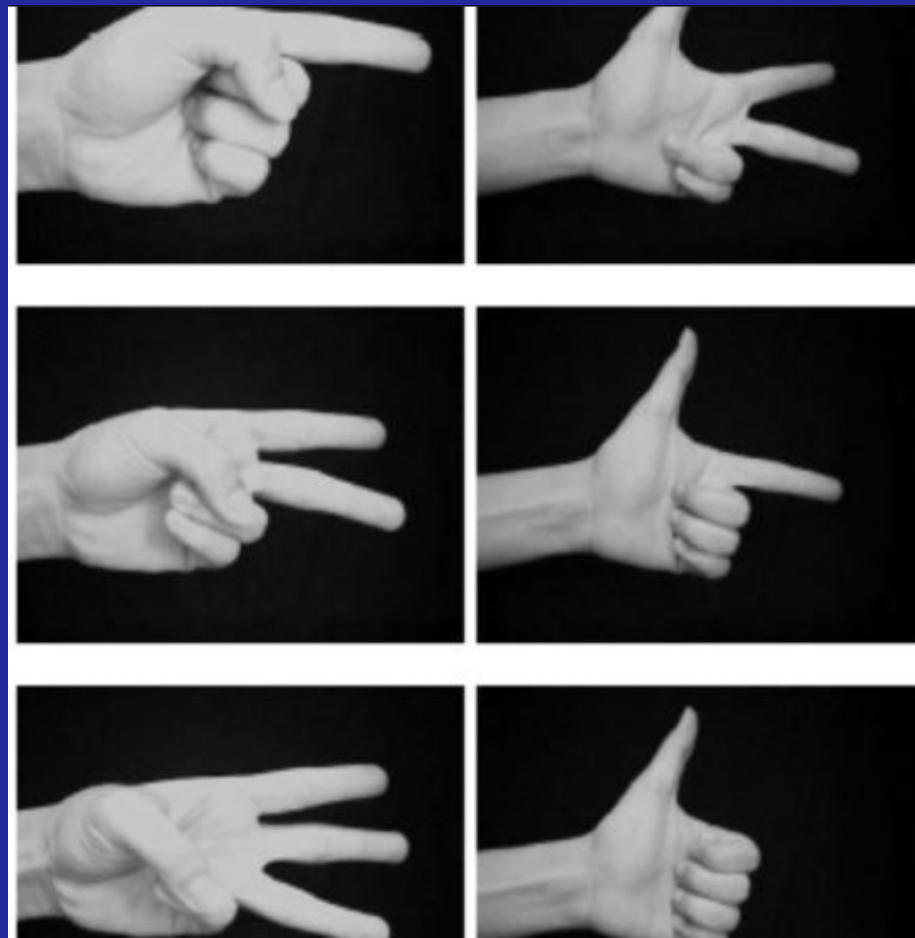**Multimedia Application**

Term Project

The Approaching
to Hand Gesture Recognition, Vehicle Color Recognition,
and Handwritten digit recognition

Done by: Ulugbek Shernazarov - U1910253
Professor: Uddin Ahmed Minhaz

Approaching to

Hand Gesture Recognition (HGR)

Vehicle Color Recognition (VCR)

Handwritten Digit Recognition (HDR)

Hand Gesture Recognition (HGR):

Dataset

Acquired by Leap Motion

10 different hand-gestures
by 5 men and 5 women

- /00 (subject with identifier 00)
  - /01_palm (images for palm gest
    - /01palm/frame197957r.png,
      gesture performed by the s
  - /02_l (images for l gesture of su
  - /10_down
- /01
- /02
- /09 (last subject with identifier 09)

Preprocessing

```
# Normalizing the data
input_data = np.array(input_data)
label = np.array(label)
input_data = input_data/255.0
input_data.shape

(20000, 50, 50)
```

```
X=(X-X.min())/(X.max()-X.min())
```

```
label = keras.utils.to_categorical(label, num_classes=10,dtype='i1')
```
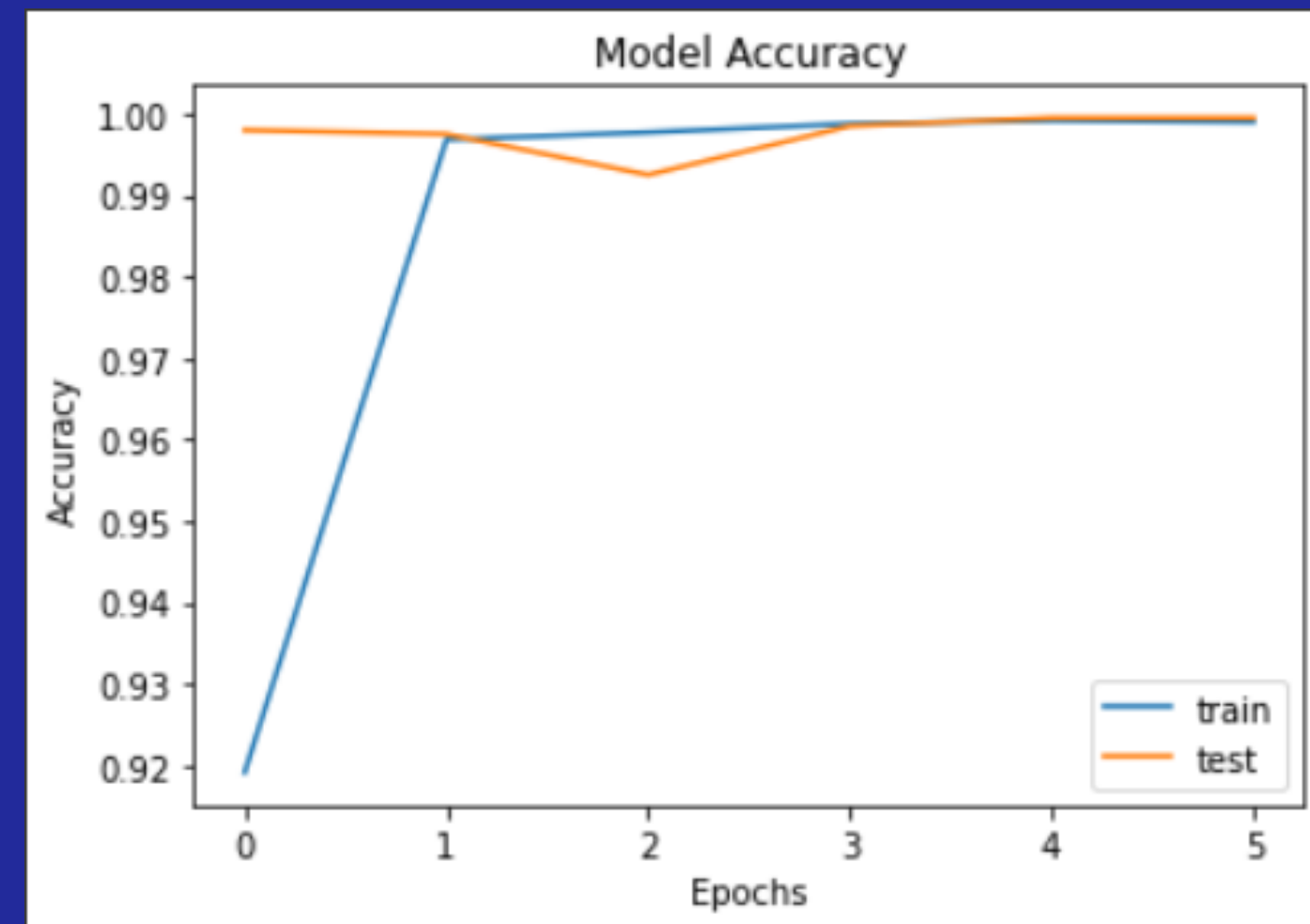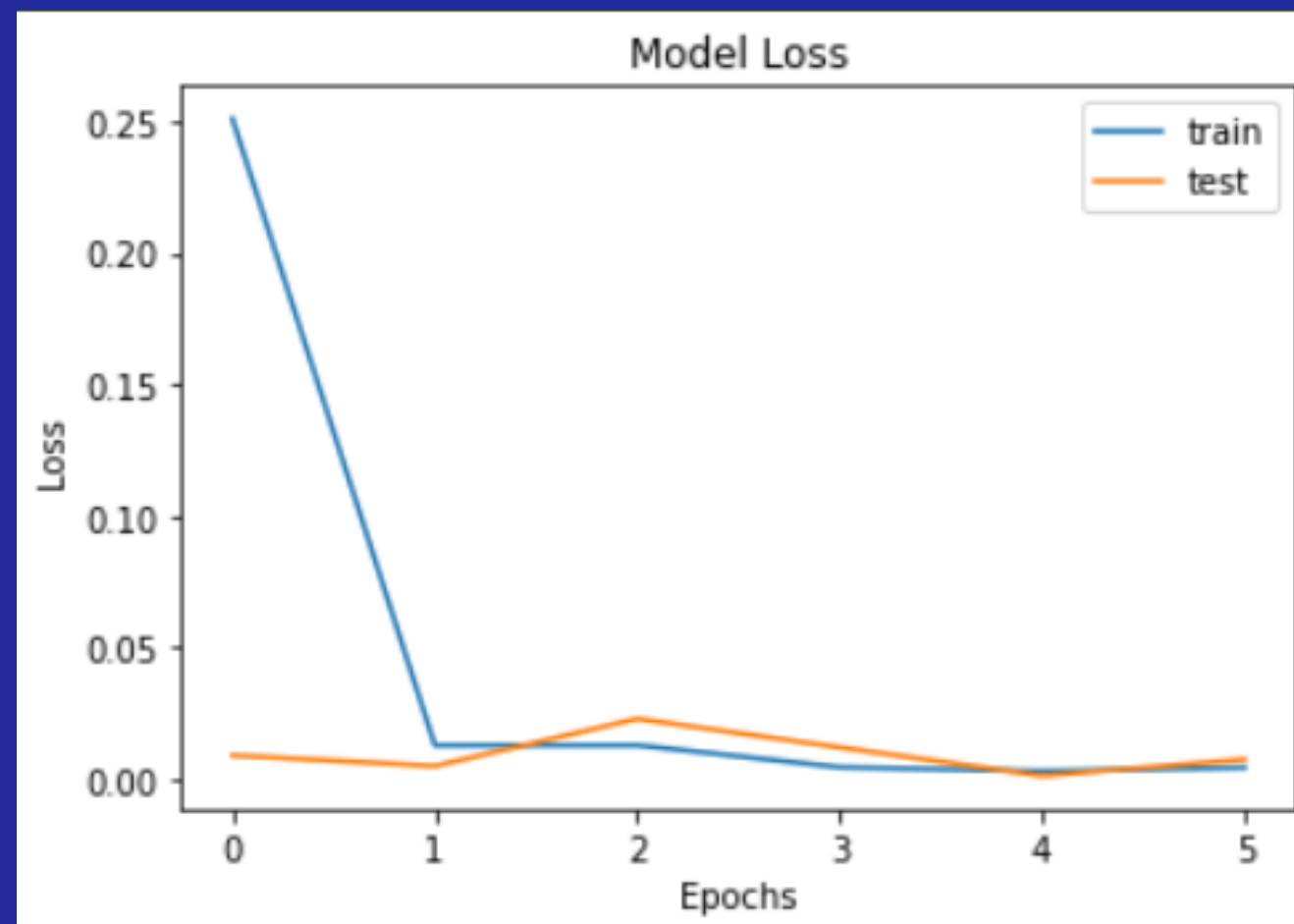
Network

conv2D - 3 layers, activation = 'relu' (between layers) and 'softmax' - output Maxpool2D - 2, Dropout - 2 (0.3), Dense - 2 layers

Results

```
network.fit(X_train, y_train, epochs = 6, batch_size=32, validation_data=(X_test, y_test))
```

Test accuracy - 99.95%

```
63/63 [==============================]
Test accuracy: 99.95%
```

Vehicle Color Recognition (VCR):

Implementation of   ## Network

Reza Fuad Rachmadi* and I Ketut Eddy Purnama[†]
Department of Multimedia and Networking Engineering
Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia 60111
Email: *fuad@its.ac.id, [†]ketut@te.its.ac.id

The CNN architecture used in our system consist 8 layers with 2 base networks with total 16 layers. First two layers and fifth layer does normalization and pooling after convolution process. The third and fourth layer does only convolution process. Before feed up to fully-connected layers, the networks do channel concatenation process. Sample of input image is taken from Chen [2] dataset
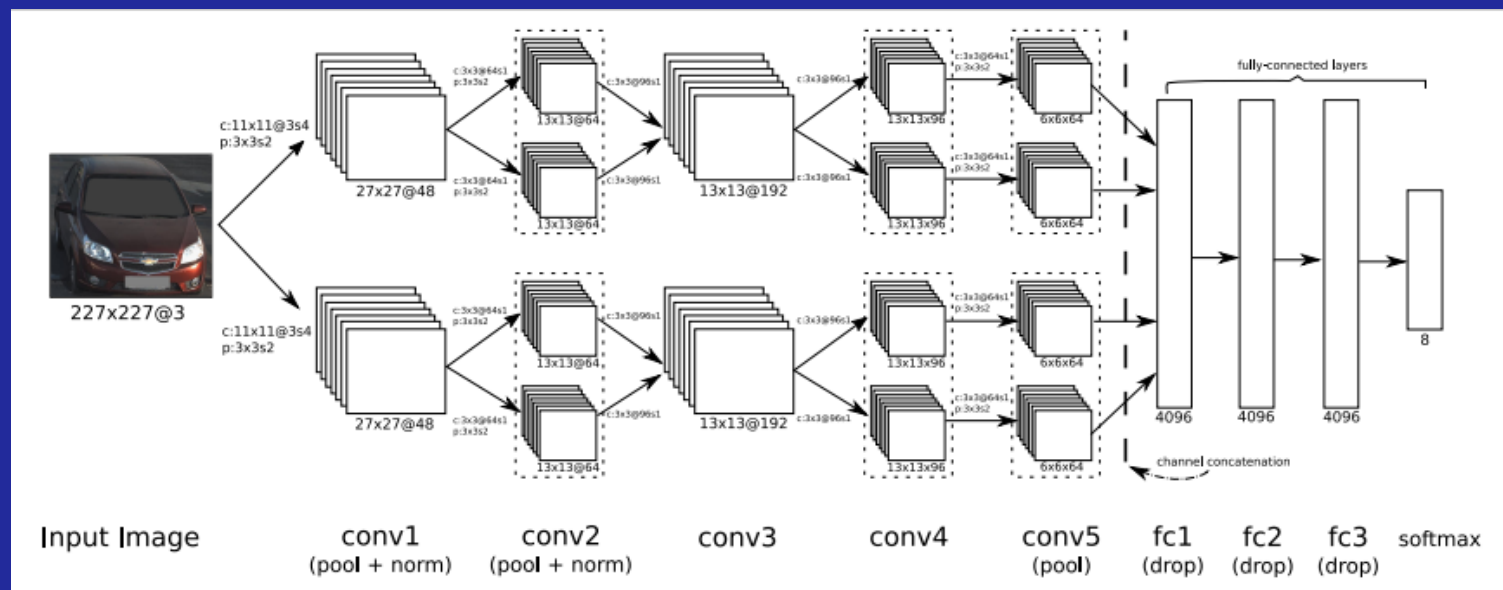
Results

```
Epoch 15: val_accuracy did not improve from 0.15815
190/190 [==============================] - 451s 2s/step - loss: 3.3017 - accuracy: 0.1542 - val_loss: 3.3459 - val_accuracy: 0.1410
```

Test accuracy - 15%
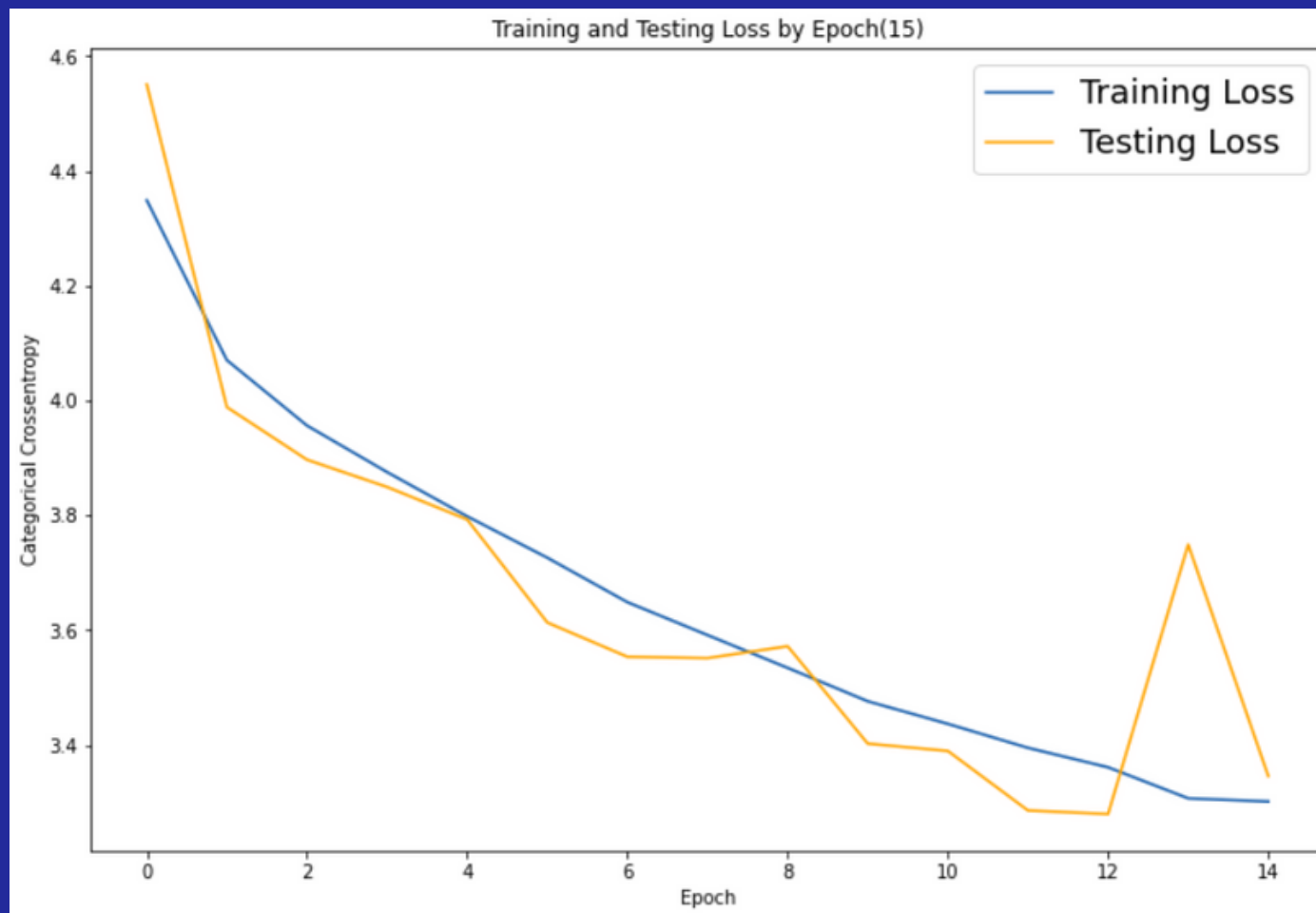
Handwritten Digit Recognition (HDR):

Dataset

Network

```python
from keras.datasets import mnist
(train_images, train_labels), (test_images, test_labels) = mnist.load_data()
```

```python
model = models.Sequential()
model.add(layers.Conv2D(32, (3, 3), activation='relu',
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
model.add(layers.MaxPooling2D((2, 2)))
model.add(layers.Conv2D(64, (3, 3), activation='relu'))
```

MNIST is a large database of small, square 28x28 pixel grayscale images of handwritten single digits between 0 and 9. It consists of a total of 70,000 handwritten images of digits, with the training set having 60,000 images and the test set having 10,000. All images are labeled with the respective digit that they represent. There are a total of 10 classes of digits (from 0 to 9).
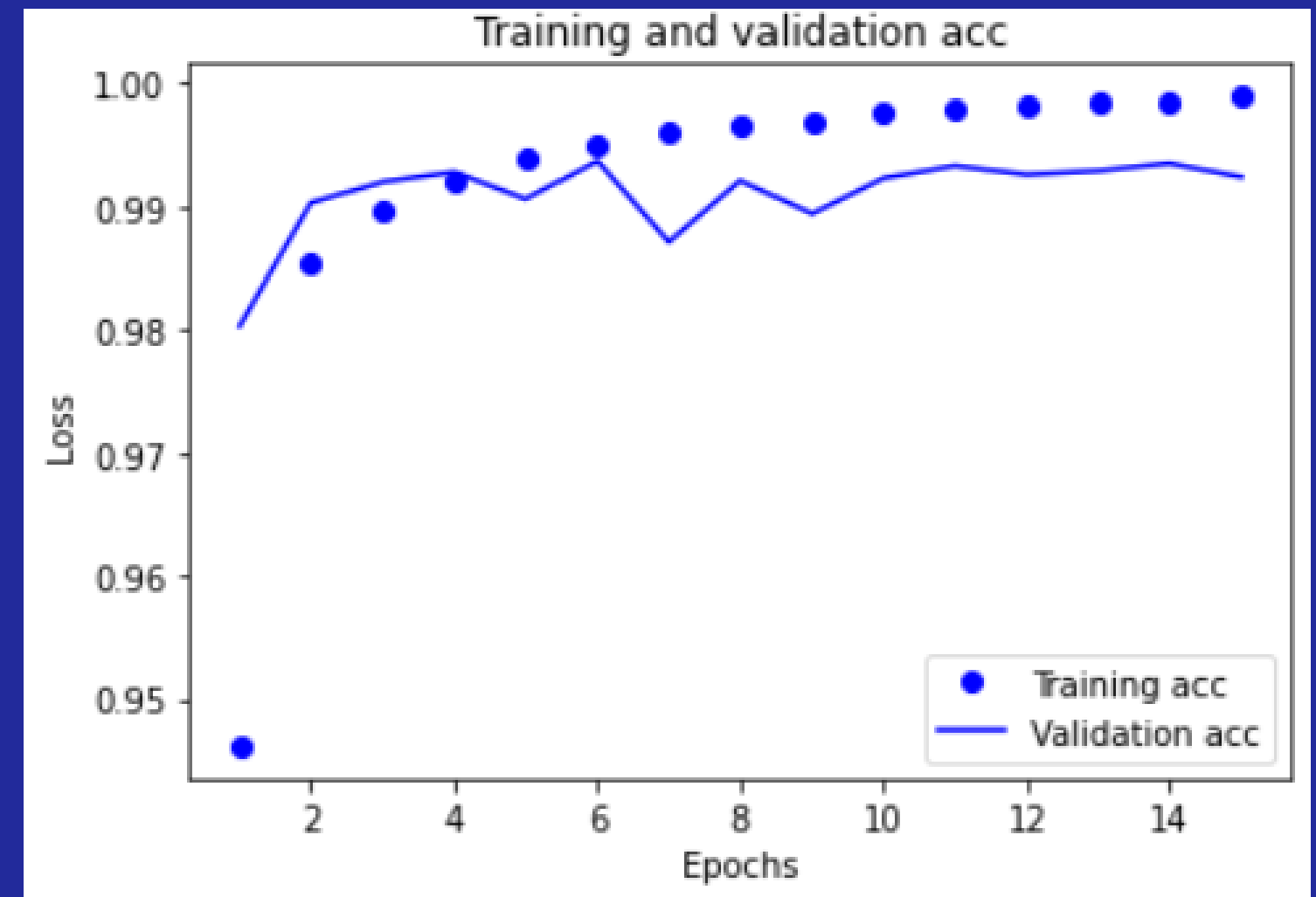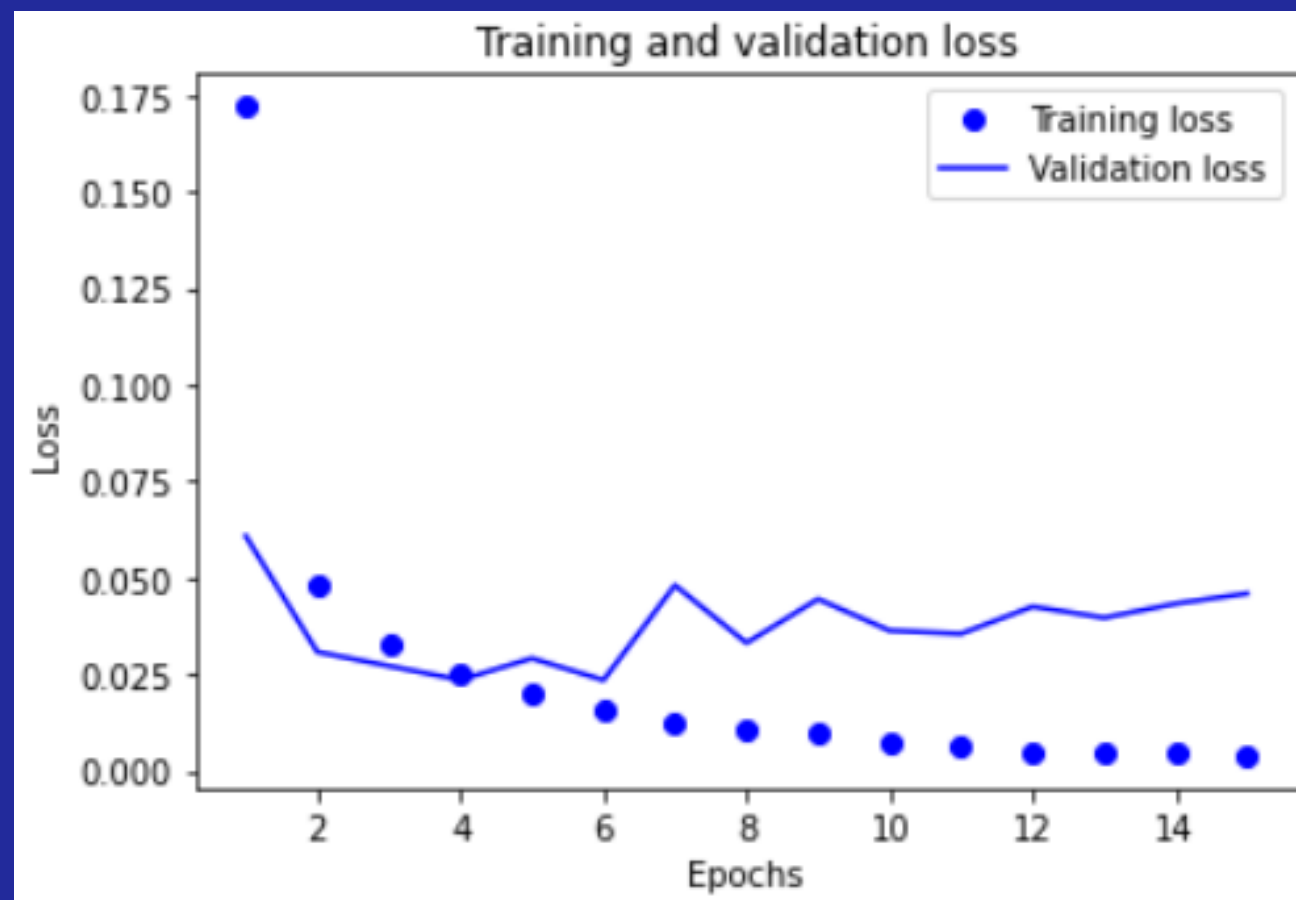
Results

```
313/313 [==============================] - 4s 12ms/step - loss: 0.0459 - accuracy: 0.9924
[0.04593425989151001, 0.9923999905586243]
```

Test accuracy - 99.24%

Thank You