Jianmeng Yu s1413557

INF-2B Natural Image Classification Task 1 Report

Task 2.2:

The 2 largest eigenvalues: 14.383095313936090, 11.384970972153763.

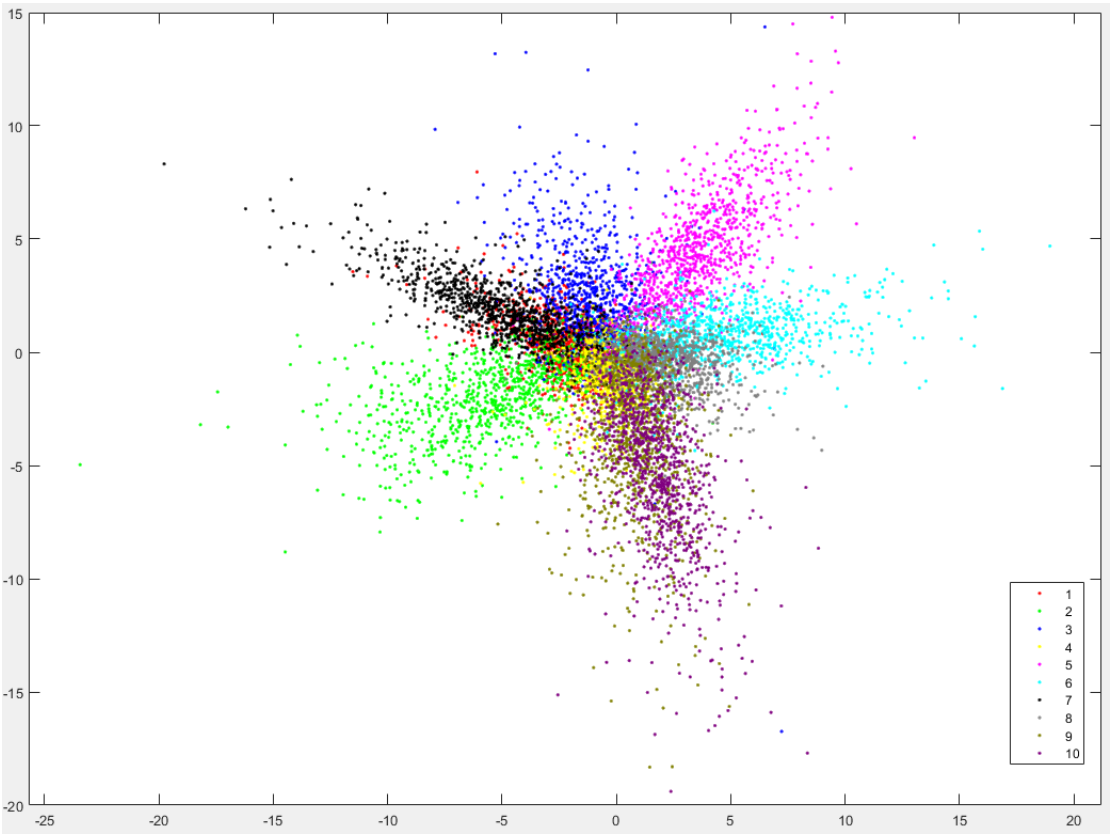
The first five rows of E2:

0.307983111873521	0.0941630560639841
-0.107190068011302	0.103257837149263
-0.0543638578929347	-0.0559819788069004
0.0102186827971333	-0.0400129159196361
-0.0135682600619131	0.0421653772983205

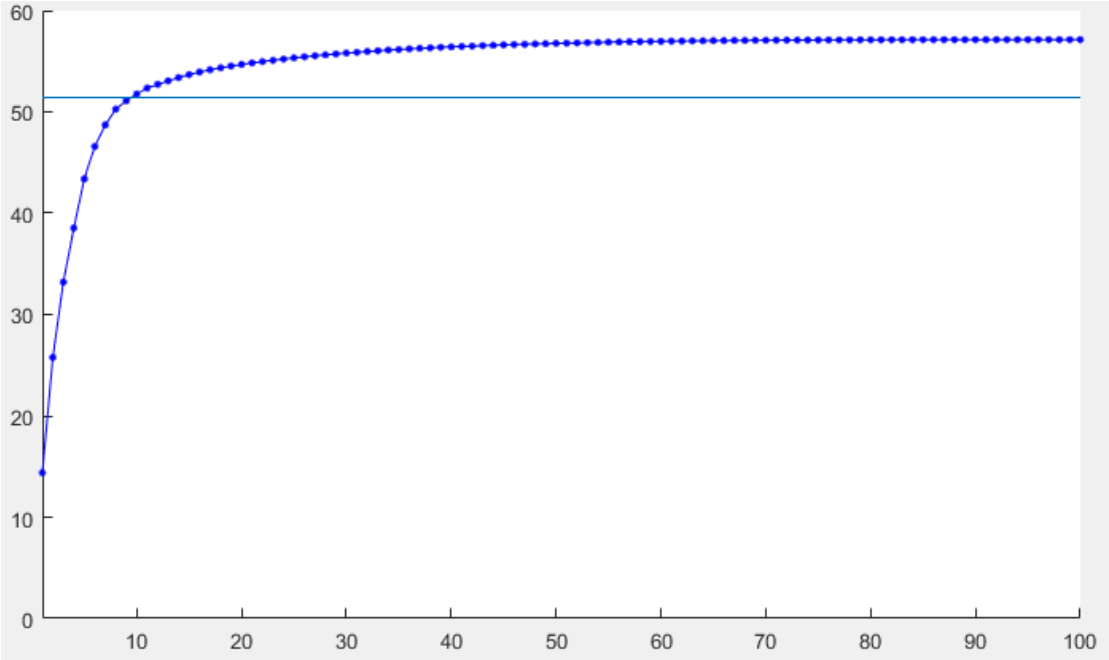
The first five rows of XPCA:

-1.77617463233227	0.300201757575069
-1.98030687916530	0.190675055342996
9.87152759338393	2.11902255496648
0.836502916842919	-0.384584141380611
1.38410762664313	-1.95957428471935

Task 2.3:



Task 2.4:



As the cumulative variance shows, we need at least 10 variance to show 90% of the total variance.

Task 2.5:

Since the rest 90% of the variance only expresses 10% of the total variance, when we classify the test feature in future tasks, we can decrease the runtime significantly by only using 10% of the most significant features, in exchange the rate of approximately 10% loss in accuracy.

Also, in the scatter graph, we can see that some of the classes have almost same direction of distribution (class 1 and 7, 9 and 10), with almost every class have clustered in the center, making classification of the ‘fuzzy’ pictures inaccurate.