HW 6
Machine Learning
Image regression - CNN

HW 6 Problem Statement (provided by Dr. Han Hu, MEEG 491V/591V-028, Fall 2021)
Problem 6-1 Image Regression:

The vapor fraction (second) column of the data file "DS-1_36W_vapor_fraction.txt" under /ocean/projects/mch210006p/shared/HW5 are the labels of the images under the folder "DS-1_36W_images." Train a convolutional neural networks (CNN) model to predict the vapor fraction of the images and compare the model prediction against the true data.

Code Running Instructions
The environment is default.

Coding explanation
(Data loading) Firstly, images need to be matched with vapor factions. Then, construct a CNN structure. Since this is a regression task, the last dense is 1. Trian a CNN model, X dataset is the images, y is the vapor faction. The CNN structure is as below:

Model: "sequential"

Layer (type)	*Output Shape*	*Param #*
conv2d (Conv2D)	*(None, 238, 238, 32)*	*320*
max_pooling2d (MaxPooling2D)	*(None, 119, 119, 32)*	*0*
conv2d_1 (Conv2D)	*(None, 117, 117, 64)*	*18496*
max_pooling2d_1 (MaxPooling2	*(None, 58, 58, 64)*	*0*
dropout (Dropout)	*(None, 58, 58, 64)*	*0*
flatten (Flatten)	*(None, 215296)*	*0*
dense (Dense)	*(None, 256)*	*55116032*
dropout_1 (Dropout)	*(None, 256)*	*0*
dense_1 (Dense)	*(None, 1)*	*257*

Total params: 55,135,105
Trainable params: 55,135,105
Non-trainable params: 0

(evaluation) The result for the image regression was evaluated by the validation loss and validation mean square error.

Results

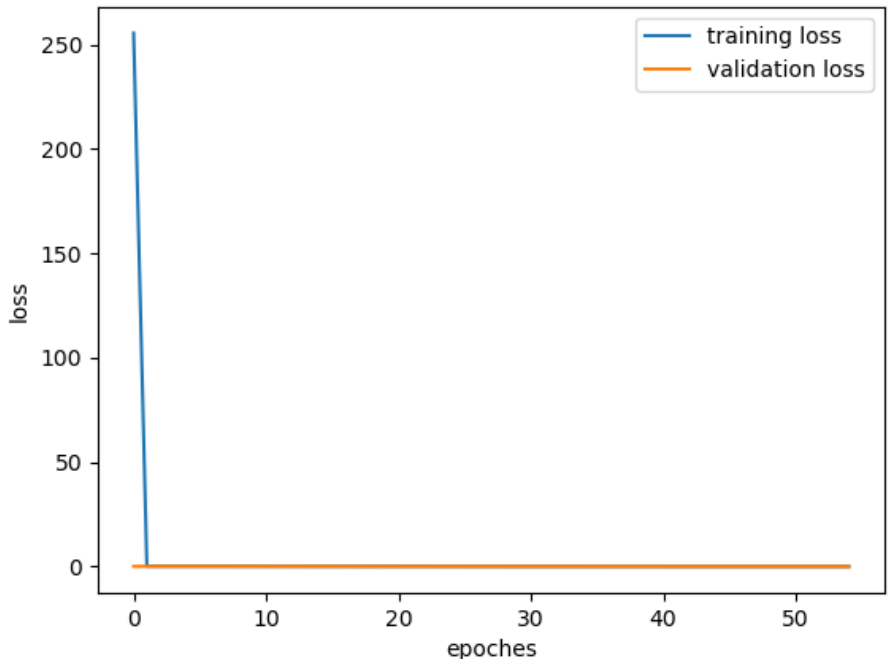

Fig 1. Training loss and validation loss for 100 epochs.

Figure 1 is the graph for the training loss and validation loss for 100 epochs. Since the early stopping criteria is used, it stopped at epoch 55.

The result of each epoch is shown as follows:

Train on 3599 samples, validate on 900 samples

Epoch 1/100

3599/3599 - 5s - loss: 255.4924 - mse: 255.4924 - val_loss: 0.0781 - val_mse: 0.0781

Epoch 2/100

3599/3599 - 3s - loss: 0.1621 - mse: 0.1621 - val_loss: 0.0732 - val_mse: 0.0732

Epoch 3/100

3599/3599 - 3s - loss: 0.1489 - mse: 0.1489 - val_loss: 0.0706 - val_mse: 0.0706

Epoch 4/100

3599/3599 - 3s - loss: 0.1414 - mse: 0.1414 - val_loss: 0.0701 - val_mse: 0.0701

Epoch 5/100

3599/3599 - 3s - loss: 0.1392 - mse: 0.1392 - val_loss: 0.0685 - val_mse: 0.0685

Epoch 6/100

3599/3599 - 3s - loss: 0.1394 - mse: 0.1394 - val_loss: 0.0670 - val_mse: 0.0670

Epoch 7/100

3599/3599 - 3s - loss: 0.1343 - mse: 0.1343 - val_loss: 0.0658 - val_mse: 0.0658

Epoch 8/100

3599/3599 - 3s - loss: 0.1309 - mse: 0.1309 - val_loss: 0.0639 - val_mse: 0.0639

Epoch 9/100

3599/3599 - 3s - loss: 0.1265 - mse: 0.1265 - val_loss: 0.0623 - val_mse: 0.0623

Epoch 10/100

3599/3599 - 3s - loss: 0.1241 - mse: 0.1241 - val_loss: 0.0602 - val_mse: 0.0602

Epoch 11/100

3599/3599 - 3s - loss: 0.1201 - mse: 0.1201 - val_loss: 0.0582 - val_mse: 0.0582

Epoch 12/100

3599/3599 - 3s - loss: 0.1144 - mse: 0.1144 - val_loss: 0.0568 - val_mse: 0.0568

Epoch 13/100

3599/3599 - 3s - loss: 0.1121 - mse: 0.1121 - val_loss: 0.0546 - val_mse: 0.0546

Epoch 14/100

3599/3599 - 3s - loss: 0.1071 - mse: 0.1071 - val_loss: 0.0525 - val_mse: 0.0525

Epoch 15/100

3599/3599 - 3s - loss: 0.1038 - mse: 0.1038 - val_loss: 0.0497 - val_mse: 0.0497

Epoch 16/100

3599/3599 - 3s - loss: 0.0969 - mse: 0.0969 - val_loss: 0.0478 - val_mse: 0.0478

Epoch 17/100

3599/3599 - 3s - loss: 0.0943 - mse: 0.0943 - val_loss: 0.0457 - val_mse: 0.0457

Epoch 18/100

3599/3599 - 3s - loss: 0.0905 - mse: 0.0905 - val_loss: 0.0429 - val_mse: 0.0429

Epoch 19/100

3599/3599 - 3s - loss: 0.0847 - mse: 0.0847 - val_loss: 0.0410 - val_mse: 0.0410

Epoch 20/100

3599/3599 - 3s - loss: 0.0769 - mse: 0.0769 - val_loss: 0.0384 - val_mse: 0.0384

Epoch 21/100

3599/3599 - 3s - loss: 0.0751 - mse: 0.0751 - val_loss: 0.0366 - val_mse: 0.0366

Epoch 22/100

3599/3599 - 3s - loss: 0.0724 - mse: 0.0724 - val_loss: 0.0336 - val_mse: 0.0336

Epoch 23/100

3599/3599 - 3s - loss: 0.0645 - mse: 0.0645 - val_loss: 0.0315 - val_mse: 0.0315

Epoch 24/100

3599/3599 - 3s - loss: 0.0613 - mse: 0.0613 - val_loss: 0.0294 - val_mse: 0.0294

Epoch 25/100

3599/3599 - 3s - loss: 0.0553 - mse: 0.0553 - val_loss: 0.0274 - val_mse: 0.0274

Epoch 26/100

3599/3599 - 3s - loss: 0.0520 - mse: 0.0520 - val_loss: 0.0251 - val_mse: 0.0251

Epoch 27/100

3599/3599 - 3s - loss: 0.0485 - mse: 0.0485 - val_loss: 0.0229 - val_mse: 0.0229

Epoch 28/100

3599/3599 - 3s - loss: 0.0438 - mse: 0.0438 - val_loss: 0.0209 - val_mse: 0.0209

Epoch 29/100

3599/3599 - 3s - loss: 0.0394 - mse: 0.0394 - val_loss: 0.0190 - val_mse: 0.0190

Epoch 30/100

3599/3599 - 3s - loss: 0.0361 - mse: 0.0361 - val_loss: 0.0169 - val_mse: 0.0169

Epoch 31/100

3599/3599 - 3s - loss: 0.0330 - mse: 0.0330 - val_loss: 0.0152 - val_mse: 0.0152

Epoch 32/100

3599/3599 - 3s - loss: 0.0287 - mse: 0.0287 - val_loss: 0.0135 - val_mse: 0.0135

Epoch 33/100

3599/3599 - 3s - loss: 0.0257 - mse: 0.0257 - val_loss: 0.0118 - val_mse: 0.0118

Epoch 34/100

3599/3599 - 3s - loss: 0.0230 - mse: 0.0230 - val_loss: 0.0106 - val_mse: 0.0106

Epoch 35/100

3599/3599 - 3s - loss: 0.0196 - mse: 0.0196 - val_loss: 0.0092 - val_mse: 0.0092

Epoch 36/100

3599/3599 - 3s - loss: 0.0173 - mse: 0.0173 - val_loss: 0.0080 - val_mse: 0.0080

Epoch 37/100

3599/3599 - 3s - loss: 0.0148 - mse: 0.0148 - val_loss: 0.0070 - val_mse: 0.0070

Epoch 38/100

3599/3599 - 3s - loss: 0.0125 - mse: 0.0125 - val_loss: 0.0058 - val_mse: 0.0058

Epoch 39/100

3599/3599 - 3s - loss: 0.0113 - mse: 0.0113 - val_loss: 0.0050 - val_mse: 0.0050

Epoch 40/100

3599/3599 - 3s - loss: 0.0091 - mse: 0.0091 - val_loss: 0.0044 - val_mse: 0.0044

Epoch 41/100

3599/3599 - 3s - loss: 0.0079 - mse: 0.0079 - val_loss: 0.0037 - val_mse: 0.0037

Epoch 42/100

3599/3599 - 3s - loss: 0.0065 - mse: 0.0065 - val_loss: 0.0031 - val_mse: 0.0031

Epoch 43/100

3599/3599 - 3s - loss: 0.0057 - mse: 0.0057 - val_loss: 0.0027 - val_mse: 0.0027

Epoch 44/100

3599/3599 - 3s - loss: 0.0047 - mse: 0.0047 - val_loss: 0.0023 - val_mse: 0.0023

Epoch 45/100

3599/3599 - 3s - loss: 0.0040 - mse: 0.0040 - val_loss: 0.0020 - val_mse: 0.0020

Epoch 46/100

3599/3599 - 3s - loss: 0.0034 - mse: 0.0034 - val_loss: 0.0017 - val_mse: 0.0017

Epoch 47/100

3599/3599 - 3s - loss: 0.0032 - mse: 0.0032 - val_loss: 0.0016 - val_mse: 0.0016

Epoch 48/100

3599/3599 - 3s - loss: 0.0027 - mse: 0.0027 - val_loss: 0.0015 - val_mse: 0.0015

Epoch 49/100

3599/3599 - 3s - loss: 0.0025 - mse: 0.0025 - val_loss: 0.0014 - val_mse: 0.0014

Epoch 50/100

3599/3599 - 3s - loss: 0.0024 - mse: 0.0024 - val_loss: 0.0014 - val_mse: 0.0014

Epoch 51/100

3599/3599 - 3s - loss: 0.0021 - mse: 0.0021 - val_loss: 0.0014 - val_mse: 0.0014

Epoch 52/100

3599/3599 - 3s - loss: 0.0020 - mse: 0.0020 - val_loss: 0.0014 - val_mse: 0.0014

Epoch 53/100

3599/3599 - 3s - loss: 0.0019 - mse: 0.0019 - val_loss: 0.0014 - val_mse: 0.0014

Epoch 54/100

3599/3599 - 3s - loss: 0.0019 - mse: 0.0019 - val_loss: 0.0014 - val_mse: 0.0014

Epoch 55/100

Restoring model weights from the end of the best epoch.

3599/3599 - 3s - loss: 0.0020 - mse: 0.0020 - val_loss: 0.0014 - val_mse: 0.0014

Epoch 00055: early stopping

test mean square error: 0.0013

The validation loss is 0.0014, and validation mean square error is 0.0014.

Challenges

The first challenge is that when I used “glob”, it will sort the data automatically. I got a not good result with validation loss 11.9630. Then I checked the data and noticed the matching is wrong. After I corrected it, matching images with vapor factions, the validation loss gets smaller.

The second challenge is that I didn’t do data normalization at first, then the training model is overfitting, which the validation loss is comparably much bigger than the training loss. Then I normalized the datasets and change the optimizer a little bit smaller than before.

References

[1] https://www.kaggle.com/guidosalimbeni/regression-with-convolutional-neural-network-keras.

[2] https://stackoverflow.com/questions/51853454/image-regression-with-cnn.