EECS 545 Machine Learning: Homework #4

Due on March 17, 2022 (2 days free late.) at 11:59pm

Professor Honglak Lee Section A

Yuang Huang
yahuang@umich.edu

Problem 1

Neural Network Layer Implementation

Solution

Part a:

$$\begin{aligned}
\frac{\partial L}{\partial W_{ij}} &= \sum_{n=1}^{N}\sum_{m=1}^{D_{out}} \frac{\partial L}{\partial Y_m^n}\frac{\partial Y_m^n}{\partial W_{ij}} \\
Y = XW + B \Rightarrow Y_m^n &= \sum_{i=1}^{D_{out}} X_i^n W_{m,i} + b_m \\
\Rightarrow \frac{\partial L}{\partial W_{ij}} &= \sum_{n=1}^{N} \frac{\partial L}{\partial Y_i^n}\frac{Y_i^n}{\partial W_{ij}} \\
&= \sum_{n=1}^{N} \frac{\partial L}{\partial Y_i^n} X_j^n \\
\Rightarrow \frac{\partial L}{\partial W} &= \mathbf{X}^T \frac{\partial L}{\partial Y}
\end{aligned} \tag{1}$$

.

$$\begin{aligned}
\frac{\partial L}{\partial b_j} &= \sum_{n=1}^{N}\sum_{m=1}^{D_{out}} \frac{\partial L}{\partial Y_m^n}\frac{\partial Y_m^n}{\partial b_j} \\
Y = XW + B \Rightarrow Y_m^n &= \sum_{i=1}^{D_{out}} X_i^n W_{m,i} + b_m \\
\Rightarrow \frac{\partial L}{\partial b_j} &= \sum_{n=1}^{N} \frac{\partial L}{\partial Y_j^n}\frac{Y_j^n}{\partial b_j} \\
&= \sum_{n=1}^{N} \frac{\partial L}{\partial Y_i^n} \cdot 1 \\
\Rightarrow \frac{\partial L}{\partial b} &= \sum_{n=1}^{N} \frac{\partial L}{\partial Y}
\end{aligned} \tag{2}$$

.

$$\begin{aligned}
\frac{\partial L}{\partial X_i^n} &= \sum_{n=1}^{N}\sum_{m=1}^{D_{out}} \frac{\partial L}{\partial Y_m^n}\frac{\partial Y_m^n}{\partial X_i^n} \\
Y = XW + B \Rightarrow Y_m^n &= \sum_{i=1}^{D_{out}} X_i^n W_{m,i} + b_m \\
\Rightarrow \frac{\partial L}{\partial X_i^n} &= \sum_{n=1}^{N} \frac{\partial L}{\partial Y_i^n}\frac{Y_i^n}{\partial X_i^n} \\
&= \sum_{m=1}^{D_{out}} \frac{\partial L}{\partial Y_m^n} W_{m,i} \\
\Rightarrow \frac{\partial L}{\partial X} &= \frac{\partial L}{\partial Y}\mathbf{W}^T
\end{aligned} \tag{3}$$

.

Part b:

Because

$$\frac{\partial Y}{\partial X} = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{4}$$

$$\frac{\partial L}{\partial X} = \frac{\partial L}{\partial Y}\frac{\partial Y}{\partial X} = \begin{cases} \frac{\partial L}{\partial Y}, & x \geq 0 \\ 0, & x < 0 \end{cases} \tag{5}$$

Part c:

Problem 2

Multi-class classification with Softmaxs

Solution
Part a:

Part b:

Part c:
- hidden-dim = 10

train acc = 0.937000 val acc = 0.951800
- hidden-dim = 50

train acc = 0.983000 val acc = 0.972400
- hidden-dim = 100

train acc = 0.985000 val acc = 0.977200
- hidden-dim = 250

train acc = 0.985000 val acc = 0.978400
- hidden-dim = 500

train acc = 0.988000 val acc = 0.981400
- hidden-dim = 800

train acc = 0.988000 val acc = 0.983600
- hidden-dim = 1200

train acc = 0.987000 val acc = 0.981400
- hidden-dim = 1600

train acc = 0.990000 val acc = 0.982600

We will find that the best setting for thr number of hidden units based on the performance on validation set is hidden-dim = 800.

Problem 2 continued on next page...

Figure 1: Training loss history

Figure 1 shows the training loss history where hidden-dim = 800. Using the optimal setting from the previous step, train my network again, the accuracy obtained on the test set is 0.9793.

Problem 3

Convolutional Neural Network for multi-class classification

Solution
Part a:

Part b:

Part c:
The accuracy obtained on the test set is 0.9771 (hidden-dim = 100).

Problem 4

Convolutional Neural Network for multi-class classification

Solution
Part a:

Part b:

Part c:

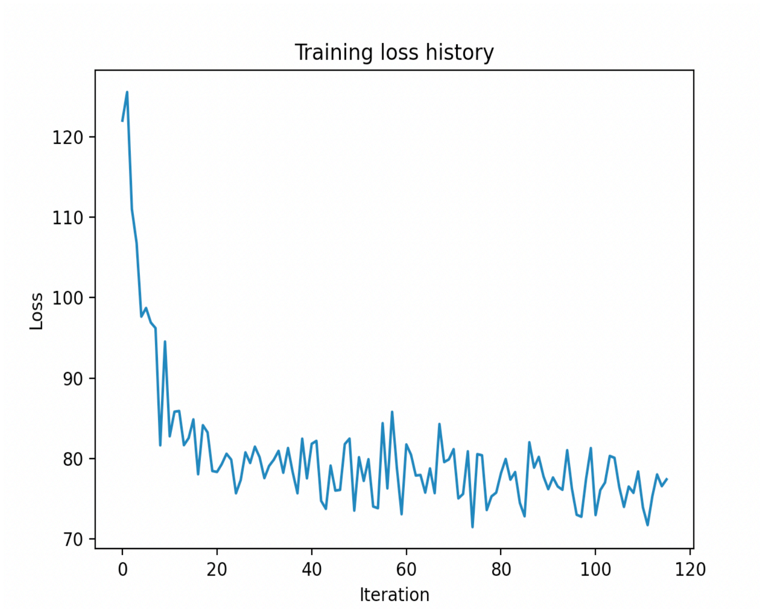


Figure 2: Training loss history

Figure 2 shows the learning curves of training loss.

Problem 4 continued on next page...

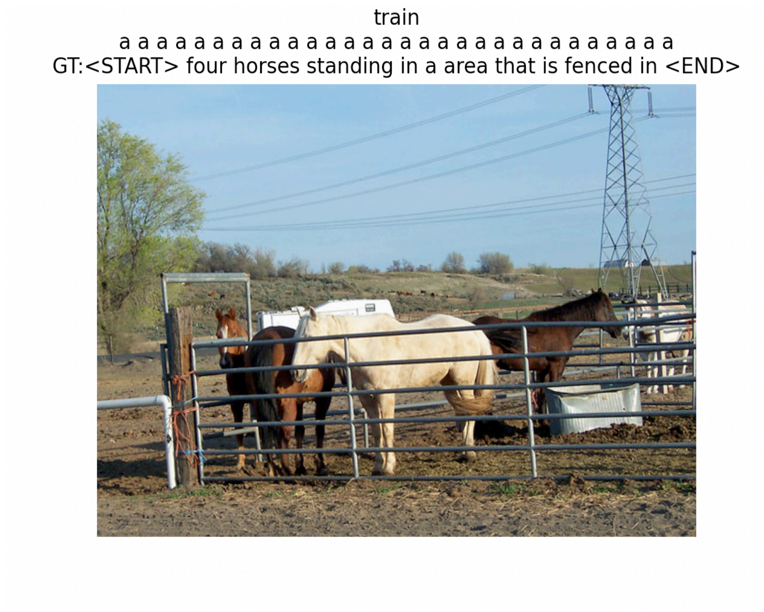


Figure 3: caption samples of horses

Figure 4: caption samples of train

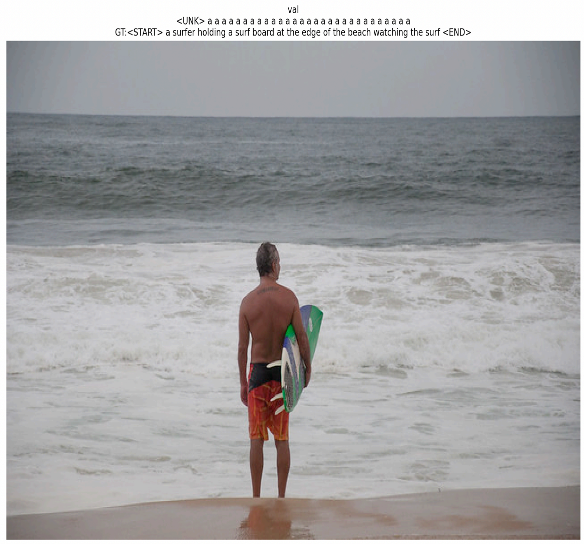


Figure 5: caption samples of surf

Figure 6: caption samples of tower

Figure 3, 4, 5, 6 show the caption samples based on my well-trained network.

Problem 5

Convolutional Neural Network for multi-class classification

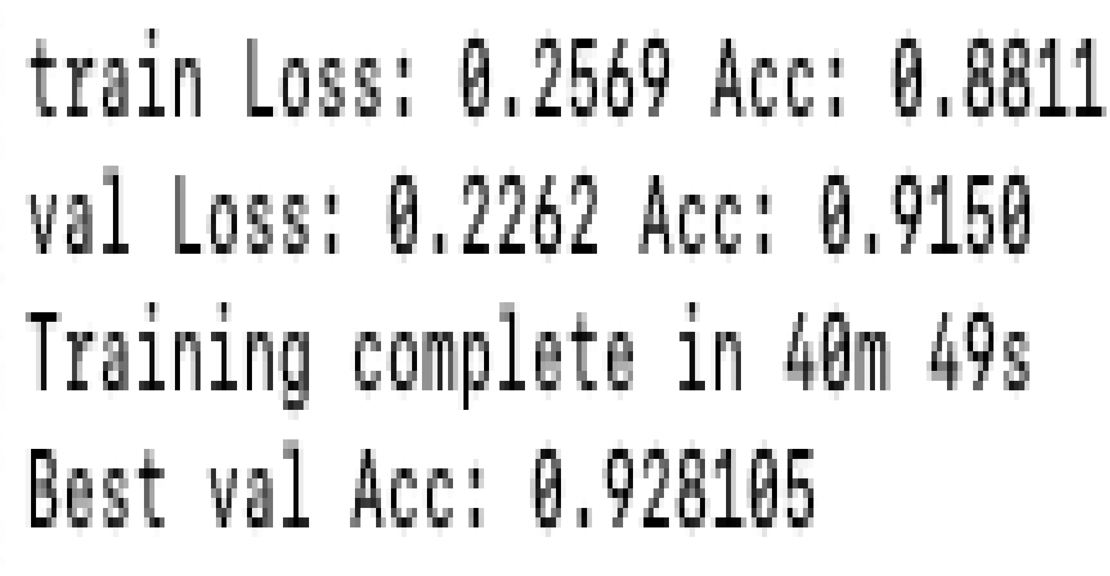
Solution
Part a:

Part b:

Part c:

Part d:

Part e:



```
train Loss: 0.2569 Acc: 0.8811
val Loss: 0.2262 Acc: 0.9150
Training complete in 40m 49s
Best val Acc: 0.928105
```

Figure 7: accuracy on validation dataset

Problem 5 continued on next page...

```
train Loss: 0.3622 Acc: 0.8320
val Loss: 0.1852 Acc: 0.9412
Training complete in 28m 33s
Best val Acc: 0.954248
```

Figure 8: accuracy on validation dataset

Problem 6

Convolutional Backward Pass

Solution

$$
\begin{aligned}
Y_{n,f} &= \sum_{c=1}^{C} X_{n,c} *_{filt} K_{f,c} \\
Y_{n,f,p,q} &= \sum_{c=1}^{C} \sum_{m=1}^{H''} \sum_{n=1}^{W''} X_{n,c,p+m-1,q+n-1} \cdot K_{f,c,m,n} \\
&= \sum_{c=1}^{C} \sum_{i=p}^{p+H''-1} \sum_{j=q}^{q+W''-1} X_{n,c,i,j} \cdot K_{f,c,i-p+1,j-q+1} \\
\Rightarrow \frac{\partial Y_{n,f,p,q}}{\partial X_{n,c,i,j}} &= \sum_{c=1}^{C} K_{f,c,i-p+1,j-q+1}
\end{aligned} \tag{6}
$$

.
With hint 1,

$$
\begin{aligned}
\frac{\partial L}{\partial X_{n,c,i,j}} &= \sum_{f=1}^{F} \sum_{p=1}^{H''} \sum_{q=1}^{W''} \frac{\partial Y_{n,f,p,q}}{\partial X_{n,c,i,j}} \frac{\partial L}{\partial Y_{n,f,p,q}} \\
&= \sum_{f=1}^{F} \sum_{c=1}^{C} \sum_{p=1}^{H''} \sum_{q=1}^{W''} K_{f,c,i-p+1,j-q+1} \frac{\partial L}{\partial Y_{n,f,p,q}} \\
&= \sum_{f=1}^{F} \sum_{c=1}^{C} \sum_{m=i}^{i-H''+1} \sum_{n=j}^{j-W''+1} K_{f,c,m,n} \frac{\partial L}{\partial Y_{n,f,i-m+1,j-n+1}} \\
\Rightarrow \frac{\partial L}{\partial X_{n,c}} &= \sum_{f=1}^{F} K_{f,c} *_{full} \left(\frac{\partial L}{\partial Y_{n,f}}\right)
\end{aligned} \tag{7}
$$

.
Where $m = i - p + 1, n = j - q + 1$.

$$
\begin{aligned}
Y_{n,f,p,q} &= \sum_{c=1}^{C}\sum_{m=1}^{H''}\sum_{n=1}^{W''} X_{n,c,p+m-1,q+n-1} \cdot K_{f,c,m,n} \\
&= \sum_{c=1}^{C}\sum_{i=p}^{p+H''-1}\sum_{j=q}^{q+W''-1} X_{n,c,i,j} \cdot K_{f,c,i-p+1,j-q+1} \\
\Rightarrow \frac{\partial Y_{n,f,p,q}}{\partial K_{f,c,i,j}} &= \sum_{c=1}^{C} X_{n,c,p+i-1,q+j-1}
\end{aligned} \tag{8}
$$

.

With hint 2,

$$
\begin{aligned}
\frac{\partial L}{\partial K_{f,c,i,j}} &= \sum_{n=1}^{N}\sum_{p=1}^{H''}\sum_{q=1}^{W''}\sum_{c=1}^{C} \frac{\partial Y_{n,f,p,q}}{\partial K_{f,c,i,j}} \frac{\partial L}{\partial Y_{n,f,p,q}} \\
&= \sum_{f=1}^{N}\sum_{c=1}^{C}\sum_{p=1}^{H''}\sum_{q=1}^{W''} X_{n,c,p+i-1,q+j-1} \frac{\partial L}{\partial Y_{n,f,p,q}} \\
&= \sum_{f=1}^{N}\sum_{c=1}^{C}\sum_{m=i}^{i+H''-1}\sum_{n=j}^{j+W''-1} X_{n,c,m,n} \frac{\partial L}{\partial Y_{n,f,m-i+1,n-j+1}} \\
\Rightarrow \frac{\partial L}{\partial K_{f,c}} &= \sum_{n=1}^{N} X_{n,c} *_{filt} (\frac{\partial L}{\partial Y_{n,f}})
\end{aligned} \tag{9}
$$

.

Where $m = p + i - 1, n = q + j - 1$.

Problem 6 continued on next page...