



FaceRecognition

Performance Report

2018-02-07

OPEN AI LAB

Revision Record

Date	Rev	Change Description	Author
2018-01-11	0.1.1	Update performance on RK3399	Luo Hao /Huifang
2018-01-30	0.1.2	QA Test	Luo Hao
2018-02-07	0.1.3	QA Test	Shulian Shen

catalog

catalog	2
1 Purpose	3
2 Test Environment	3
3 Face Recognition Flow	4
4 Performance with Arm Compute Library(ACL) BYPASSACL=0	4
4.1 Single A53 CPU @1.42GHz	4
4.2 Single A72 CPU @1.8GHz	6
4.3 Two A72 CPU @1.8GHz	7
4.4 Four A53 CPU @1.42GHz	8
5 Performance with OpenBLAS (BYPASSACL=0xffff)	9
5.1 Single A53 CPU @1.42GHz	9
5.2 Single A72 CPU @1.8GHz	10
5.3 Two A72 CPU @1.8GHz	11
5.4 Four A53 CPU @1.42GHz	12
6 Accuracy Test	13
6.1 Detection test	13
6.2 Recognition test	13
7 Stress Test	15
8 Conclusion	15

1 Purpose

This Report is tested on RK3399 platform and the Caffe-HRT version is 0.5.0 with Arm Compute Library(ACL) 17.12/OpenBLAS.

2 Test Environment

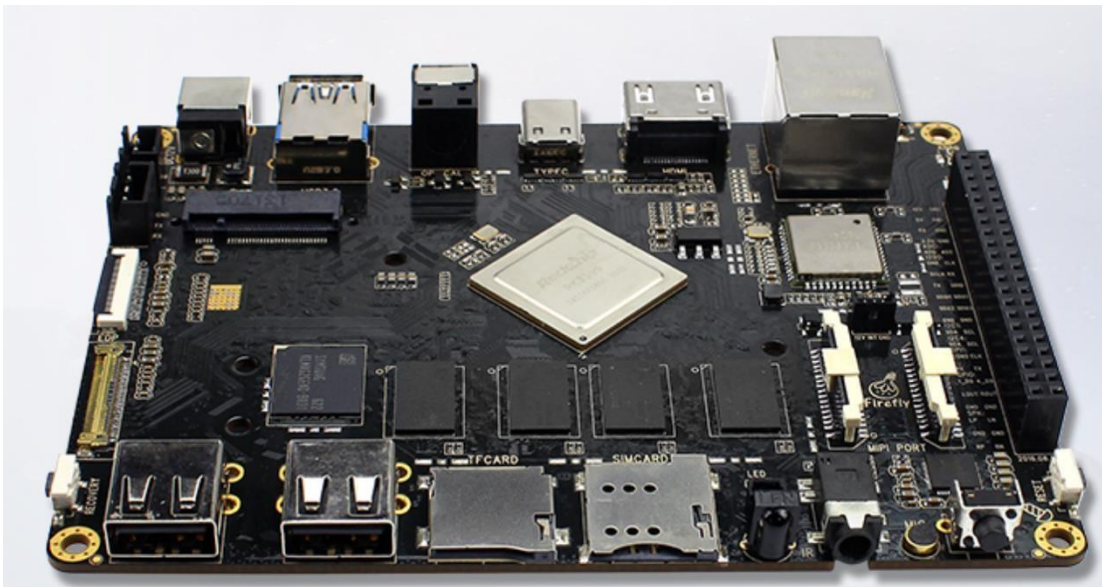
Hardware SoC: firefly

- <http://www.t-firefly.com/product/rk3399.html>
- GPU: Mali T864 (800MHz)
- ARM: 2G
- CPU: Dual-core Cortex-A72 up to 2.0GHz (real frequency is 1.8GHz); Quad-core Cortex-A53 up to 1.5GHz (real frequency is 1.4GHz)

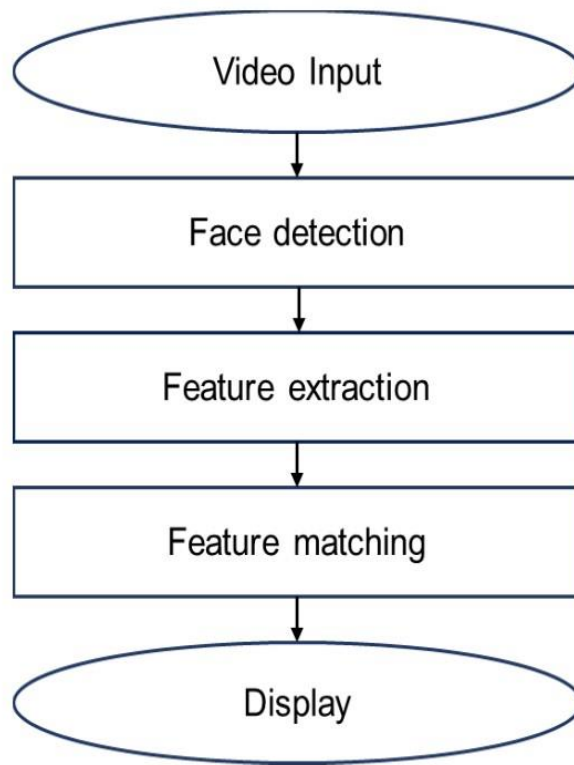
Operating System: Ubuntu 16.04

FaceRecognition vision: 0.1.3

All performance data is 100 frames average processing time.



3 Face Recognition Flow



Face detection : Detect the face from input frame

Feature extraction : Extrace features from detected face

Feature matching : search the Database according the extracted feature, find the matching face

Fig 3.1 The process of face recognition

4 Performance with Arm Compute Library(ACL) BYPASSACL=0

FaceRecognition performance is influenced by many factors, only faces, cores, minimum size of face are tested. Note: the alignment is not included. The performances with ACL are as follow.

4.1 Single A53 CPU @1.42GHz

Table 4.1 Performance of different part on Single A53

Test Case	Detection (ms/frame)	Extraction (ms/frame)	Verification (ms/frame)	Total (ms/frame)
480P 32*32 single face	540	1172	0	1715
480P 48*48 single face	273	1150	0	1425
480P 64*64 single face	185	924	0	1110
480P 32*32 two faces	573	2433	0	3006
480P 48*48 two faces	307	2291	0	2601
480P 64*64 two faces	251	2371	0	2625

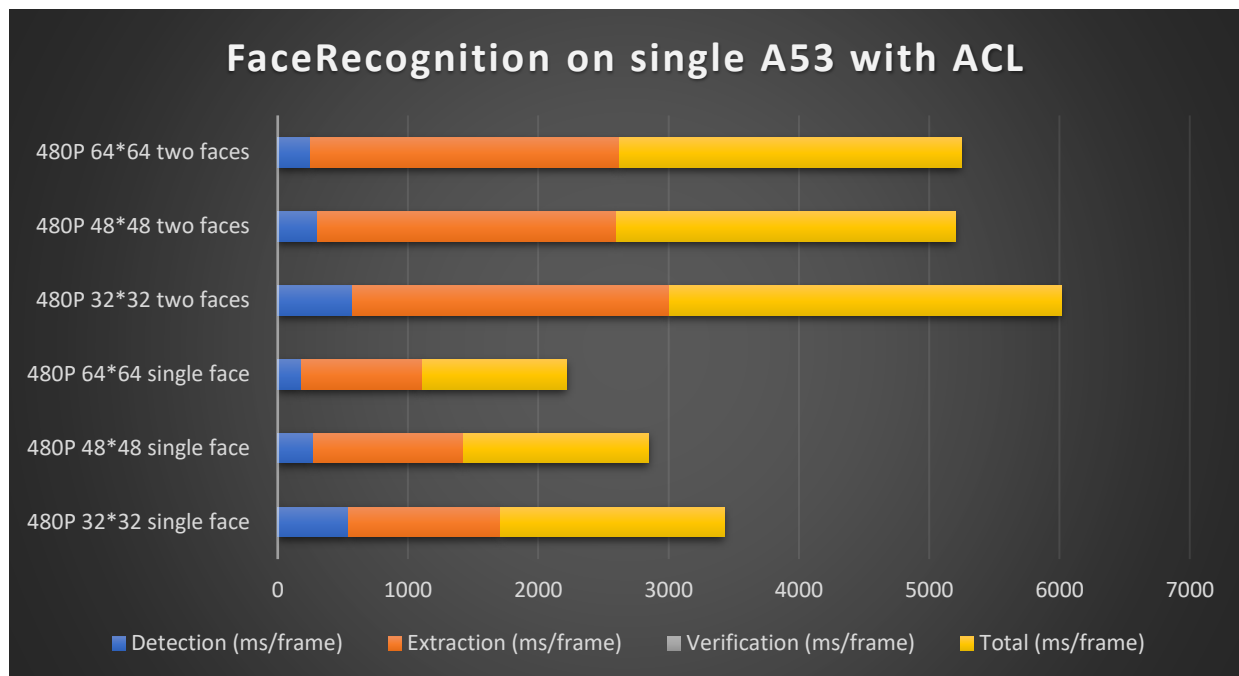


Fig 4.1 Performance Compare on Single A53

4.2 Single A72 CPU @1.8GHz

Table 4.2 Performance of different part on Single A72

Test Case	Detection (ms/frame)	Extraction (ms/frame)	Verification (ms/frame)	Total (ms/frame)
480P 32*32 single face	289	583	0	1033
480P 48*48 single face	134	747	0	882
480P 64*64 single face	74	575	0	650
480P 32*32 two faces	322	1537	0	1860
480P 48*48 two faces	171	1600	0	1773
480P 64*64 two faces	125	1699	0	1825

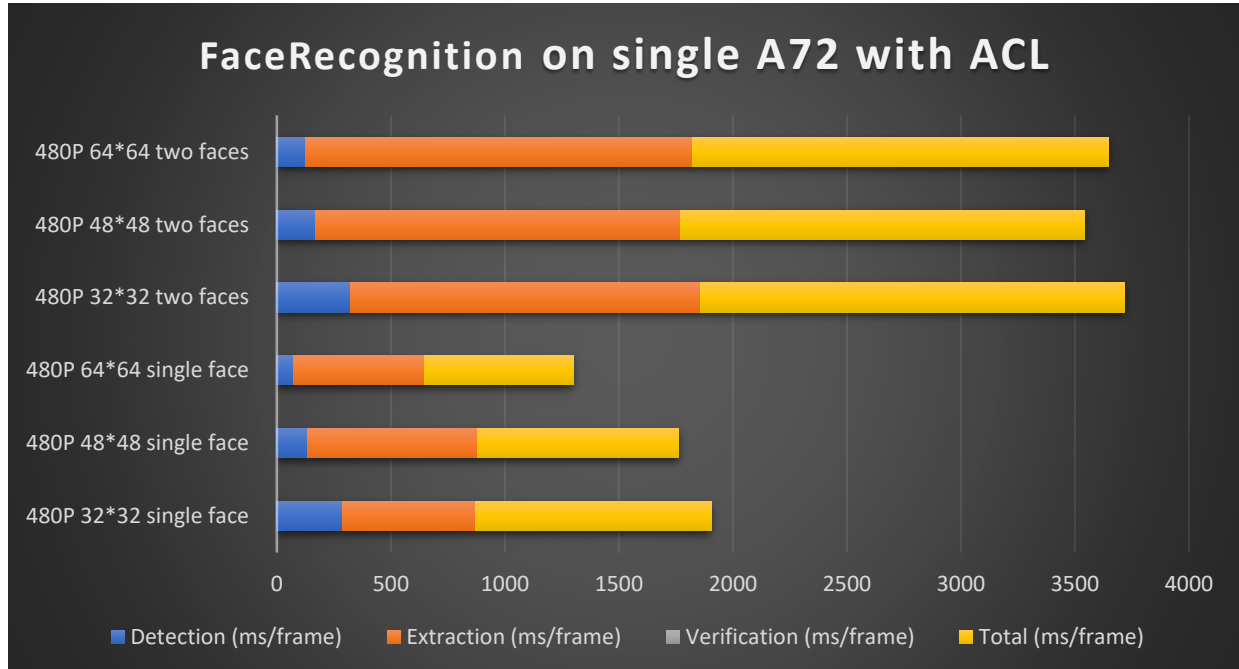


Fig 4.2 Performance Compare on Single A72

4.3 Two A72 CPU @1.8GHz

Table 4.3 Performance of different part on Two A72

Test Case	Detection (ms/frame)	Extraction (ms/frame)	Verification (ms/frame)	Total (ms/frame)
480P 32*32 single face	288	468	0	756
480P 48*48 single face	133	437	0	571
480P 64*64 single face	79	250	0	329
480P 32*32 two faces	359	926	0	1322
480P 48*48 two faces	161	1031	0	1194
480P 64*64 two faces	120	986	0	1108

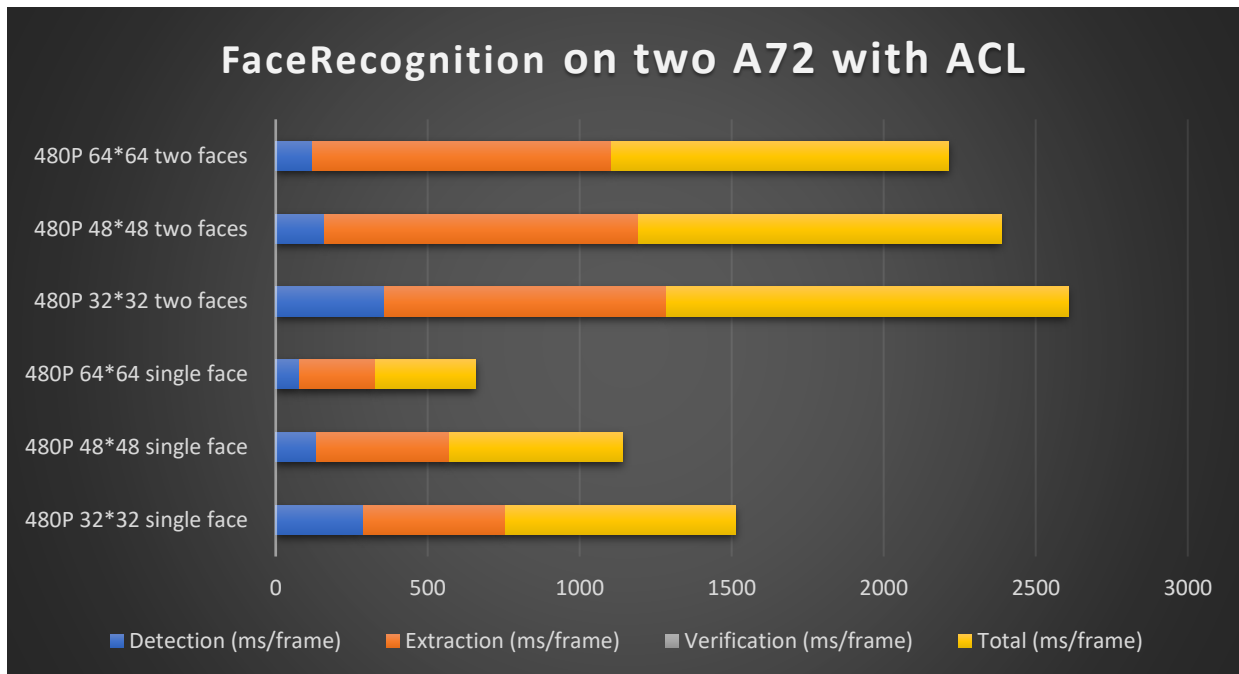


Fig 4.3 Performance Compare on two A72

4.4 Four A53 CPU @ 1.42GHz

Table 4.4 Performance of different part on Four A53

Test Case	Detection (ms/frame)	Extraction (ms/frame)	Verification (ms/frame)	Total (ms/frame)
480P 32*32 single face	395	493	0	890
480P 48*48 single face	191	440	0	633
480P 64*64 single face	126	407	0	534
480P 32*32 two faces	393	926	0	1322
480P 48*48 two faces	214	941	0	1159
480P 64*64 two faces	177	944	0	1125

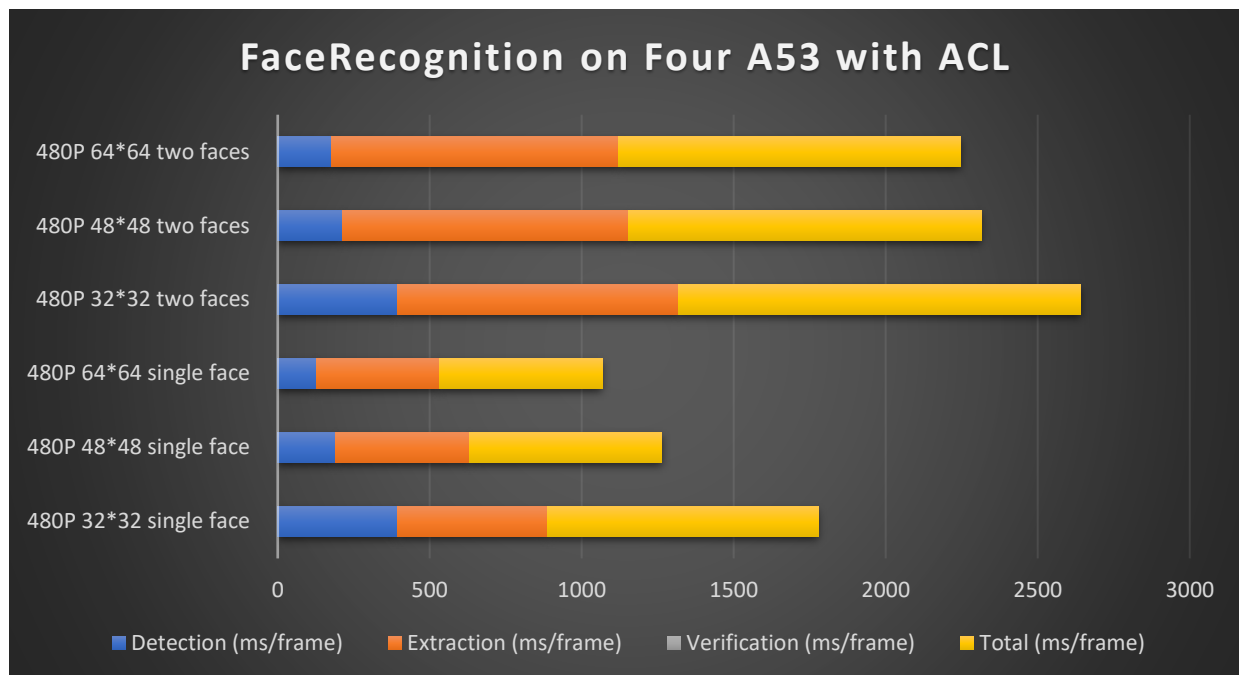


Fig 4.1 Performance Compare on Four A53

5 Performance with OpenBLAS (BYPASSACL=0Xffff)

Face recognition performance is influenced by many factors, only faces, cores, minimum size of face are test. Note: the alignment is not included. The performances with OpenBLAS are as follow.

5.1 Single A53 CPU @1.42GHz

Table 5.1 Performance of different part on Single A53

Test Case	Detection (ms/frame)	Extraction (ms/frame)	Verification (ms/frame)	Total (ms/frame)
480P 32*32 single face	351	851	0	1205
480P 48*48 single face	179	821	0	1001
480P 64*64 single face	116	460	0	577
480P 32*32 two faces	373	1651	0	2028
480P 48*48 two faces	217	1649	0	1867
480P 64*64 two faces	202	1642	0	1848

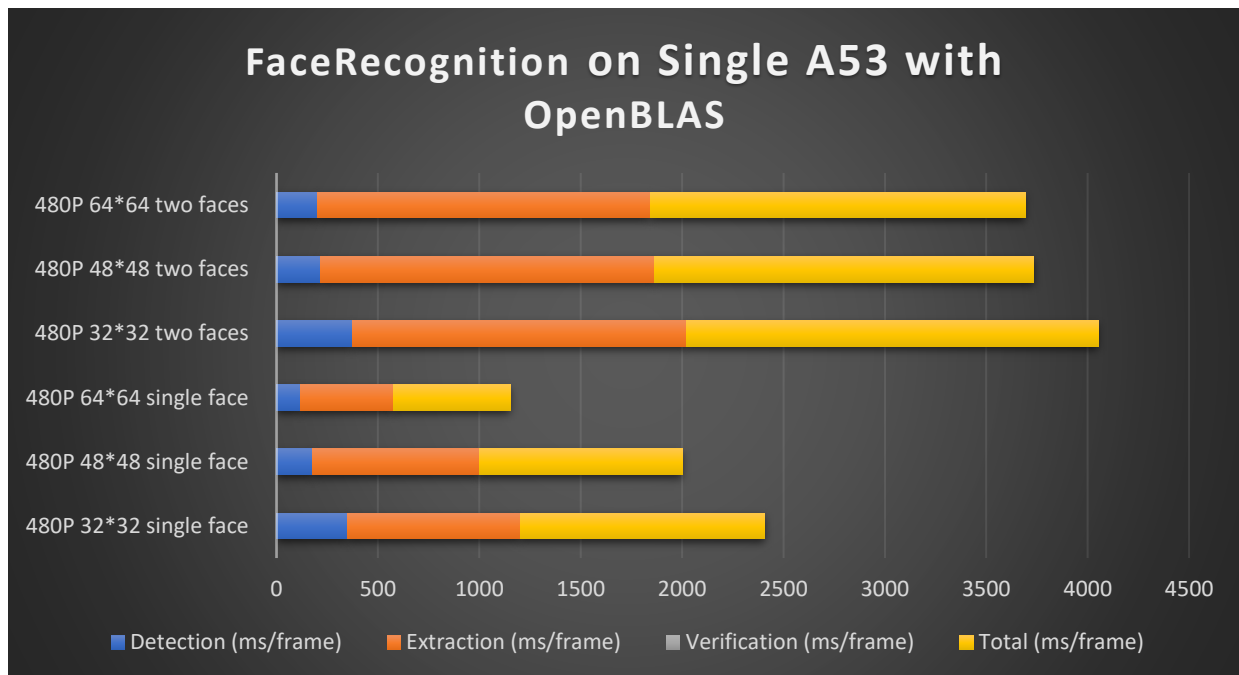


Fig 5.1 Performance Compare on Single A53

5.2 Single A72 CPU @1.8GHz

Table 5.2 Performance of different part on Single A72

Test Case	Detection (ms/frame)	Extraction (ms/frame)	Verification (ms/frame)	Total (ms/frame)
480P 32*32 single face	122	313	0	436
480P 48*48 single face	67	315	0	383
480P 64*64 single face	43	182	0	225
480P 32*32 two faces	159	714	0	875
480P 48*48 two faces	92	737	0	831
480P 64*64 two faces	84	726	0	812

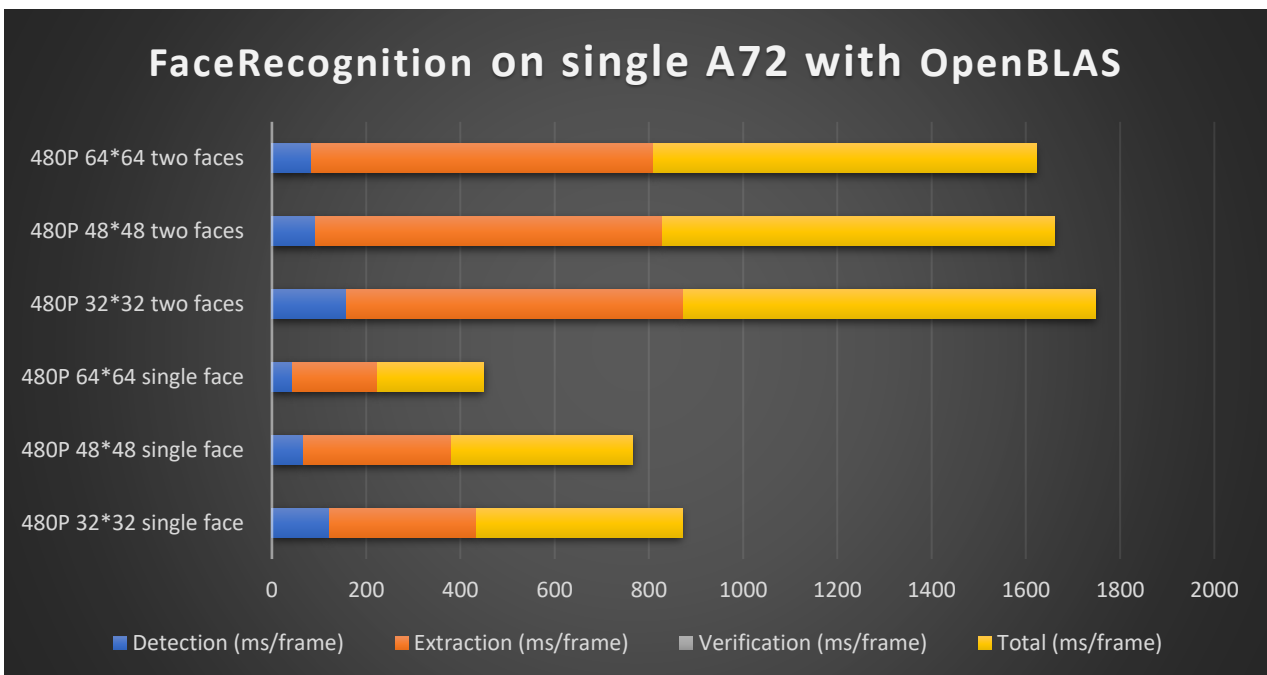


Fig 5.2 Performance Compare on Single A72

5.3 Two A72 CPU @1.8GHz

Table 5.3 Performance of different part on Two A72

Test Case	Detection (ms/frame)	Extraction (ms/frame)	Verification (ms/frame)	Total (ms/frame)
480P 32*32 single face	107	209	0	317
480P 48*48 single face	56	195	0	252
480P 64*64 single face	37	130	0	168
480P 32*32 two faces	151	538	0	691
480P 48*48 two faces	90	562	0	655
480P 64*64 two faces	82	554	0	638

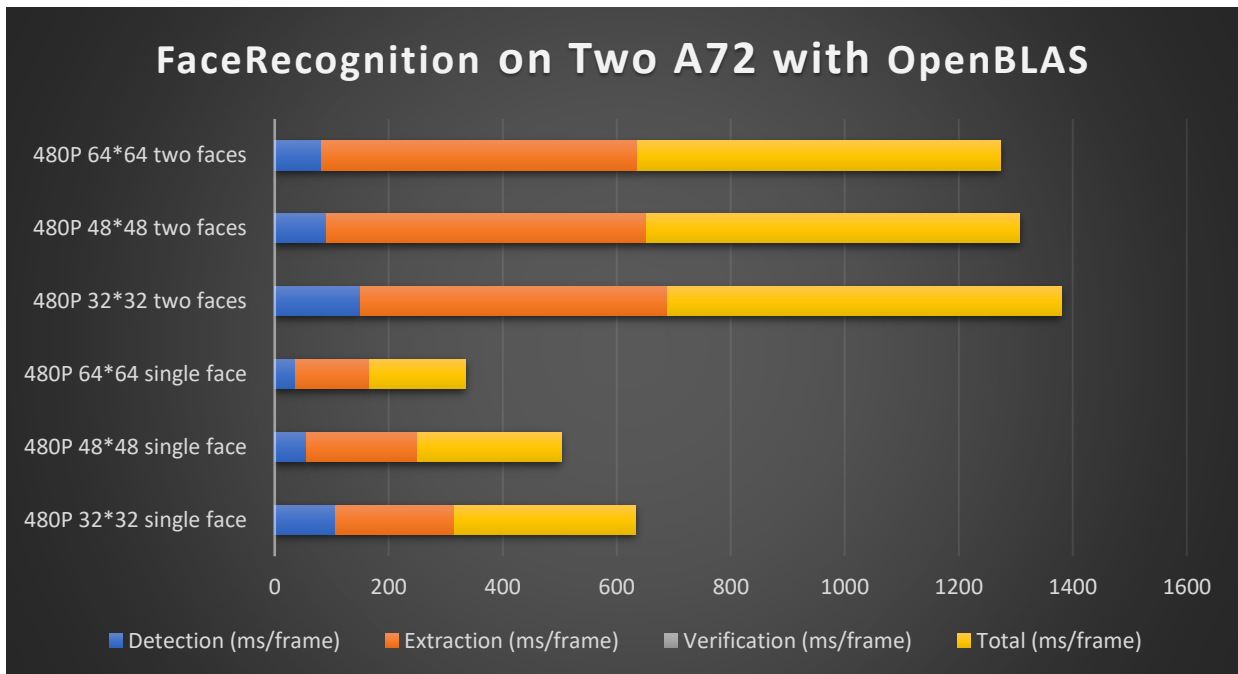


Fig 5.3 Performance Compare on two A72

5.4 Four A53 CPU @ 1.42GHz

Table 5.4 Performance of different part on Four A53

Test Case	Detection (ms/frame)	Extraction (ms/frame)	Verification (ms/frame)	Total (ms/frame)
480P 32*32 single face	218	335	0	554
480P 48*48 single face	121	329	0	452
480P 64*64 single face	85	214	0	300
480P 32*32 two faces	240	669	0	912
480P 48*48 two faces	139	673	0	816
480P 64*64 two faces	131	664	0	798

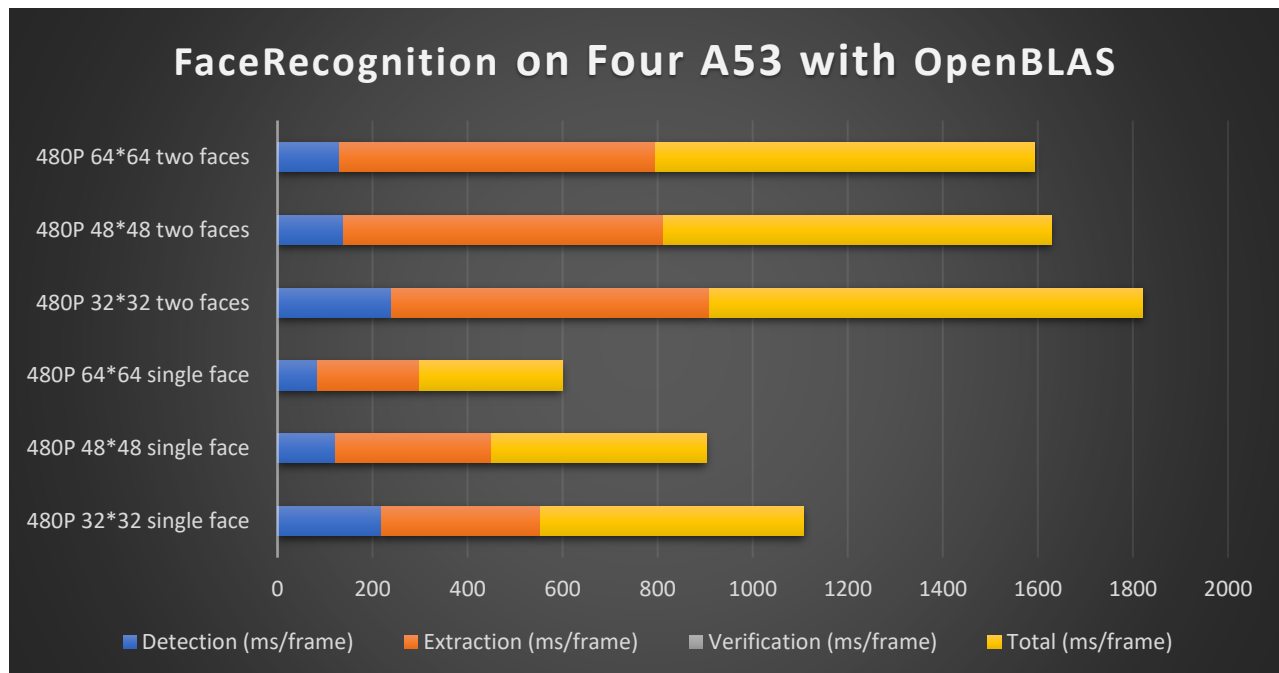


Fig 5.1 Performance Compare on Four A53

6 Accuracy Test

To test accuracy of the model, images with labelled ground true is needed. For detection test, we use a subset of CelebA dataset containing 10000 images. For recognition test, we use the full LFW dataset.

6.1 Detection test

To evaluate detection accuracy, following indices are used:

- Recall: Correctly detected faces / all faces in test dataset
- Precision: Correctly detected faces / all detected faces
- Stability: The average IOU of detected face and labelled face.
- IOU: Intersection over Union, intersection means the overlap area of detection and ground truth, union mean the union area of detection and ground truth

A correct face is defined when $\text{IOU} \geq 0.5$. The accuracy results of detection are given below:

Table 7.1 Accuracy results of detection.

Recall	Precision	Stability
0.80	0.77	0.65

6.2 Recognition test

To evaluate recognition accuracy, following indices are used:

- ROC curve: The ROC curve is created by plotting the true positive rate (TPR) against the false positive rate (FPR) at various threshold settings
- TPR/FPR: TPR(true-positive rate) is also known as sensitivity or probability of detection. FPR(false-positive rate) is also known as the fall-out or probability of false alarm. Calculation defined in the below.
- PR curve: P(Precision), R(Recall).

FaceRecognition Performance Report

		True condition	
		Condition positive	Condition negative
Predicted condition	Predicted condition positive	True positive, Power	False positive, Type I error
	Predicted condition negative	False negative, Type II error	True negative
		True positive rate (TPR), Recall, Sensitivity, probability of detection = $\frac{\sum \text{True positive}}{\sum \text{Condition positive}}$	False positive rate (FPR), Fall-out, probability of false alarm = $\frac{\sum \text{False positive}}{\sum \text{Condition negative}}$
		False negative rate (FNR), Miss rate = $\frac{\sum \text{False negative}}{\sum \text{Condition positive}}$	True negative rate (TNR), Specificity (SPC) = $\frac{\sum \text{True negative}}{\sum \text{Condition negative}}$

Fig 7.1 TP/FP/FN/TN

For security application, we used a relatively higher threshold, the average accuracy in the LFW dataset is: 0.73.

For different thresholds, the ROC curve and PR curve are given below:

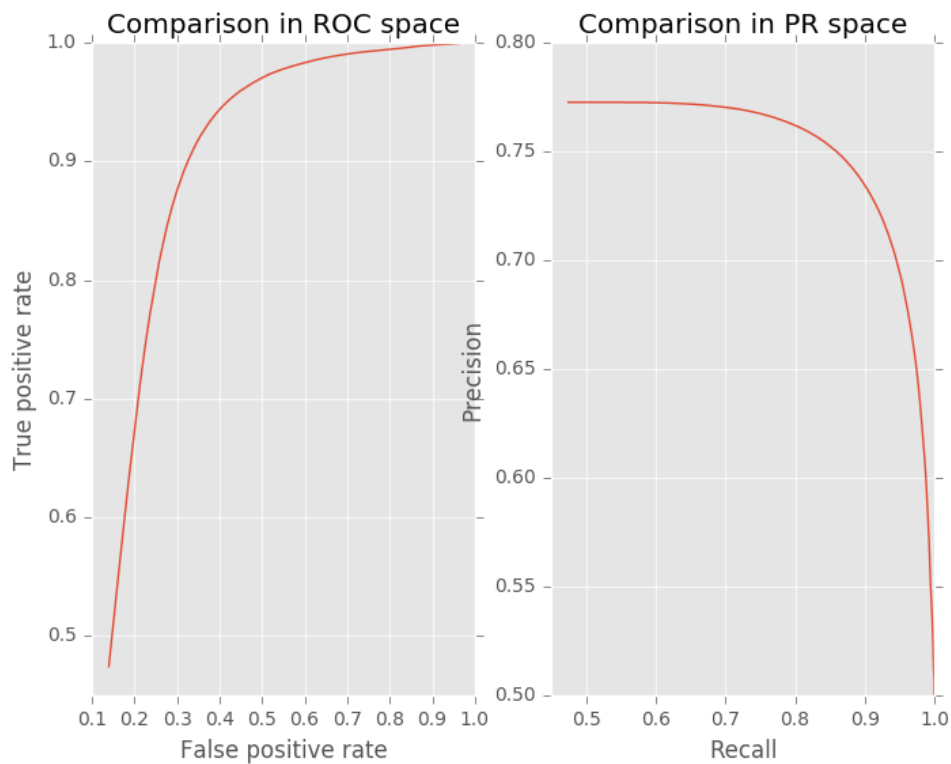


Fig 7.2 ROC curve and PR curve

7 Stress Test

The face detection passed over night (about 16 hours) test with one or two faces.

8 Conclusion

From the above test cases, we can deduce that:

- Detection time is influenced by minimum face size, in general, detection speed of min face size set to 64x64 is faster than 32x32 and 48x48, so if you don't need the small face detect you can increase face size to speed up.
- The feature extraction time increases with the number of faces.
- The performance on two A72 cores is better than on four A53 cores.