CSE 469 Final Project
K-means Clustering for Image Processing
Sarah Schwartz and Goktug Geneci

Image processing is an important step in many modern high-impact data mining and machine learning problems. To be able to complete any type of object detection in an image, that image must first be segmented into discrete parts which can then be processed by a machine learning algorithm or other program. There are a variety of real world impacts to this problem, including optical character recognition and autonomous vehicular controls. For a task like OCR, the key task is to separate the text from the background, so that the text can then be read and transcribed. The implications of this in areas like educational accessibility are potentially massive both overall, and on a smaller level, for example, lecture transcription. For autonomous vehicle controls, a car must be able to recognize signage in order to drive safely and predictably. This means that as its surroundings are being processed, it must be able to differentiate segments of the field to make determinations. We have approached this problem on a smaller scale by applying k-means clustering to colors in am image to segment it into its relevant parts.

For the clustering task, our input data was a set of three images, and our output was the same three images, but with clustering of colored portions. We used a picture of text on a background to simulate the character recognition task, a picture of a stop sign to simulate the vehicle environment task, and a picture of the interlocking UB group. We preprocessed our images by converting each one into a multidimensional matrix of the RGB values for each pixel. These RGB values were later used in cluster assignment. We applied the K-means clustering algorithm to these matrices with a variety of cluster numbers to evaluate which number of clusters was optimal for the given image.

Figure 1: Original interlocking UB photo

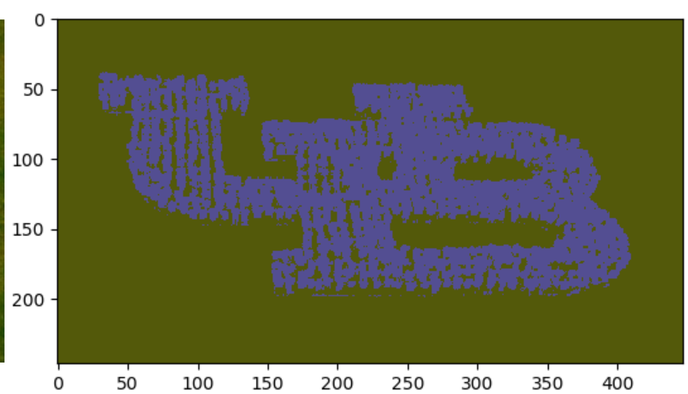

Figure 2: Interlocking UB with two clusters

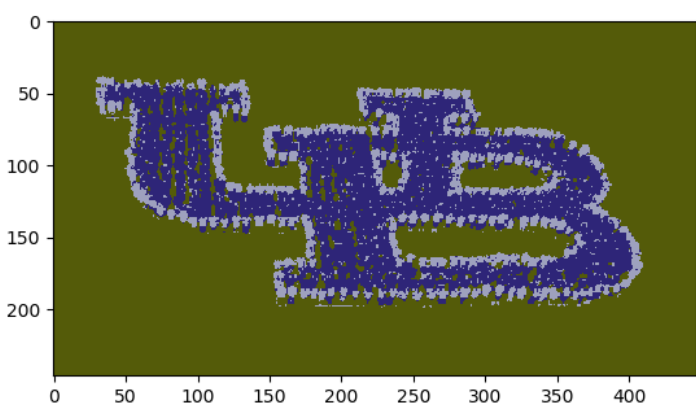

Figure 3: Interlocking UB with three clusters

In the case of the interlocking UB image, we found that three clusters was the lowest number of clusters we could have while still getting a good result. In the two cluster case, the colors are hard to differentiate and it's nearly impossible to tell that the interlocking UB is made up of people. In the three cluster case; however, we see the white border and can differentiate the symbol, as well as the people to some extent. Next, we look at our text processing results.

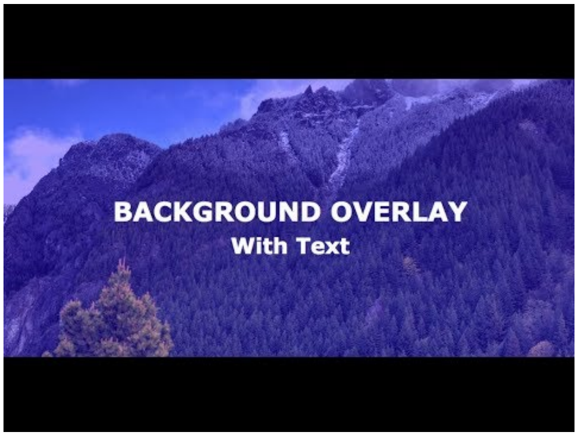

Figure 4: Original text example

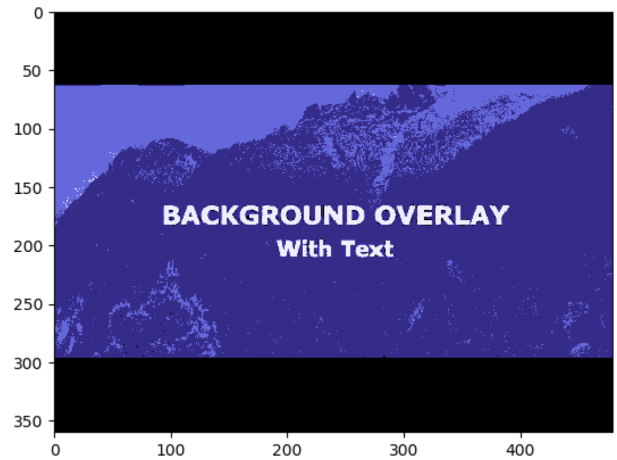

Figure 5: Text example with four clusters

Again, in this case we see that it takes multiple clusters to fully differentiate the target region. An initial attempt at clustering using three clusters yielded a result where the text and the sky were in the same cluster, and difficult to differentiate. With four clusters, the text is clear and is in a cluster by itself. Finally, we look at the traffic sign case.

Figure 6: Original traffic sign example

Figure 5: Traffic sign example with five clusters

In this case, we focused on separating the stop sign from its surrounding context. This was one of the more surprising results, as it took five clusters to split it effectively. We would have anticipated fewer, given the relatively high differential between the color of the sign and the colors in the image; however, this did not end up being the case.

In order to identify a successful result, our metric was whether or not the area of interest was easily differentiable from the other areas of the image, and specifically, whether or not that area of interest was in its own cluster. In each case, we were able to get a successful result, but we had to slightly modify our approach to fit each case. Based on our results, we can draw several meaningful conclusions that can be abstracted out to the image processing problem as a whole. We highlight the key conclusion here. In each case we evaluated, it took a different number of clusters to appropriately isolate the object of interest. Because of this, any approach to the image segmentation step of the image processing problem should be adaptable to the case of interest. It would be unwise to be overly general in this step, as the implications of not identifying an object like a stop sign in a timely and appropriate manner are serious. It would, and should, require a different number of clusters to pull black text from a white background than to separate an object from its complex background.