

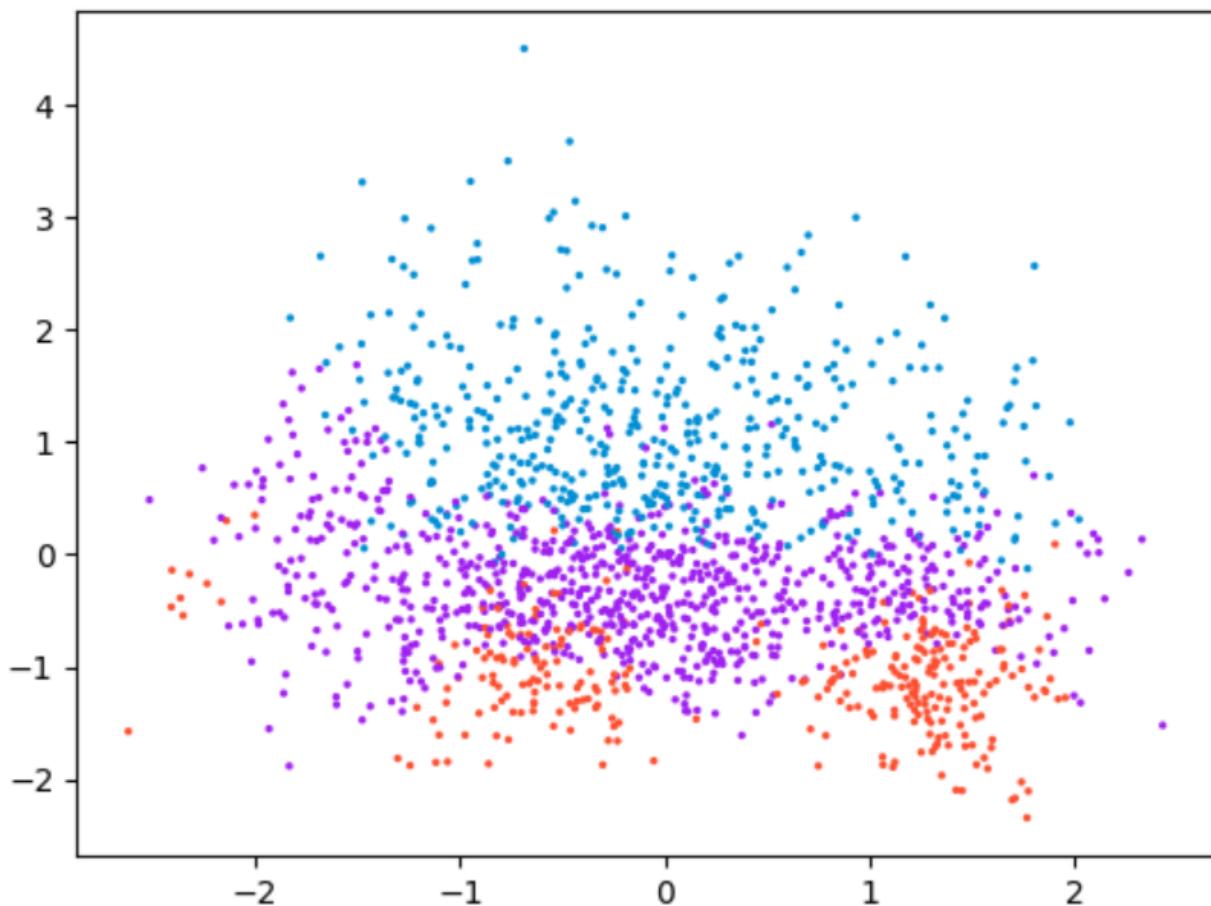
COMP 448/548
Medical Image Analysis

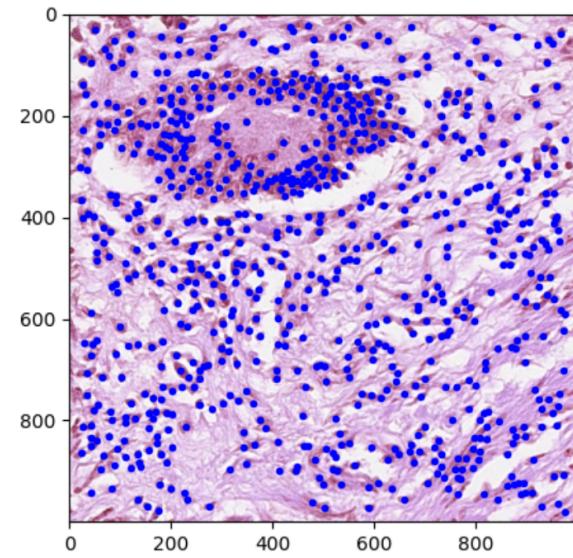
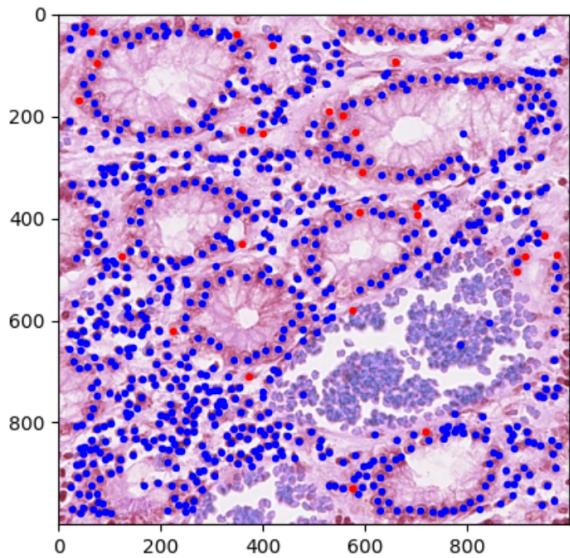
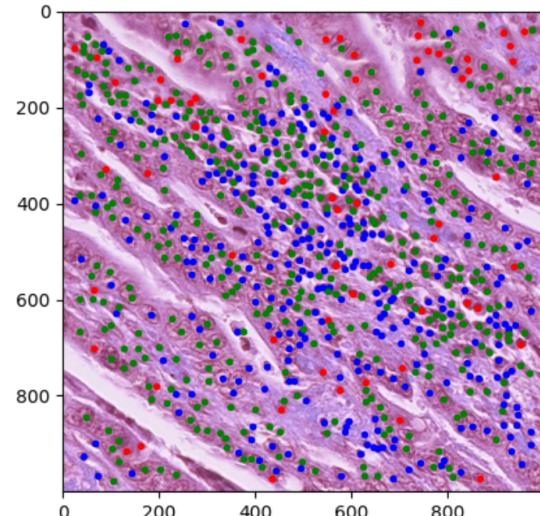
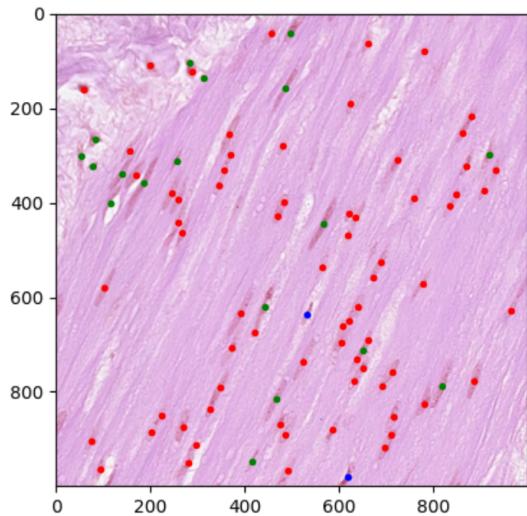
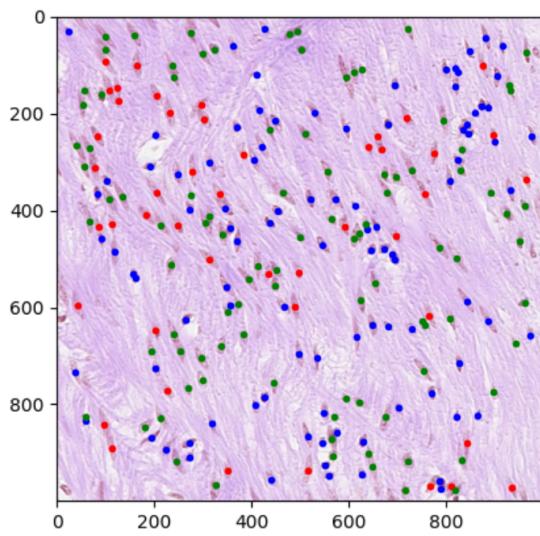
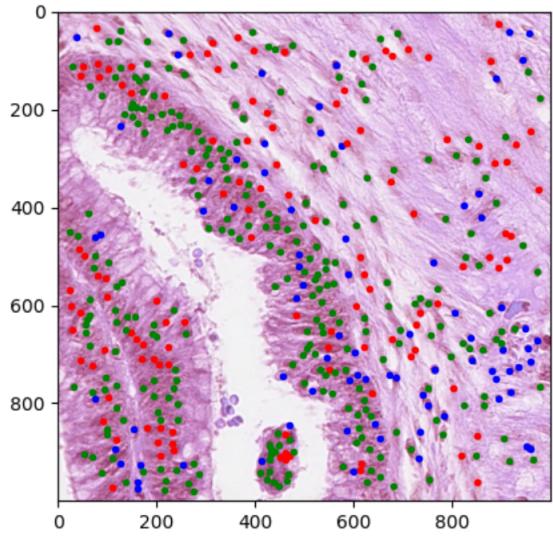
Spring 23
Homework #2

Yiğit Can Ateş 75617 - Melis Oktayoğlu 64388

Results

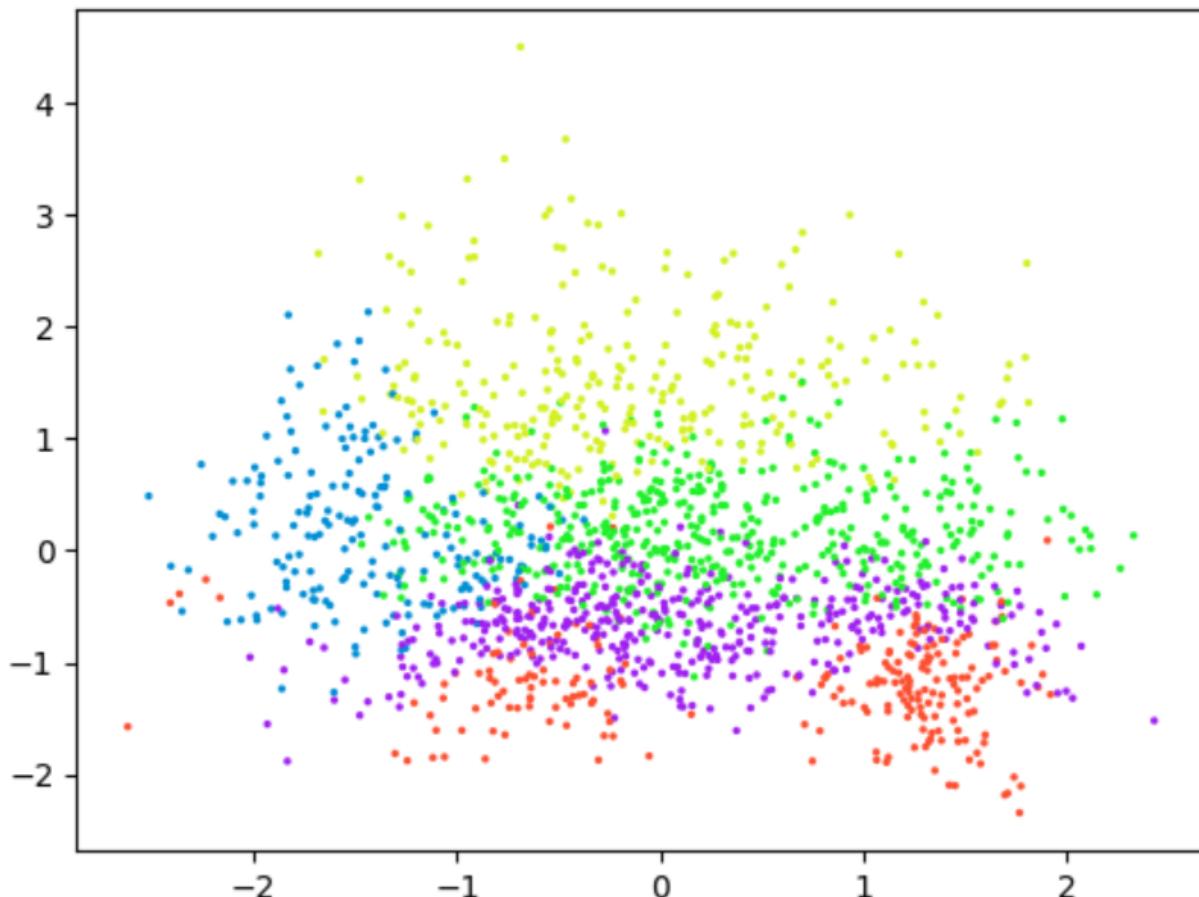
patch_size = 16, binNumber = 5, dist = 2, k=3							
Training Set				Test Set			
	Epithelial	Spindle	Inflammation		Epithelial	Spindle	Inflammation
Cluster1	0.28	0.50	0.22	Cluster1	0.41	0.24	0.35
Cluster2	0.31	0.64	0.05	Cluster2	0.38	0.62	0
Cluster3	0.45	0.42	0.13	Cluster3	0	0	0

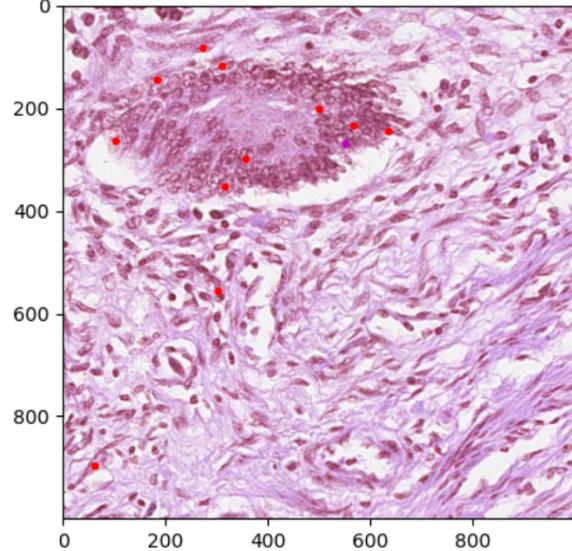
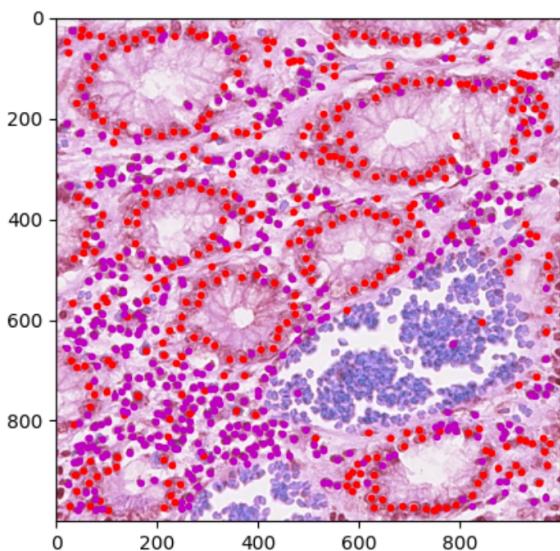
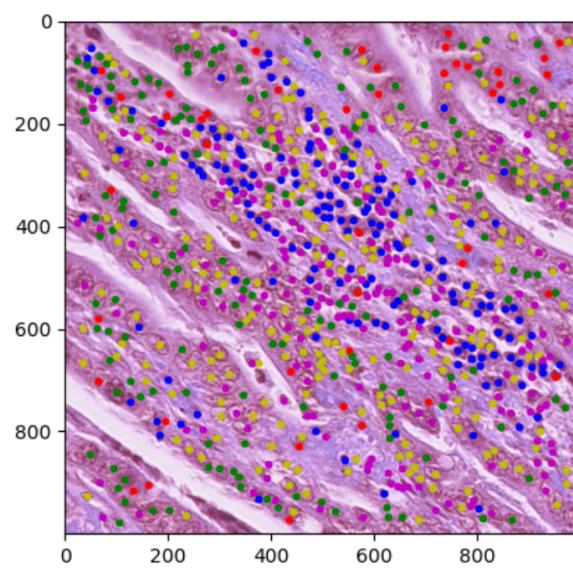
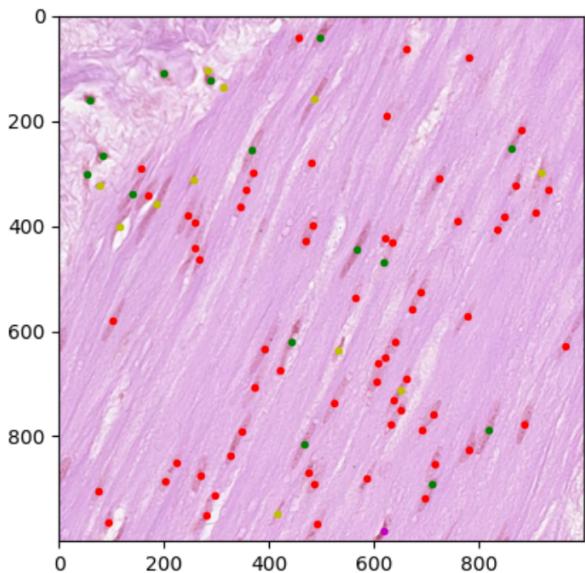
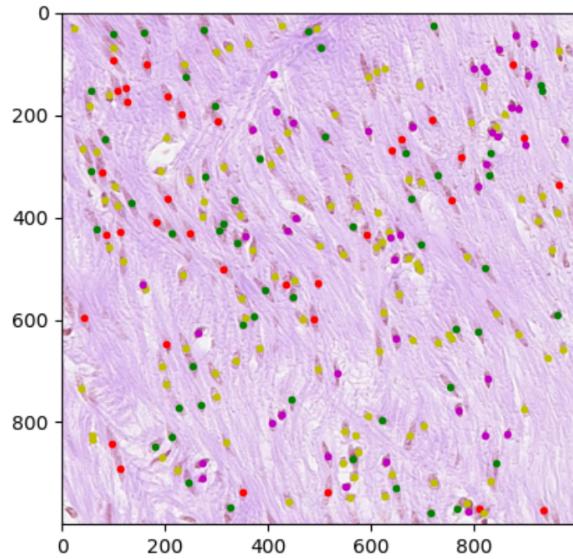
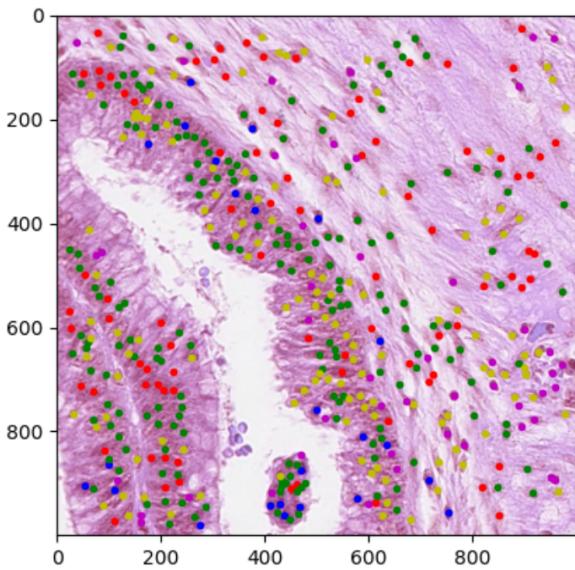




patch_size = 16, binNumber = 5, dist = 2, k=5

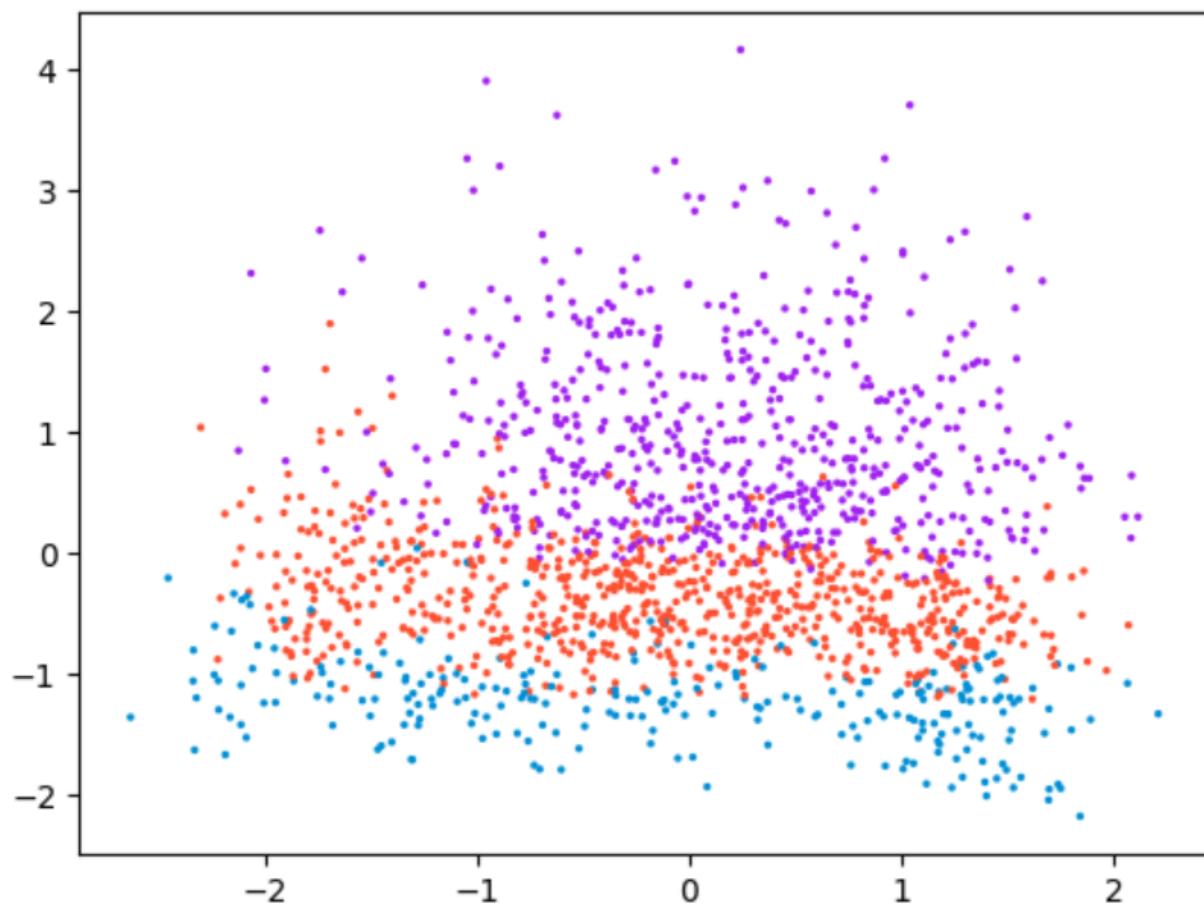
Training Set				Test Set			
	Epithelial	Spindle	Inflammation		Epithelial	Spindle	Inflammation
Cluster1	0.2	0.23	0.57	Cluster1	0.68	0.28	0.04
Cluster2	0.27	0.70	0.03	Cluster2	0.16	0.23	0.61
Cluster3	0.55	0.43	0.02	Cluster3	0	0	0
Cluster4	0.41	0.55	0.04	Cluster4	0	0	0
Cluster5	0.27	0.46	0.27	Cluster5	0	0	0

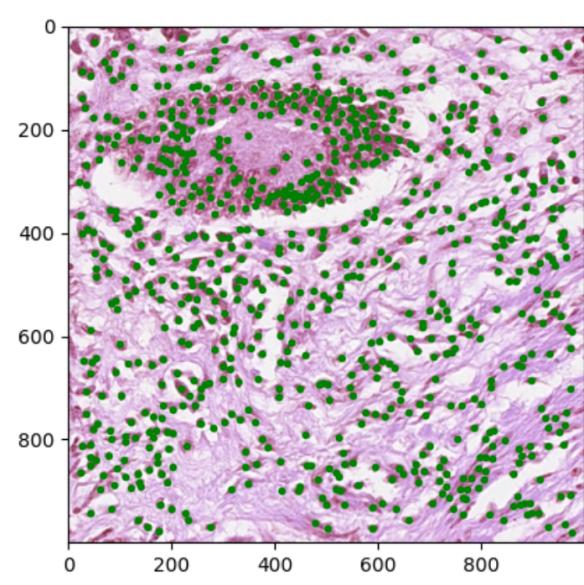
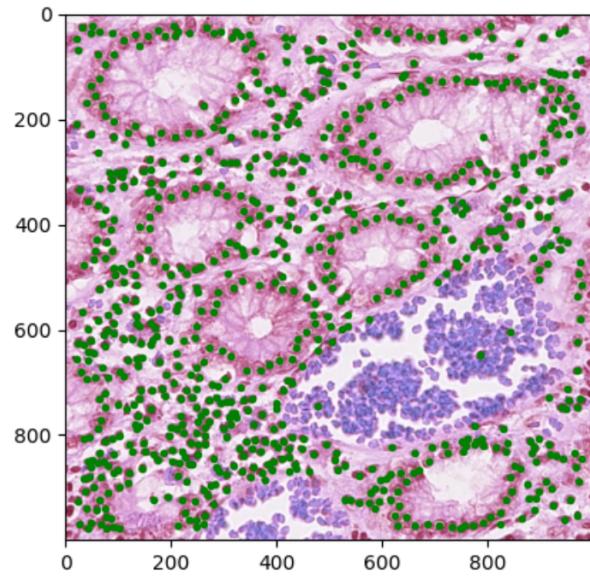
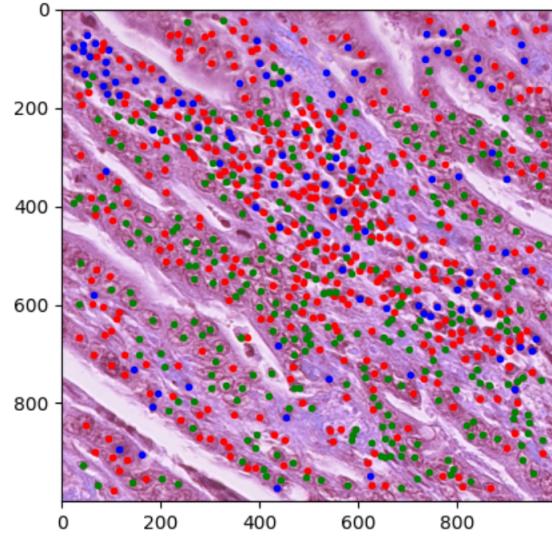
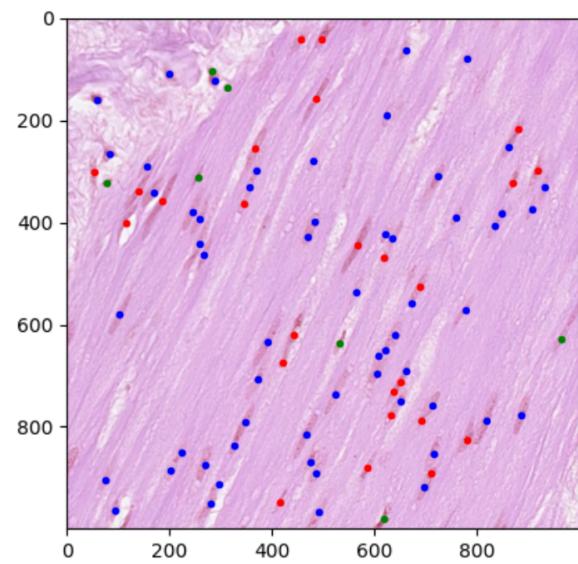
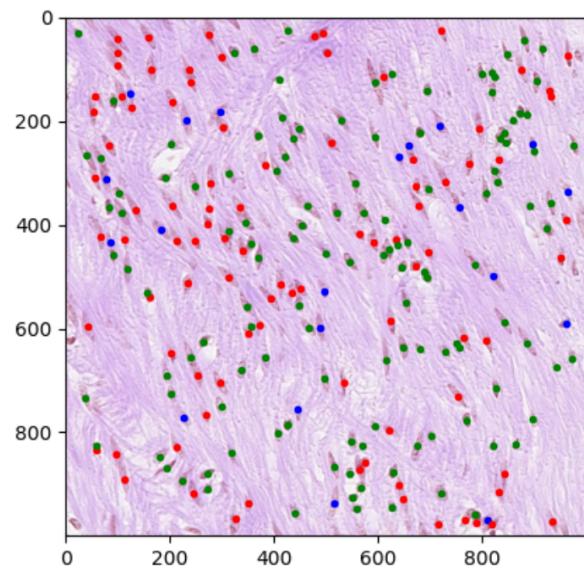
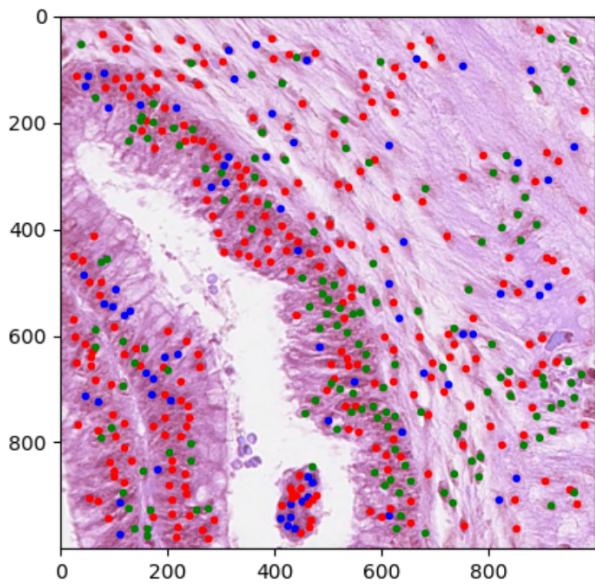




patch_size = 12, binNumber = 16, dist = 4, k=3

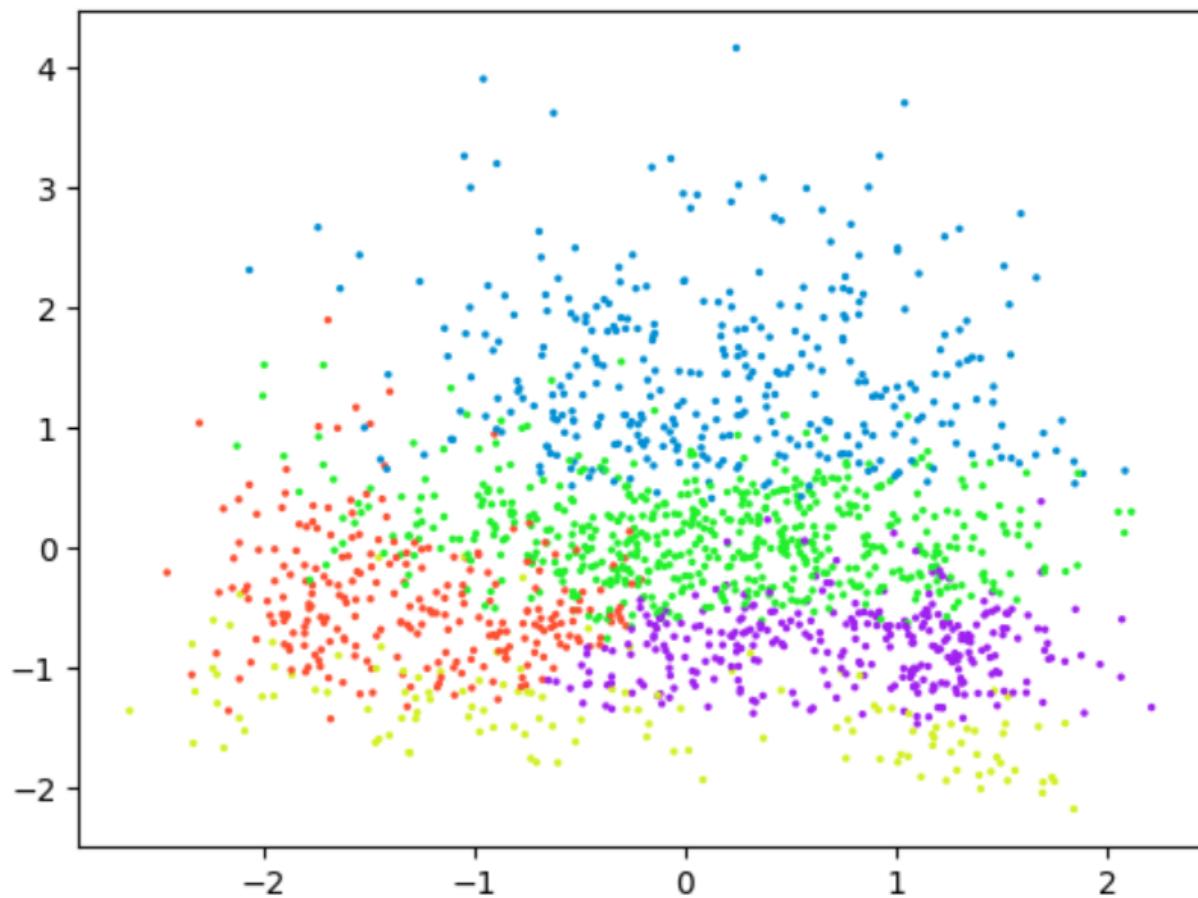
Training Set				Test Set			
	Epithelial	Spindle	Inflammation		Epithelial	Spindle	Inflammation
Cluster1	0.28	0.52	0.20	Cluster1	0.33	0.46	0.21
Cluster2	0.38	0.45	0.17	Cluster2	0	0	0
Cluster3	0.41	0.51	0.08	Cluster3	0	0	0

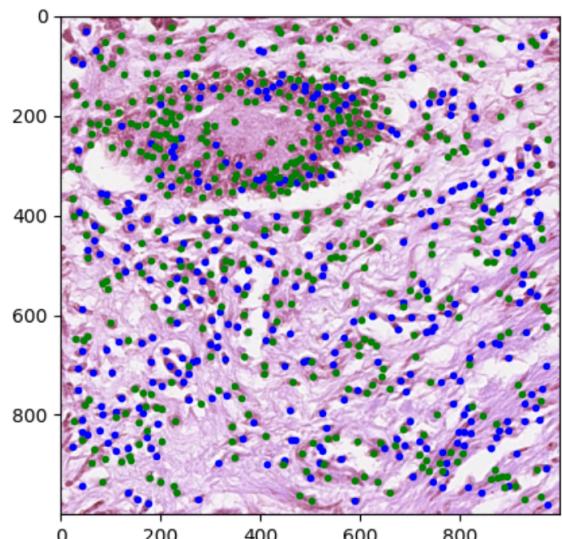
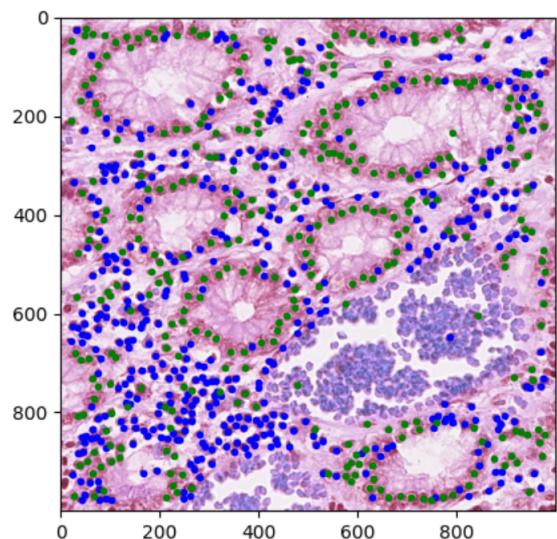
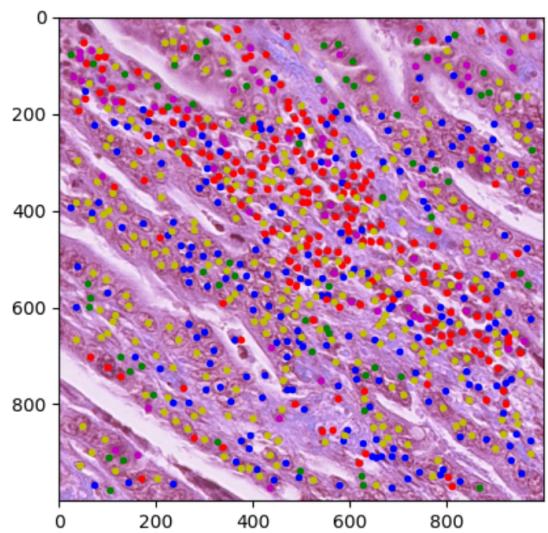
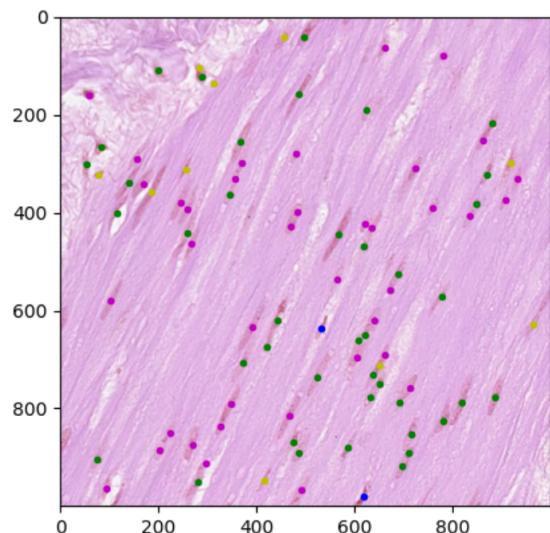
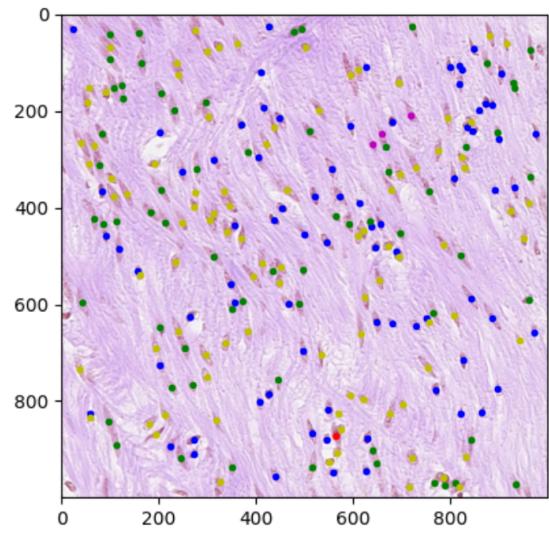
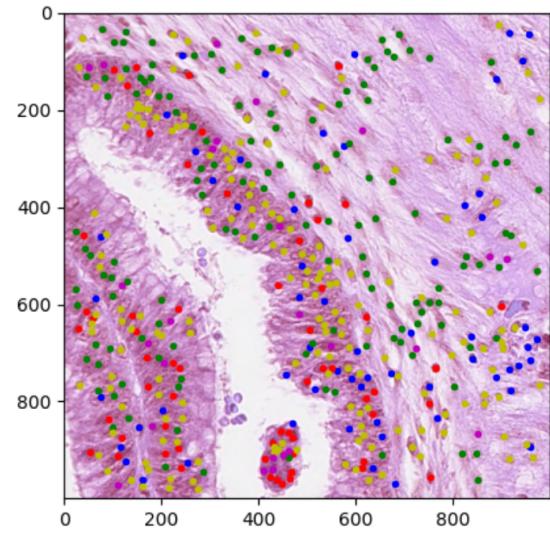




patch_size = 12, binNumber = 16, dist = 4, k=5

Training Set				Test Set			
	Epithelial	Spindle	Inflammation		Epithelial	Spindle	Inflammation
Cluster1	0.39	0.53	0.08	Cluster1	0.20	0.41	0.39
Cluster2	0.31	0.20	0.49	Cluster2	0.45	0.51	0.04
Cluster3	0.32	0.68	0	Cluster3	0	0	0
Cluster4	0.45	0.48	0.07	Cluster4	0	0	0
Cluster5	0.27	0.5	0.23	Cluster5	0	0	0





Discussion

Quantitative and visual results for corresponding parameter combinations are given above. We see that increasing patch size from 12 to 16 slightly improves clustering performance. We chose these values based on visual investigation of sizes of individual cells in the images provided.

Increasing cluster number from 3 to 5 allowed us to get clusters with high numbers of a particular type of cell (for example 0.70 spindle). However, we found no dominating type in most of the clusters.

We see that it is particularly challenging to separate inflammation cells from other types, due to their complex phenotype and low number. Low number of inflammation cells causes class imbalance problem. In order to overcome this, we used weighted k-means clustering where weights of classes were determined based on their relative ratios. This increased our clustering performance compared to regular k-means clustering. We normalized features by removing the mean and scaling to unit variance.