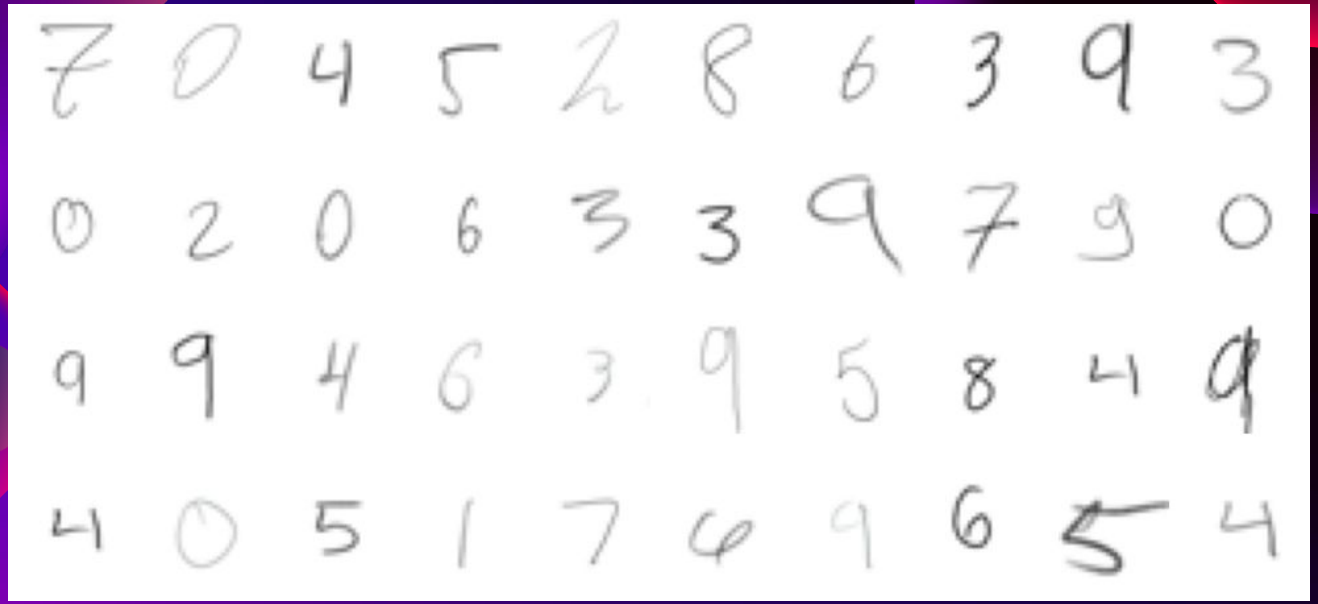
MSU-NIST



Ling Thang and Joaquin Trujillo

Data Collection

- Collected handwritten digits using a Ipad
- Notability for easy file management
- Asked students if they wanted to help me collect data
- CS Lounge, Classes, and library

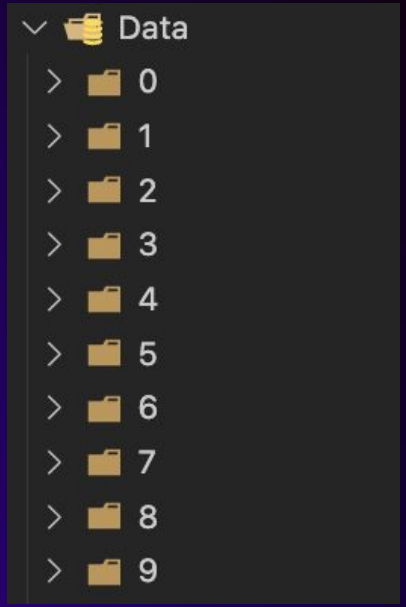

100 participants 1000 total digits

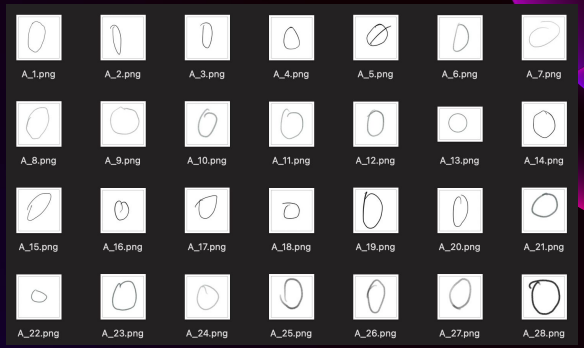
Data Processing

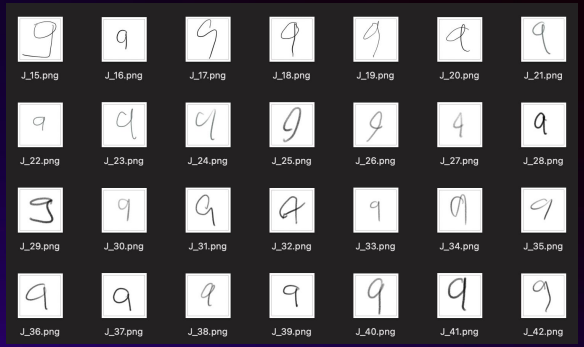
Cropped these images to 1:1 scale

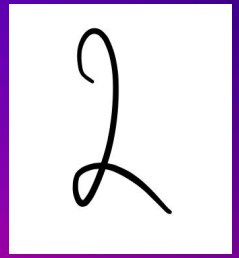
Stored each image in a folder

Converting to npy files to use in other workspaces

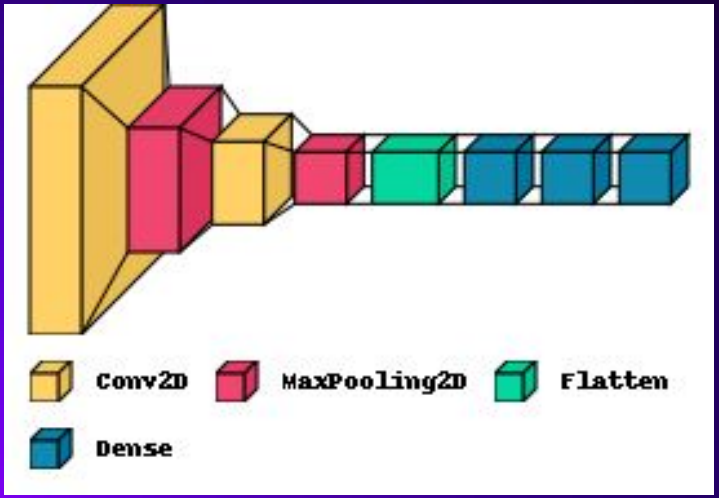








CNN v1



With keras MNIST data set
This model had a 90% accuracy

First architecture just to see where our data stacked up.

```
******* evaluating network *******
20/20 [==============================] - 0s 3ms/step
              precision    recall  f1-score   support
...
    accuracy                           0.74       200
   macro avg       0.77      0.77      0.75       200
weighted avg       0.79      0.74      0.74       200
```

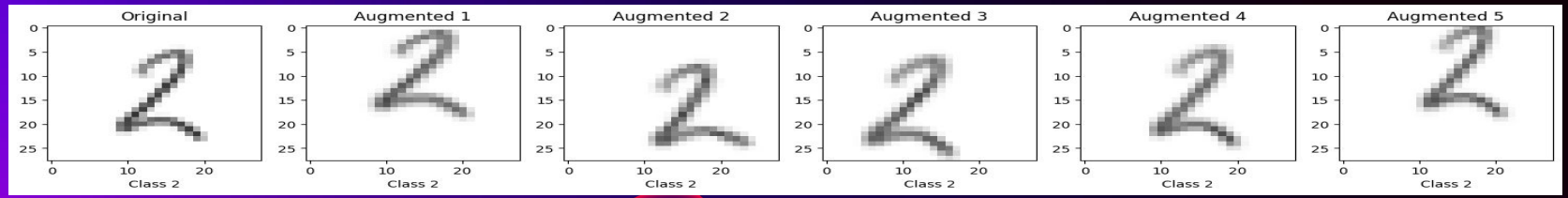
Data Augmentation

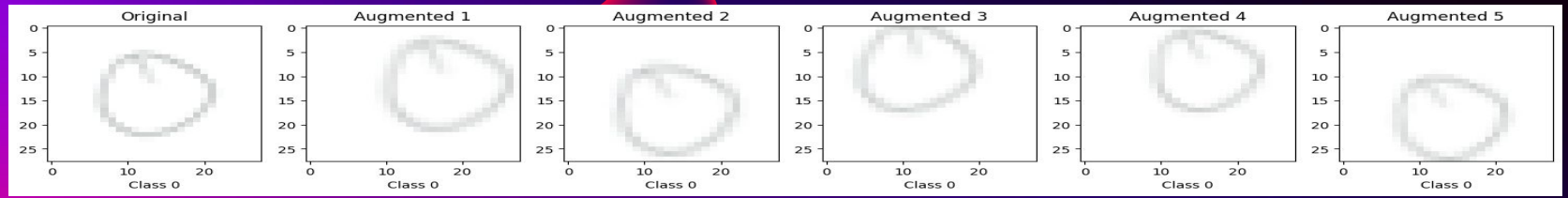
Parameters for augmentation

```
# data augmentation
datagen = ImageDataGenerator(
        rotation_range=10,
          zoom_range = 0.10,
          width_shift_range=0.1,
          height_shift_range=0.1)
```

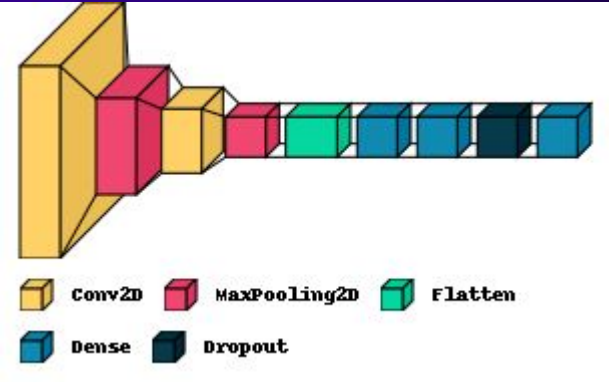
New Size of our dataset

```
Shape of augmented dataset: (4000, 28, 28, 4)
Shape of augmented labels: (4000, 10)
0: 400
1: 435
2: 365
3: 395
4: 425
5: 390
6: 375
7: 435
8: 385
9: 395
```





CNN V2

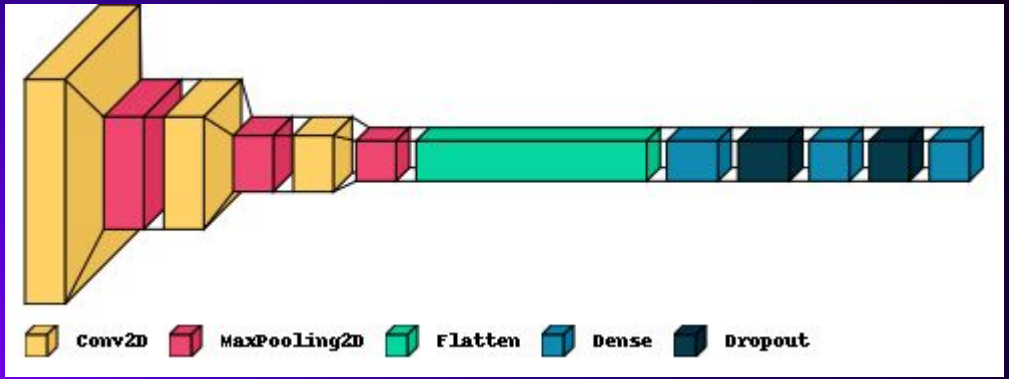



```
******* evaluating network *******
4/4 [==============================] - 0s 8ms/step
              precision    recall  f1-score   support

           0       0.69      0.55      0.61        20
           1       1.00      1.00      1.00        13
           2       0.96      0.93      0.94        27
           3       0.80      0.95      0.87        21

...
    accuracy                           0.81       200
   macro avg       0.81      0.81      0.81       200
weighted avg       0.81      0.81      0.81       200
```

CNN v3




```
******* evaluating network *******
4/4 ━━━━━━━━━━━━━━━━━━━━ 0s 5ms/step
              precision    recall  f1-score   support

           0       0.91      1.00      0.95        20
           1       0.93      1.00      0.96        13
           2       0.96      0.89      0.92        27
           3       0.91      0.95      0.93        21
...
    accuracy                           0.95       200
   macro avg       0.95      0.96      0.95       200
weighted avg       0.95      0.95      0.95       200
```

Key Insights and Reflections

Data Collection

The process of collecting

Data Ingestion

Processing the images so that they were workable

Data Cleaning

Making sure that our augmented data was clean to use for our model

Hardware

Slight Variation In Accuracy Depending On What Machine We Were On

Data Augmentation

Augmenting the data to increase the size of a training set

Channels

Probably the most annoying thing we ran into