National Tsing Hua University
Fall 2023 11210IPT 553000
Deep Learning in Biomedical Optical Imaging
Homework 2

WISARUD YONGBANJERD

109003888:

1. Task A: Performance between BCE loss and CE loss.

1.1 BCE loss

With one hidden unit layer the model performance in terms of accuracy and loss are shown below in Fig. 1.

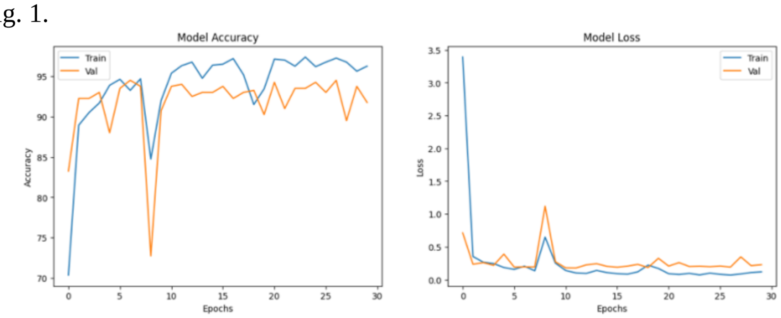

Fig. 1. Graph showing model's accuracy (left) and model's loss (right) from using BCE loss function.

From Fig.1 (left), training accuracy (blue) starts very high at close to 95% but drops significantly after the initial epochs. After this drop, it rises again and then oscillates, staying within a range of around 85% to 93%. While validation accuracy (Orange) also starts very high but experiences a significant drop, similar to the training accuracy. It then rises, but unlike the training accuracy, it tends to oscillate more and has a few more noticeable drops and peaks.

From Fig.1. (right), training loss (Blue) starts high and drops rapidly in the initial epochs. After this sharp decline, it stabilizes and oscillates at a much lower range, indicating that the model is learning. While validation loss (Orange) mirrors the training loss in the beginning, with a sharp decrease. After the decrease, however, the validation loss exhibits more fluctuations compared to the training loss.

You can observe that the model doesn't show a clear sign of overfitting throughout its training. Typically, overfitting is indicated when training accuracy continues to increase (or training loss continues to decrease) while validation accuracy decreases (or validation loss increases). In this case, both training and validation metrics follow somewhat similar trajectories. The sharp drop in accuracy and the spike in loss for both training and validation in the early epochs is intriguing. This could be attributed to a number of reasons, including the learning rate being too high initially or some random initialization that the model quickly adjusted from. The oscillations in both the accuracy and loss, especially in the validation curves, might suggest that the learning rate is still a bit high, causing the model to "jump around" in the loss landscape. By the end of the training (around epoch 30), both the training and validation curves seem to be stabilizing. The model might have reached a point where it won't benefit much from further training with the current settings.

With the model evaluation, the model is 73.1% confident that the scan is abnormal and 26.9% confident that the scan is normal. The model result is prediction is abnormal. While using test set to perform interference and evaluation, test accuracy is 72.0% meaning that model correctly classifies 72% of the unseen samples. It can be seen that the test accuracy is close to the training and validation evaluation, then it's a sign that the model generalizes well.

1.2 CE loss

With one hidden unit layer the model perform in terms of accuracy and loss are shown below in Fig. 2.

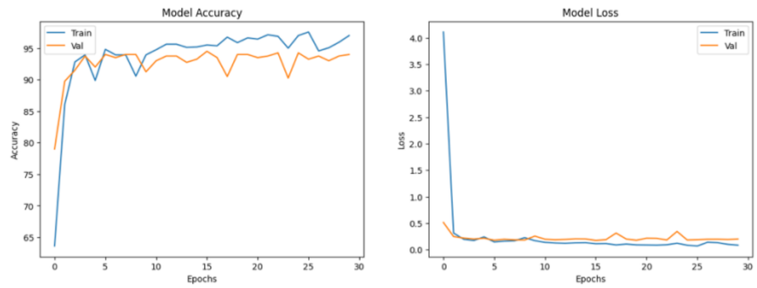

Fig. 2. Graph showing model's accuracy (left) and model's loss (right) from using CE loss function.

From Fig.2. (left), training and validation accuracies are both steadily increasing, which is a good sign. By the end of the training, both training and validation accuracies seem to converge at around 93-95%, indicating a well-generalizing model with minimal overfitting.

From Fig.2. (right), both training and validation losses decrease substantially in the initial epochs and then stabilize. The training and validation losses converge towards the end, further confirming that overfitting is minimal.

With the model evaluation, this model is 78.18% confident that the scan is normal. The model result is prediction is normal. While using test set to perform interference and evaluation, test accuracy is 73.0% meaning that the model correctly predicts the class of the image 73.0% of the time.

Both models perform well on the given data, with the CE loss model showing slightly better and more consistent performance. The CE loss model's accuracy and stability across epochs make it a preferable choice between the two. However, it's essential to note that model performance can also be influenced by factors other than just the loss function, such as initialization, learning rate, data augmentation, etc.

1.3 Comparing CE Loss Model to BEC Loss Model in the prediction confidence and overall accuracy

1.3.1 Single Image Prediction Confidence:

<u>BCE Loss:</u> The model has a 73.1% confidence that the scan is abnormal, and this is correct.

<u>CE Loss:</u> The model is 78.18% confident that the scan is normal, and this is correct.

Both models have made correct predictions on their respective images. However, the CE model shows a slightly higher confidence in its prediction compared to the BCE model. The difference in confidence may be attributed to the inherent characteristics of the images themselves, or it might be related to the loss function's influence during training.

1.3.2 Overall Test Set Accuracy:

<u>BCE Loss:</u> 72.0%

<u>CE Loss:</u> 73.0%

The accuracy between the two models is very close. The CE loss model has a slightly higher accuracy by 1%. They are nearly equivalent in terms of accuracy.

The single image prediction from the model evaluation gives you insight into how the model evaluates individual examples, but the test set accuracy is a better measure of overall model performance.

2. Task B: Performance between Different Hyperparameters.

I will be choosing learning rate and number of epochs to see the performance between different hyperparameters. The CE loss with 3 hidden layers and 1 output layer for model structure will be used to train the model. Table 1. shows the values of various numbers used in learning rate and number of epochs. I trained by varies each of the parameter at a time.

Table 1. Various Number of Learning Rate and Number of Epochs used for Training the Model

Items	Values
Learning Rate with the number of epochs = 30	0.0001
	0.0005
	0.01
Number of Epochs with the learning rate = 0.001	10
	50
	100

From the Learning rate 0.0001 the test accuracy is 75.5%. Take # 252 as test image. This model is 98.81% confident that the scan is normal. Prediction is normal and the result is correct. The model accuracy and the model loss are shown in Fig. 3.

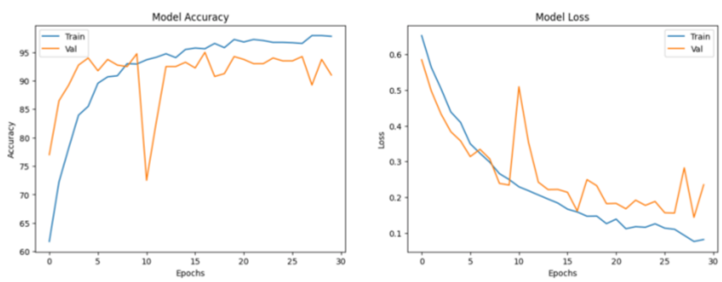

Fig. 3. Showing the result of the model accuracy and loss when set the learning rate = 0.0001

Model Accuracy Plot: Train and validation accuracy are close, but the validation accuracy has some fluctuations.
Model Loss Plot: Loss is decreasing consistently for both train and validation sets.

From learning rate 0.0005 the test accuracy is 70.75%. Take # 289 as test image. This model is 83.53% confident that the scan is normal. Prediction is normal and the result is correct. The model accuracy and the model loss are shown in Fig. 4.

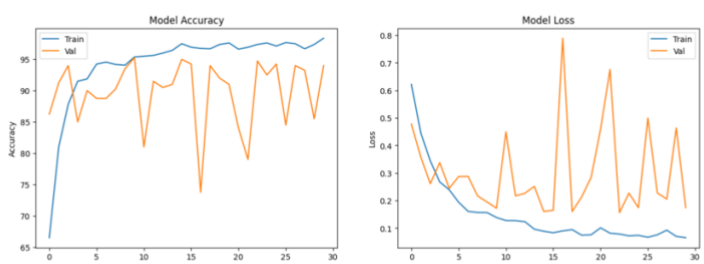

Fig. 4. Showing the result of the model accuracy and loss when set the learning rate = 0.0005

Model Accuracy Plot: There are high fluctuations in validation accuracy, indicating potential instability during training.

Model Loss Plot: The validation loss has significant spikes, suggesting the model might be having difficulties generalizing to new data.

From the Learning rate 0.01 the test accuracy is 75%. Take # 207 as test image. This model is 99.97% confident that the scan is normal. Prediction is normal and the result is correct. The model accuracy and the model loss are shown in Fig. 5

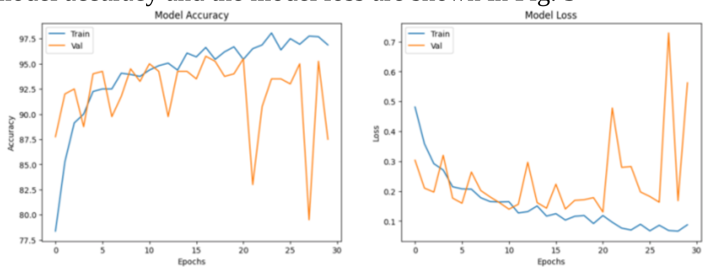

Fig. 5. Showing the result of the model accuracy and loss when set the learning rate = 0.01

Model Accuracy Plot: Train and validation accuracy are close with minor fluctuations in the validation set.
Model Loss Plot: Validation loss has sharp peaks, particularly towards the end, which might suggest the model is overfitting to the training data.

The model with a learning rate of 0.0001 seems to be the most stable among the three in terms of training dynamics. The accuracy and loss plots show consistent learning and generalization to the validation set. The model trained with 0.0005 as the learning rate demonstrates instability, especially in the validation metrics. This might suggest that this learning rate could be on the edge of being too high, leading to difficulties in converging. With the learning rate of 0.01, we see the model can achieve a high test accuracy similar to the first one, but the sharp peaks in the validation loss towards the end of training might indicate potential overfitting.

From the number epochs = 10 the test accuracy is 75.75%. Take # 182 as test image. This model is 88.37% confident that the scan is abnormal. Prediction is abnormal and the result is correct. The model accuracy and the model loss are shown in Fig. 6.

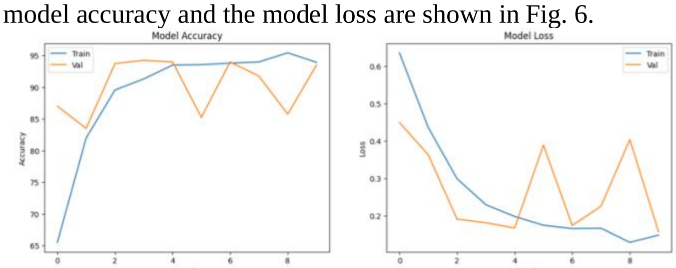

Fig. 6. Showing the result of the model accuracy and loss when set the number of epochs = 10

Model Accuracy Plot: The training accuracy is generally higher than the validation accuracy, which is typical. There is some fluctuation in the validation accuracy, indicating that the model might be starting to overfit as it captures the training data's noise.
Model Loss Plot: The training loss decreases steadily, while the validation loss has some fluctuations. This again could hint towards the onset of overfitting.

From the number epochs = 50 the test accuracy is 71.25%. Take # 370 as test image. This model is 99.97% confident that the scan is normal. Prediction is abnormal and the result is correct. The model accuracy and the model loss are shown in Fig. 7.

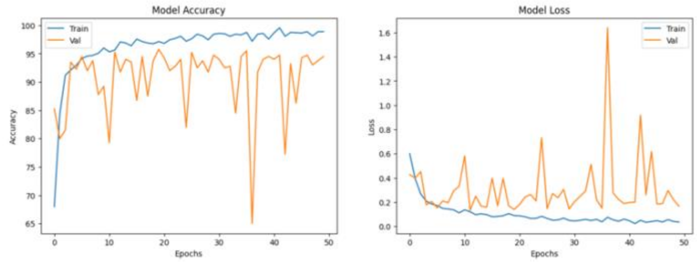

Fig. 7. Showing the result of the model accuracy and loss when set the number of epochs = 50

Model Accuracy Plot: The gap between the training and validation accuracy has increased, especially towards the later epochs. This is a stronger sign of overfitting. Essentially, the model is becoming too specialized in the training data and performs less well on validation data.
Model Loss Plot: The validation loss starts to increase around epoch 20, while the training loss continues to decrease. This divergence is a hallmark of overfitting.

From the number epochs = 100 the test accuracy is 74.0%. Take # 178 as test image. This model is 99.91% confident that the scan is abnormal. Prediction is abnormal and the result is correct. The model accuracy and the model loss are shown in Fig. 8.

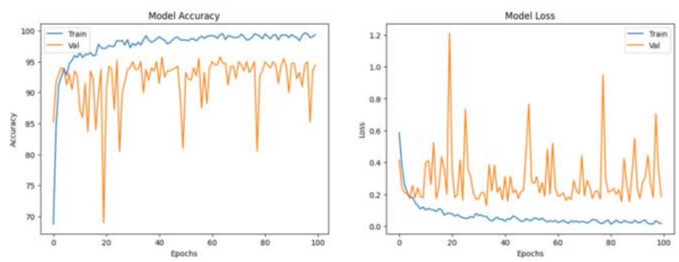

Fig. 8. Showing the result of the model accuracy and loss when set the number of epochs = 100

Model Accuracy Plot: The trend seen in the model trained for 50 epochs continues here. The training accuracy is consistently high, hovering near 100%, while the validation accuracy fluctuates considerably. Again, this model appears to overfit as the epochs increase.
Model Loss Plot: The validation loss shows significant oscillations, indicating instability in the learning process. The training loss remains relatively low, reiterating the model's tendency to overfit.

As the number of epochs increases, the model tends to overfit more, which is expected behavior. Overfitting happens when a model learns the training data too well, including its noise and outliers, making it perform less optimally on unseen data. In conclusion, among the three models, the one trained with 10 epochs seems to be the most balanced in terms of overfitting.

References

1. De Jun Huang.(Jun 12 2021). *Learning Day57/Practical 5:Loss function - CrossEntropyLoss vs BCELoss in Pytorch;Softmax vs sigmoid;Loss calculation.* https://medium.com/dejunhuang/learning-day-57-practical-5-loss-function-crossentropyloss-vs-bceloss-in-pytorch-softmax-vs-bd866c8a0d23K. Gallo and G. Assanto, "All-optical diode based on second-harmonic generation in an asymmetric waveguide," J. Opt. Soc. Am. B 16(2), 267–269 (1999).
2. Pytorch Contributors. (2023). Crossentropyloss. https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html
3. Ng, A. (n.d.). Deep Learning Specialization. Coursera. Retrieved [October 12, 2023], from https://www.coursera.org/specializations/deep-learning