Hand gesture model report

This hand gesture model did not apply on the product. The following will include, expected input and output, data acquisition, data preprocessing, modeling and finding.

Expected input and output

|  |  |  |
| --- | --- | --- |
|  | Class Palm | Class fist |
| Input |  |  |
| Output | Palm [1] | Fist [0] |

Data acquisition

1. Custom data  
   3 of the groupmates namely KaOn, Ali, KaHo have created about ~~3000~~ 2000fist and palm gesture images
2. Online data

* figshare dataset includes 4 categories of hand gestures: open hand, fist, right hand, left hand + one negative class [1]
* Hand Images Databases have 3000 palm images [2]
* rock-paper-scissors datasets [3]
* manually download some image data form image provider such as google image, unsplash, pixabay.

1. possible dataset (**did not use in this model training**)

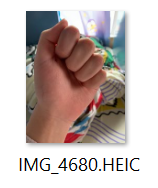
* <https://www.kaggle.com/gti-upm/leapgestrecog/version/1>
* <https://www.gti.ssr.upm.es/data/HandGesture_database.html>
* <https://www.kaggle.com/datamunge/sign-language-mnist>
* <https://www.kaggle.com/ardamavi/sign-language-digits-dataset>
* <https://www-prima.inrialpes.fr/FGnet/data/10-Gesture/gestures/main.html>
* <https://sites.google.com/view/11khands>

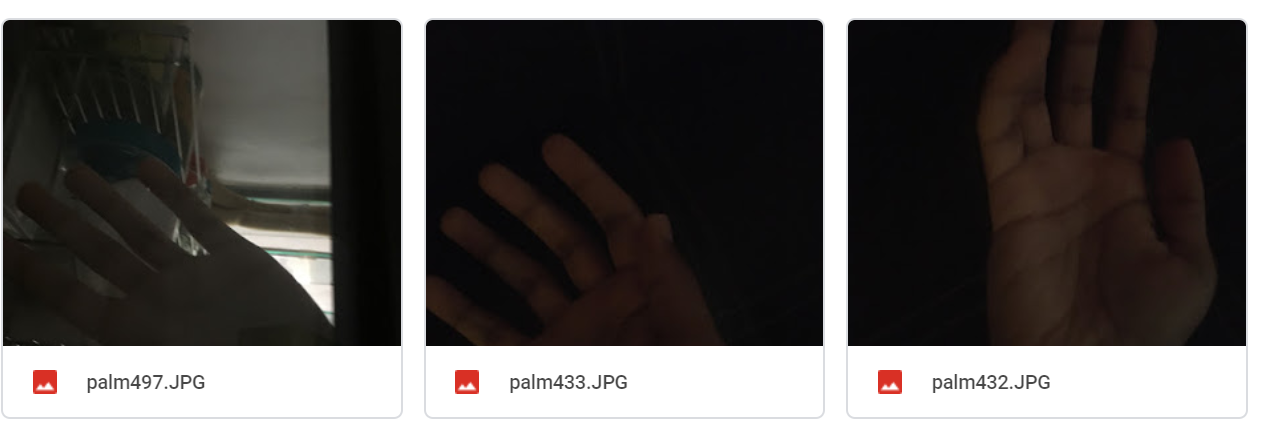
Above possible dataset are not suitable because of the following factor.

* + The data were created using motion sensor
  + Images contain the whole human body instead of just hand

Some of the above possible dataset have useable palm image, however the background is so clean that without noise and also we have collected enough data to make the whole dataset balance, increasing more palm image may case more data imbalance problem so it is not ideal.

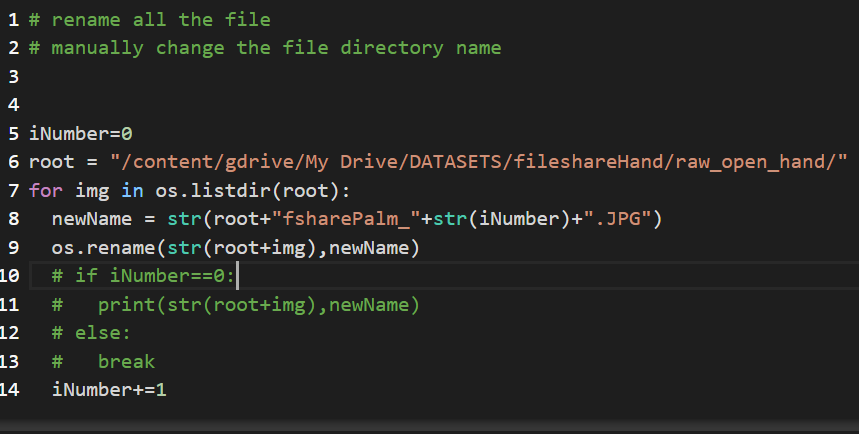
Data cleaning

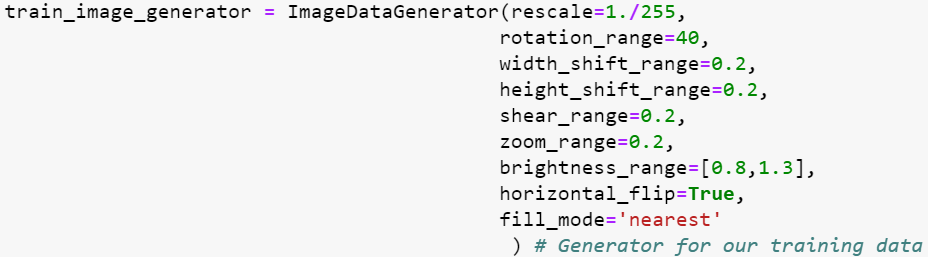


1. Remove .HEIC image file that iphone uses because opencv has trouble processing it. There was a small amount of .HEIC file, only about 15 images   
   in custom made image dataset.
2. Remove dark image which also contain about 10 images in custom made image dataset.  
   

Data preprocessing

1. The train test split process was carried out. First, we shuffle the data, then multiplying total number by 0.8 in order to split the data into training set.  
     
   After that ,use os.rename() function to move training set to training file directory.
2. Rename some of the image for clarification.   
     
   For example, In second round of creating custom data I rename the data that I created as palmv2\_01jpg, fistv2\_01.jpg.

kaho’s second round image is named as KHpalmv2\_01.jpg, KHfistv2\_01.jpg.  
  
data from the fileshare is named fshareFist\_01.jpg, fsharePalm\_01.jpg  


1. Implementing image ~~argumentation~~ augmentation  
     
   code of implementing data augmentation  
   Data augmentation provide more new image especially rotated, shifted, flipped etc for training the model. Also, it is able to avoid overfitting.

Modeling

**Version name: training\_8\_LeakyRelu**

Used equal size training, validation, and test dataset

**Architecture** :

model= Sequential([

Conv2D(32, 3, padding='same', activation=LeakyReLU(alpha=0.01), input\_shape=(IMG\_HEIGHT, IMG\_WIDTH ,3)),

#BatchNormalization(),

MaxPooling2D(pool\_size=(2, 2)),

Conv2D(64, 3, padding='same', activation=LeakyReLU(alpha=0.01)),

#BatchNormalization(),

MaxPooling2D(pool\_size=(2, 2)),

#Dropout(rate=0.2),

Conv2D(128, 3, padding='same', activation=LeakyReLU(alpha=0.01)),

#BatchNormalization(),

MaxPooling2D(pool\_size=(2, 2)),

#Dropout(rate=0.2),

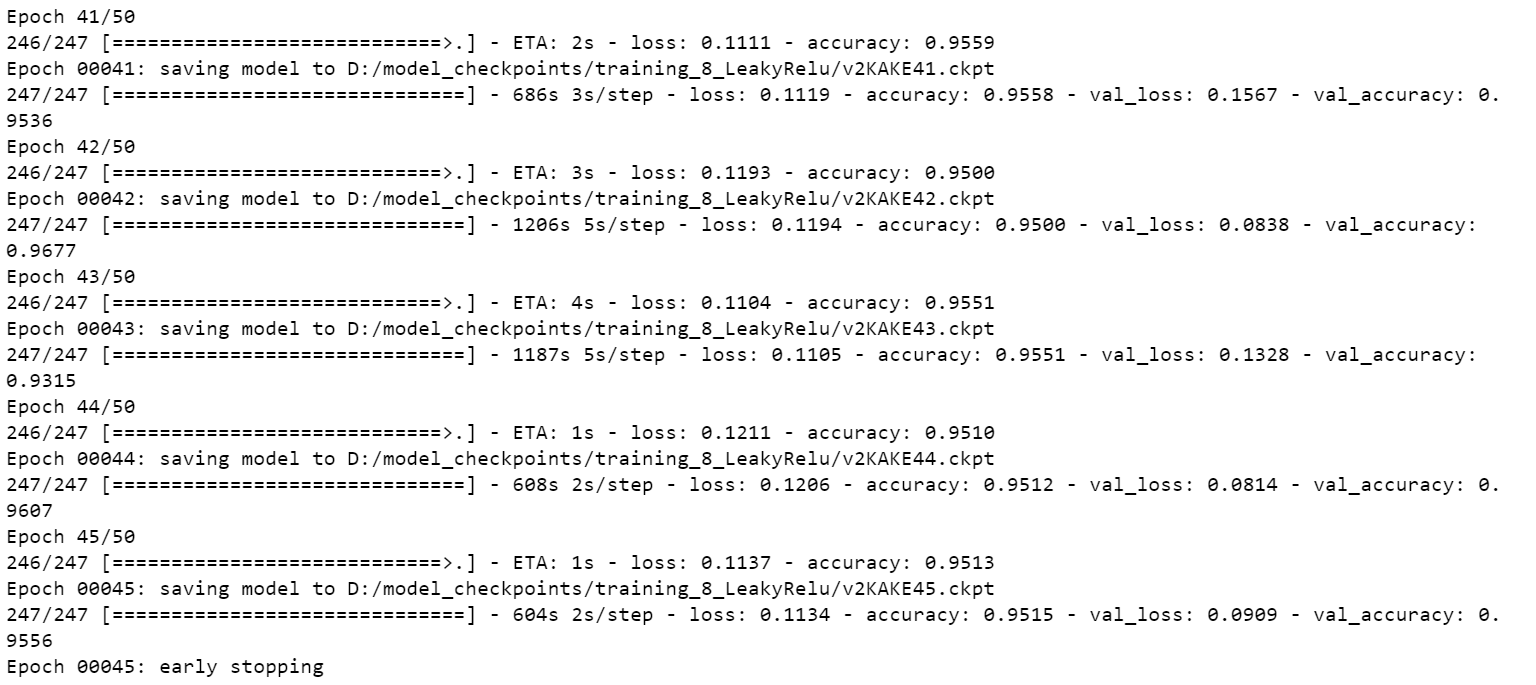
Flatten(),

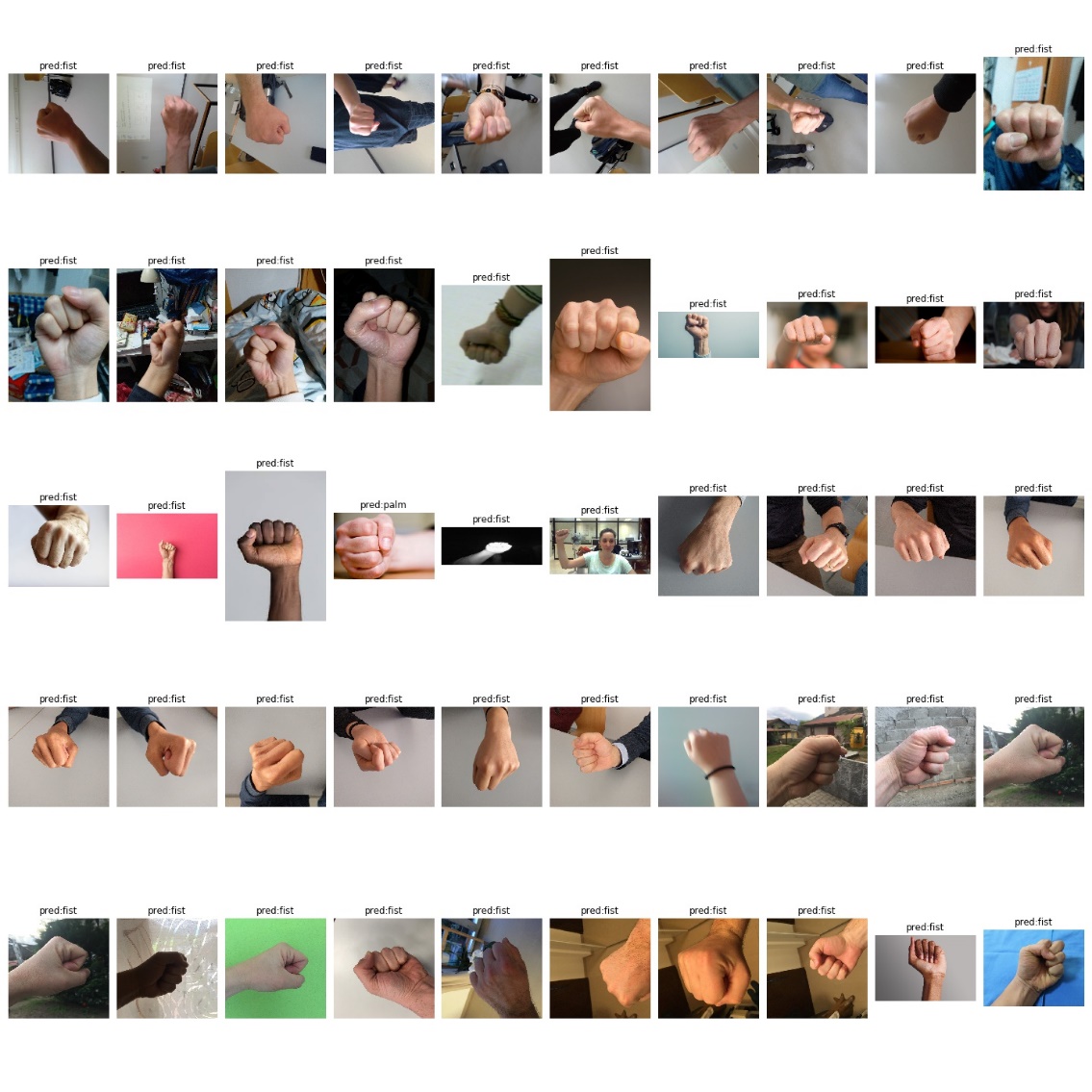
Dropout(rate=0.2),

Dense(512, activation=LeakyReLU(alpha=0.01)),

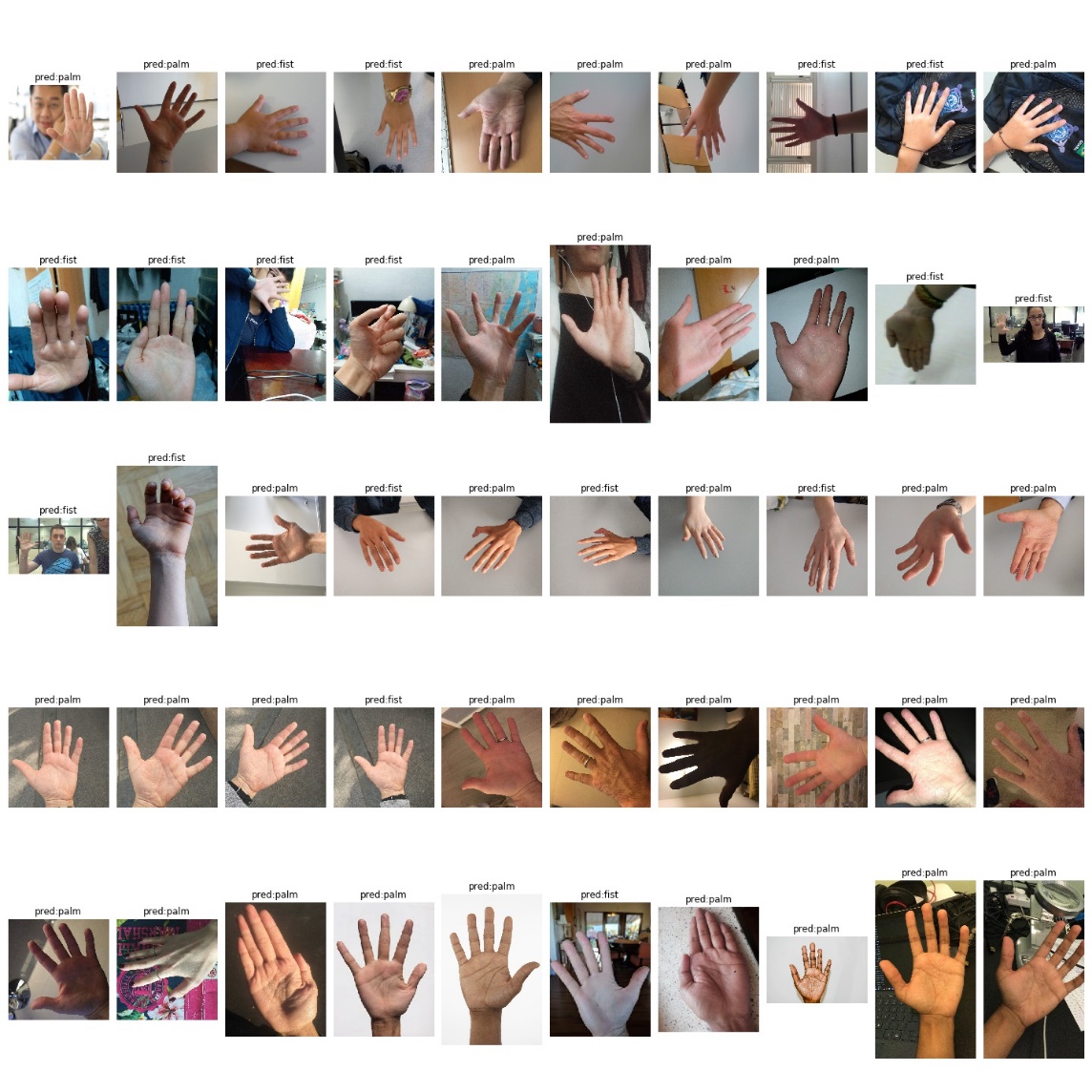
Dense(1)

])

Result  


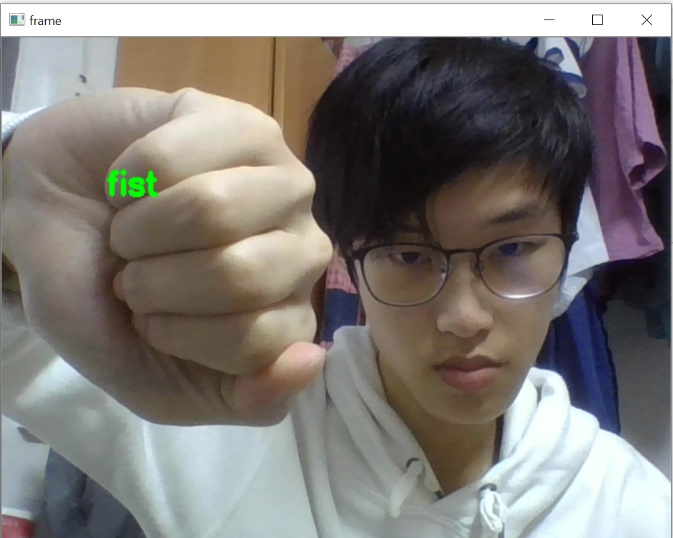
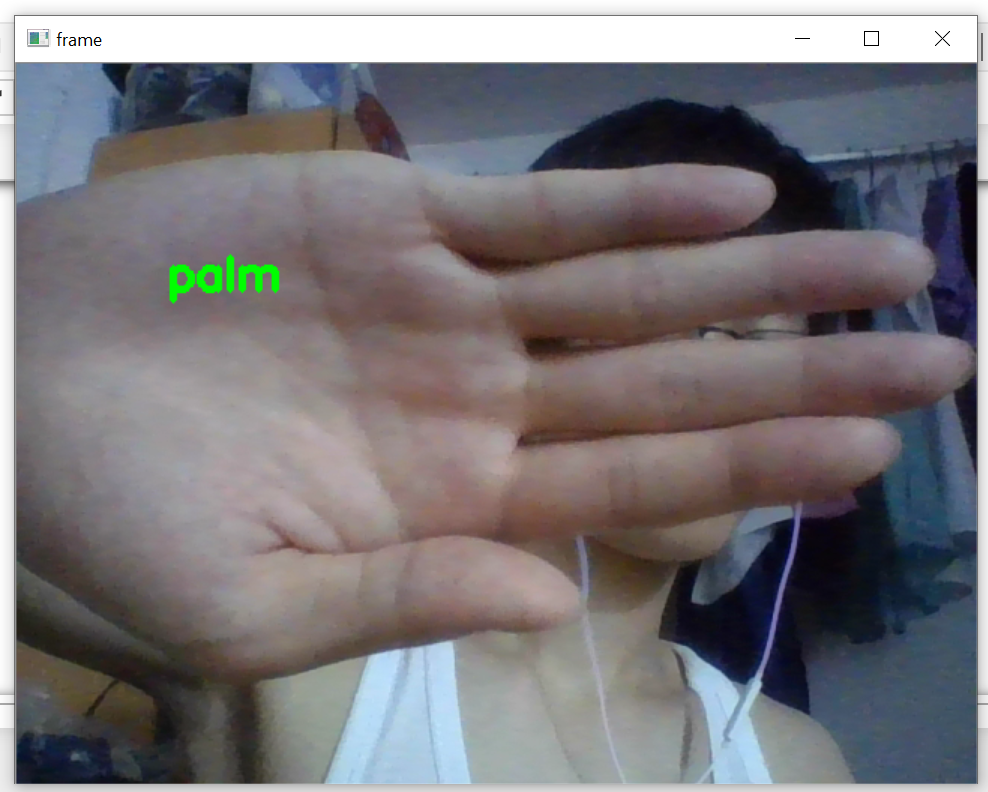
The model achieved 0.95 validation accuracy in epoch 45. Further investigation has conducted in the following to know how well the model is.

Test set fist image model fit result

Test set palm image model fit result

As shown above, the fist class achieved 49/50 correct while palm class achieved 33/50 correct. This imbalance result may cause by overfit to class 0 (fist), since no measurement namely dropout layer, batchnormalize layer, l1 & l2 regularization were implemented except data augmentation.   
  
  
  
  


Validation set palm image model fit result

Predicted 20 as fist out of 50, 30 as palm out of 50 which is 60% correct here, it is hard to see any patterns among those incorrect labeled images.  
  


Perform prediction real time using opencv video capture. Please refer to attachment.

The model is able to distinguish roughly palm and fist because the image of my hand are in the training dataset. Despite that, the model cannot recognize palm if it is an image that has low similarity. Test1edit.mp4 have shown this issue.   
  
Finding

The above tests show that the model is not reliable for the real-world production. Even though the model is capable of distinguishing it is a palm of not, but it is not accurate. After many times of trial-and-error, we can conclude that the model need to apply regularization or adding drop out layer to address the overfitting problem because simply data augmentation is not enough. In addition, adding more noise into the palm class data as it has a cleaner background compare to fist class data which may enhance model generalization. In other words, replace palm images that has similar or pure background among the whole palm training set.

Reference

[1] Magaud Alexandre, D. (2019): Hand gestures raw images. Zenodo. Dataset. <https://doi.org/10.5281/zenodo.3271625>  
Available: <https://figshare.com/articles/Hand_gestures_raw_images/11394585/1>

[2] Ahmad B.A. Hassanat .(2020): Hand Images Databases. Available: <https://xwww.mutah.edu.jo/biometrix/hand-images-databases.html>

[3] A. Giusti, D. Huber, L. M. Gambardella .(2018):” Introducing Machine Learning Concepts by Training a Neural Network to Recognize Hand Gestures" Available: <https://github.com/alessandro-giusti/rock-paper-scissors>