

Capstone Project 1 - Categorize Clothes 1

What is the problem that you want to solve?

Organizing dishes and clothes are some of the most time intensive tasks that are still done manually in the home. I hope to take a step in the direction of alleviating those tasks by creating an automated system for clothes categorization. New items of clothing must also be manually categorized when being added to online stores and clothing categorization can save time for that as well.

Who is your client and why do they care about this problem?

Washer and dryer makers and indirectly home owners. It has the potential to move towards saving millions of people hours per week. As a result a folding machine should be made if the categorization is right a high percentage of the time. Another segment of customers which will be reached first is online clothes sales.

What data are you going to use for this?

Fashion.mnist (<https://www.kaggle.com/zalando-research/fashionmnist/data>) will be the main data I use to explore image processing further and train on. At the end I will test against photos I take and may have to improve it further by using another dataset such as Deep Fashion (<http://mmlab.ie.cuhk.edu.hk/projects/DeepFashion.html>).

Outline your approach to solving this problem.

First I will examine histograms and image statistics to determine if image preprocessing and normalization is likely to improve the results of the models. I will then apply any changes or filters as necessary. I will use different machine learning models such as MLP, SVM, Logistic Regression and then move on to CNN. I will move onto CNN last, because deep learning models like CNNs which can aid in automated feature extraction\engineering and also building the classifier. The fashion.mnist is a grayscale and I am not sure how well it would apply to actual images that are filtered similarly. Fashion.mnist is also limited in that it has “T-shirt/top”, “Trouser”, “pullover”, “Dress” and “Coat” of washable clothes.

What are your deliverables?

The results of this project will be a model which can automatically categorize a piece of clothing based on the various clothing categories used in my historical training dataset. The hard outcomes will be code (jupyter notebooks), reports, visualizations and presentations to support my analysis. I will also generally test the results in terms of real life by feeding in a couple images I have taken and testing the on the model trained by the stock photos.