

Object Recognition Results

Feature detector	Feature descriptor	Feature matcher	Classifier	Vocabulary build time (1000 word size - 177 images)	Training samples build time (177 images)	Training samples retrieved (from 177 images)	Classifier training time	Classifier test time - 177 images (sliding window with 482 ROIs per image)	Accuracy	Precision	Recall
STAR	SIFT	FLANN	Artificial Neural Network	00m31.204s	00m44.265s	403	00m00.028s	15m14.323s	0.874	0.234	0.162
STAR	SURF	FLANN	Support Vector Machine	00m21.251s	00m17.901s	403	00m38.217s	03m02.452s	0.855	0.271	0.214
STAR	SURF	BFMatcher	Support Vector Machine	00m20.932s	00m17.985s	403	00m37.934s	03m33.083s	0.854	0.299	0.234
STAR	SIFT	FLANN	Support Vector Machine	00m31.204s	00m44.265s	403	00m36.318s	09m43.652s	0.847	0.306	0.362
STAR	BRIEF	FLANN	Support Vector Machine	00m20.131s	00m20.105s	403	00m35.184s	03m46.283s	0.841	0.276	0.277
ORB	ORB	FLANN	Artificial Neural Network	01m25.694s	00m43.962s	447	00m00.188s	17m04.451s	0.839	0.206	0.195
STAR	FREAK	FLANN	Support Vector Machine	00m20.824s	00m24.739s	403	00m36.273s	05m22.562s	0.815	0.274	0.279
SURF	SURF	FLANN	Artificial Neural Network	00m37.574s	00m35.434s	457	00m00.201s	13m03.423s	0.815	0.168	0.202
SIFT	SIFT	BFMatcher	Artificial Neural Network	01m46.338s	01m32.902s	488	00m00.234s	43m00.362s	0.794	0.217	0.296
SIFT	SIFT	BFMatcher	Support Vector Machine	01m40.631s	01m30.025s	488	00m49.265s	41m43.748s	0.784	0.242	0.385
SIFT	SIFT	FLANN	Support Vector Machine	01m46.338s	01m32.902s	488	00m50.727s	42m32.801s	0.776	0.251	0.411
ORB	ORB	FLANN	Support Vector Machine	01m25.695s	00m43.966s	447	00m44.078s	16m56.037s	0.739	0.239	0.549
SIFT	SURF	FLANN	Support Vector Machine	01m17.674s	00m43.966s	488	00m51.802s	27m05.743s	0.714	0.219	0.543
SIFT	SURF	BFMatcher	Support Vector Machine	01m11.727s	00m39.477s	488	00m50.481s	26m21.736s	0.705	0.213	0.528
GFTT	FREAK	FLANN	Support Vector Machine	01m01.011s	00m40.011s	497	00m50.479s	40m07.149s	0.699	0.201	0.478
MSER	SURF	FLANN	Support Vector Machine	00m22.772s	00m20.369s	464	00m47.321s	07m41.181s	0.672	0.241	0.735
FAST	FREAK	FLANN	Support Vector Machine	00m56.567s	01m49.256s	514	00m54.863s	51m38.865s	0.666	0.204	0.596
BRISK	BRISK	FLANN	Support Vector Machine	00m21.704s	00m30.616s	464	00m47.038s	13m12.818s	0.661	0.213	0.682
SIFT	BRIEF	FLANN	Support Vector Machine	01m03.294s	00m47.819s	488	00m48.438s	29m37.773s	0.616	0.187	0.661
SIFT	FREAK	BFMatcher	Support Vector Machine	01m08.355s	00m38.618s	488	00m49.269s	25m22.225s	0.606	0.188	0.696
SIFT	FREAK	FLANN	Support Vector Machine	01m06.325s	01m00.102s	488	00m53.349s	35m35.147s	0.605	0.191	0.717
SIFT	BRIEF	BFMatcher	Support Vector Machine	01m05.877s	00m38.599s	488	00m50.382s	25m04.586s	0.601	0.191	0.732
BRISK	FREAK	FLANN	Support Vector Machine	00m30.058s	00m29.093s	464	00m45.131s	11m03.882s	0.579	0.191	0.801
SURF	SURF	FLANN	Decision Tree	00m37.188s	00m34.271s	457	00m00.064s	18m05.666s	0.578	0.175	0.648
SURF	SURF	FLANN	Random Tree	00m37.073s	00m43.967s	457	00m00.199s	16m17.609s	0.503	0.172	0.847
SURF	SURF	FLANN	Boosting Tree	00m37.495s	00m43.962s	457	00m00.956s	15m41.621s	0.499	0.171	0.845
SURF	SURF	FLANN	Extremely Random Tree	00m35.759s	00m43.969s	457	00m00.491s	18m33.911s	0.469	0.167	0.864
ORB	ORB	FLANN	Normal Bayes Classifier	01m24.585s	00m26.650s	487	00m05.779s	27m22.274s	0.446	0.165	0.886
SURF	SURF	FLANN	Gradient Boosting Tree	00m37.207s	00m43.964s	457	00m04.295s	17m23.841s	0.423	0.161	0.897
SIFT	BRISK	FLANN	Support Vector Machine	01m08.126s	01m00.105s	488	00m49.559s	45m40.242s	0.421	0.159	0.889

Accuracy: $(\text{truePositives} + \text{trueNegatives}) / (\text{truePositives} + \text{trueNegatives} + \text{falsePositives} + \text{falseNegatives})$

Precision: $\text{truePositives} / (\text{truePositives} + \text{falsePositives})$

Recall: $\text{truePositives} / (\text{truePositives} + \text{falseNegatives})$

Measures obtained comparing the objects ground truth masks with the recognition results

Recognition results were computed using a sliding window, that gathered the most probable image regions where the target object could be located

These regions were computed with a voting mask, and the final results were thresholded to remove regions with small number of votes