Classification of Clothing Patterns and Colors

Type B: Application

Alec Resha

Description

The goal of this project is to classify the pattern and color of clothing and rename a file to reflect the classification. Images of clothing will be classified based on pattern and color.

Motivation

The motivation for this is part of a larger system. I have a larger project in mind for a program to manage clothes in a closet to make it easier to put together outfits, know what hasn't been worn in a while to donate it, learn preferences, etc. The work addressed in this specific project would be to automatically classify the pattern and color of new/existing articles of clothing to aid in the data entering process. It would make it easier to let the user take pictures of their clothes and have it generate the information for it. This would likely be released open source, but at a minimum I would use it make mornings easier.

Data Set

The dataset is available here https://data.world/crowdflower/categorization-dress-patterns for pattern training based on images. The dataset consists of 15,702 rows which contain an ID, pattern category, the confidence in the pattern (for less exact patterns), and url to the image. Each image is an overall picture of an outfit with a redsquare to note the article of clothing. All the images are of dresses, but the pattern recognition should transfer to other articles of clothing.

A smaller dataset is also available at https://github.com/lstearns86/clothing-pattern-dataset if analyzing the uncropped images proves inconsistent or inaccurate.

Algorithm Design

This would likely use a CNN since the data is all images. It will also need to be constrained to the red square in each image. If this proves to be too problematic, the other dataset can be used since they are upclose images of patterns.

For color identification, KMeans will be used to cluster the colors in each image to find the most common colors, smaller values will be filtered out as they are likely to be from image backgrounds. This will not be ML based, just a simple clustering algorithm but it will be used to aid in classifying the images.

Evaluation Plan

This will be evaluated with a train/test data split as well as manually evaluating results.

Task Division

All work will be done by me since the project is being done alone.