Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_1 (Batch (None, 28, 28, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 14, 14, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_2 (Batch (None, 14, 14, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 3136) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 128) 401536

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_3 (Batch (None, 128) 512

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 1290

=================================================================

Total params: 423,050

Trainable params: 422,602

Non-trainable params: 448

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/10

- 195s - loss: 0.3332 - acc: 0.8935 - val\_loss: 0.0682 - val\_acc: 0.9795

Epoch 2/10

- 201s - loss: 0.1372 - acc: 0.9579 - val\_loss: 0.0483 - val\_acc: 0.9852

Epoch 3/10

- 195s - loss: 0.1080 - acc: 0.9662 - val\_loss: 0.0387 - val\_acc: 0.9881

Epoch 4/10

- 184s - loss: 0.0962 - acc: 0.9700 - val\_loss: 0.0356 - val\_acc: 0.9902

Epoch 5/10

- 186s - loss: 0.0885 - acc: 0.9725 - val\_loss: 0.0379 - val\_acc: 0.9881

Epoch 6/10

- 179s - loss: 0.0837 - acc: 0.9746 - val\_loss: 0.0359 - val\_acc: 0.9886

Epoch 7/10

- 180s - loss: 0.0769 - acc: 0.9766 - val\_loss: 0.0296 - val\_acc: 0.9900

Epoch 8/10

- 179s - loss: 0.0745 - acc: 0.9770 - val\_loss: 0.0297 - val\_acc: 0.9919

Epoch 9/10

- 179s - loss: 0.0738 - acc: 0.9775 - val\_loss: 0.0251 - val\_acc: 0.9933

Epoch 10/10

- 178s - loss: 0.0690 - acc: 0.9794 - val\_loss: 0.0277 - val\_acc: 0.9924

Final loss: 0.027665, final accuracy: 0.992381

Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_1 (Batch (None, 28, 28, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 28, 28, 32) 25632

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_2 (Batch (None, 28, 28, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 14, 14, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_3 (Batch (None, 14, 14, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_4 (Conv2D) (None, 14, 14, 64) 36928

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_4 (Batch (None, 14, 14, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 3136) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 128) 401536

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_5 (Batch (None, 128) 512

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 1290

=================================================================

Total params: 485,994

Trainable params: 485,354

Non-trainable params: 640

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/10

- 543s - loss: 0.3112 - acc: 0.9054 - val\_loss: 0.0539 - val\_acc: 0.9833

Epoch 2/10

- 541s - loss: 0.1126 - acc: 0.9651 - val\_loss: 0.0638 - val\_acc: 0.9819

Epoch 3/10

- 579s - loss: 0.0840 - acc: 0.9740 - val\_loss: 0.0379 - val\_acc: 0.9881

Epoch 4/10

- 564s - loss: 0.0678 - acc: 0.9793 - val\_loss: 0.0339 - val\_acc: 0.9893

Epoch 5/10

- 550s - loss: 0.0661 - acc: 0.9804 - val\_loss: 0.0355 - val\_acc: 0.9893

Epoch 6/10

- 560s - loss: 0.0585 - acc: 0.9818 - val\_loss: 0.0295 - val\_acc: 0.9902

Epoch 7/10

- 549s - loss: 0.0568 - acc: 0.9823 - val\_loss: 0.0223 - val\_acc: 0.9929

Epoch 8/10

- 566s - loss: 0.0516 - acc: 0.9836 - val\_loss: 0.0240 - val\_acc: 0.9929

Epoch 9/10

- 557s - loss: 0.0465 - acc: 0.9855 - val\_loss: 0.0280 - val\_acc: 0.9924

Epoch 10/10

- 551s - loss: 0.0481 - acc: 0.9851 - val\_loss: 0.0226 - val\_acc: 0.9926

Final loss: 0.022632, final accuracy: 0.992619

Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_1 (Batch (None, 28, 28, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

leaky\_re\_lu\_1 (LeakyReLU) (None, 28, 28, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 14, 14, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_2 (Batch (None, 14, 14, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

leaky\_re\_lu\_2 (LeakyReLU) (None, 14, 14, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 3136) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 128) 401536

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_3 (Batch (None, 128) 512

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

leaky\_re\_lu\_3 (LeakyReLU) (None, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 1290

=================================================================

Total params: 423,050

Trainable params: 422,602

Non-trainable params: 448

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/10

- 211s - loss: 0.4393 - acc: 0.8661 - val\_loss: 0.0760 - val\_acc: 0.9781

Epoch 2/10

- 210s - loss: 0.1825 - acc: 0.9448 - val\_loss: 0.0602 - val\_acc: 0.9810

Epoch 3/10

- 193s - loss: 0.1422 - acc: 0.9547 - val\_loss: 0.0395 - val\_acc: 0.9883

Epoch 4/10

- 192s - loss: 0.1180 - acc: 0.9644 - val\_loss: 0.0395 - val\_acc: 0.9869

Epoch 5/10

- 191s - loss: 0.1100 - acc: 0.9659 - val\_loss: 0.0323 - val\_acc: 0.9886

Epoch 6/10

- 192s - loss: 0.0990 - acc: 0.9698 - val\_loss: 0.0281 - val\_acc: 0.9902

Epoch 7/10

- 191s - loss: 0.0895 - acc: 0.9722 - val\_loss: 0.0283 - val\_acc: 0.9910

Epoch 8/10

- 190s - loss: 0.0867 - acc: 0.9718 - val\_loss: 0.0317 - val\_acc: 0.9898

Epoch 9/10

- 204s - loss: 0.0843 - acc: 0.9742 - val\_loss: 0.0295 - val\_acc: 0.9907

Epoch 10/10

- 201s - loss: 0.0792 - acc: 0.9755 - val\_loss: 0.0226 - val\_acc: 0.9926

Final loss: 0.022646, final accuracy: 0.992619

Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_1 (Batch (None, 28, 28, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

leaky\_re\_lu\_1 (LeakyReLU) (None, 28, 28, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 28, 28, 32) 25632

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_2 (Batch (None, 28, 28, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

leaky\_re\_lu\_2 (LeakyReLU) (None, 28, 28, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 14, 14, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_3 (Batch (None, 14, 14, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

leaky\_re\_lu\_3 (LeakyReLU) (None, 14, 14, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_4 (Conv2D) (None, 14, 14, 64) 36928

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_4 (Batch (None, 14, 14, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

leaky\_re\_lu\_4 (LeakyReLU) (None, 14, 14, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 3136) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 256) 803072

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_5 (Batch (None, 256) 1024

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

leaky\_re\_lu\_5 (LeakyReLU) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 2570

=================================================================

Total params: 889,322

Trainable params: 888,426

Non-trainable params: 896

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/10

- 605s - loss: 0.3439 - acc: 0.8933 - val\_loss: 0.0636 - val\_acc: 0.9810

Epoch 2/10

- 613s - loss: 0.1231 - acc: 0.9629 - val\_loss: 0.0488 - val\_acc: 0.9864

Epoch 3/10

- 604s - loss: 0.0982 - acc: 0.9706 - val\_loss: 0.0321 - val\_acc: 0.9886

Epoch 4/10

- 598s - loss: 0.0858 - acc: 0.9733 - val\_loss: 0.0267 - val\_acc: 0.9902

Epoch 5/10

- 600s - loss: 0.0731 - acc: 0.9774 - val\_loss: 0.0366 - val\_acc: 0.9895

Epoch 6/10

- 601s - loss: 0.0693 - acc: 0.9788 - val\_loss: 0.0220 - val\_acc: 0.9929

Epoch 7/10

- 600s - loss: 0.0656 - acc: 0.9796 - val\_loss: 0.0242 - val\_acc: 0.9926

Epoch 8/10

- 563s - loss: 0.0607 - acc: 0.9808 - val\_loss: 0.0289 - val\_acc: 0.9905

Epoch 9/10

- 564s - loss: 0.0571 - acc: 0.9824 - val\_loss: 0.0339 - val\_acc: 0.9871

Epoch 10/10

Epoch 00010: ReduceLROnPlateau reducing learning rate to 0.0005000000237487257.

- 564s - loss: 0.0549 - acc: 0.9831 - val\_loss: 0.0431 - val\_acc: 0.9864

Final loss: 0.043084, final accuracy: 0.986429

Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 28, 28, 32) 25632

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 14, 14, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_4 (Conv2D) (None, 14, 14, 64) 36928

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 3136) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 256) 803072

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 2570

=================================================================

Total params: 887,530

Trainable params: 887,530

Non-trainable params: 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/20

- 334s - loss: 0.3965 - acc: 0.8730 - val\_loss: 0.0518 - val\_acc: 0.9848

Epoch 2/20

- 313s - loss: 0.1344 - acc: 0.9596 - val\_loss: 0.0417 - val\_acc: 0.9869

Epoch 3/20

- 320s - loss: 0.1004 - acc: 0.9705 - val\_loss: 0.0484 - val\_acc: 0.9852

Epoch 4/20

- 321s - loss: 0.0891 - acc: 0.9741 - val\_loss: 0.0388 - val\_acc: 0.9895

Epoch 5/20

- 314s - loss: 0.0826 - acc: 0.9764 - val\_loss: 0.0399 - val\_acc: 0.9888

Epoch 6/20

- 329s - loss: 0.0812 - acc: 0.9764 - val\_loss: 0.0439 - val\_acc: 0.9886

Epoch 7/20

- 318s - loss: 0.0760 - acc: 0.9787 - val\_loss: 0.0330 - val\_acc: 0.9900

Epoch 8/20

- 321s - loss: 0.0751 - acc: 0.9788 - val\_loss: 0.0383 - val\_acc: 0.9900

Epoch 9/20

- 332s - loss: 0.0785 - acc: 0.9780 - val\_loss: 0.0329 - val\_acc: 0.9902

Epoch 10/20

- 339s - loss: 0.0795 - acc: 0.9773 - val\_loss: 0.0324 - val\_acc: 0.9905

Epoch 11/20

- 354s - loss: 0.0798 - acc: 0.9784 - val\_loss: 0.0284 - val\_acc: 0.9907

Epoch 12/20

- 334s - loss: 0.0817 - acc: 0.9785 - val\_loss: 0.0383 - val\_acc: 0.9905

Epoch 13/20

- 339s - loss: 0.0842 - acc: 0.9774 - val\_loss: 0.0298 - val\_acc: 0.9907

Epoch 14/20

- 346s - loss: 0.0840 - acc: 0.9779 - val\_loss: 0.0380 - val\_acc: 0.9874

Epoch 15/20

- 335s - loss: 0.0841 - acc: 0.9785 - val\_loss: 0.0314 - val\_acc: 0.9917

Epoch 16/20

- 332s - loss: 0.0862 - acc: 0.9782 - val\_loss: 0.0304 - val\_acc: 0.9898

Epoch 17/20

- 333s - loss: 0.0895 - acc: 0.9764 - val\_loss: 0.0311 - val\_acc: 0.9919

Epoch 18/20

- 320s - loss: 0.0911 - acc: 0.9768 - val\_loss: 0.0394 - val\_acc: 0.9898

Epoch 19/20

- 318s - loss: 0.0939 - acc: 0.9760 - val\_loss: 0.0333 - val\_acc: 0.9902

Epoch 20/20

- 326s - loss: 0.0944 - acc: 0.9758 - val\_loss: 0.0323 - val\_acc: 0.9905

Final loss: 0.032315, final accuracy: 0.990476

Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 14, 14, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 7, 7, 128) 73856

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 7, 7, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 6272) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 128) 802944

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_4 (Dropout) (None, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 1290

=================================================================

Total params: 897,418

Trainable params: 897,418

Non-trainable params: 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/15

- 101s - loss: 0.5046 - acc: 0.8355 - val\_loss: 0.0818 - val\_acc: 0.9755

Epoch 2/15

- 100s - loss: 0.1582 - acc: 0.9514 - val\_loss: 0.0479 - val\_acc: 0.9883

Epoch 3/15

- 103s - loss: 0.1234 - acc: 0.9633 - val\_loss: 0.0389 - val\_acc: 0.9898

Epoch 4/15

- 106s - loss: 0.1075 - acc: 0.9677 - val\_loss: 0.0337 - val\_acc: 0.9905

Epoch 5/15

- 102s - loss: 0.0985 - acc: 0.9711 - val\_loss: 0.0325 - val\_acc: 0.9912

Epoch 6/15

- 101s - loss: 0.0881 - acc: 0.9736 - val\_loss: 0.0347 - val\_acc: 0.9905

Epoch 7/15

- 100s - loss: 0.0886 - acc: 0.9736 - val\_loss: 0.0326 - val\_acc: 0.9905

Epoch 8/15

- 104s - loss: 0.0866 - acc: 0.9748 - val\_loss: 0.0307 - val\_acc: 0.9914

Epoch 9/15

- 100s - loss: 0.0850 - acc: 0.9743 - val\_loss: 0.0287 - val\_acc: 0.9921

Epoch 10/15

- 102s - loss: 0.0863 - acc: 0.9759 - val\_loss: 0.0337 - val\_acc: 0.9898

Epoch 11/15

- 102s - loss: 0.0837 - acc: 0.9753 - val\_loss: 0.0319 - val\_acc: 0.9914

Epoch 12/15

- 105s - loss: 0.0887 - acc: 0.9753 - val\_loss: 0.0331 - val\_acc: 0.9912

Epoch 13/15

Epoch 00013: ReduceLROnPlateau reducing learning rate to 0.0005000000237487257.

- 112s - loss: 0.0894 - acc: 0.9749 - val\_loss: 0.0327 - val\_acc: 0.9914

Epoch 14/15

- 122s - loss: 0.0767 - acc: 0.9791 - val\_loss: 0.0318 - val\_acc: 0.9921

Epoch 15/15

- 130s - loss: 0.0716 - acc: 0.9793 - val\_loss: 0.0306 - val\_acc: 0.9921

Final loss: 0.030594, final accuracy: 0.992143

Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 1600

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 14, 14, 64) 51264

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 7, 7, 128) 73856

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 7, 7, 128) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 6272) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 256) 1605888

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_4 (Dropout) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 2570

=================================================================

Total params: 1,735,178

Trainable params: 1,735,178

Non-trainable params: 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/15

- 185s - loss: 0.4168 - acc: 0.8663 - val\_loss: 0.0671 - val\_acc: 0.9769

Epoch 2/15

- 183s - loss: 0.1285 - acc: 0.9611 - val\_loss: 0.0589 - val\_acc: 0.9840

Epoch 3/15

- 179s - loss: 0.0975 - acc: 0.9708 - val\_loss: 0.0461 - val\_acc: 0.9831

Epoch 4/15

- 195s - loss: 0.0877 - acc: 0.9734 - val\_loss: 0.0410 - val\_acc: 0.9883

Epoch 5/15

- 164s - loss: 0.0789 - acc: 0.9767 - val\_loss: 0.0332 - val\_acc: 0.9905

Epoch 6/15

- 164s - loss: 0.0769 - acc: 0.9775 - val\_loss: 0.0351 - val\_acc: 0.9898

Epoch 7/15

- 167s - loss: 0.0750 - acc: 0.9787 - val\_loss: 0.0299 - val\_acc: 0.9910

Epoch 8/15

- 165s - loss: 0.0710 - acc: 0.9795 - val\_loss: 0.0332 - val\_acc: 0.9917

Epoch 9/15

- 162s - loss: 0.0716 - acc: 0.9793 - val\_loss: 0.0313 - val\_acc: 0.9898

Epoch 10/15

- 166s - loss: 0.0724 - acc: 0.9804 - val\_loss: 0.0300 - val\_acc: 0.9924

Epoch 11/15

- 168s - loss: 0.0754 - acc: 0.9795 - val\_loss: 0.0313 - val\_acc: 0.9919

Epoch 12/15

- 173s - loss: 0.0713 - acc: 0.9798 - val\_loss: 0.0317 - val\_acc: 0.9910

Epoch 13/15

- 184s - loss: 0.0808 - acc: 0.9791 - val\_loss: 0.0340 - val\_acc: 0.9914

Epoch 14/15

Epoch 00014: ReduceLROnPlateau reducing learning rate to 0.0005000000237487257.

- 188s - loss: 0.0785 - acc: 0.9786 - val\_loss: 0.0357 - val\_acc: 0.9900

Epoch 15/15

- 188s - loss: 0.0625 - acc: 0.9828 - val\_loss: 0.0318 - val\_acc: 0.9910

Final loss: 0.031797, final accuracy: 0.990952

precision recall f1-score support

0 1.00 1.00 1.00 408

1 0.99 1.00 0.99 471

2 1.00 0.98 0.99 420

3 1.00 0.99 0.99 506

4 0.99 0.99 0.99 397

5 0.99 0.98 0.98 339

6 0.99 1.00 0.99 402

7 0.98 0.99 0.99 438

8 0.99 0.99 0.99 403

9 0.99 0.99 0.99 416

avg / total 0.99 0.99 0.99 4200

Total params: 889,322

Trainable params: 888,426

Non-trainable params: 896

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/30

- 555s - loss: 0.2635 - acc: 0.9164 - val\_loss: 0.0935 - val\_acc: 0.9717

Epoch 2/30

- 554s - loss: 0.0964 - acc: 0.9705 - val\_loss: 0.0357 - val\_acc: 0.9886

Epoch 3/30

- 567s - loss: 0.0766 - acc: 0.9763 - val\_loss: 0.0393 - val\_acc: 0.9883

Epoch 4/30

- 565s - loss: 0.0610 - acc: 0.9815 - val\_loss: 0.0319 - val\_acc: 0.9914

Epoch 5/30

- 579s - loss: 0.0556 - acc: 0.9829 - val\_loss: 0.0312 - val\_acc: 0.9907

Epoch 6/30

- 615s - loss: 0.0524 - acc: 0.9844 - val\_loss: 0.0233 - val\_acc: 0.9921

Epoch 7/30

- 555s - loss: 0.0450 - acc: 0.9862 - val\_loss: 0.0256 - val\_acc: 0.9914

Epoch 8/30

- 543s - loss: 0.0469 - acc: 0.9857 - val\_loss: 0.0211 - val\_acc: 0.9924

Epoch 9/30

- 541s - loss: 0.0403 - acc: 0.9874 - val\_loss: 0.0210 - val\_acc: 0.9917

Epoch 10/30

- 541s - loss: 0.0409 - acc: 0.9874 - val\_loss: 0.0231 - val\_acc: 0.9933

Epoch 11/30

- 541s - loss: 0.0389 - acc: 0.9882 - val\_loss: 0.0244 - val\_acc: 0.9929

Epoch 12/30

- 542s - loss: 0.0375 - acc: 0.9888 - val\_loss: 0.0214 - val\_acc: 0.9938

Epoch 13/30

- 541s - loss: 0.0345 - acc: 0.9890 - val\_loss: 0.0265 - val\_acc: 0.9931

Epoch 14/30

- 542s - loss: 0.0371 - acc: 0.9894 - val\_loss: 0.0219 - val\_acc: 0.9940

Epoch 15/30

- 541s - loss: 0.0338 - acc: 0.9895 - val\_loss: 0.0186 - val\_acc: 0.9938

Epoch 16/30

- 540s - loss: 0.0328 - acc: 0.9899 - val\_loss: 0.0202 - val\_acc: 0.9938

Epoch 17/30

- 541s - loss: 0.0330 - acc: 0.9901 - val\_loss: 0.0239 - val\_acc: 0.9938

Epoch 18/30

- 541s - loss: 0.0340 - acc: 0.9902 - val\_loss: 0.0179 - val\_acc: 0.9943

Epoch 19/30

- 542s - loss: 0.0326 - acc: 0.9899 - val\_loss: 0.0162 - val\_acc: 0.9948

Epoch 20/30

- 541s - loss: 0.0304 - acc: 0.9911 - val\_loss: 0.0217 - val\_acc: 0.9926

Epoch 21/30

- 540s - loss: 0.0320 - acc: 0.9904 - val\_loss: 0.0173 - val\_acc: 0.9940

Epoch 22/30

- 543s - loss: 0.0303 - acc: 0.9909 - val\_loss: 0.0175 - val\_acc: 0.9931

Epoch 23/30

Epoch 00023: ReduceLROnPlateau reducing learning rate to 0.0005000000237487257.

- 541s - loss: 0.0275 - acc: 0.9918 - val\_loss: 0.0199 - val\_acc: 0.9938

Epoch 24/30

- 549s - loss: 0.0231 - acc: 0.9931 - val\_loss: 0.0144 - val\_acc: 0.9948

Epoch 25/30

- 542s - loss: 0.0240 - acc: 0.9927 - val\_loss: 0.0152 - val\_acc: 0.9945

Epoch 26/30

- 540s - loss: 0.0218 - acc: 0.9934 - val\_loss: 0.0145 - val\_acc: 0.9955

Epoch 27/30

- 541s - loss: 0.0219 - acc: 0.9932 - val\_loss: 0.0139 - val\_acc: 0.9952

Epoch 28/30

- 541s - loss: 0.0213 - acc: 0.9935 - val\_loss: 0.0168 - val\_acc: 0.9955

Epoch 29/30

- 540s - loss: 0.0230 - acc: 0.9930 - val\_loss: 0.0151 - val\_acc: 0.9945

Epoch 30/30

Epoch 00030: ReduceLROnPlateau reducing learning rate to 0.0002500000118743628.

- 540s - loss: 0.0226 - acc: 0.9931 - val\_loss: 0.0137 - val\_acc: 0.9950

Final loss: 0.013653, final accuracy: 0.995000

precision recall f1-score support

0 1.00 0.99 1.00 408

1 0.99 0.99 0.99 471

2 1.00 1.00 1.00 420

3 1.00 1.00 1.00 506

4 0.99 0.99 0.99 397

5 0.99 0.99 0.99 339

6 0.99 1.00 0.99 402

7 1.00 1.00 1.00 438

8 0.99 0.99 0.99 403

9 1.00 1.00 1.00 416

avg / total 1.00 0.99 0.99 4200

Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_1 (Batch (None, 28, 28, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

activation\_1 (Activation) (None, 28, 28, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 28, 28, 32) 25632

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_2 (Batch (None, 28, 28, 32) 128

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

activation\_2 (Activation) (None, 28, 28, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 14, 14, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_3 (Batch (None, 14, 14, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

activation\_3 (Activation) (None, 14, 14, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_4 (Conv2D) (None, 14, 14, 64) 36928

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_4 (Batch (None, 14, 14, 64) 256

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

activation\_4 (Activation) (None, 14, 14, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 3136) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 256) 803072

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

batch\_normalization\_5 (Batch (None, 256) 1024

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

activation\_5 (Activation) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 2570

=================================================================

Total params: 889,322

Trainable params: 888,426

Non-trainable params: 896

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/30

- 579s - loss: 0.2397 - acc: 0.9253 - val\_loss: 0.0516 - val\_acc: 0.9840

Epoch 2/30

- 563s - loss: 0.0940 - acc: 0.9704 - val\_loss: 0.0551 - val\_acc: 0.9831

Epoch 3/30

- 564s - loss: 0.0740 - acc: 0.9772 - val\_loss: 0.0259 - val\_acc: 0.9912

Epoch 4/30

- 571s - loss: 0.0641 - acc: 0.9802 - val\_loss: 0.0385 - val\_acc: 0.9883

Epoch 5/30

- 564s - loss: 0.0599 - acc: 0.9821 - val\_loss: 0.0285 - val\_acc: 0.9910

Epoch 6/30

- 562s - loss: 0.0566 - acc: 0.9829 - val\_loss: 0.0340 - val\_acc: 0.9907

Epoch 7/30

- 587s - loss: 0.0513 - acc: 0.9843 - val\_loss: 0.0251 - val\_acc: 0.9921

Epoch 8/30

- 563s - loss: 0.0493 - acc: 0.9848 - val\_loss: 0.0310 - val\_acc: 0.9919

Epoch 9/30

- 559s - loss: 0.0488 - acc: 0.9855 - val\_loss: 0.0222 - val\_acc: 0.9933

Epoch 10/30

- 568s - loss: 0.0479 - acc: 0.9858 - val\_loss: 0.0239 - val\_acc: 0.9931

Epoch 11/30

- 564s - loss: 0.0445 - acc: 0.9867 - val\_loss: 0.0314 - val\_acc: 0.9917

Epoch 12/30

- 607s - loss: 0.0453 - acc: 0.9867 - val\_loss: 0.0410 - val\_acc: 0.9919

Epoch 13/30

- 607s - loss: 0.0455 - acc: 0.9865 - val\_loss: 0.0225 - val\_acc: 0.9938

Epoch 14/30

- 599s - loss: 0.0427 - acc: 0.9877 - val\_loss: 0.0214 - val\_acc: 0.9936

Epoch 15/30

- 551s - loss: 0.0398 - acc: 0.9885 - val\_loss: 0.0337 - val\_acc: 0.9914

Epoch 16/30

- 553s - loss: 0.0397 - acc: 0.9880 - val\_loss: 0.0226 - val\_acc: 0.9940

Epoch 17/30

- 565s - loss: 0.0386 - acc: 0.9886 - val\_loss: 0.0223 - val\_acc: 0.9943

Epoch 18/30

- 565s - loss: 0.0384 - acc: 0.9887 - val\_loss: 0.0273 - val\_acc: 0.9921

Epoch 19/30

- 592s - loss: 0.0367 - acc: 0.9894 - val\_loss: 0.0883 - val\_acc: 0.9802

Epoch 20/30

- 605s - loss: 0.0398 - acc: 0.9894 - val\_loss: 0.0183 - val\_acc: 0.9950

Epoch 21/30

- 615s - loss: 0.0394 - acc: 0.9888 - val\_loss: 0.0227 - val\_acc: 0.9943

Epoch 22/30

- 650s - loss: 0.0363 - acc: 0.9895 - val\_loss: 0.0239 - val\_acc: 0.9936

Epoch 23/30

- 598s - loss: 0.0363 - acc: 0.9898 - val\_loss: 0.0208 - val\_acc: 0.9940

Epoch 24/30

Epoch 00024: ReduceLROnPlateau reducing learning rate to 0.0005000000237487257.

- 618s - loss: 0.0383 - acc: 0.9894 - val\_loss: 0.0147 - val\_acc: 0.9948

Epoch 25/30

- 588s - loss: 0.0322 - acc: 0.9902 - val\_loss: 0.0182 - val\_acc: 0.9948

Epoch 26/30

- 557s - loss: 0.0315 - acc: 0.9911 - val\_loss: 0.0165 - val\_acc: 0.9957

Epoch 27/30

- 552s - loss: 0.0316 - acc: 0.9919 - val\_loss: 0.0216 - val\_acc: 0.9943

Epoch 28/30

- 544s - loss: 0.0298 - acc: 0.9917 - val\_loss: 0.0183 - val\_acc: 0.9940

Epoch 29/30

- 543s - loss: 0.0306 - acc: 0.9913 - val\_loss: 0.0177 - val\_acc: 0.9948

Epoch 30/30

Epoch 00030: ReduceLROnPlateau reducing learning rate to 0.0002500000118743628.

- 541s - loss: 0.0278 - acc: 0.9921 - val\_loss: 0.0192 - val\_acc: 0.9940

Final loss: 0.019178, final accuracy: 0.994048

precision recall f1-score support

0 1.00 1.00 1.00 408

1 0.99 1.00 1.00 471

2 1.00 1.00 1.00 420

3 1.00 0.99 1.00 506

4 1.00 0.99 1.00 397

5 0.99 0.99 0.99 339

6 0.99 1.00 1.00 402

7 0.99 0.99 0.99 438

8 0.99 1.00 0.99 403

9 0.99 1.00 0.99 416

avg / total 0.99 0.99 0.99 4200

Using TensorFlow backend.

train shape: (42000, 785)

test shape: (28000, 784)

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Layer (type) Output Shape Param #

=================================================================

conv2d\_1 (Conv2D) (None, 28, 28, 32) 832

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_2 (Conv2D) (None, 28, 28, 32) 25632

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_1 (MaxPooling2 (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_1 (Dropout) (None, 14, 14, 32) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_3 (Conv2D) (None, 14, 14, 64) 18496

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

conv2d\_4 (Conv2D) (None, 14, 14, 64) 36928

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

max\_pooling2d\_2 (MaxPooling2 (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_2 (Dropout) (None, 7, 7, 64) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

flatten\_1 (Flatten) (None, 3136) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_1 (Dense) (None, 256) 803072

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dropout\_3 (Dropout) (None, 256) 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

dense\_2 (Dense) (None, 10) 2570

=================================================================

Total params: 887,530

Trainable params: 887,530

Non-trainable params: 0

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Epoch 1/10

- 328s - loss: 0.4030 - acc: 0.8699 - val\_loss: 0.0670 - val\_acc: 0.9805

Epoch 2/10

- 327s - loss: 0.1383 - acc: 0.9590 - val\_loss: 0.0592 - val\_acc: 0.9843

Epoch 3/10

- 331s - loss: 0.1089 - acc: 0.9673 - val\_loss: 0.0300 - val\_acc: 0.9910

Epoch 4/10

- 344s - loss: 0.0904 - acc: 0.9737 - val\_loss: 0.0350 - val\_acc: 0.9902

Epoch 5/10

- 356s - loss: 0.0805 - acc: 0.9761 - val\_loss: 0.0295 - val\_acc: 0.9914

Epoch 6/10

- 361s - loss: 0.0758 - acc: 0.9774 - val\_loss: 0.0278 - val\_acc: 0.9900

Epoch 7/10

- 334s - loss: 0.0683 - acc: 0.9798 - val\_loss: 0.0219 - val\_acc: 0.9936

Epoch 8/10

- 336s - loss: 0.0630 - acc: 0.9810 - val\_loss: 0.0297 - val\_acc: 0.9921

Epoch 9/10

- 334s - loss: 0.0609 - acc: 0.9818 - val\_loss: 0.0258 - val\_acc: 0.9919

Epoch 10/10

- 352s - loss: 0.0587 - acc: 0.9830 - val\_loss: 0.0231 - val\_acc: 0.9933

Final loss: 0.023150, final accuracy: 0.993333