1)

Here is the screenshot which shows the gradient updates

```
Epoch 1/10
860/860 [=
                           =] - ETA: 0s - loss: 0.5649 - accuracy: 0.7933
860/860 [==:
:============ ] - 118s 137ms/step - loss: 0.3458 - accuracy: 0.8735 - val_loss: 0.2734 - val_accuracy: 0.8970
Epoch 4/10
860/860 [==========================] - ETA: 0s - loss: 0.3166 - accuracy: 0.8816
Epoch 4: val_loss improved from 0.27342 to 0.26447, saving model to model.weights.best.hdf5
                          ==] - 121s 141ms/step - loss: 0.3166 - accuracy: 0.8816 - val_loss: 0.2645 - val_accuracy: 0.8996
Epoch 5/10
Epoch 6/10
#859/860 [==================].] - ETA: 0s - loss: 0.2804 - accuracy: 0.8965

Epoch 6: val_loss improved from 0.24558 to 0.23302, saving model to model.weights.best.hdf5

860/860 [==========] - 124s 144ms/step - loss: 0.2804 - accuracy: 0.8965 - val_loss: 0.2330 - val_accuracy: 0.9114
860/860 [============] - ETA: 05 - loss: 0.2682 - accuracy: 0.9008
Epoch 7: val_loss improved from 0.23302 to 0.22831, saving model to model.weights.best.hdf5
860/860 [==========] - 121s 140ms/step - loss: 0.2682 - accuracy: 0.9008 - val_loss: 0.2283 - val_accuracy: 0.9094
.
860/860 [==
                         ===] - 118s 137ms/step - loss: 0.2557 - accuracy: 0.9054 - val_loss: 0.2223 - val_accuracy: 0.9162
Epoch 9/10
Epoch 10/10
<keras.callbacks.History at 0x7ff8e1c2a8b0>
```

The accuracy of test set is shown below

Test accuracy: 0.9171000123023987

2)