TMA AI Center

Cloud & Open Source Solutions for Computer Vision

Thang Tran, Tech Lead

TMA AI Center

Features Comparison

Features	Amazon	Google	Microsoft	IBM
Object Detection	✅	✅	✅	✅
Facial Analysis	✅	✅	✅	✅
Facial Recognition	✅	❌	✅	❌
Web Detection	❌	✅	❌	❌
Optical Character Recognition (OCR)	✅	✅	✅	❌

Object Detection

can detect and extract multiple objects in an image.

Amazon

Google

Microsoft

Facial Analysis

detect human faces and face attributes (age, emotion, gender...)

Amazon

looks like a face	99.9 %
appears to be female	99.9 %
age range	23 - 38 years old
not smiling	78.4 %
not wearing glasses	99.7 %
not wearing sunglasses	99.9 %
eyes are open	99.9 %
mouth is closed	94.8 %
does not have a mustache	99.6 %
does not have a beard	99.7 %

Google
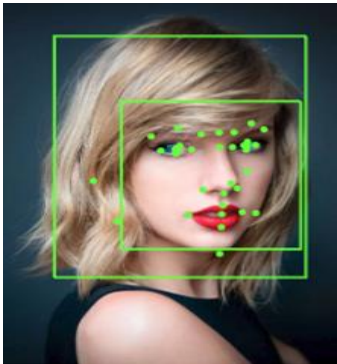
Joy	Very Unlikely
Sorrow	Very Unlikely
Anger	Very Unlikely
Surprise	Very Unlikely
Exposed	Very Unlikely
Blurred	Very Unlikely
Headwear	Very Unlikely
Confidence	99%

Microsoft
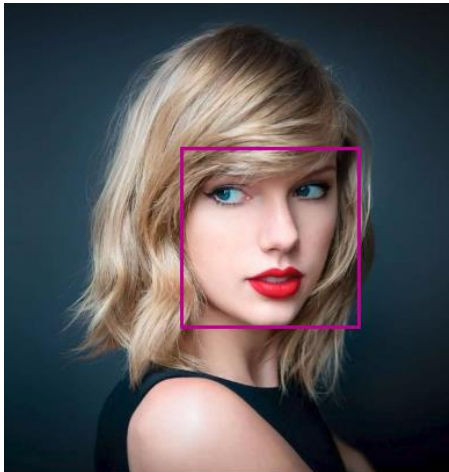
GENDER	female
AGE	22
GLASSES	NoGlasses
EMOTION	Neutral

IBM
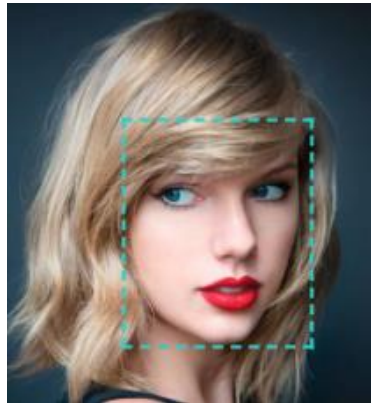
Face 1	
age 19-21	1.00
FEMALE	1.00

4

Facial Recognition

TMA Innovation Center

Compare faces to see how closely they match based similarity percentage

Amazon

Microsoft

Verification result: The two faces belong to the same person. **Confidence is 0.55786.**

Web Detection

detects web references to an image

Input Image

Google

Result

Optical Character Recognition (OCR)

Optical Character Recognition (OCR)

Handwritten Text
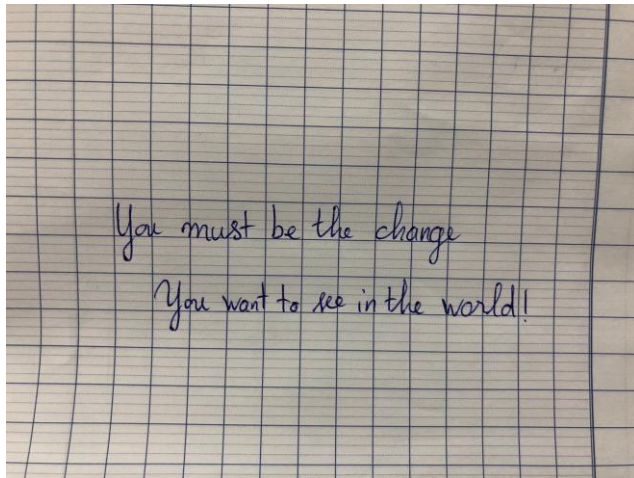
Google
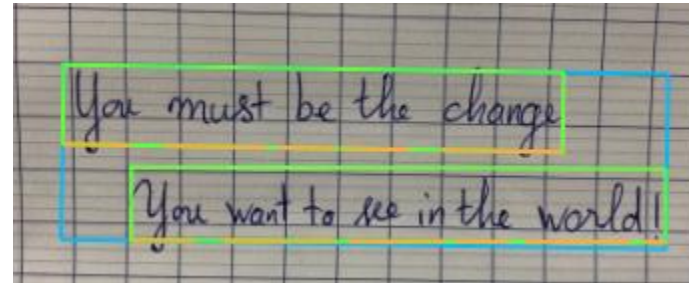
+Page 1

+Block 1

You must be the change

You want to see in the world !

Microsoft

you must be the chang

you want to see in the world

megaremains

Optical Character Recognition (OCR)

	Google	Amazon	Microsoft
Boxing style	**Group**	**Word**	**Sentence**
Language Recognition	**Y**	**N**	**Y**
Typewritten Text	**~81%**	**~97%**	**~97%**
Handwritten Text	**~91%**	**-**	**~83%**

Open Source Summary

Features	Available
Facial Recognition	✅
Facial Detection	✅
Object Detection	✅
Optical Character Recognize (OCR)	✅
Web Detection	❌

Facial Recognition

Open Source	face_recognition	OpenFace	SphereFace
Train Dataset	-	**- FaceSrub (106,863 image)** **- CASIA-WebFace (494,414 image)**	**CASIA-WebFace (494,414 image)**
Test Dataset	**Labeled Faces in the Wild (13,000 image)**	**Labeled Faces in the Wild (13,000 image)**	**Labeled Faces in the Wild (13,000 image)**
Accuracy	**99.38%**	**92.92%**	**99.30%**
Test on TIC Dataset (39 person, ~ 500 images)	**- Accuracy: 91,56%**	**- Accuracy: 92,24%**	**- Accuracy: 98,86%**

Facial Detection

Open Source	ocv-DNN	ocv-HAAR	dlib-HOG	dlib-CNN	face_recognition-HOG	face_recognition-CNN	MTCNN
Test on GTX1080	18ms	42ms	109ms	42ms	108ms	40ms	334ms

Object Detection
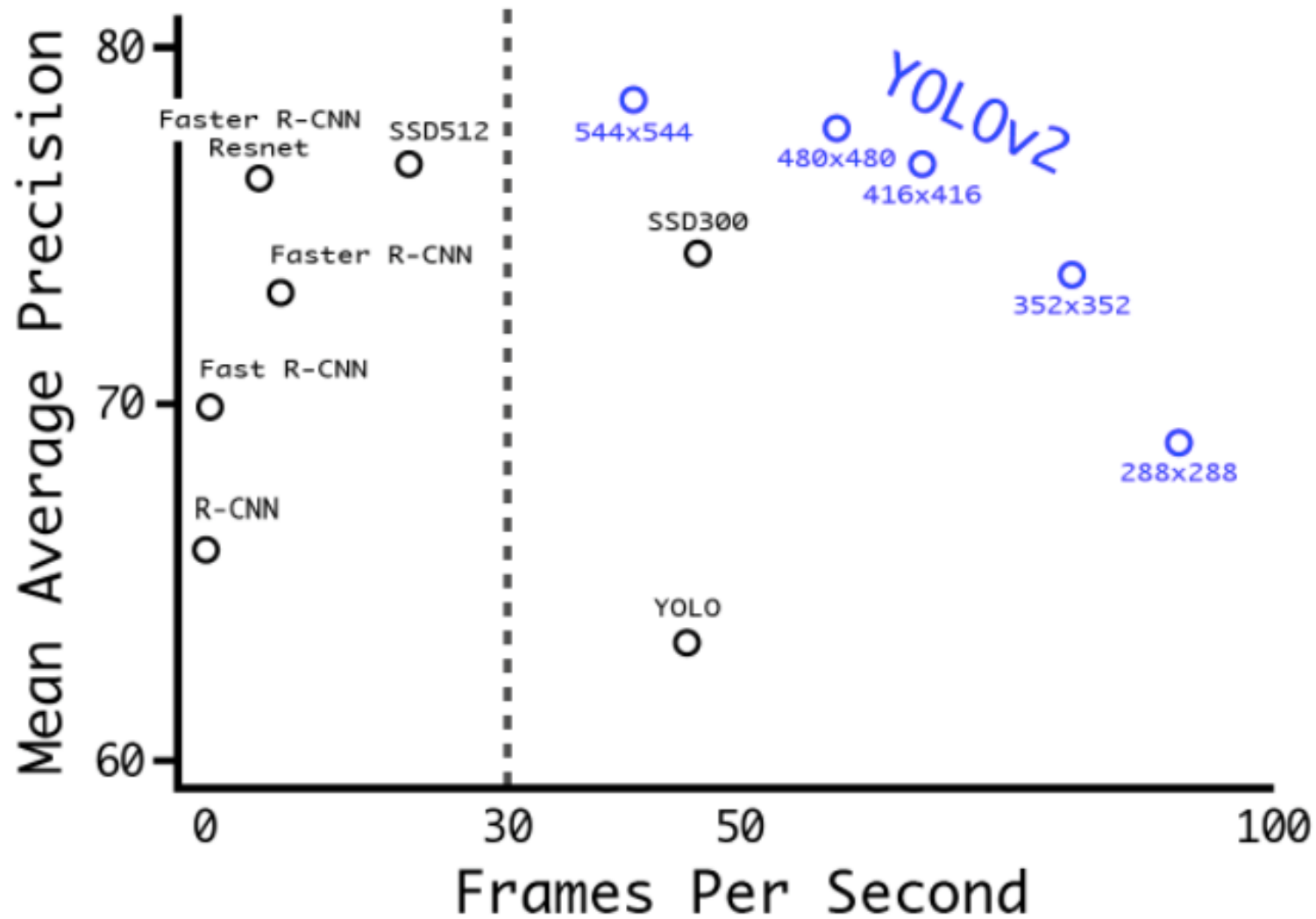
Object Detection
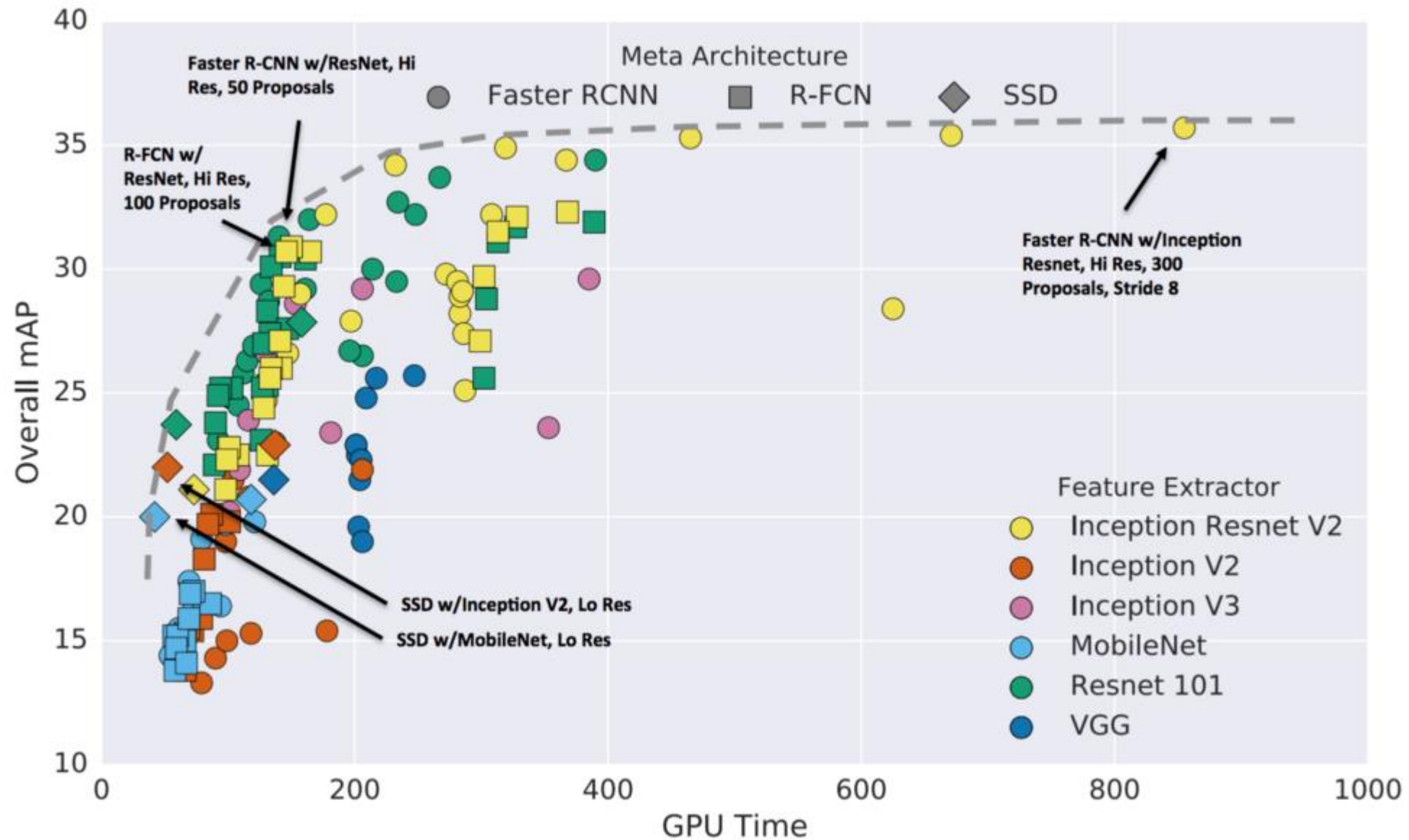
Optical Character Recognize (OCR)

Open Source	tesseract
Train Dataset	About 500,000 lines per language
Test Dataset	1000 lines per language
Accuracy	different language have different accuracy

Accuracy

Lang	Truth Words	Char Error Rates			Word Error Rates		
		Base	T-LSTM	LSTM+Dict	Base	T-LSTM	LSTM+Dict
Czech	19319	3.08	1.86	1.81	12.69	8.26	7.58
English	21543	2.51	1.95	1.76	7.58	7.18	5.77
French	20746	5.98	3.1	2.98	16.63	11.82	10.47
Hungarian	17977	3.75	2.95	2.86	14.61	12.26	11.45
Indonesian	16616	2.77	1.96	2.26	9.45	8.64	7.51
Dutch	18878	6.23	6.65	5.97	16.38	17.55	15.75
Norwegian	18129	2.19	1.74	1.81	7.73	5.87	5.7
Portuguese	17726	2.91	1.9	1.87	10.09	7.46	6.75
Spanish	20333	4.28	2.5	2.31	11.63	8.55	7.45
Vietnamese	22472	4.04	3.55	2.61	10.89	10.05	7.93

Open Source - Evaluation

Features	Available	Evaluation
Face Recognition	✅	- face_recognition - SphereFace
Face Detection	✅	- ocv-DNN
Object Detection	✅	
Optical Character Recognize (OCR)	✅	- Tesseract
Web Detection	❌	