Project 03

Dogs, Fried Chicken, or Blueberry Muffins?

Team 8:
Hongyu Li
Jia Zheng
Jiaqi Dong
Xiaochen Fan

Feature extraction:

SIFT, RGB

HOG (histogram of oriented gradients)—similar with RGB, extract features as black and white

ORB (Oriented FAST and rotated BRIEF)—similar with SIFT, a fast and efficient alternative to SIFT

SURF (speeded up robust features)—classic method of feature extraction

Model:

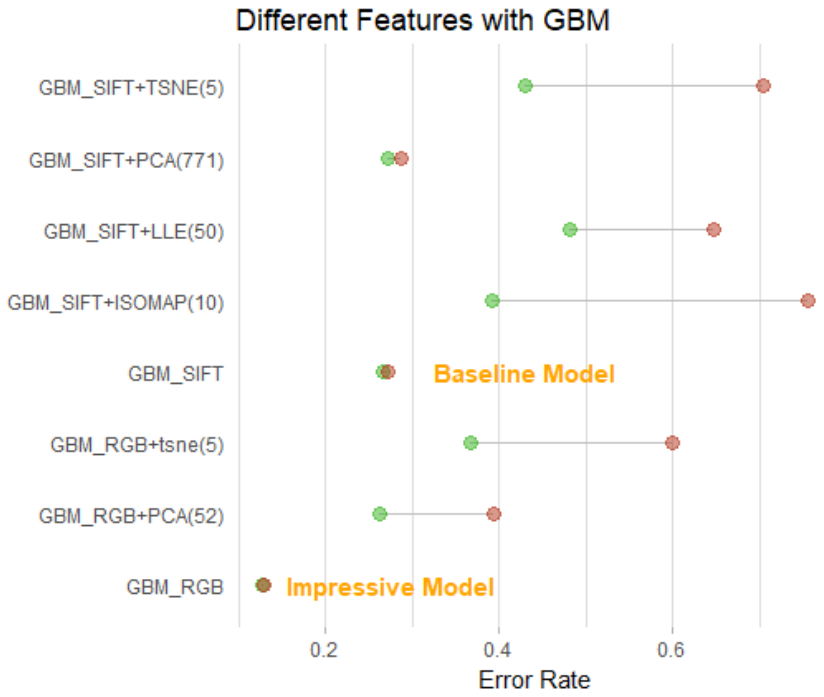
GBM

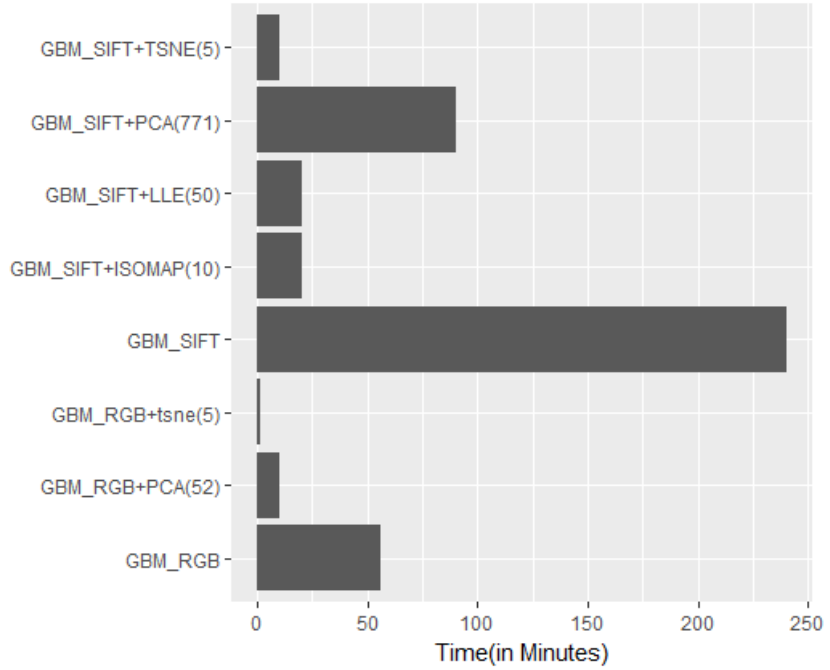
Xgboost—similar with GBM, a boosting algorithm

Random Forest—decision tree algorithm, not additive model

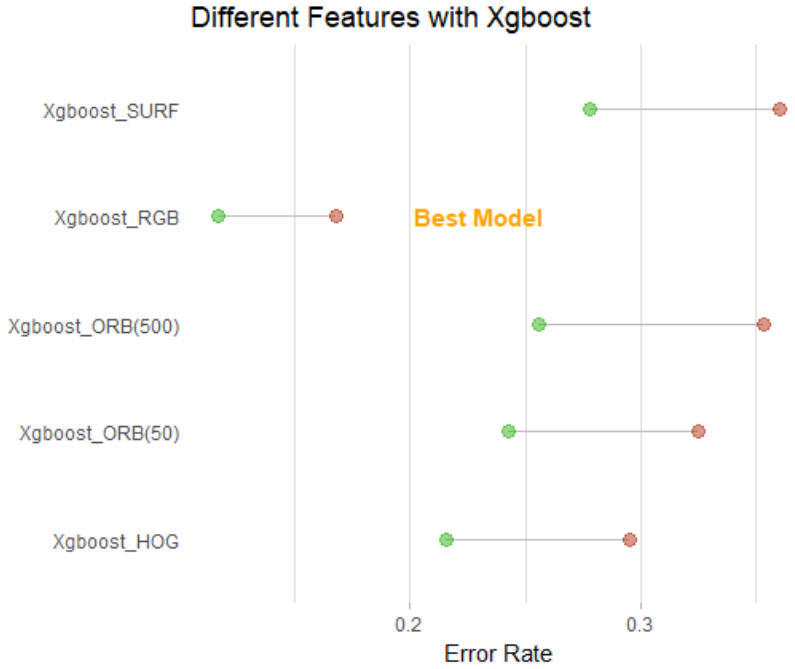
SVM (supporting vector machine)—when feature extraction is done, seems like supporting vector; therefore, try SVM

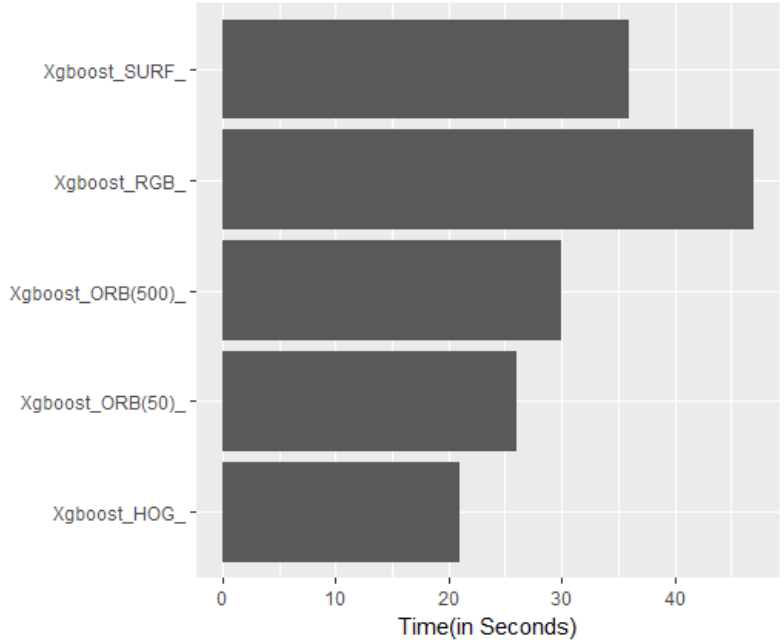
Model	Feature extraction	CV error	Test error	Time
GBM	**SIFT**	**0.2658**	**0.2716**	**252**
GBM	SIFT+PCA	0.2725	0.2883	102
GBM	SIFT+TSNE	0.43	0.7033	22
GBM	SIFT+LLE	0.4821	0.6467	32
GBM	SIFT+ISOMAP	0.3912	0.755	32
GBM	**RGB**	**0.1279**	**0.13**	**68**
GBM	RGB+TSNE	0.3667	0.6	18
GBM	RGB+PCA	0.2625	0.3933	27


Different Features with GBM
GBM_SIFT+TSNE(5)
GBM_SIFT+PCA(771)
GBM_SIFT+LLE(50)
GBM_SIFT+ISOMAP(10)
GBM_SIFT
GBM_RGB+tsne(5)
GBM_RGB+PCA(52)
GBM_RGB
Baseline Model
Impressive Model
0.2
0.4
0.6
Error Rate

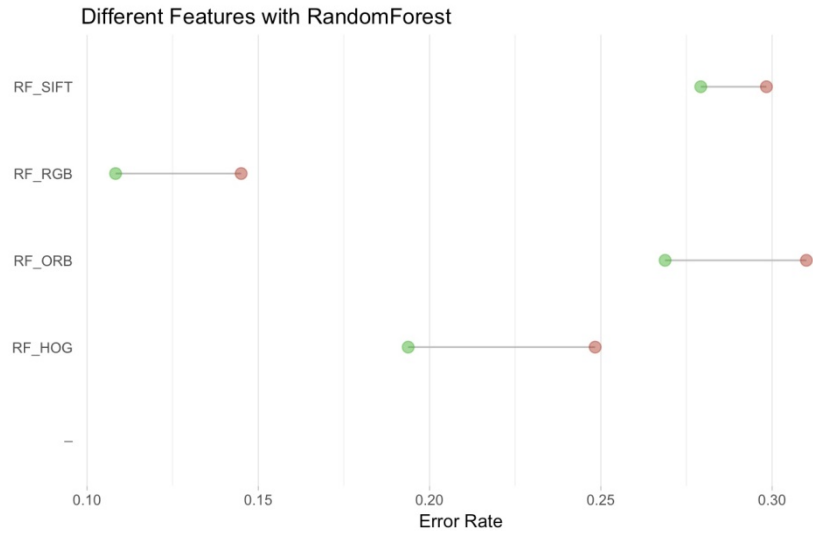

GBM_SIFT+TSNE(5)
GBM_SIFT+PCA(771)
GBM_SIFT+LLE(50)
GBM_SIFT+ISOMAP(10)
GBM_SIFT
GBM_RGB+tsne(5)
GBM_RGB+PCA(52)
GBM_RGB
0
50
100
150
200
250
Time(in Minutes)

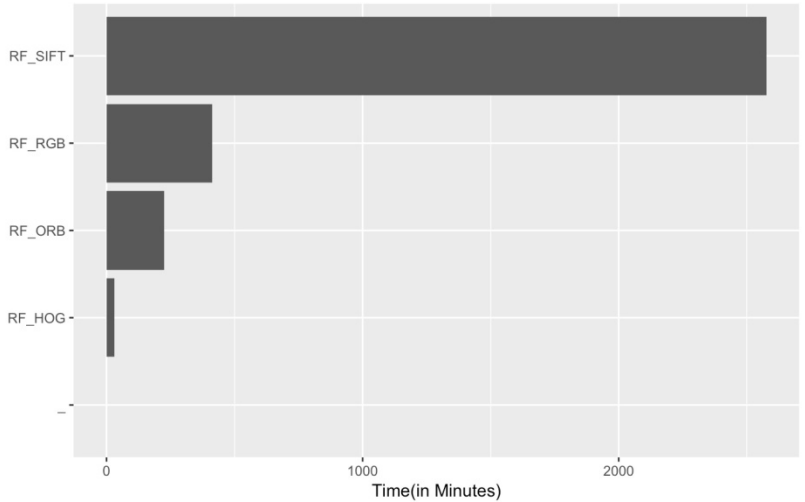
Model	Feature extraction	CV error	Test error	Time
Xgboost	**RGB**	**0.117**	**0.1683**	**18**
Xgboost	HOG	0.216	0.295	5
Xgboost	ORB(500)	0.256	0.3533	6
Xgboost	ORB(50)	0.243	0.325	6
Xgboost	SURF	0.278	0.36	11


Different Features with Xgboost
Xgboost_SURF
Xgboost_RGB
Best Model
Xgboost_ORB(500)
Xgboost_ORB(50)
Xgboost_HOG
0.2
0.3
Error Rate


Xgboost_SURF_
Xgboost_RGB_
Xgboost_ORB(500)_
Xgboost_ORB(50)_
Xgboost_HOG_
0
10
20
30
40
Time(in Seconds)

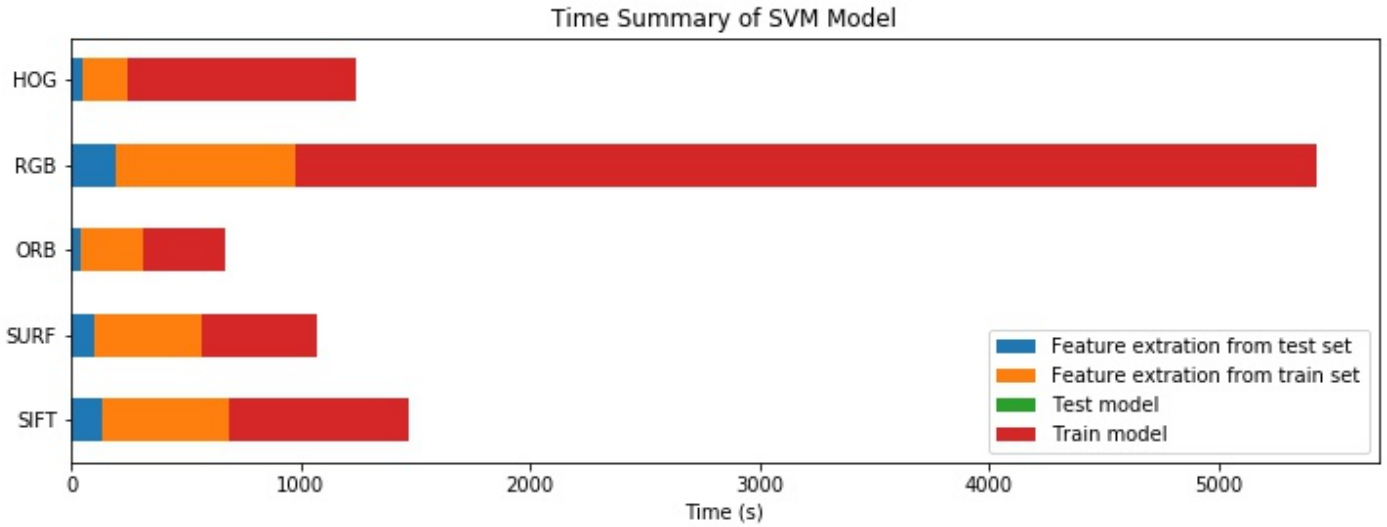
Model	Feature extraction	CV error	Test error	Time
RF	SIFT	0.279167	0.298333	55
RF	**RGB**	**0.108333**	**0.145**	**24**
RF	HOG	0.19375	0.248333	5
RF	ORB	0.26875	0.31	9

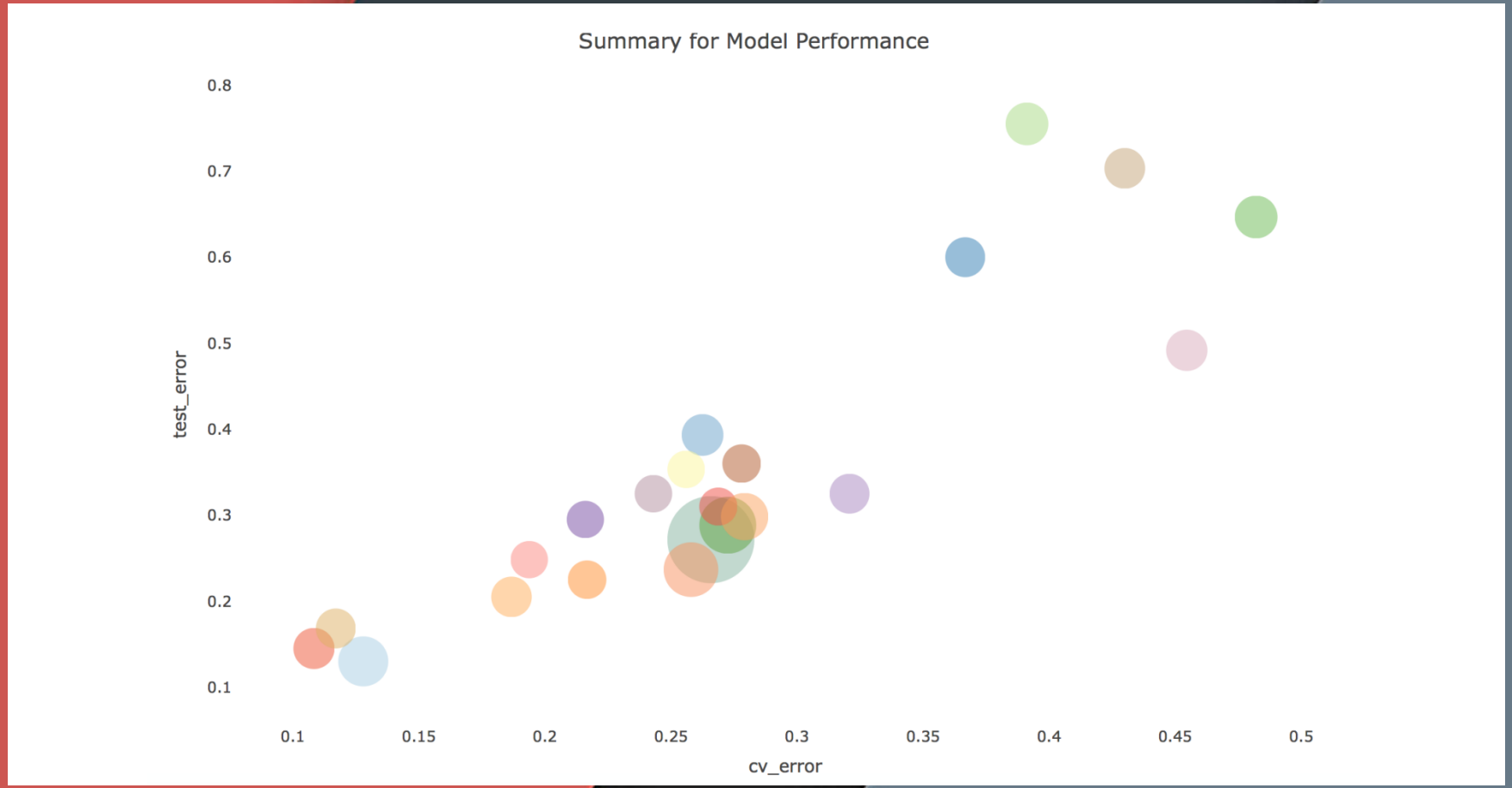

Different Features with RandomForest
RF_SIFT
RF_RGB
RF_ORB
RF_HOG
_
0.10
0.15
0.20
0.25
0.30
Error Rate


RF_SIFT
RF_RGB
RF_ORB
RF_HOG
_
0
1000
2000
Time(in Minutes)

Model	Feature extraction	CV error	Test error	Time
SVM	SIFT	0.4546	0.4917	25
SVM	SURF	0.3208	0.325	18
SVM	ORB	0.2167	0.225	11
SVM	RGB	0.2579	0.2367	91
SVM	**HOG**	**0.1867**	**0.205**	**21**


Error Summary of SVM Model
CV Error
Test Error
SIFT
SURF
ORB
RGB
HOG


Time Summary of SVM Model
HOG
RGB
ORB
SURF
SIFT
0
1000
2000
3000
4000
5000
Time (s)
Feature extration from test set
Feature extration from train set
Test model
Train model


Summary for Model Performance
test_error
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.1
0.15
0.2
0.25
0.3
0.35
0.4
0.45
0.5
cv_error

Model	Feature extraction	CV error	Test error	Time
GBM	**SIFT**	**0.2658**	**0.2716**	**252**
GBM	**RGB**	**0.1279**	**0.13**	**68**
Xgboost	**RGB**	**0.117**	**0.1683**	**18**
RF	**RGB**	**0.108333**	**0.145**	**24**
SVM	**HOG**	**0.1867**	**0.205**	**21**

Compared to baseline (GBM+SIFT), Xgboost+RGB is better in each aspects.