

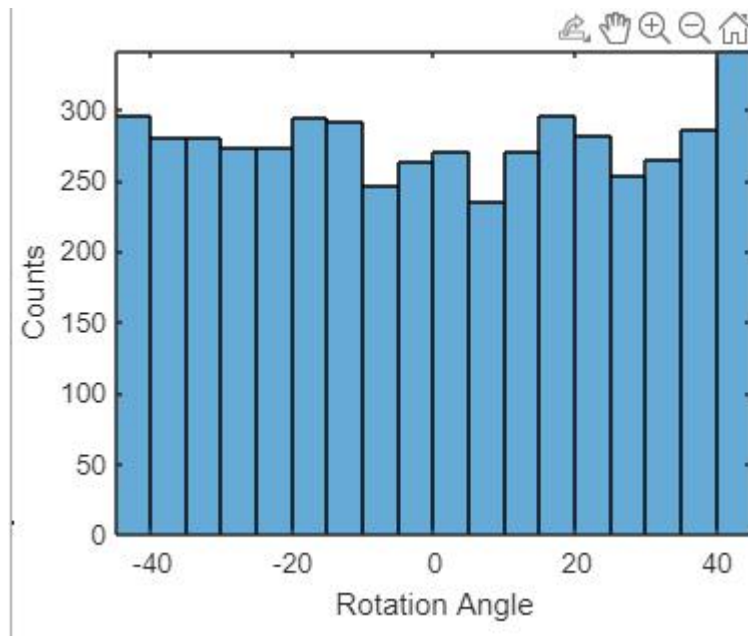
## JichenDai\_HW2

### Build a Convolutional Neural Network

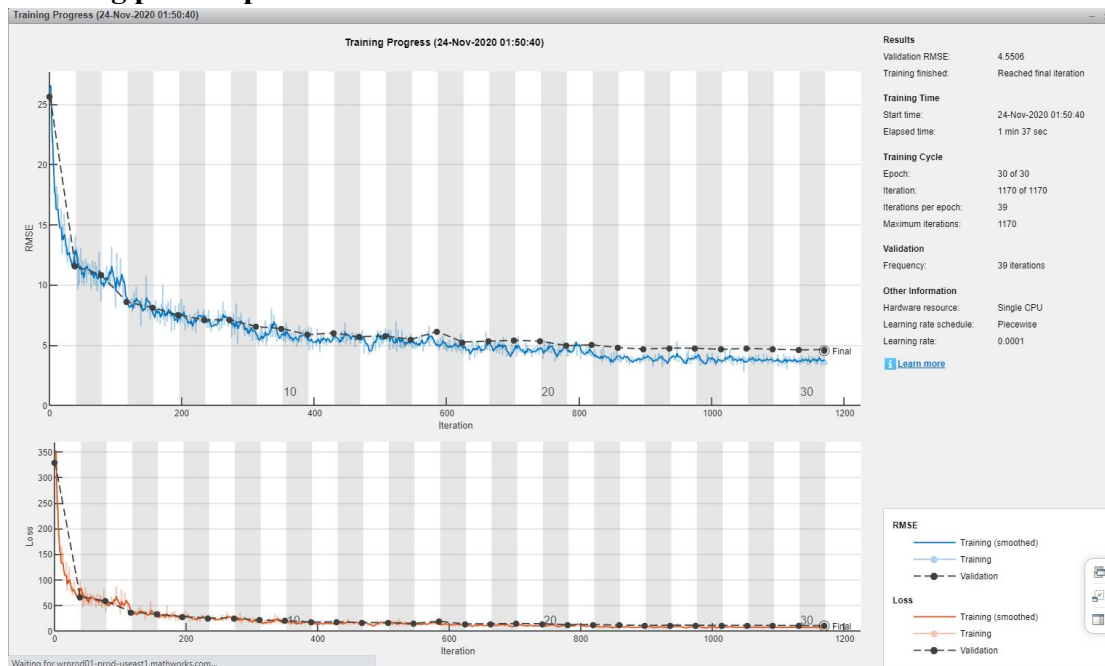
Code can be seen in [CNN.m](#)

Accuracy = 0.9736

#### Normalization Plot:



#### Training process plot:



## Five Worst predicted samples



### Why failed predict these samples:

Because these 6s and 0s are too similar, the tails of these 6's are not obvious, like a scribbled 0.

For this 4, I think it is because the upper open is too big.

## Comparison with Random Learning:

Code can be seen in [Random\\_Learning.m](#)

**Training speed:** The training speed for Random Learning is extremely slow, and I can hardly see a tendency to converge.

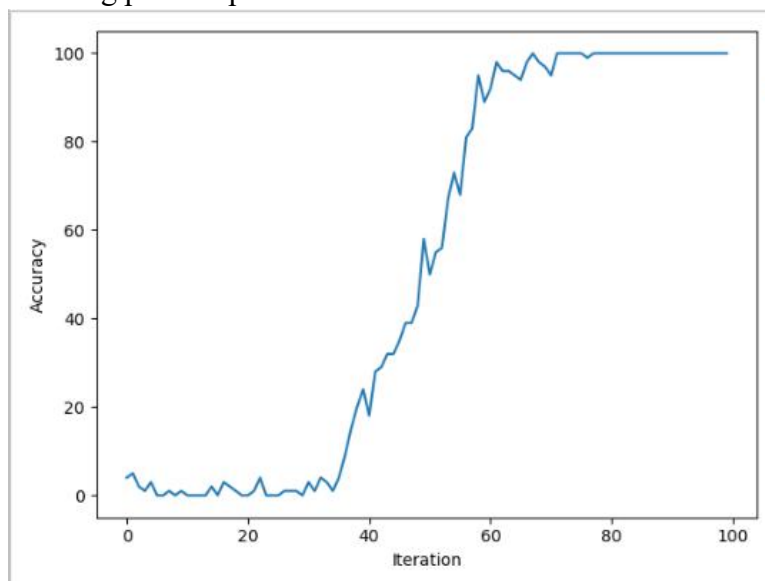
**Accuracy:** Since the model cannot converge, I was unable to compare the accuracy of random Learning and Gradient descent.

## Bonus:

### 1. Binary Addition

Code can be seen in [Binary\\_Addition.py](#)

Training process plot:



## 2. LSTM Gradient

(a)

$$\overline{O_t} = \tanh(C_t) \overline{h_t}$$

$$\overline{C_t} = (1 - \tanh(C_t)^2) O_t \overline{h_t}$$

$$\overline{f_t} = C_{t-1} \overline{C_t}$$

$$\overline{C_{t-1}} = f_t \cdot \overline{C_t}$$

$$\overline{i_t} = g_t \overline{C_t}$$

$$\overline{g_t} = i_t \overline{C_t}$$

(b)  $\overline{W_{ix}} = \sum_t i_t (1 - i_t) x_t \overline{i_t}$

