Machine Learning - Deep Learning

Prof. Lev Muchnik

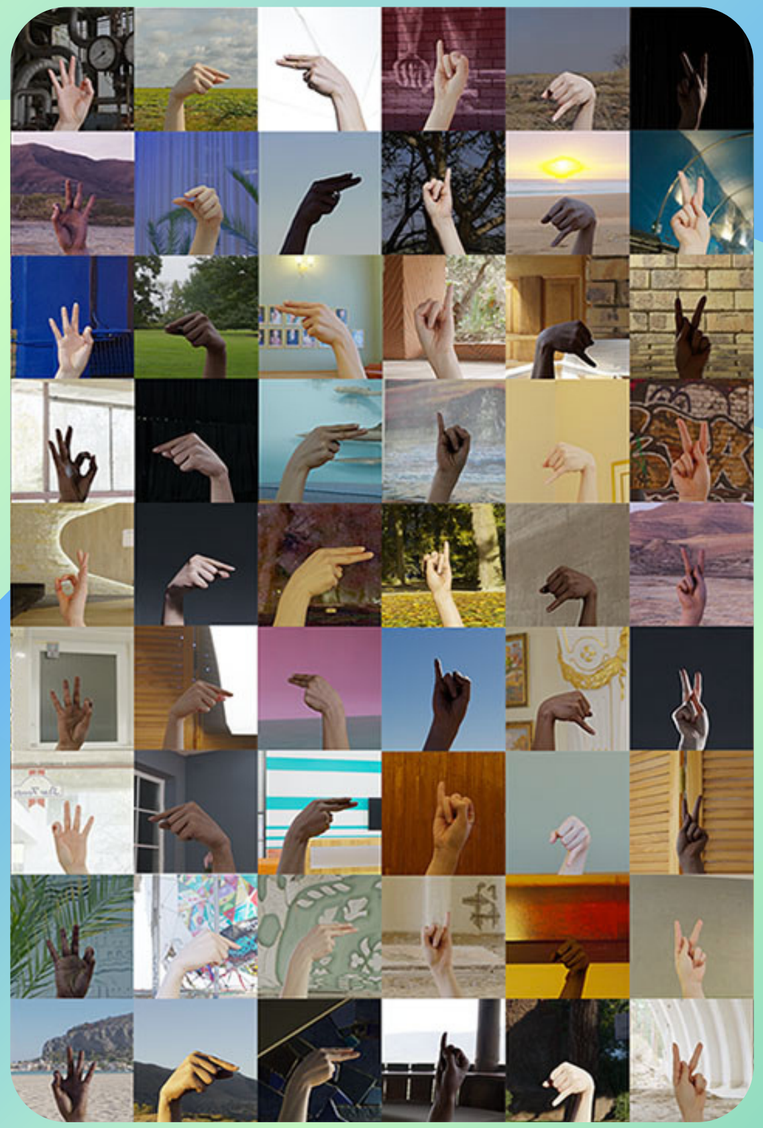
Final Project

Sign Language Symbol Detection

Itai Alush | Gal Eichenbaum | Ben Tobin

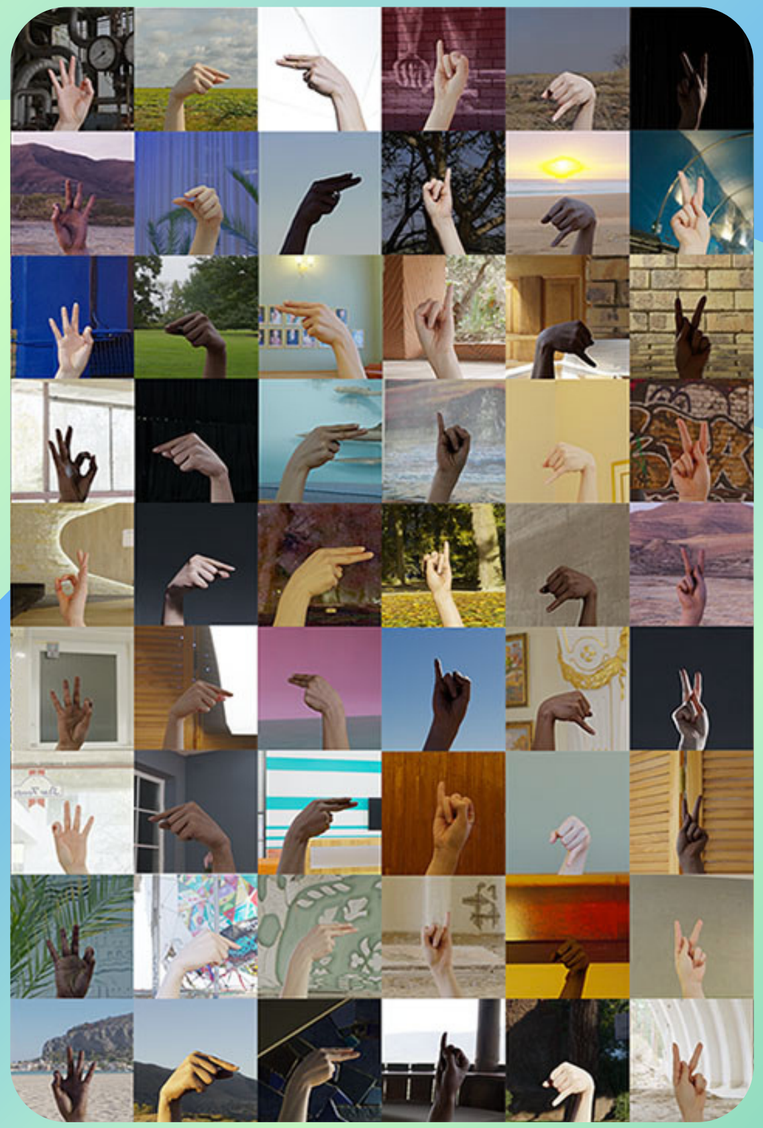
Agenda →

1) Problem we're facing
2) Some General Statistics
3) Our Data
4) Modelling
5) Results
6) Conclusions



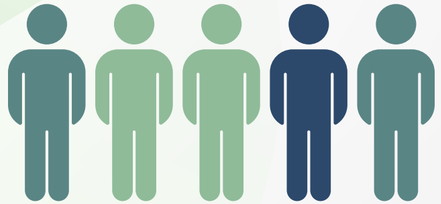
SIGN LANGUAGE IS NOT A POPULAR LANGUAGE

THE MOTIVATION BEHIND THIS PROJECT IS TO HELP REDUCE THE COMMUNICATION GAP WITH THE HEARING IMPARED



CAN DEEP LEARNING METHODS HELP IDETIFY SIGN LANGUAGE?

GENERAL STATISTICS →



Disabling hearing loss	466M
Deaf people worldwide	72M
Access to education	2%
Lack Communication	77%
Difficulty communication with drivers	51%

The Dataset Source →

- Dataset represent different angles, background and other sentiments of sign language letters.
- Created in 2015.
- Containing 24,300 Training images and 2,700 Test images, divided into 26 categories (for each letter) and 1 blank category.

CREATING OUR DATA →

01

LOWERING THE TRAINING SET TO 4850 IMAGES

02

CREATING VALIDATION SET OUT OF OUR TRAINING IMAGES

03

RESIZING IMAGES FROM 513X512 TO 48X48.

04

NORMALAIZING THE DATE

Our Data Set

Training Images per category

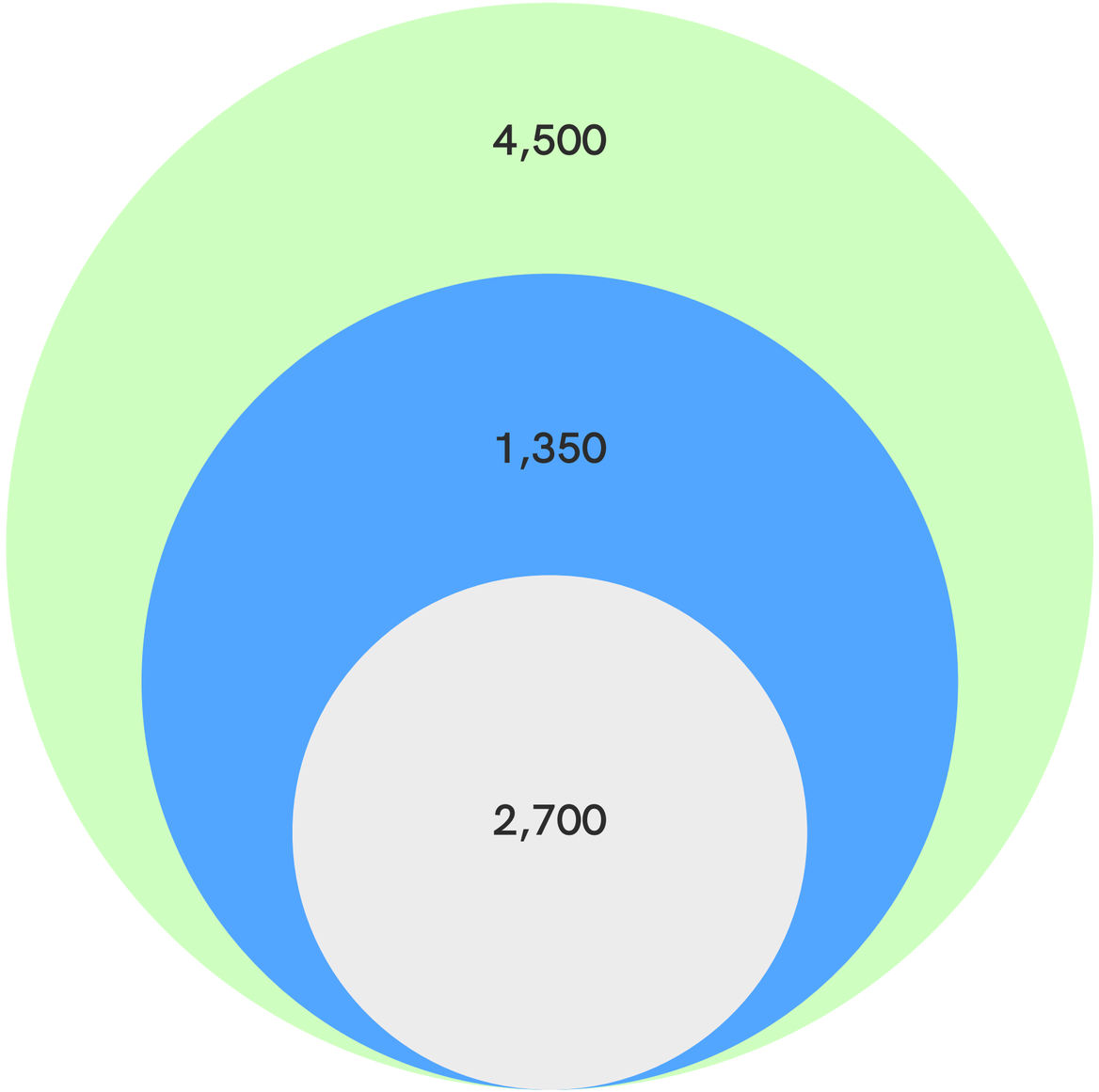
150

Validation Images per category

50

Test Images per caetgopry

100



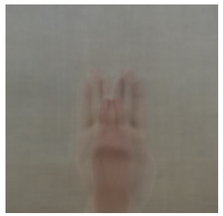
EDA → IMAGE SHARPNESS

Category	Average Sharpness	STD of Sharpness
A	0.001609	0.001165
B	0.001987	0.001258
C	0.002089	0.001386
D	0.002306	0.001528
E	0.00159	0.000886
F	0.002452	0.002004
G	0.002424	0.001606
H	0.002694	0.002343
I	0.00171	0.001462
J	0.001715	0.000954
K	0.001903	0.001126
L	0.002006	0.001324
M	0.001936	0.001247

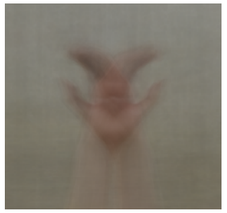
Category	Average Sharpness	STD of Sharpness
N	0.001603	0.001111
O	0.002145	0.001498
P	0.001911	0.001218
Q	0.001823	0.001703
R	0.002112	0.00089
S	0.001651	0.001095
T	0.001663	0.000987
U	0.001762	0.001039
v	0.001739	0.001081
W	0.002122	0.001081
X	0.002077	0.001419
Y	0.001908	0.001329
Z	0.001807	0.001146
Blank	0.001332	0.00121

EDA → IMAGE AVERAGES

A

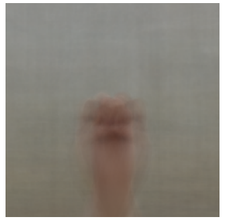


B



C

D



E

F

G

H

I

J

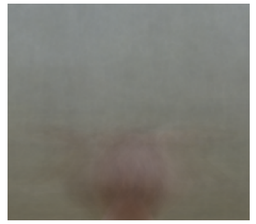
K

L

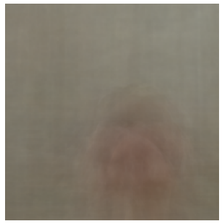
M

N

O

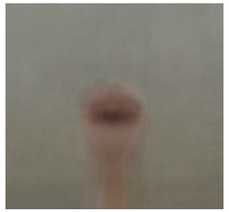


P

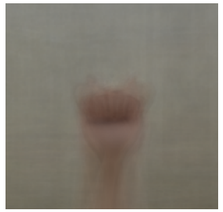


Q

R



S



T

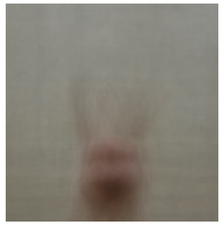
U

V

W

X

Y



Z

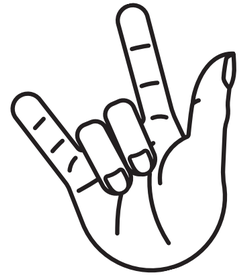
Blank

Project Limitations →

Given various computing and data limitations we think there is much room to increase accuracy in the future using this model.

- Small training set size.
- Reduced image quality.
- Various backgrounds that may cause noise.
- Different symbols with high similarities.

MODEL ARCHITECHTUERS



	CNN	LINEAR LAYERS	BATCH NORM	DROP OUT
MODEL 1				
MODEL 2				
MODEL 3				
MODEL 4				
MODEL 5				
MODEL 6				

MODEL LAYERS

INPUT LAYER

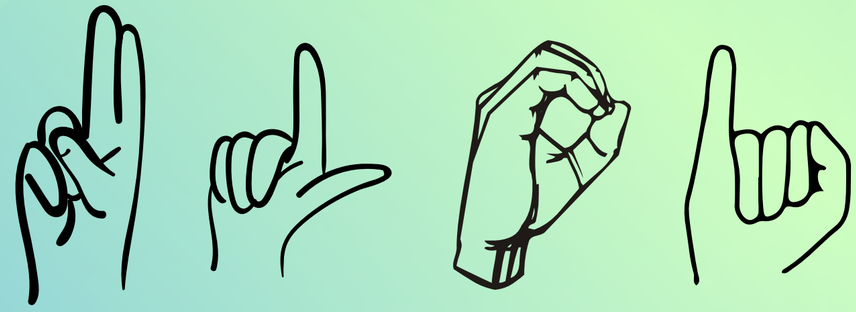
48*48*3
RESIZED PICTURES IN RGB

HIDDEN LAYERS

VARIOUS AMOUNTS OF LAYERS AND SIZES

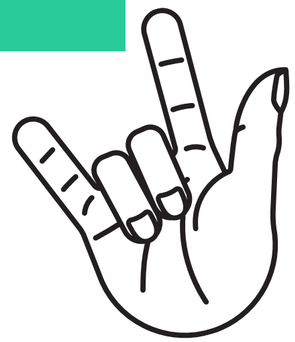
OUTPUT LAYER

27
AMOUNT OF DIFFERENT SYMBOLS IN THE DATA

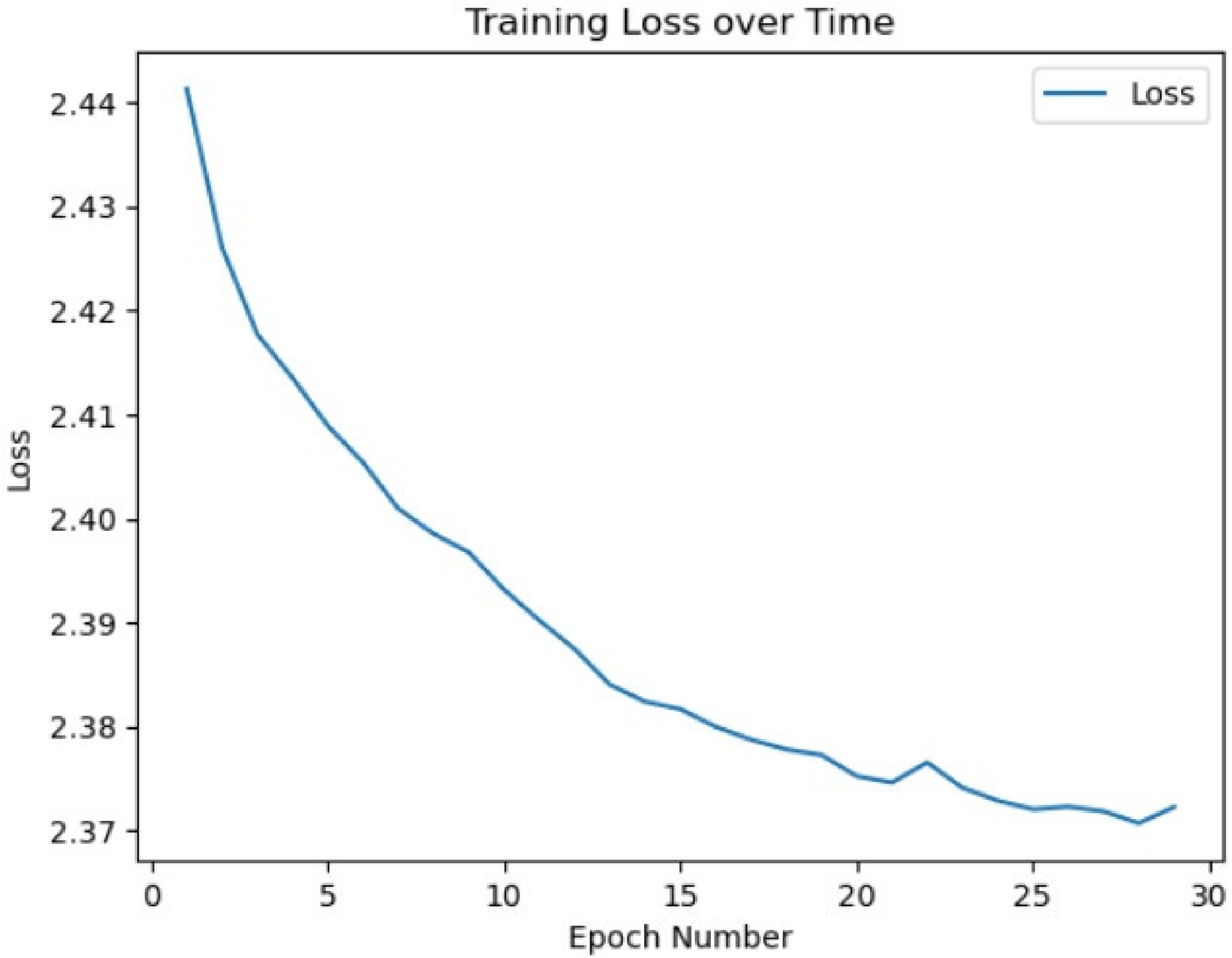


BEST MODEL RESULTS

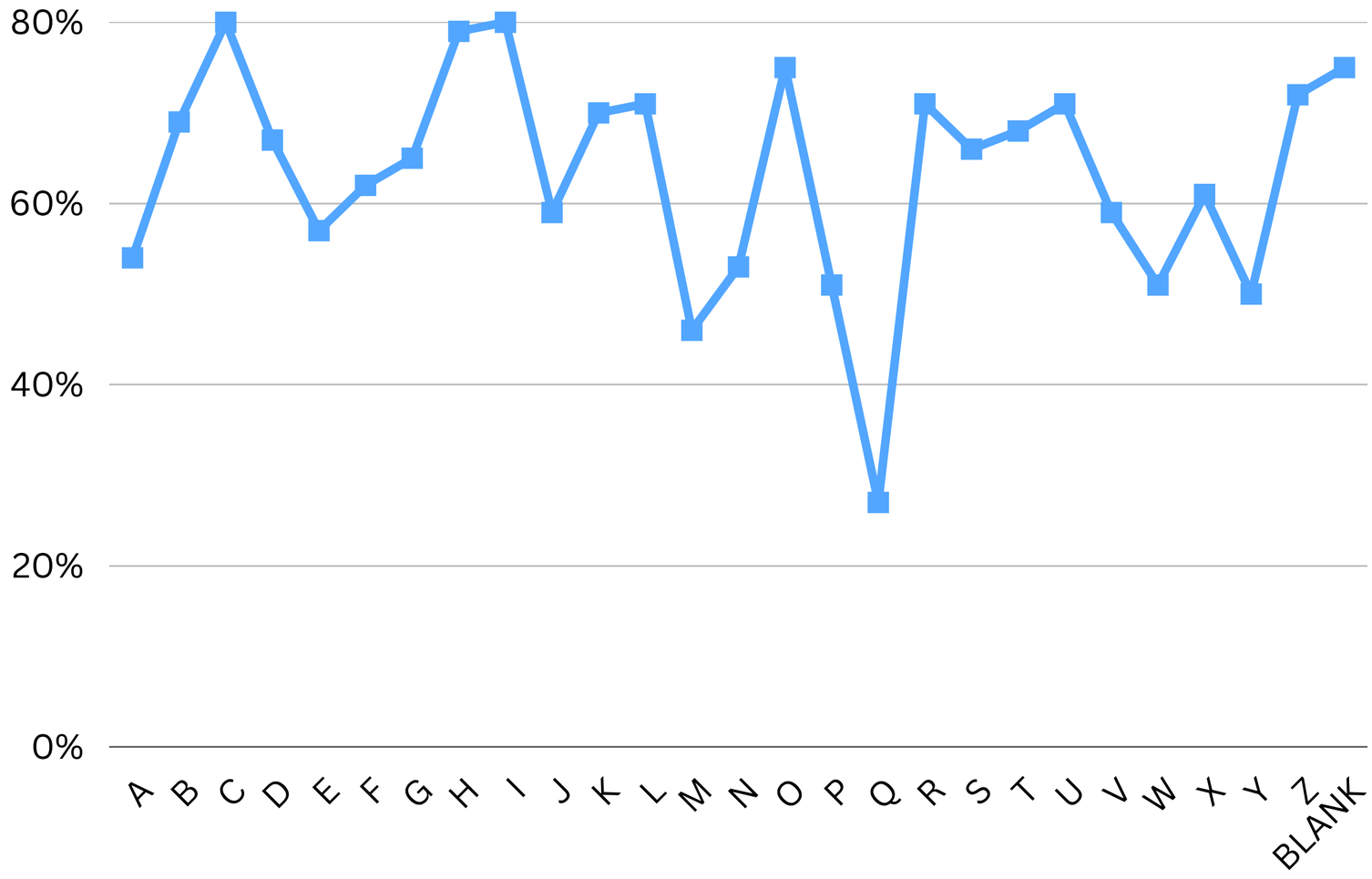
MODEL ARCHITECTURE	INPUT LAYER	HIDDEN LAYERS	OUTPUT LAYER	VALIDATION ACCURACY	TEST ACCURACY
LINEAR MODEL + BATCH NORMALIZATION	48 * 48 * 3	[3469,1728]	27	62%	63%



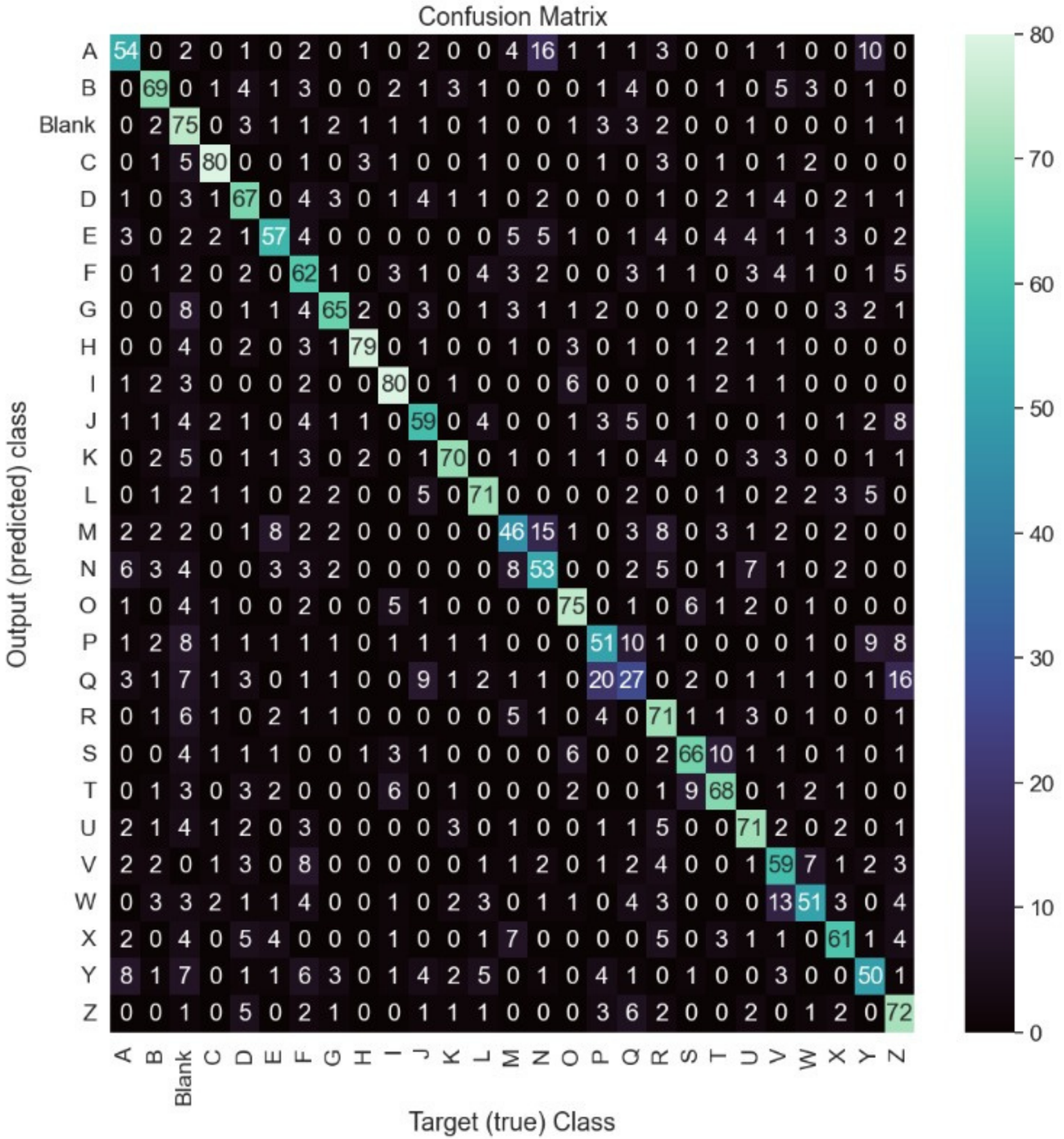
Visualizing Loss



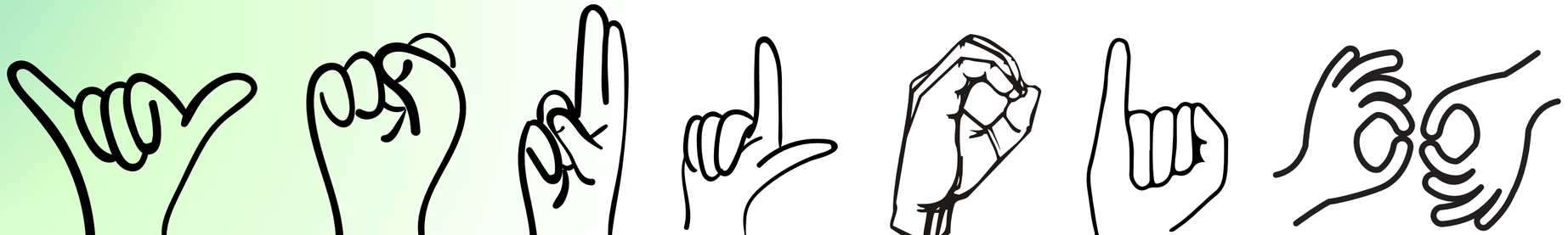
Accuracy Per Symbol



Symbol Confusion Matrix



PROJECT CONCLUSIONS	01 63% Accuracy > 3.7% Baseline
02 10 symbols with over 70% Accuracy	03 Only 2 Symbols with under 50% Accuracy



A LITTLE TEST WE RAN

OUTPUT

Your Word is: DATA

Whats Next?

60% is not an exceptional accuracy but we think this is a great starting point for a field that has barely been explored.

- Increase computing power to evolve the model.
- Figure out how to separate and optimize specific symbols.
- Help Implement model into educational tools.
- Evolve the model to take video as an input.

THANK YOU :)