# EECS 545: Homework #4

Mingliang Duanmu
duanmuml@umich.edu

March 16, 2022

## 1 Neural Network Layer Implementation

**Fully-Connected Layer**

$$\frac{\partial Y_m^{(n)}}{\partial W_{ij}} = \frac{\partial}{\partial W_{ij}}(\sum_l W_{nl}X_{lm} + b_m) = X_{nj}\delta_{ni}\delta_{mj}$$

$$\frac{\partial Y_m^{(n)}}{\partial b_j} = \frac{\partial}{\partial b_j}(\sum_l W_{nl}X_{lm} + b_m) = \delta_{mj}$$

$$\frac{\partial Y_m^{(n)}}{\partial X_{ij}} = \frac{\partial}{\partial X_{ij}}(\sum_l W_{nl}X_{lm} + b_m) = W_{nj}\delta_{ni}\delta_{mj}$$

where $\delta_{ij}$ is Kronecker Delta, $\delta_{ij} = 1$ if $i = j$, else $0$. By vectorizing the result, we get

$$\frac{\partial L}{\partial W} = \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial W} = \frac{\partial L}{\partial Y}X^T$$

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial b} = \sum_i \frac{\partial L}{\partial Y_i}$$

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial X} = W^T\frac{\partial L}{\partial Y}$$
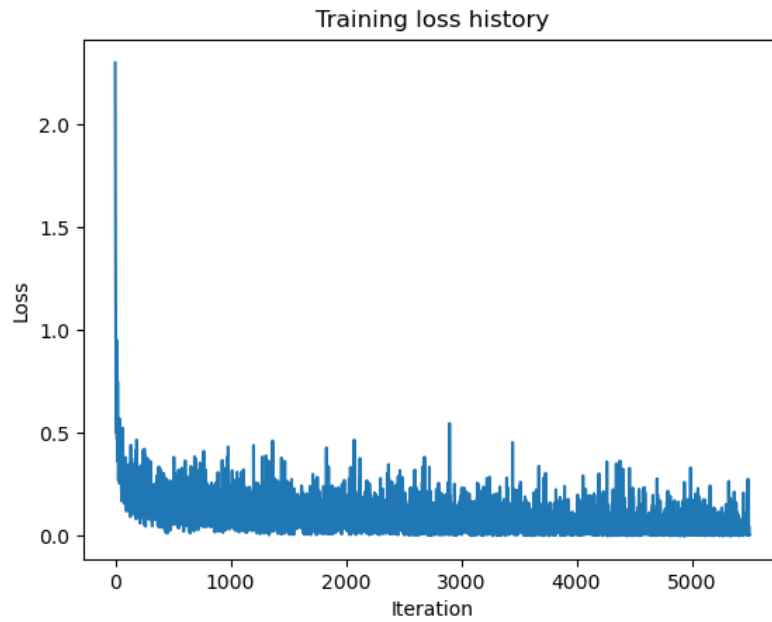
**ReLU**

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial X} = \frac{\partial L}{\partial Y} \circ \Delta(X)$$

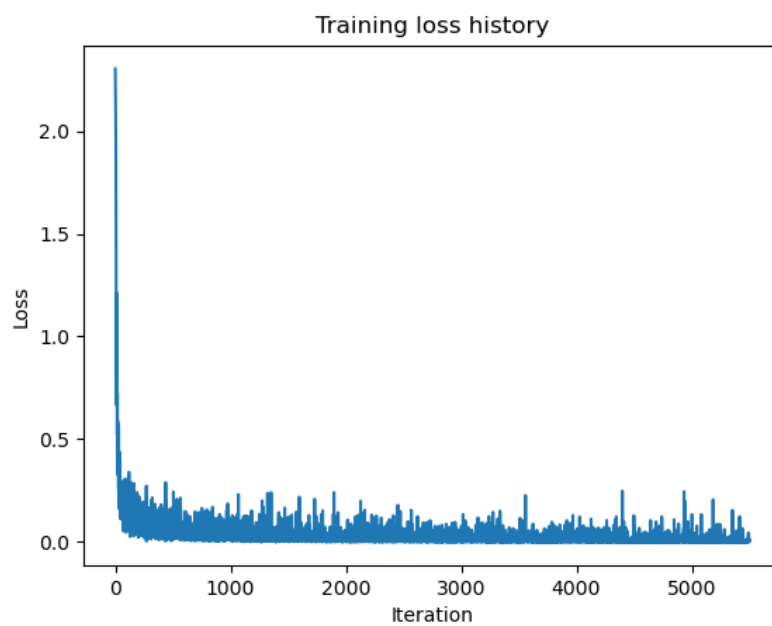where $\circ$ is element-wise product and $\Delta(X)_{ij} = I(x_{ij} > 0)$.

# 2 Multi-class classification with Softmax

```
hidden_dim = 80, accuracy = 0.9695
hidden_dim = 100, accuracy = 0.9698
hidden_dim = 128, accuracy = 0.9697
hidden_dim = 192, accuracy = 0.9678
hidden_dim = 200, accuracy = 0.972
hidden_dim = 256, accuracy = 0.97
```
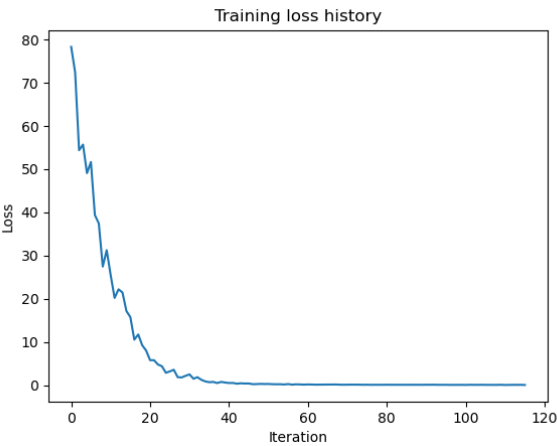
# 3 Convolutional Neural Network for multi-class classification

`accuracy = 0.9803`

# 4 Application to Image Captioning



Training loss history



train
four horses standing in a area that is fenced in <END>
GT:<START> four horses standing in a area that is fenced in <END>



train
a passenger train is parked while a person <UNK> to board it <END>
:<START> a passenger train is parked while a person <UNK> to board it <EI



val
a player of a <UNK> with a a <UNK> <END>
ART> a surfer holding a surf board at the edge of the beach watching the su



val
a round clock with <UNK> <UNK> hanging on the side of a building <END>
GT:<START> a large white clock is on a tower <END>

# 5 Transfer Learning

Finetune the pre-trained model: `accuracy` = 0.954248
Freeze the parameters in pre-trained model and train the final fc layer: `accuracy` = 0.960784