

1

num_train_batches: 5

num_test_batches: 1

num_img_per_batch: 10000

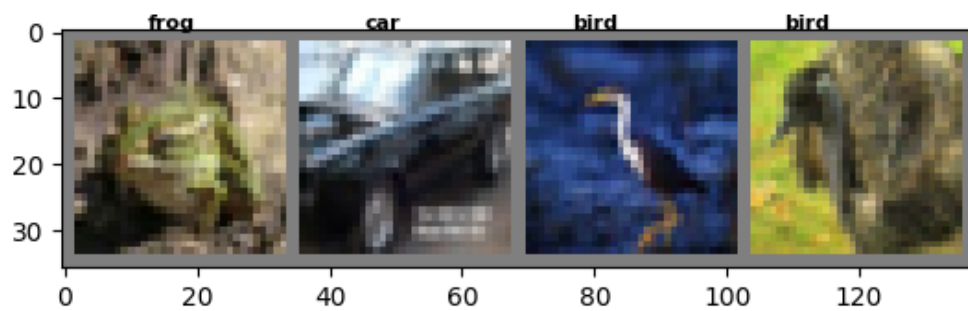
num_train_img: 50000

num_test_img: 10000

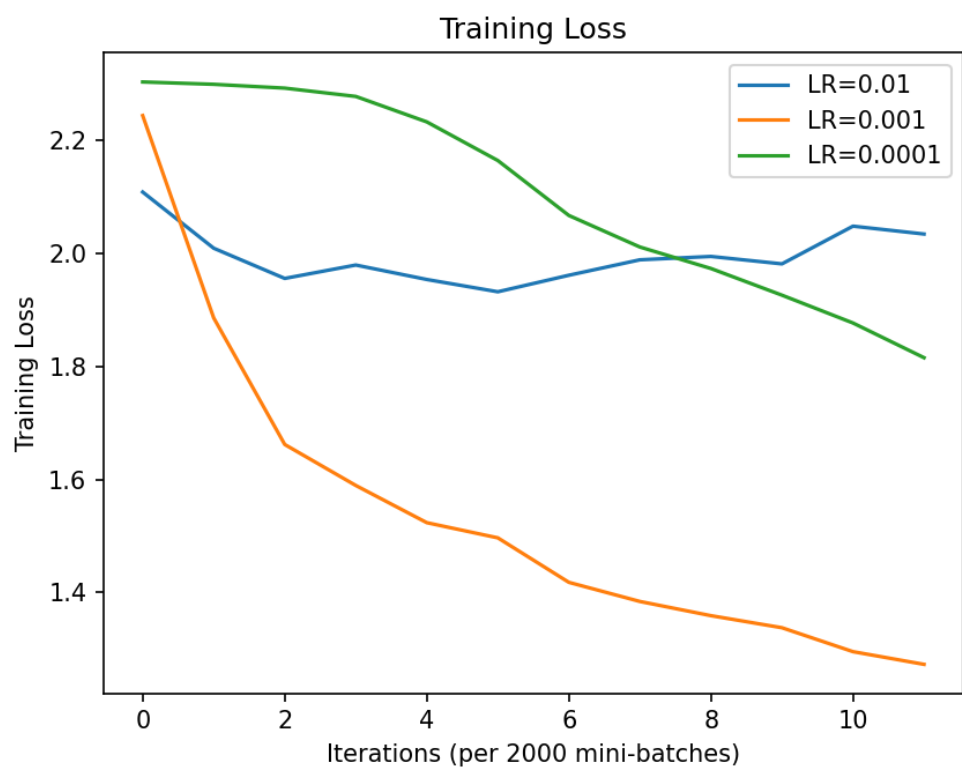
size_batch_bytes: 30730

size_image_bytes: 3

size_batchimage_bytes: 30000



4





d.

The learning rate of 0.01 is too large for minimizing the loss. The model oscillates between peaks, so after many iterations, the loss curve looks flat. On the other hand, 0.0001 seems too small, causing the loss to decrease very slowly.

In contrast, 0.001 is the most appropriate value. The loss continues to decrease steadily, and it achieves the best performance on both the training and test error compared to the other learning rates.