



أكاديمية سدايا **SDAIA Academy**

Deep Learning Project
American Sign Language
Recognition Using Hand Tacking

Elyas Maghrabi

Outline



MOTIVATION



METHODOLOGY



RESULT AND
DEMO



FUTURE WORK

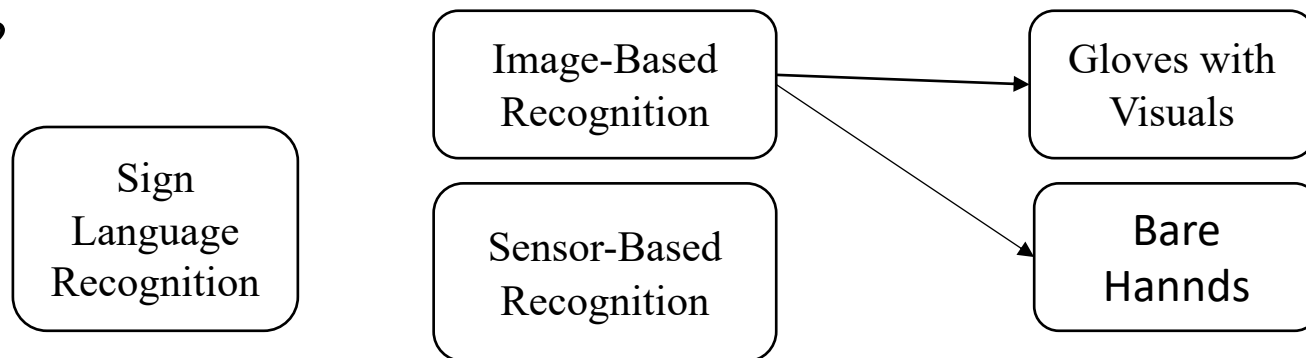


CONCLUSION

Motivation

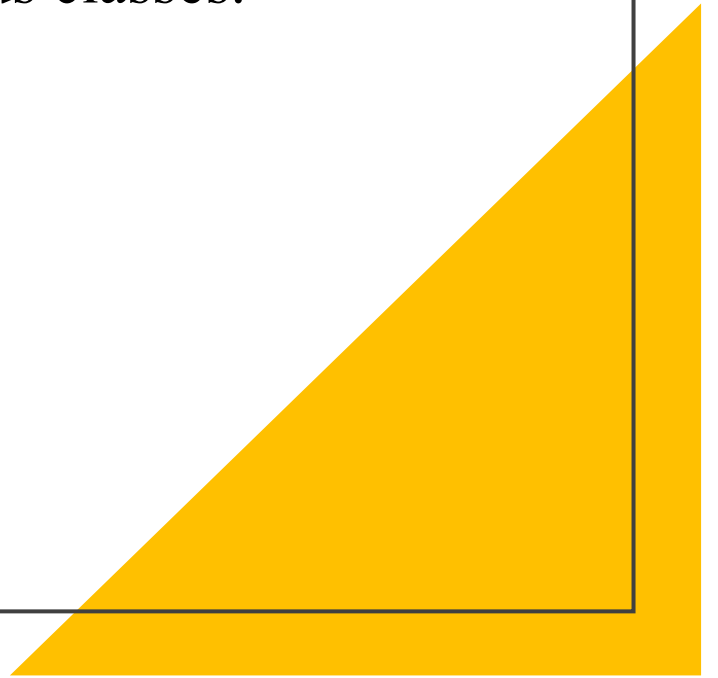
- Deaf and speaking impaired people cannot use speaking to communicate with others.
- Number of Deaf and hearing among the world, It is estimated that will be by 2050 over 700 million people according to world health organization.

Solution?



Dataset contains

The data set is a collection of images of alphabets from the American Sign Language, separated in 26 folders which represent the various classes.



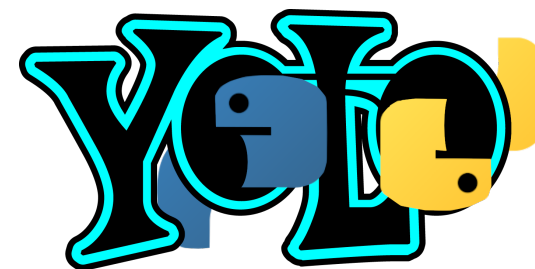
Libraries



pandas



pillow



NumPy



Keras



TensorFlow

The problem

Preprocessing

- How to deal with large Dataset
- Limited time(12 hours)
- Labeling

What features?

- Real time (ASL) detection based on gesture made by user.
- Customized gesture generation.

Prediction Model

- Try to build a model to predict ASL detection based. (image and real-time)

Modelling

Model: MobileNetV2

Two layers

activation= relu

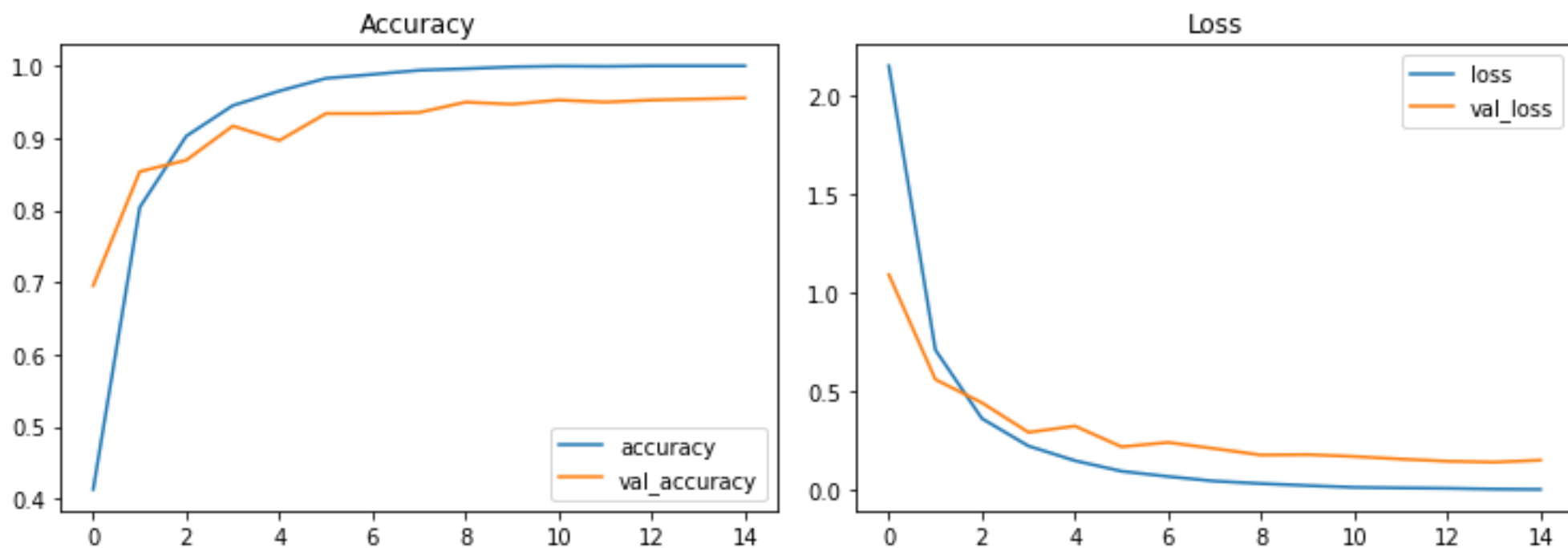
Loss=categorical_crossentropy

Output

activation=softmax

Optimizer: adam

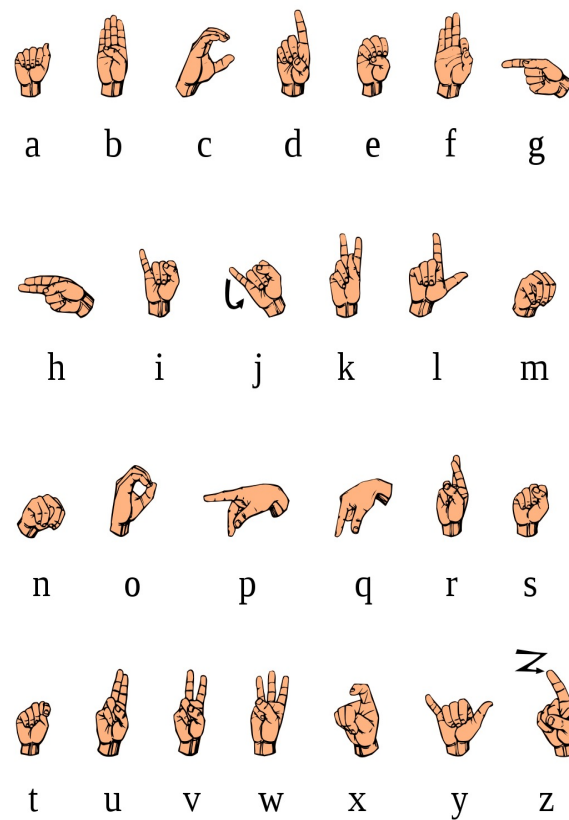
Modeling Result



Test Loss: 0.19315
Test Accuracy: 94.14%

	precision	recall	f1-score	support
A	0.96	0.96	0.96	26
B	0.94	0.97	0.95	32
C	0.90	0.96	0.93	27
D	0.86	1.00	0.93	25
E	0.93	0.93	0.93	29
F	0.96	0.93	0.95	28
G	0.83	1.00	0.91	29
H	1.00	0.86	0.93	36
I	1.00	0.88	0.94	33
J	0.96	0.86	0.91	29
K	0.97	0.97	0.97	33
L	1.00	1.00	1.00	39
M	0.85	0.92	0.88	37
N	0.93	0.84	0.88	31
O	0.97	1.00	0.99	34
P	1.00	1.00	1.00	22
Q	1.00	1.00	1.00	36
R	0.96	0.96	0.96	23
S	1.00	0.71	0.83	31
T	0.89	0.92	0.91	26
U	0.97	0.92	0.95	38
V	0.84	0.88	0.86	24
W	1.00	0.95	0.97	19
X	0.87	0.96	0.91	27
Y	0.88	0.96	0.92	23
Z	0.93	1.00	0.96	26
del	1.00	1.00	1.00	35
nothing	0.93	1.00	0.96	37
space	1.00	0.97	0.99	35
accuracy			0.94	870
macro avg	0.94	0.94	0.94	870
weighted avg	0.95	0.94	0.94	870

Result



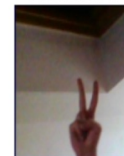
True: M
Predicted: M



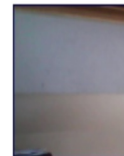
True: T
Predicted: T



True: V
Predicted: V



True: nothing
Predicted: nothing



True: D
Predicted: D



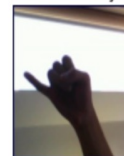
True: M
Predicted: M



True: L
Predicted: L



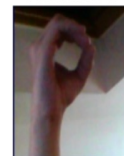
True: J
Predicted: J



True: L
Predicted: L



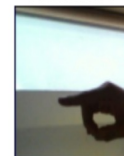
True: O
Predicted: O



True: Z
Predicted: Z



True: P
Predicted: P





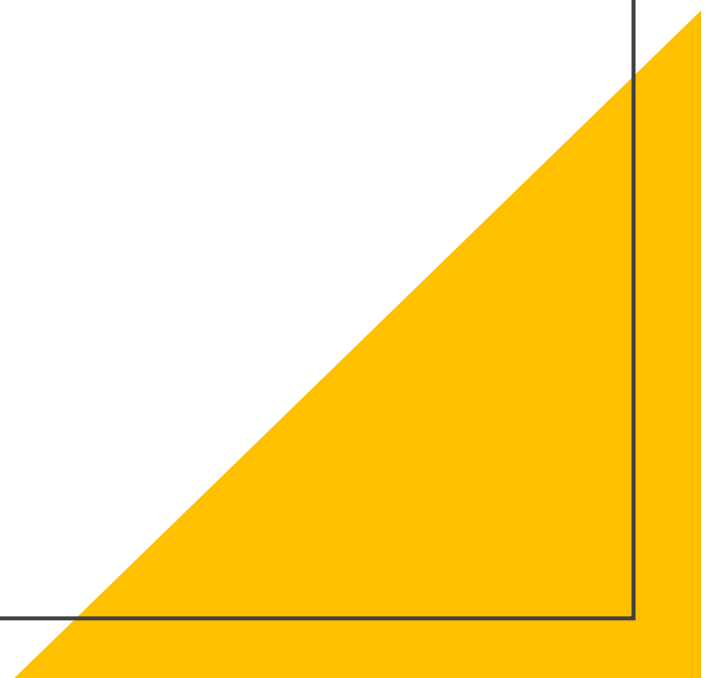
Demo



Future Work

- I planning to achieve higher accuracy real time even in case of complex backgrounds by trying out various background subtraction algorithms.
- Word and sentence formation.
- Better hand tacking.

Conclusion





Thank You

