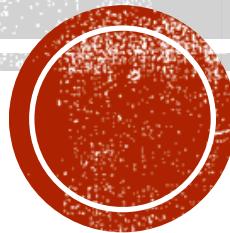


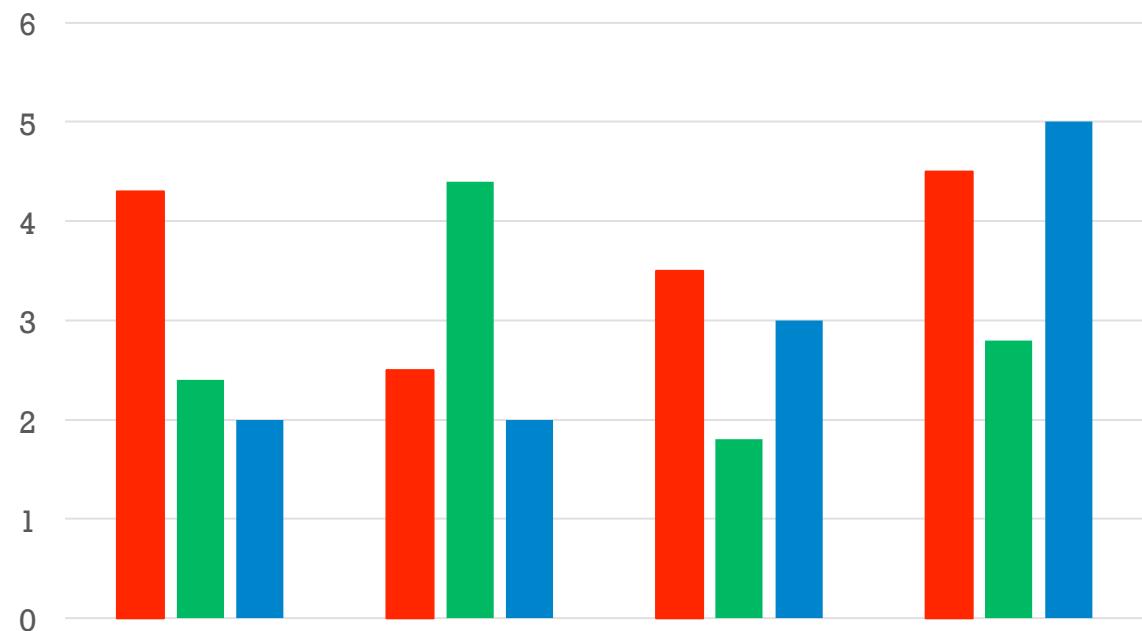
CATS VS DOGS

Group 3

Chenlu Ji, Chi Li, Danmo Wang, Haoyang Chen, Ruixiong Shi



BASE MODEL



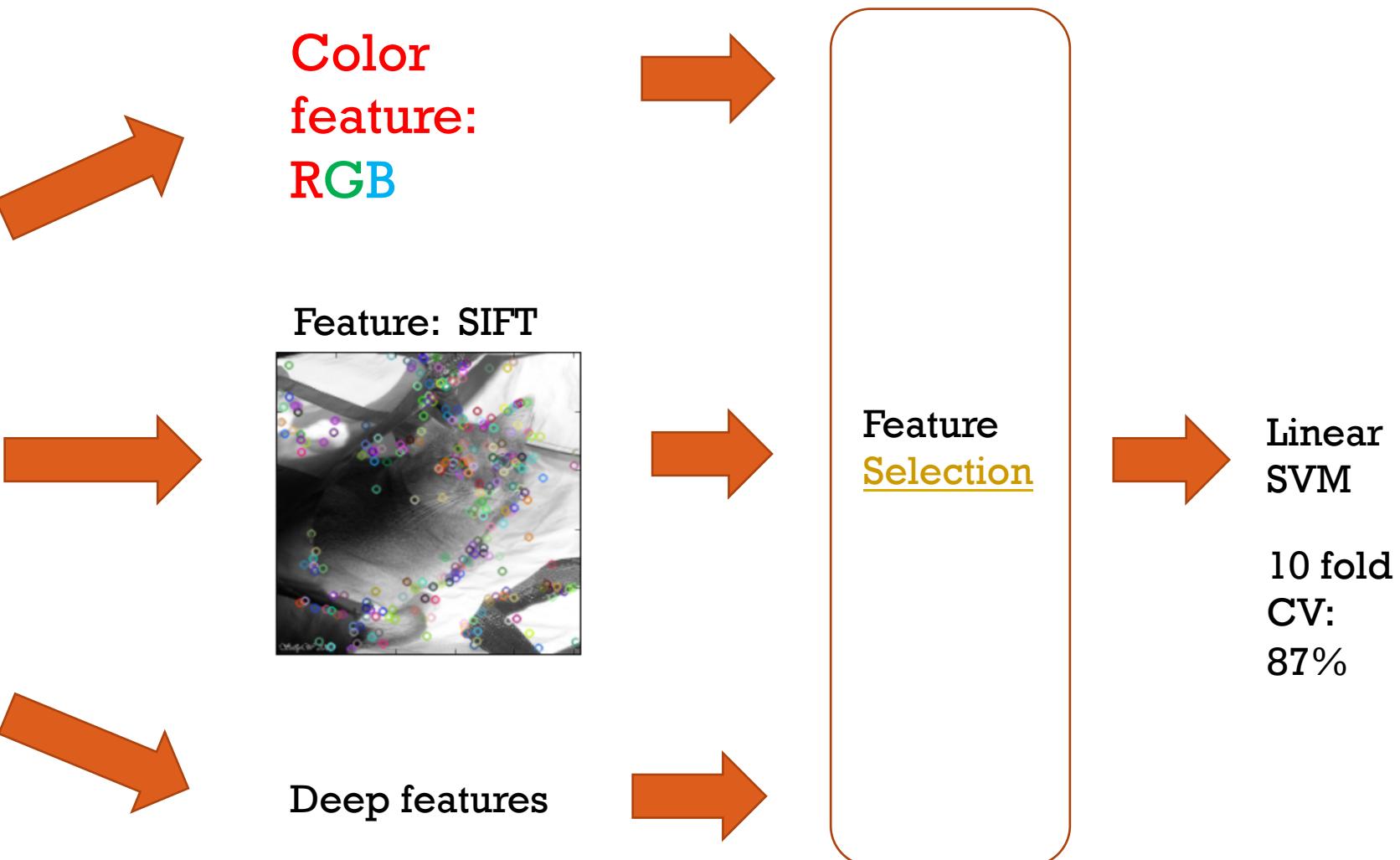
BASE MODEL

- We divided red, green and blue channel into 8 intervals
- Count frequency
- Train a linear SVM with $C = 1$ (default) give us 60.3% accuracy with 10 fold cross-validation

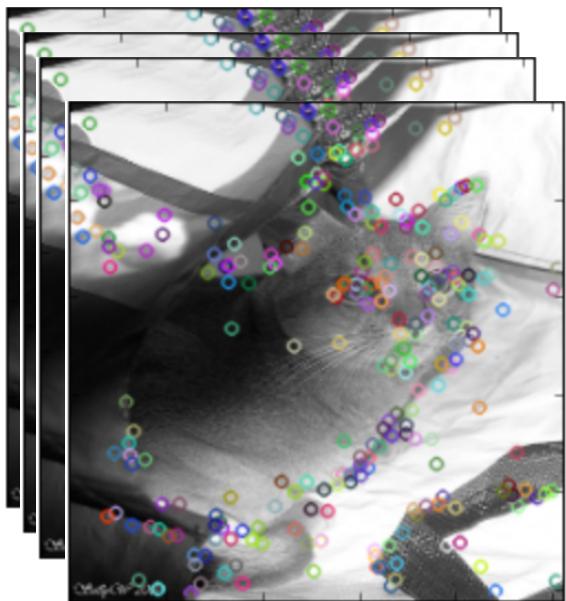


THE FINAL MODEL





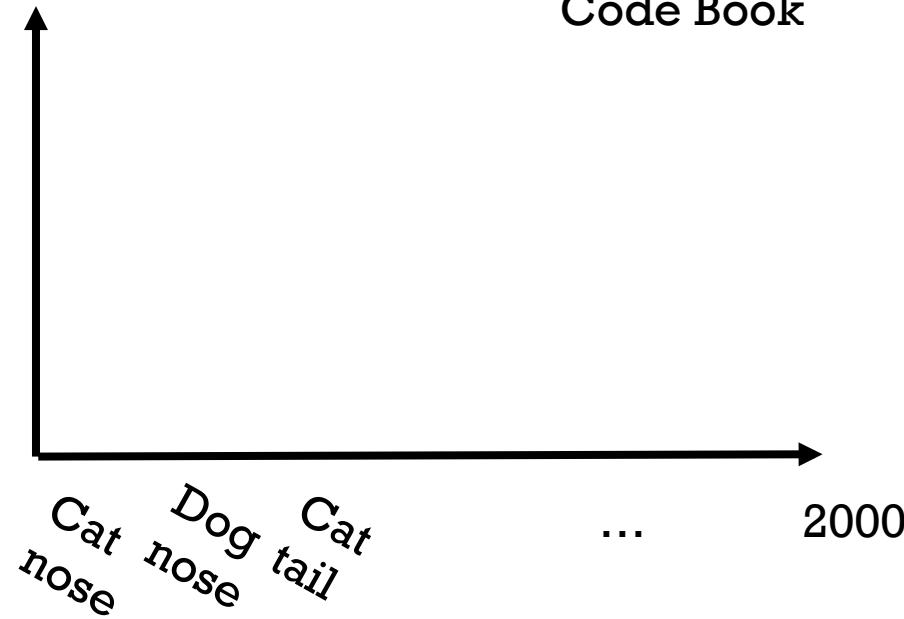
SIFT

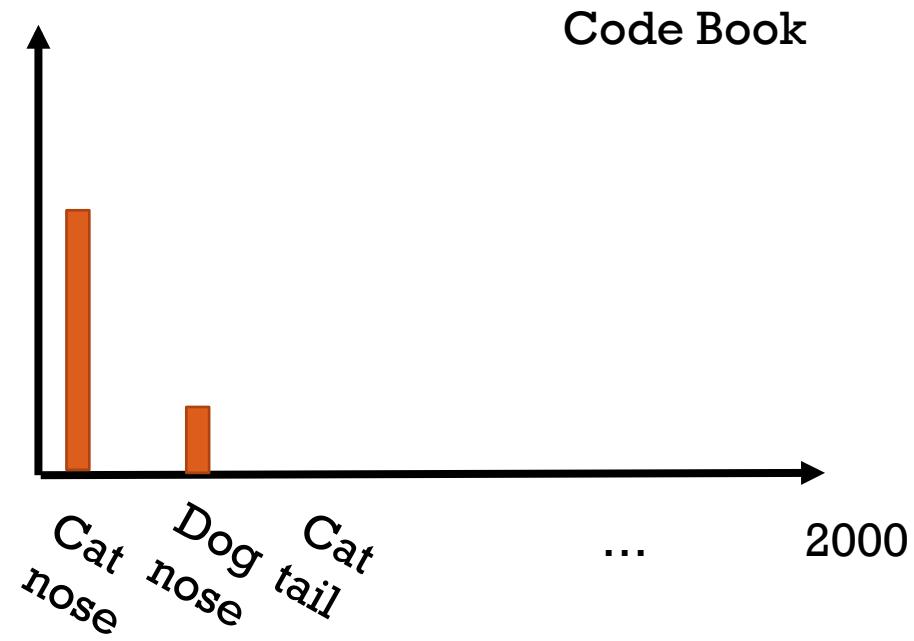
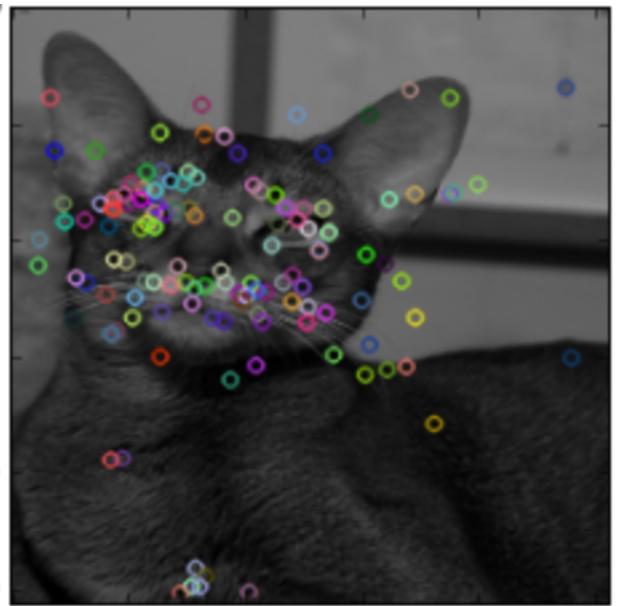


K-means with
2000 centroids



Code Book





COMPARING WITH BASE MODEL

Feature	Classifier	Accuracy (cv= 10)
RGB	Linear SVM C=1	60.3%
SIFT 1000 words Iteration = 10	Linear SVM C=1	68%
SIFT 2000 word, iteration = 10	Linear SVM C=1	71%
SIFT 2000 word, iteration = 30	Linear SVM C=1	74%
SIFT 2000 words iteration = 30	SVM chi2 kernel C = 1 Gamma = 1	79%

[Back](#)

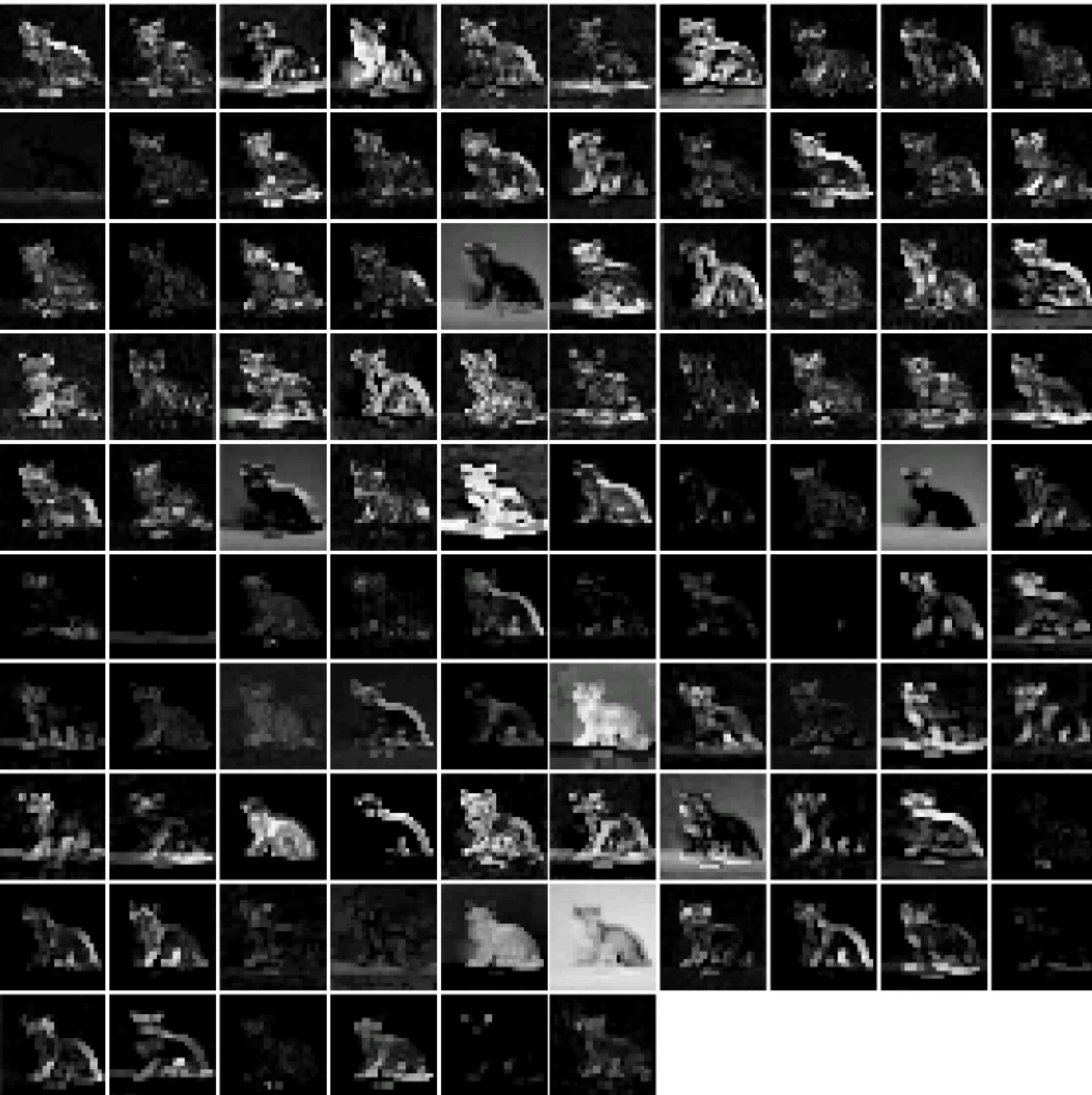


First layer

96 kernels

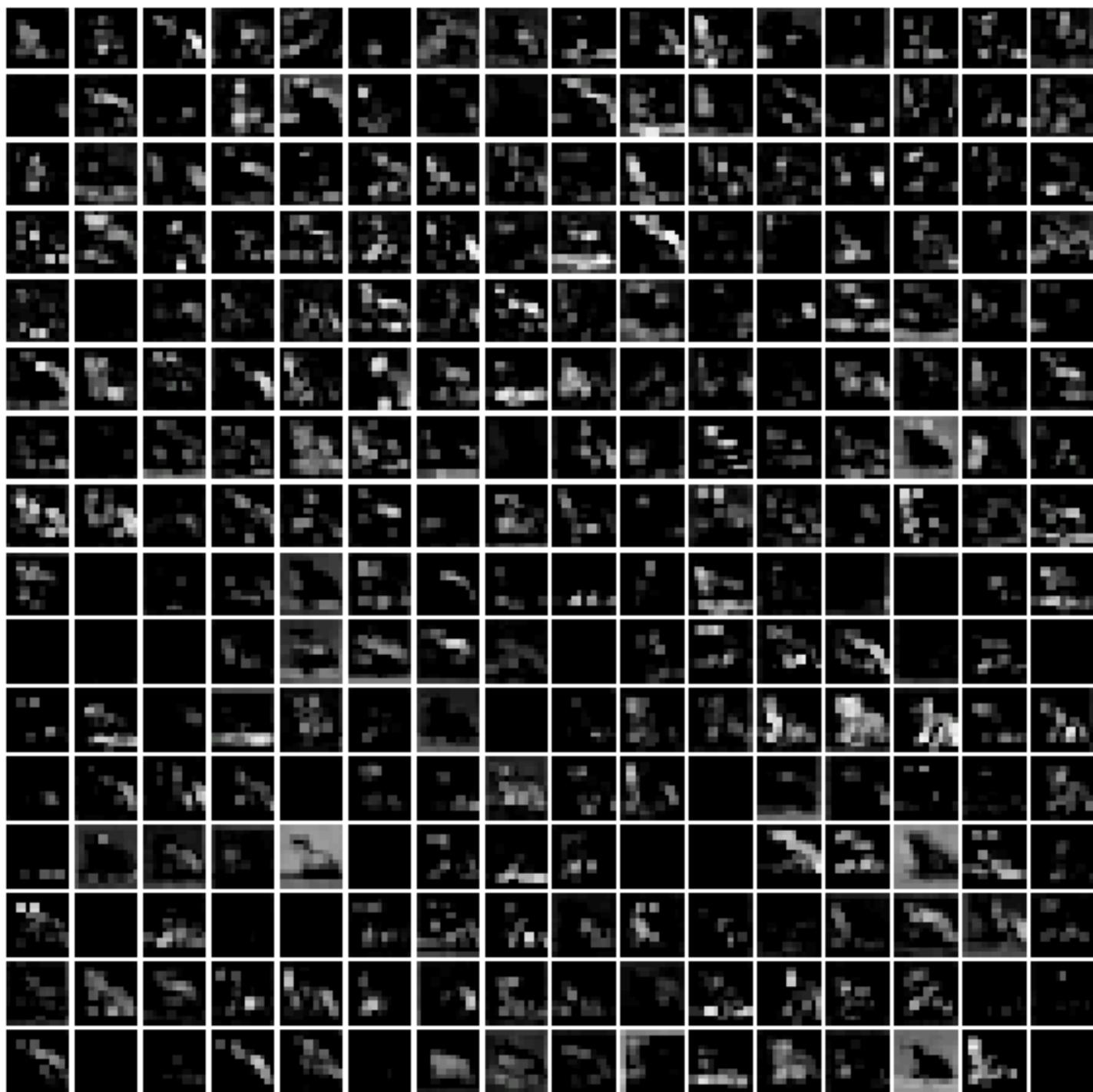


Shape



- The head and body (key info) of the cat is captured
- The relative position of these information is also been captured

Second
Layer



COMPARING WITH BASE MODEL AND SIFT

Feature	Classifier	Accuracy
RGB	Linear SVM C=1	60.3%
SIFT 2000 words iteration = 30	SVM chi2 kernel C = 1 Gamma = 1	79%
Deep feature Layer 1	Linear SVM C = 1	73%
Deep feature Layer 2	Linear SVM C = 1	81% (5 min)
Deep feature Layer 2	SGD classifier	80.5% (10 sec)

FEATURE SELECTION

Ensemble different features with different classifiers and vote for final prediction

- Performance is unstable (2% - 3%)
- R:
 - no customize kernel
 - no SGDclassifier
 - slow SVM training with huge dimensional data (nearly 50000 features)
 - Can not solve problem with rPython

Stack all the features together

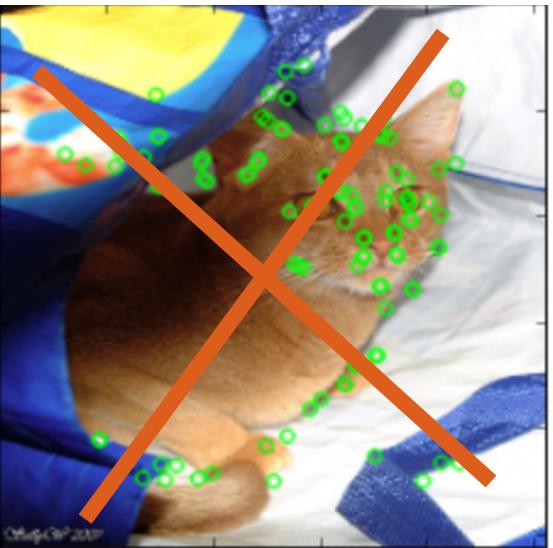
- Random forest to evaluate importance of feature
- Dimension reduction (9000)
- 5%-6%



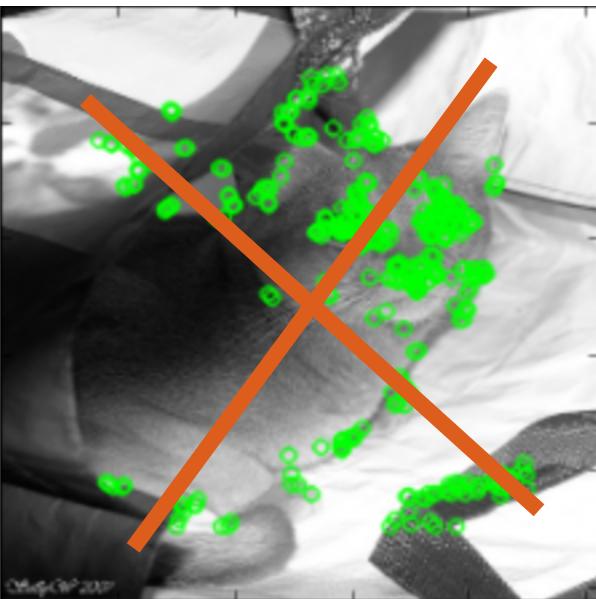
**OTHER THING WE HAVE
TRIED**



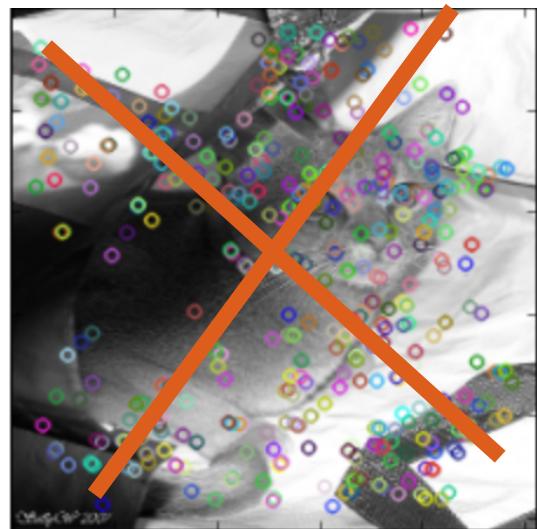
AKAZE



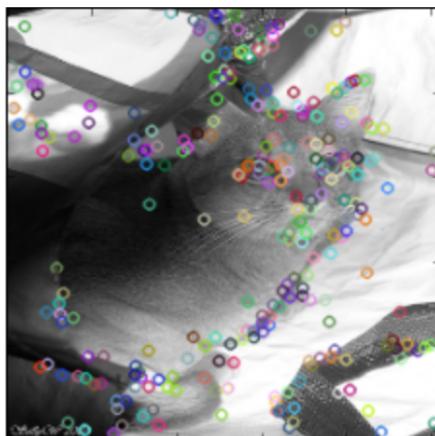
ORB



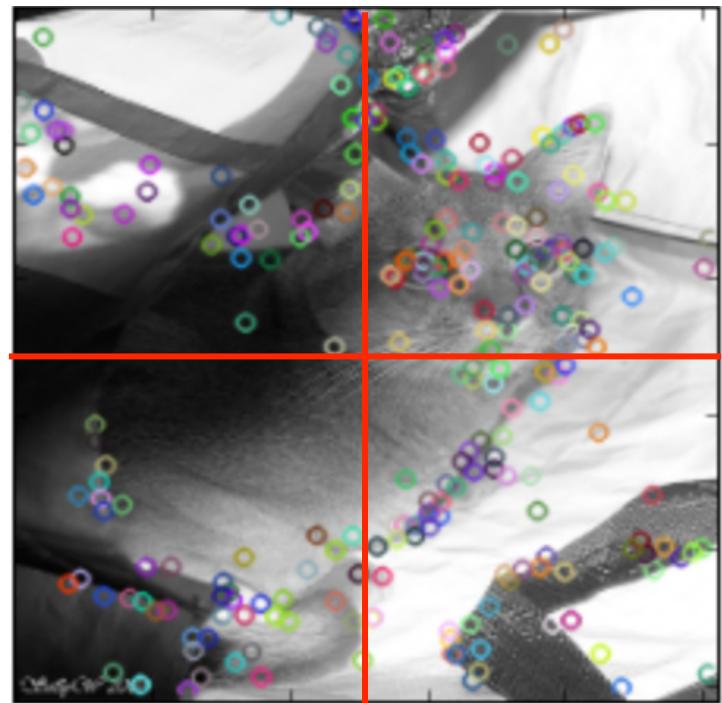
SURF



SIFT

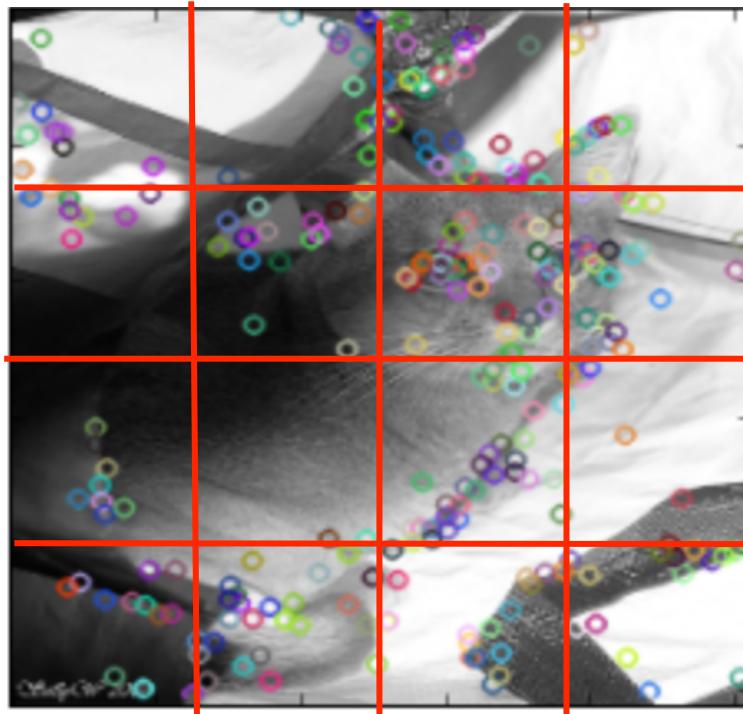


Level 1 SIFT

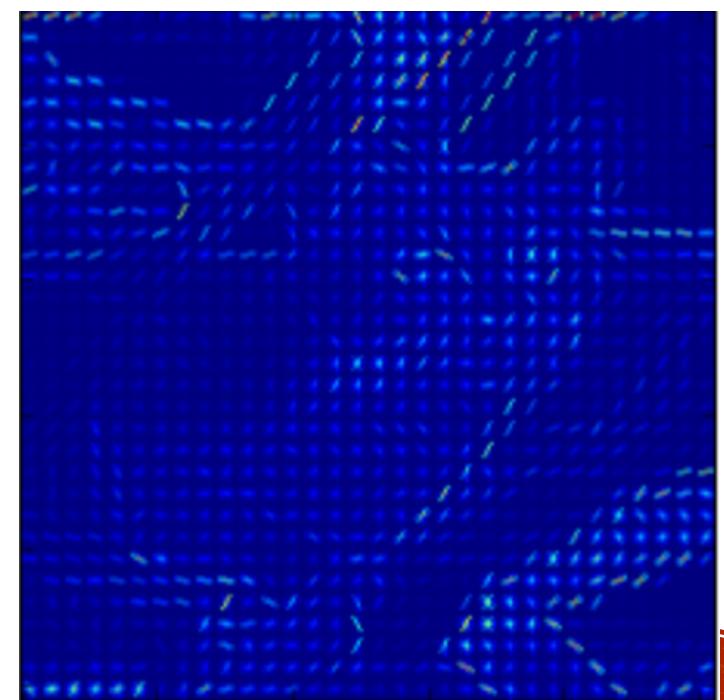


HSV

Leevl 2 SIFT



HOG



NORMALIZATION

- For the SIFT code book
- L1 normalization
- TF-IDF normalization



THINGS WE CAN DO BETTER

- Preprocessing:
 - Deep feature: extract mean, resize to 227*227
 - SIFT: resize to 256*256
 - RGB: none
 - Keep Ratio, crop at center
- Do not use R for training
 - 50000 → 9000, lose information
 - Faster training algorithms in Python
- Tune final SVM classifier harder (current C = 1000)



THANK YOU!

We are group 3!

