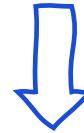


BAGS OF VISUAL WORDS (BOVW)

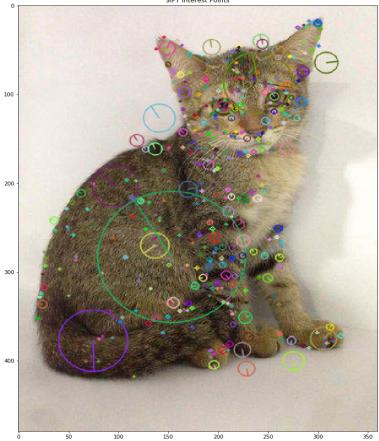
- ❖ Represent an image by frequency of features (words)
- ❖ Images share dissimilarities in the count of selected features
- ❖ Applications
 - Image matching
 - Image annotation
 - Image classification



“visual words”: people, tree, leaves, ground, building, sky, cloud...

Feature extraction

SIFT

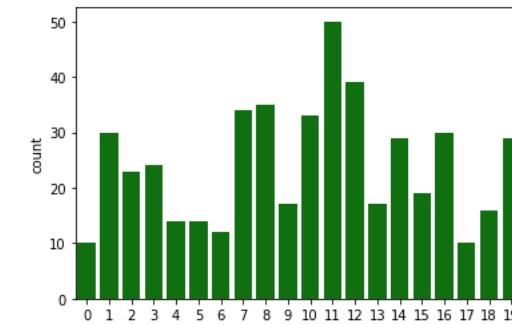
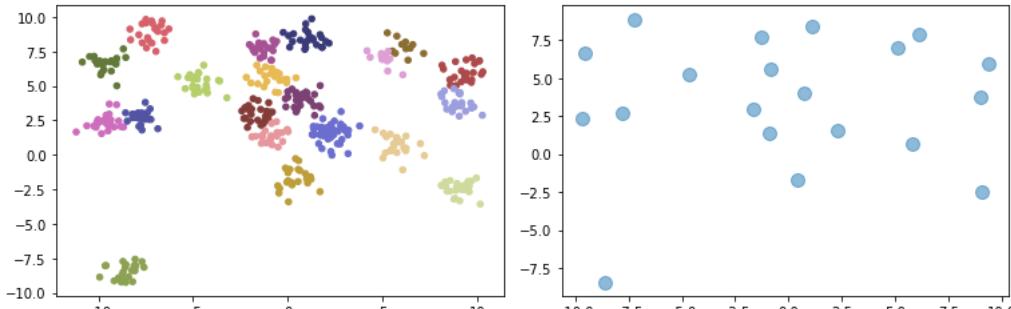


	0	1	2	3	4	5	...	458	459	460	461	462	463
0	36.0	30.0	8.0	16.0	36.0	25.0	...	36.0	0.0	23.0	1.0	1.0	11.0
1	33.0	4.0	1.0	27.0	23.0	3.0	...	3.0	1.0	87.0	23.0	61.0	80.0
2	2.0	2.0	0.0	9.0	4.0	0.0	...	6.0	135.0	4.0	29.0	50.0	16.0
3	0.0	3.0	0.0	0.0	2.0	0.0	...	43.0	135.0	0.0	58.0	29.0	0.0
4	0.0	0.0	0.0	0.0	1.0	0.0	...	45.0	1.0	0.0	58.0	17.0	0.0
...
123	0.0	0.0	0.0	0.0	0.0	0.0	...	8.0	17.0	0.0	0.0	0.0	0.0
124	0.0	0.0	0.0	0.0	0.0	0.0	...	8.0	29.0	0.0	0.0	0.0	0.0
125	0.0	0.0	0.0	0.0	0.0	0.0	...	56.0	13.0	0.0	0.0	0.0	0.0
126	1.0	3.0	2.0	2.0	0.0	0.0	...	34.0	7.0	0.0	0.0	0.0	1.0
127	20.0	15.0	9.0	13.0	1.0	0.0	...	6.0	0.0	33.0	98.0	92.0	27.0

descriptor

Bags of visual words

K-Means Clustering



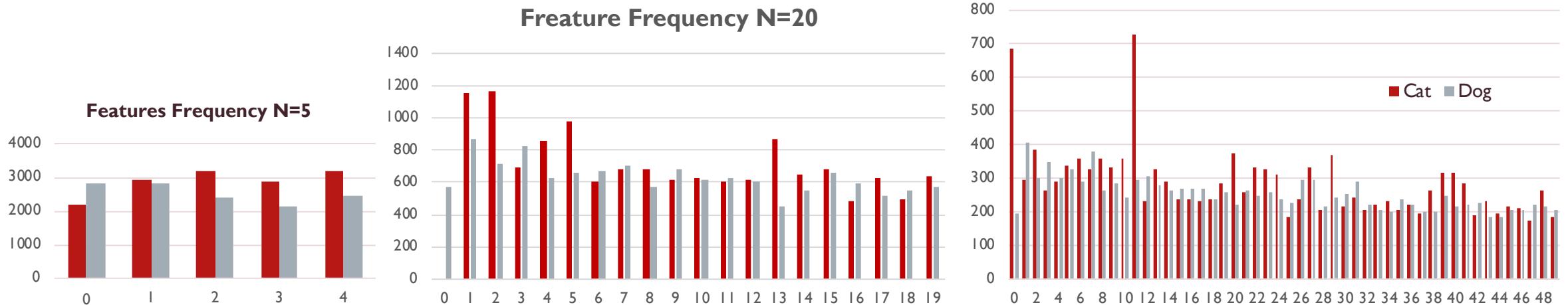
Vector quantization

	0	1	2	3	4	5	6	7	8	...	13	14	15	16	17	18	19	target
0	17	17	32	19	16	25	12	26	34	...	28	22	16	22	20	18	23	cat
1	6	9	8	4	4	11	6	7	11	...	10	4	4	9	3	3	2	dog
2	17	27	16	27	19	17	20	25	14	...	22	20	21	14	16	17	24	dog
3	31	26	40	28	24	27	39	32	43	...	36	25	41	33	25	28	23	cat
4	3	7	4	5	2	7	8	6	10	...	9	5	6	9	9	7	6	dog
5	10	30	23	24	14	14	12	34	35	...	17	29	19	30	10	16	29	cat
6	20	5	20	16	18	30	23	12	9	...	32	11	9	11	13	13	6	cat
7	24	40	11	35	17	36	31	82	13	...	20	36	18	13	11	17	23	dog

Classification

K nearest neighbors, decision tree, support vector machine, random forest

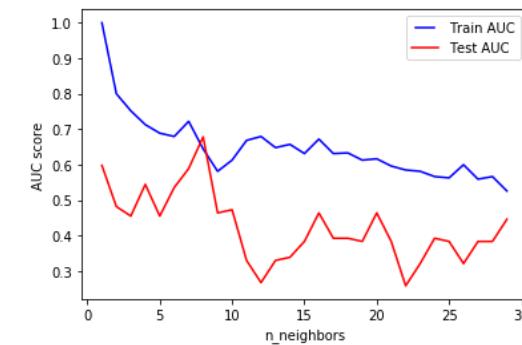
FINDINGS



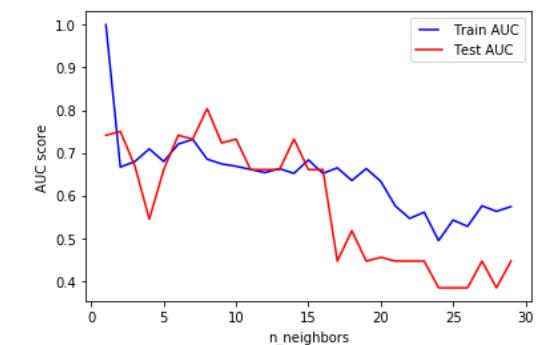
Changing K-means number of cluster within a range:

1. Preserves the proportion difference in feature frequency between classes
2. Only affects KNN performance

N=20; K=7
Test acc. = 0.6



N=50; K=8
Test acc. = 0.88

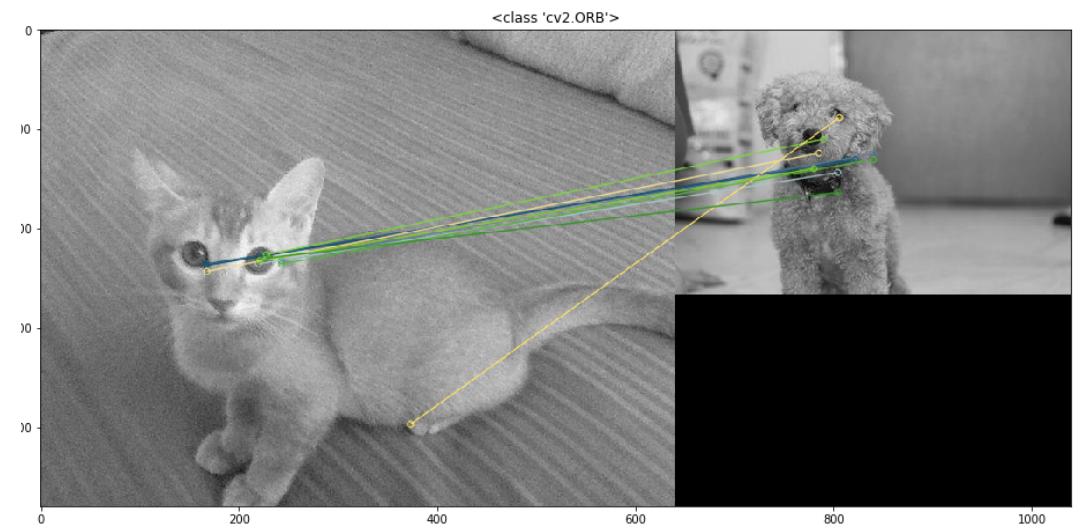
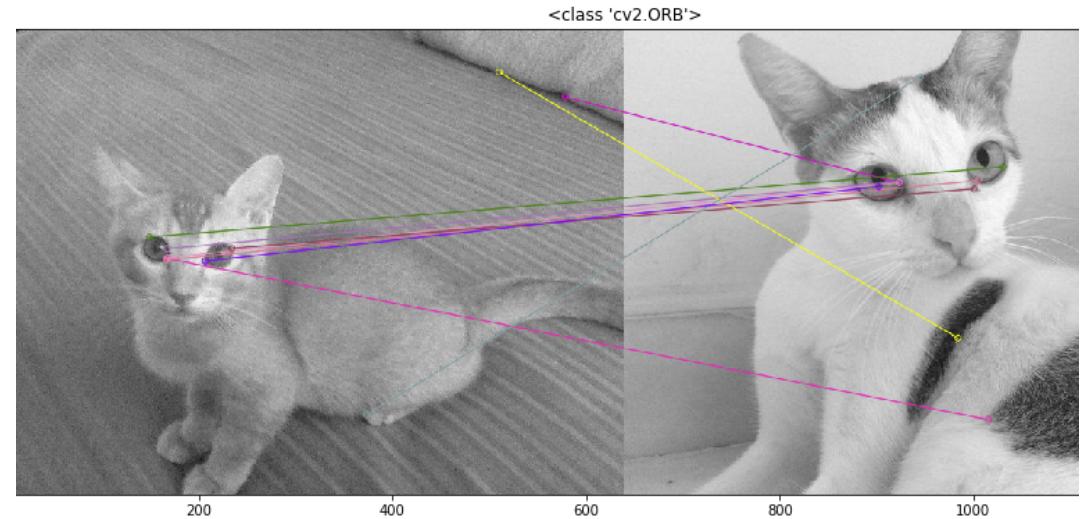
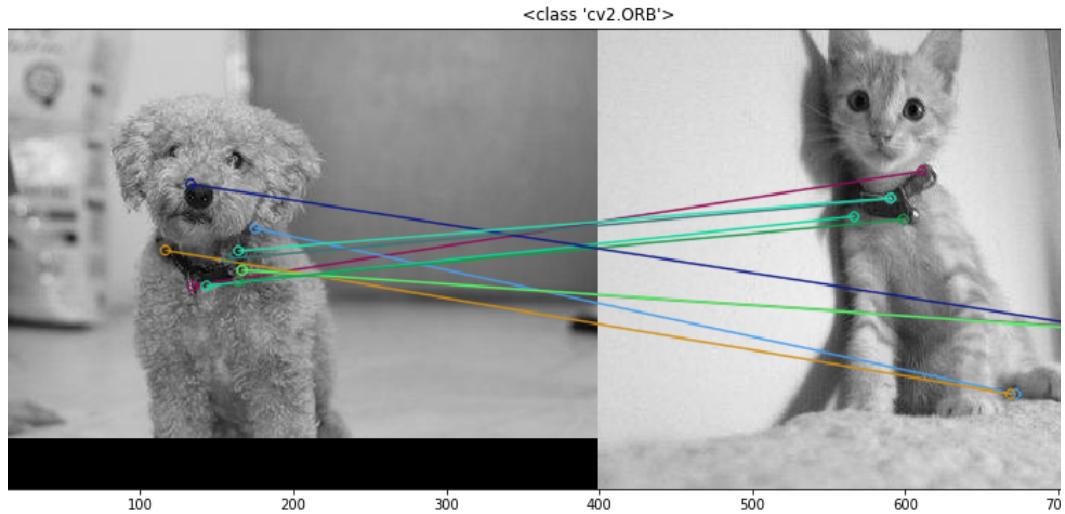


RESULTS

Number of clusters=20		K nearest neighbors K=7	Decision tree	SVM	Random forest N=100
Test set	Accuracy	0.6	0.667	0.533	0.733
	AUC	0.6	0.661	0.5	0.723
CV=5 (+/- 1 std.)	Accuracy	0.46 (+/- 0.13)	0.61 (+/-0.16)	0.53 (+/-0.01)	0.55 (+/- 0.18)
	AUC	0.49 (+/-0.21)	0.61 (+/-0.15)	0.58 (+/- 0.1)	0.58 (+/- 0.22)

- Decision tree is robust to change in clusters
- Sample size is small

Image matching



FUTURE WORK



Object localization



Natural language models



Feature engineering