



I swear I'm not a robot!

Type the characters above:

I might be a robot

Go

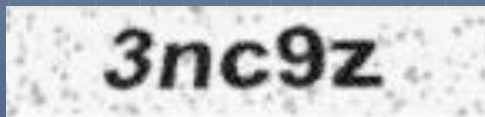
Breaking CAPTCHA

By,
Ray Zhao

Background

What is CAPTCHA?

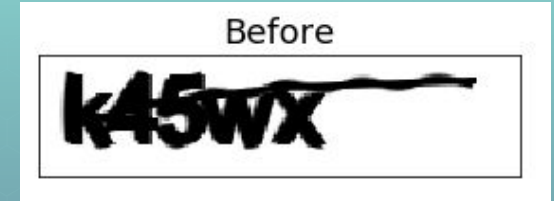
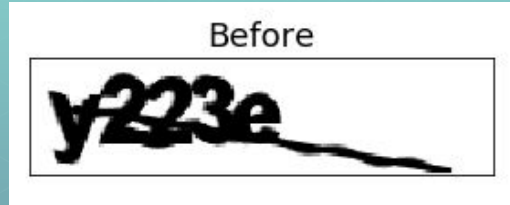
- CAPTCHA is an acronym for Completely Automated Public Turing test to tell Computers and Humans Apart.



- Can we teach a machine to read and predict the text-based CAPTCHA?



Captcha Data



- This version of CAPTCHA consists on 5 characters of either numbers or lowercase alphabet letters.
- Placement of characters are the same throughout the data
- Rotate and shift CAPTCHAs
- Prevents model from memorizing character location



Hand Drawn CAPTCHA

- The EMNIST dataset is an extension of the MNIST digits dataset that has alphabetical letters.

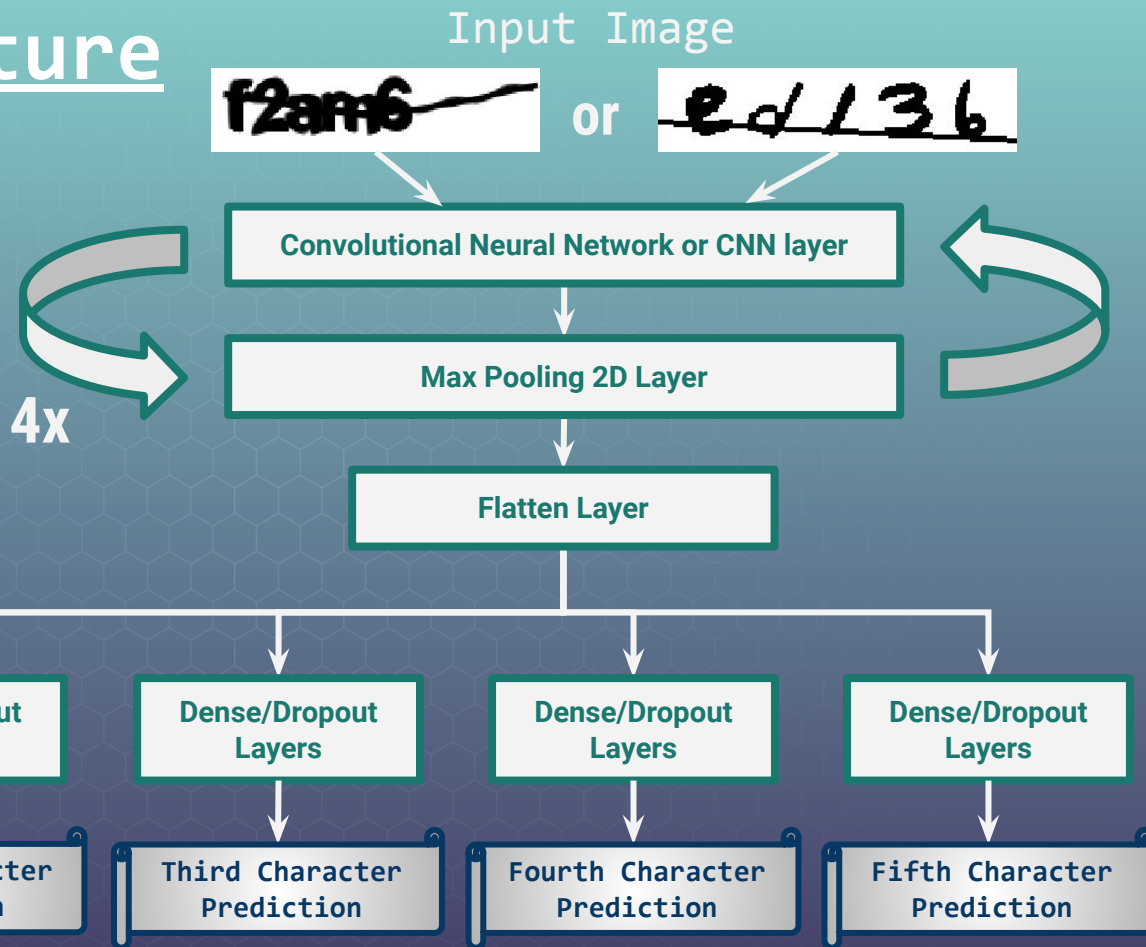


- Hand drawn CAPTCHAs were created by concatenating random individual handwritten numbers and letters.
- Then a line was added at a random angle

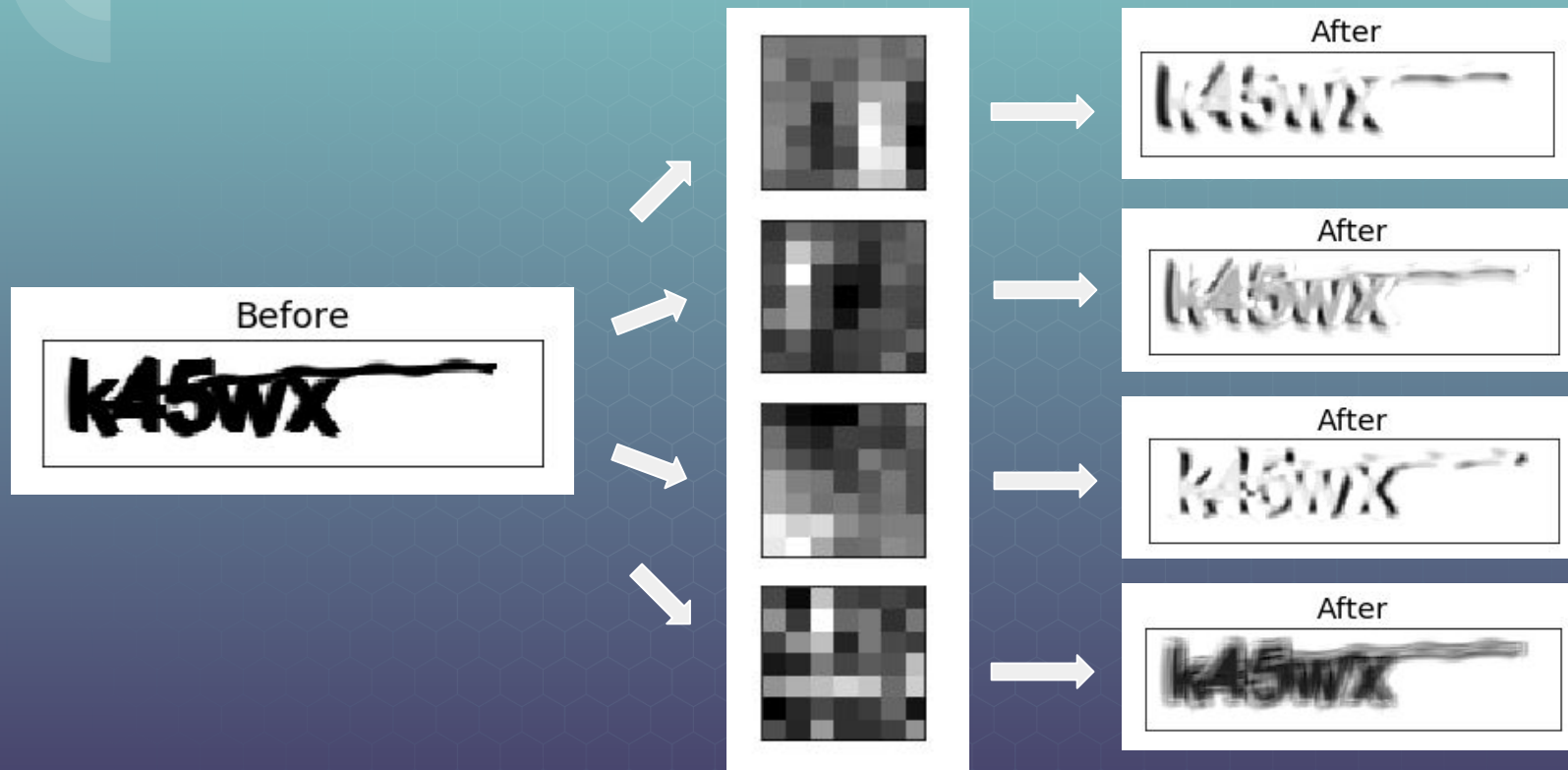


Model Structure

The CNN and Max Pooling layers are repeated 4 times. Each CNN has 32, 64, 32, 32 filters.



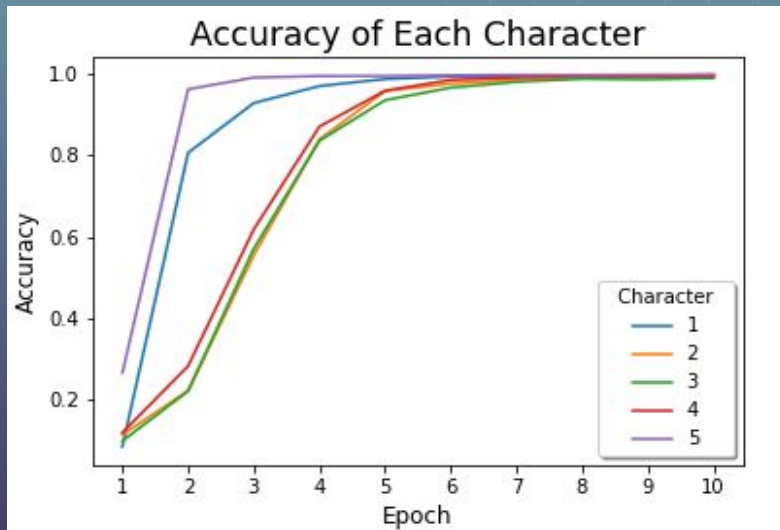
What's happening in the CNN layers?



Model Results

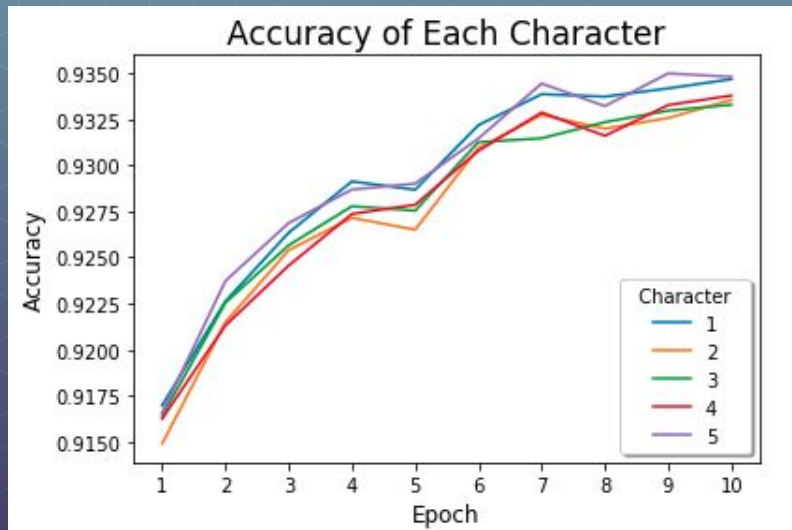
CAPTCHA Data

99.8%, 99.5%, 98.9%, 99.6%, 99.8%



'Handwritten' CAPTCHA Data

93.5%, 93.4%, 93.3%, 93.4%, 93.5%



Dash Web App

- Hand draw a captcha for the model to predict
- Press Save to see what you've drawn so far.
- Once you're satisfied, push BREAK to have the model predict.

Breaking CAPTCHA

Draw a CAPTCHA and press SAVE to see work

← → SAVE

Run CAPTCHA Breaking Model

BREAK

Breaking CAPTCHA

Draw a CAPTCHA and press SAVE to see work

6 k 3 4 f

← → SAVE

6 k 3 4 f

Run CAPTCHA Breaking Model

BREAK

Predicted Text: 6k34f

Breaking CAPTCHA

Draw a CAPTCHA and press SAVE to see work

~~6 k 3 4 f~~

← → SAVE

~~6 k 3 4 f~~

Run CAPTCHA Breaking Model

BREAK

Predicted Text: 6k34f

Conclusion

- Building models to predict CAPTCHA is dependent on the data it is trained on.
- This leads to websites creating new versions of CAPTCHAs to combat people training models to predict CAPTCHAs.
- [Link to Web App](#)



Tech Stack



Dash

by plotly



The friendly PIL fork



TensorFlow



OpenCV

