Parker Williamson

2/5/2018

Springboard Data Science Career Track

Capstone Project 1 – Data Wrangling

Image data needs to be broken down by image and category for meaningful analysis. ~~Histograms can show the pixel range used~~ and intermediate images of the CNN models can show important features. ~~I will also look at the mean and median values of each image~~. ~~How preprocessing changes the image of the biggest difference~~. ~~Show an example of each category and figure out which~~ categories are the most commonly confused.

Mean and median values start to show some of the differences between categories. Mode wouldn’t make sense to examine because it would be zero for every class since the zero is the most common value, because of the background.

An example of all 10 classes of images are on the following pages, clearly shirt, t-shirt, and pullover categories are a lot more similar than some of the other classes.

 

 

 

 

 

 

 

 

 

 



 

