	Ideal	Shen Yang Team: YS	
	Score	Score	.---------- Instructor Notes----------.
Total Points	10	8.55	Score Points:08.55
Total Percentage	100	85.5	
Exceptional Work, One idea: (required 7000) perform feature extraction upon the images using a feature extractor such as HOG (histogram of ordered gradients) or LBP (local binary pattern). Then build a nearest neighbor classifier using a method appropriate for your chosen features. You will need to investigate appropriate methods for comparisons with your chosen feature extraction technique. NOTE: this often requires some type of brute force matching per pair of images.	10	5	ORB is run on all the images, but th matching approach is only investigated for a few images. Need to run this on a large subset of the data to make conclusions.
Give an overview of the dataset. Describe the purpose of the data set you selected (i.e., why was this data collected in the first place?). Why is this data important? Once you begin modeling, how well would your prediction algorithm need to perform to be considered useful? Be specific and use your own words to describe the aspects of the data.	20	19	This is a good description. I see there is description of the costs for false positives and false negatives--excellent. However, the 10% false positive rate seems a bit arbitrary. The actual value should probably be based upon the difference between money saved (true positives and negatives) and money lost (through false positives) to get to break even point.
Read in your images data as numpy arrays. Resize and recolor images as necessary.	5	4.5	I think the grayscale conversion loses a good deal of the information for your dataset. Disease coloring is likely a nice feature to have.
Linearize the images to create a table of 1-D image features (each row should be one image).	4	4	
Visualize several images.	1	1	
Perform linear dimensionality reduction of the images using principal components analysis. Visualize the explained variance of each component. How many dimensions are required to adequately represent your image data?	5	5	
Perform non-linear dimensionality reduction of your image data.	5	5	
Compare the representation using non-linear dimensions to using linear dimensions. The method you choose to compare dimensionality methods should quantitatively explain which method is better at representing the images with fewer components. Do you prefer one method over another? Why?	20	14	The classifiers constructed are given the rconstructed images, not the low dimensional representation. I think you need to give the representation here, not the reconstructed images (as these suffer from dimensionality as well).
Perform feature extraction upon the images using any feature extraction technique (e.g., gabor filters, ordered gradients, etc.).	10	10	
Does this feature extraction method show promise for your prediction task? Why? Use visualizations to analyze this questions. For example, visualize the differences between statistics of extracted features in each target class. Another option, use a heat map of the pairwise differences (ordered by class) among all extracted features. Another option, build a nearest neighbor classifier to see actual classification performance.	20	18	Need a bit more explanation of the results here.