CS 498 AML: Homework 3

Tanishq Dubey (tdubey3)
Harrison Qu (hqu5)
Yu Wang (yuwang15)

February 21, 2017

Problem 4.10

a Our principal component methodology resulted in the following errors, recorded as averaged square distance. This means that a distance of 0 is no error, while high distance corresponds to high error.
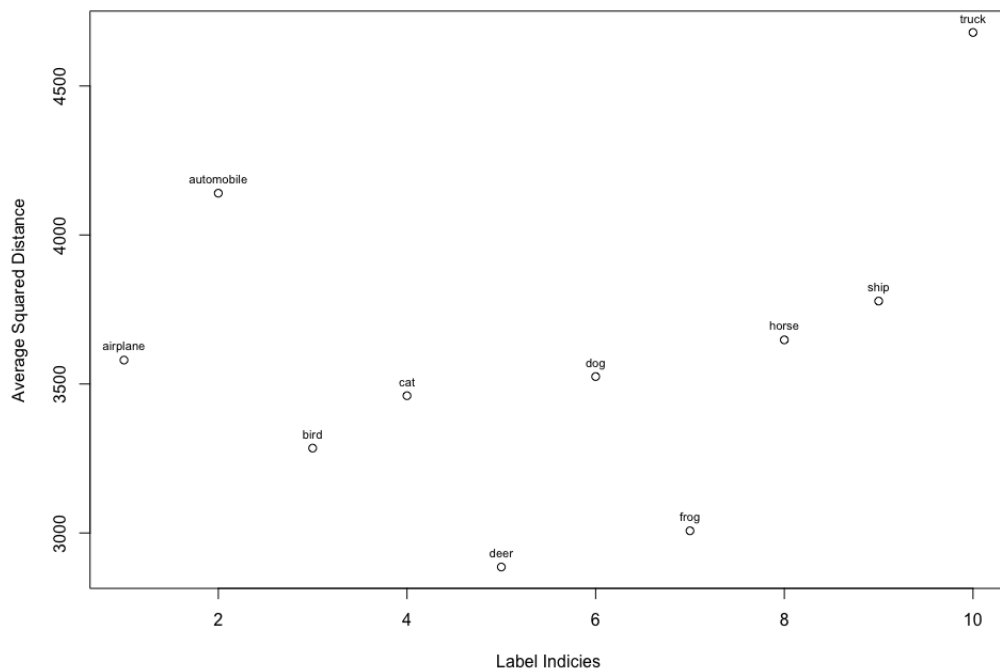

Figure 1: Plot of the errors of each label

(a)	(b)	(c)	(d)	(e)

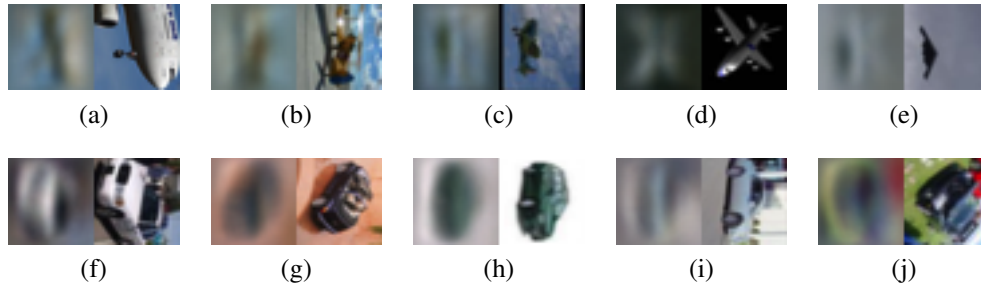|				
(f)	(g)	(h)	(i)	(j)

Figure 2: A selection of images comparing the first 20 principal components vs the actual image. The first row is 5 samples from the airplane label while the second row is 5 samples from the automobile label

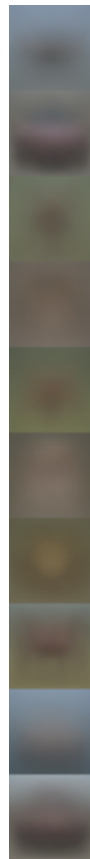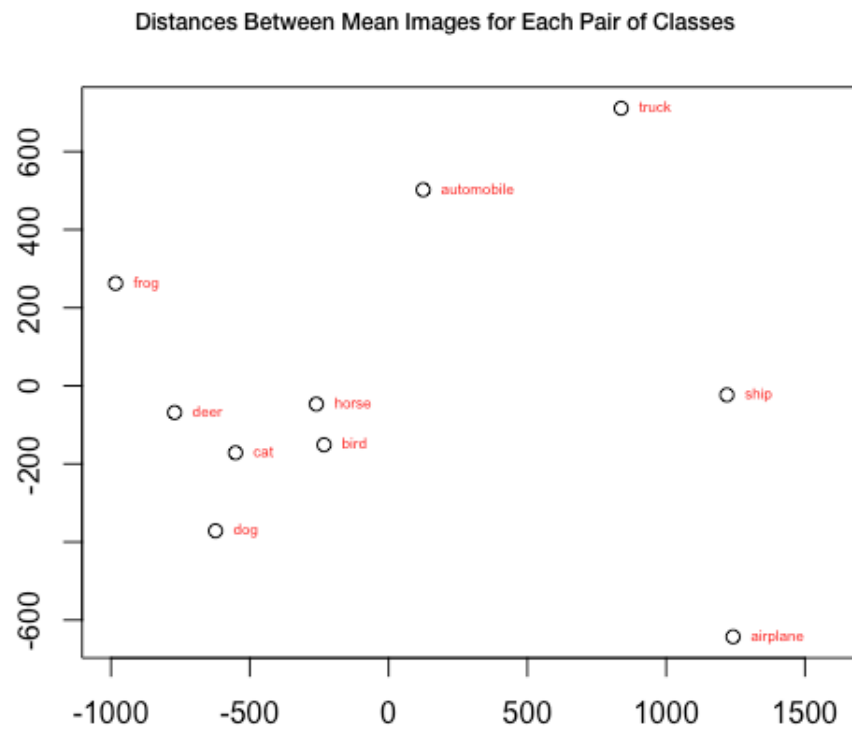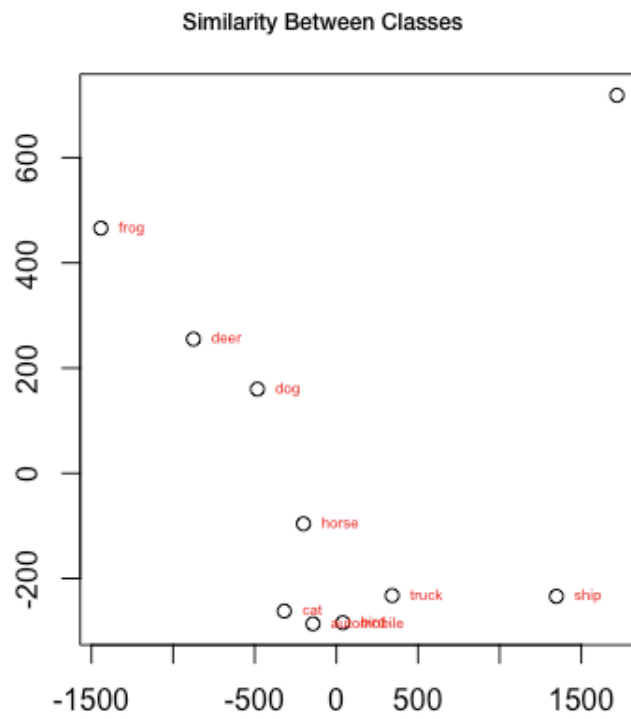In addition, the means of each label were computed and are as such:

Figure 3: The means of each label, beginning at airplane and moving down the list.

b Our distances between mean images for each pair of classes using PCA for a 2D map can be represented as such:

Distances Between Mean Images for Each Pair of Classes

c Similarity between classes is represented as such using a 2D PCA map:

Similarity Between Classes

This figure differs from the one found in part "b" due to the fact that this method is using the principal components of one class to determine the distance to another. On the other hand, the methodology in "b" is using principal coordinate analysis to find distance between classes.