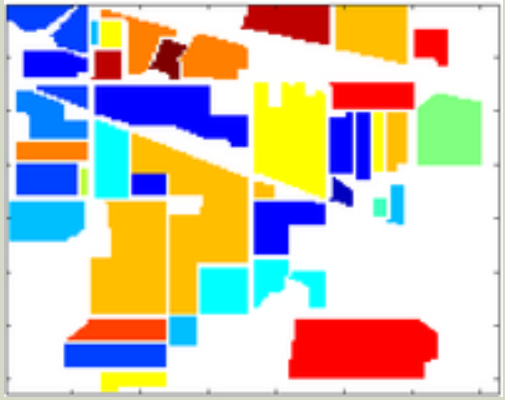
RECAP: MOTIVATION

- Investigating the impact of using multimodality in the process of image classification
- Our form of multimodality: hyperspectral imaging
  - Each spectral band forms a different mode



PROGRESS

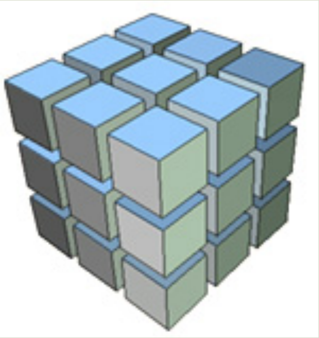
AVIRIS Indian Pines Data Set: Hyperspectral imaging



Class #	Class	# Samples
1	Alfalfa	46
2	Corn-notill	1428
3	Corn-mintill	830
4	Corn	237
5	Grass-pasture	483
6	Grass-trees	730
7	Grass-pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-notill	972
11	Soybean-mintill	2455
12	Soybean-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-Grass-Trees-Drives	386
16	Stone-Steel-Towers	93

PROCESS

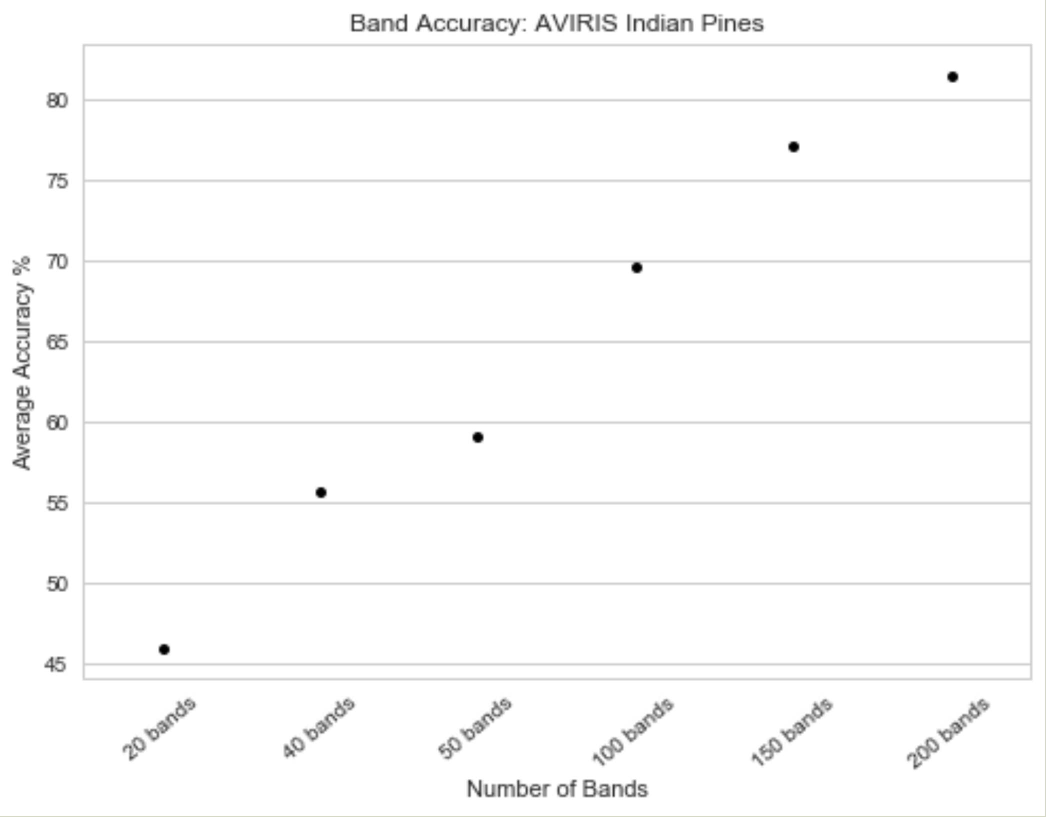
- Data set includes:
  - Image of Indian Pines is 145 x 145
  - Data is 145 x 145 x 200, 200 spectral bands
  - Each band is 10µm



- Classified the 12 largest classes (dropping the 4 smallest to improve accuracy)
  - SVM with one against rest multiclass-classification
  - 80% training, 20% testing – randomly shuffled and split

RESULTS

- Calculated accuracy of classification for 20-,40-,50-,100-,150-, and 200-band data, as well as with different ranges for each of these # of bands.

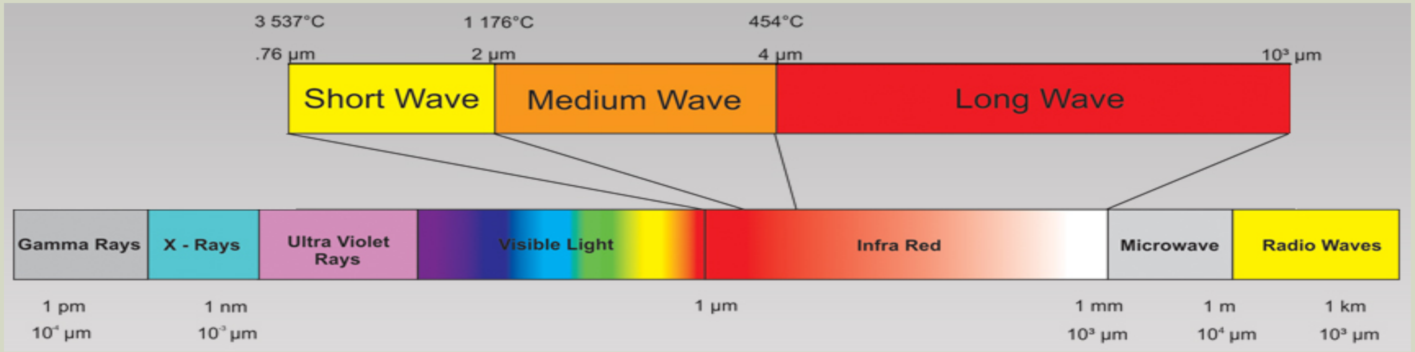


ACCURACY RESULTS

- <u>20 bands</u> ~35-54%
  - Best range: .2-.4μm, 54.975%
- <u>40 bands</u> ~49-60%
  - Best range: 1.45-1.99μm, 60.338%
- <u>50 bands</u> ~54-62%
  - Best range: 1.2-1.89μm, 62.761%
- <u>100 bands</u> ~68-71%
  - Best range: .2-1.2μm, 71.341%
- <u>150 bands</u> ~74-79%
  - Best range: .2-1.89μm, 79.72%
- <u>200 bands</u> ~79-83% (.2-2.39μm)

RESULTS

- The number of spectral bands affects accuracy
  - The more bands used – the higher accuracy



- The range of spectral bands affects accuracy
  - Certain ranges (of the same number of bands) perform slightly better than others
  - Suggests that different bands can extract more useful information from an image for classification than others

IMPORTANT OBSERVATIONS

- The hay-windrowed class consistently performed best.
  - It was shown that even as the number of bands decreased, it had high accuracy.

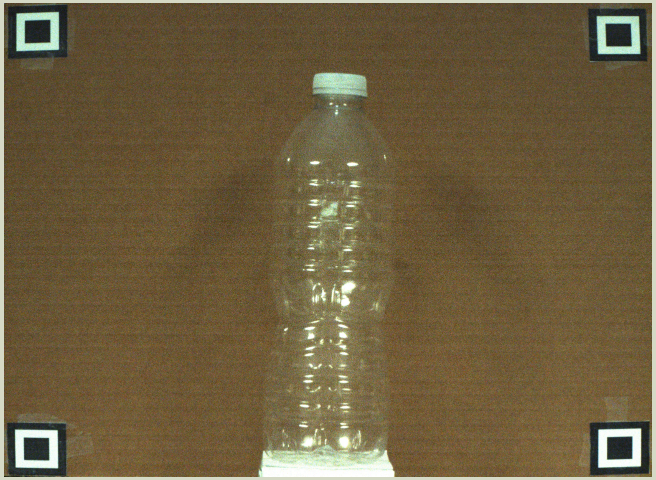
n=165	Predicted: NO	Predicted: YES	
Actual: NO	TN = 50	FP = 10	60
Actual: YES	FN = 5	TP = 100	105
	55	110	

- Additionally, the ranges excluding 1.89-2.39μm typically performed better amongst the same # of bands.
  - Indicates some ranges can identify unique qualities of each class more so than others.

MY OWN DATA

- *Classification of liquids:* 12 classes in total

- Empty, water, hydrogen peroxide, rubbing alcohol, milk, soy milk, lemonade vitamin water, cranberry juice, orange juice, Ginger Ale, Coca Cola, seltzer water



Visible light and UV light camera examples

MY OWN DATA

- Final data will be 640 x 480 x 12 cuboid per class

- Resolution and pixel size is not consistent between cameras
  - Need to match to lowest resolution
  - Covert pixel size to lowest scale (8 bit int)
  - Account for alignment issues

- 12 Bands: Achieved with several cameras and filters
  - UV light
  - Visible light
  - NIR/SWIR (4 bands w/filters)
  - MWIR (3 bands w/filters)
  - LWIR

OPEN QUESTIONS

How will overlap and discretization of the spectral ranges affect my results?

Will the results be consistent with my analysis of the AVIRIS Indian Pines data set?

Can multimodality help distinguish liquids of the same or similar color?

FINAL GOALS

- Process data (i.e. same resolution, pixel value range, and alignment)
- Classify data and analyze the results, potentially answering these research questions
- If time allows, perform the same process again with another hyperspectral data set (potentially tensor data of different dimensions)

THANK YOU

I would like to acknowledge and thank the National Science Foundation for funding this research. This work has been supported by NSF grant CCF-1559855.

An additional thanks to my mentor Dr. Waheed Bajwa, Dr. Mark Pierce, Lazaros Gallos, and the kind, informative graduate students of the INSPIRE lab for helping me succeed.