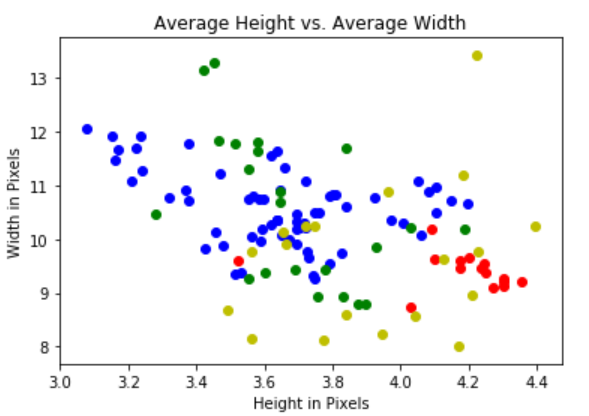
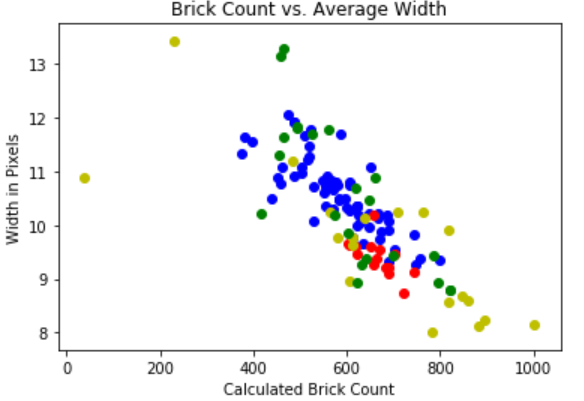
Questions:

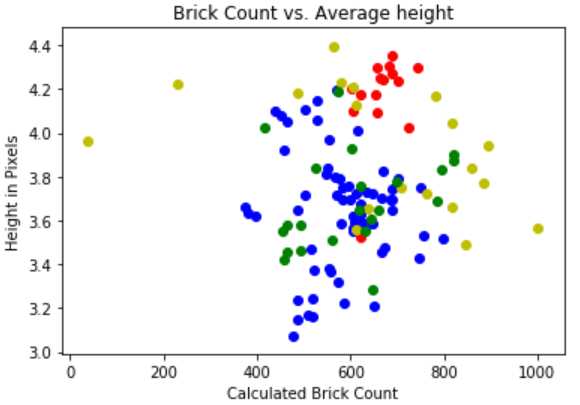
a) What features did you select and implement? Why did you select these features and how do you think they will be informative for this classification problem?

The features that I choose to extract were the average length and height of the bricks in the picture as well as a calculated number of bricks. I chose to extract these features because these features are related to how I determine what kind of brick pattern is in a picture.

Below I have plotted these three features relation to each other.

|  |
| --- |
| Legend  Red- Flemish Stretch bond Green- Stretcher Bond  Blue- English Bond Yellow- Other Brick Patterns |





\*\*Maybe the histograms instead?

As you can see from these graphs, the features of individual images do seem to group together with features of other images in the same class. I expect that the non-brick pictures will fall far outside of these groupings because those pictures will not follow the same patterns as these pictures.

For my probabilistic model, I have had to make a few assumptions or omissions to apply the model. Ignoring the individual a-priori’s on each class because anything I put here will be an assumption and I wish to see how the model does before adding in prior assumptions. I am assuming that each feature is independent of each other feature such that the multivariable probabilistic function can be separated into 3 separate probabilistic functions which are each only dependent on a single feature; as such I can multiply them together to get the overall function. I am also assuming that each feature’s distribution is Gaussian.

b) Given your data set and features, were there any outliers in your data set that needed to be removed or handled specially during pre-processing/feature extraction? If so, what was causing these images to be outliers? If not, do you think your data set is a good representation for the problem or are you missing imagery of a certain type in your data set that could impact your overall performance?

This dataset is not large enough to determine for sure if anything is truly an outlier, yet. I suspect that because this model is doing image processing to find the edge of the bricks, that any exceptionally noisy images will likely have to be ignored.

I do not have any examples of non-brick patterns so it will be interesting to see how those images work in my current model. Currently the Other Brick Patterns class has the most variation of my dataset, but I suspect that the non-brick pattern class will have even more variation.

c) How effective is your trained system (preprocessing + feature extraction + classifier)? Do you think you already have a system that would be competitive/effective for the project overall? If so, why? If not, what is needed to improve your approach?

Without the non-brick pictures class, it is impossible to know how effective it will be in the overall problem. Also, a very small dataset was used to both train and test this model; without more images to test it is impossible to say how well the model will perform in real-world applications.

testing