# Original Dataset Sample Screenshot

A picture containing text, electronics, indoor, screenshot

Description automatically generated

# First Problem: High Variance and Low Bias, Model is overfitting.

In early trials training the model, the model seemed to be overfitting.

|  |  |
| --- | --- |
| Set | Accuracy |
| Training | 99% |
| Validation | 60% |

## Successful Solution Trial

* 1. Feed the model more data.
  2. Train a bigger model by adding a 100-unit fully connected layer before the SoftMax layer.

## Result

Model performs much better on validation set. However, still not satisfactory, validation accuracy is now 85%.

# Second Problem: Model misclassifies side when face is not centered (L/R) But no False Positives.

A picture containing text, screenshot, screen, computer

Description automatically generated

Figure 1 Labelled left (appears right because image is mirrored), misclassified as Right because face is shifted to the Right with respect to the frame.

A picture containing text, screenshot, screen, computer

Description automatically generated

Figure 2 Labelled right (appears left because image is mirrored), misclassified as Left because face is shifted to the Left with respect to the frame.

## Failed Solution Trial

Tried increasing the dataset by appending augmented training data (random x-axis translations).

## Result

No noticeable change in model performance.

## Successful Solution Trial

Crop faces in dataset and before inference in runtime.

* + 1. Our first option is to use OpenCV’s DNN Face Detector which uses SSD and uses ResNet-10 as its Backbone.
    2. Our second option is to Use Haar Cascade.

## Result

First Option: Successful, but many crops with no faces (False Positives) are detected.

Second Option: Few numbers of faces are detected, not robust to partial face occlusion which is important because part of the user’s face is covered when he puts the swab in his nostril.

We used option One as it was much more accurate, yet fast enough for our model. The problem of crops with no face was solved by filtering bounding boxes with over 99% confidence and whose starting y-coordinate is before 1/3 of the height of the image frame. This was observed upon manual analysis of crops detected.

Graphical user interface, application

Description automatically generated

Figure 3 Crop 2 is an example of a no-face crop.

# Results after solving problems.

After solving the problem of overfitting and the problem where the side is misclassified, the model’s performance is now very impressive. It has low bias and low variance.

|  |  |
| --- | --- |
| Set | Accuracy |
| Training | 100% |
| Test | 94.5% |

## Tensorboard Logs for loss and accuracy.

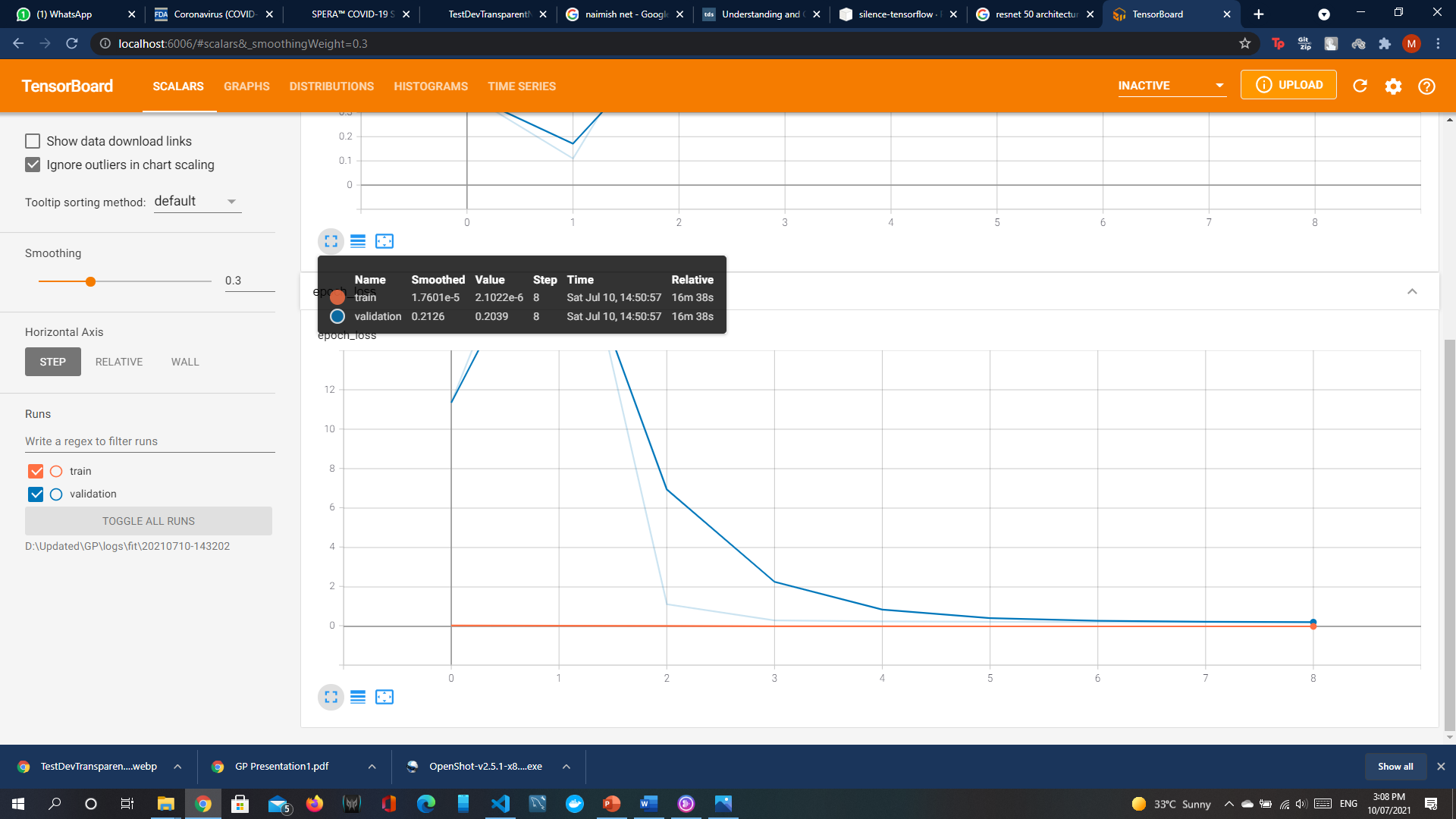


Figure 4 Epoch Loss Graph

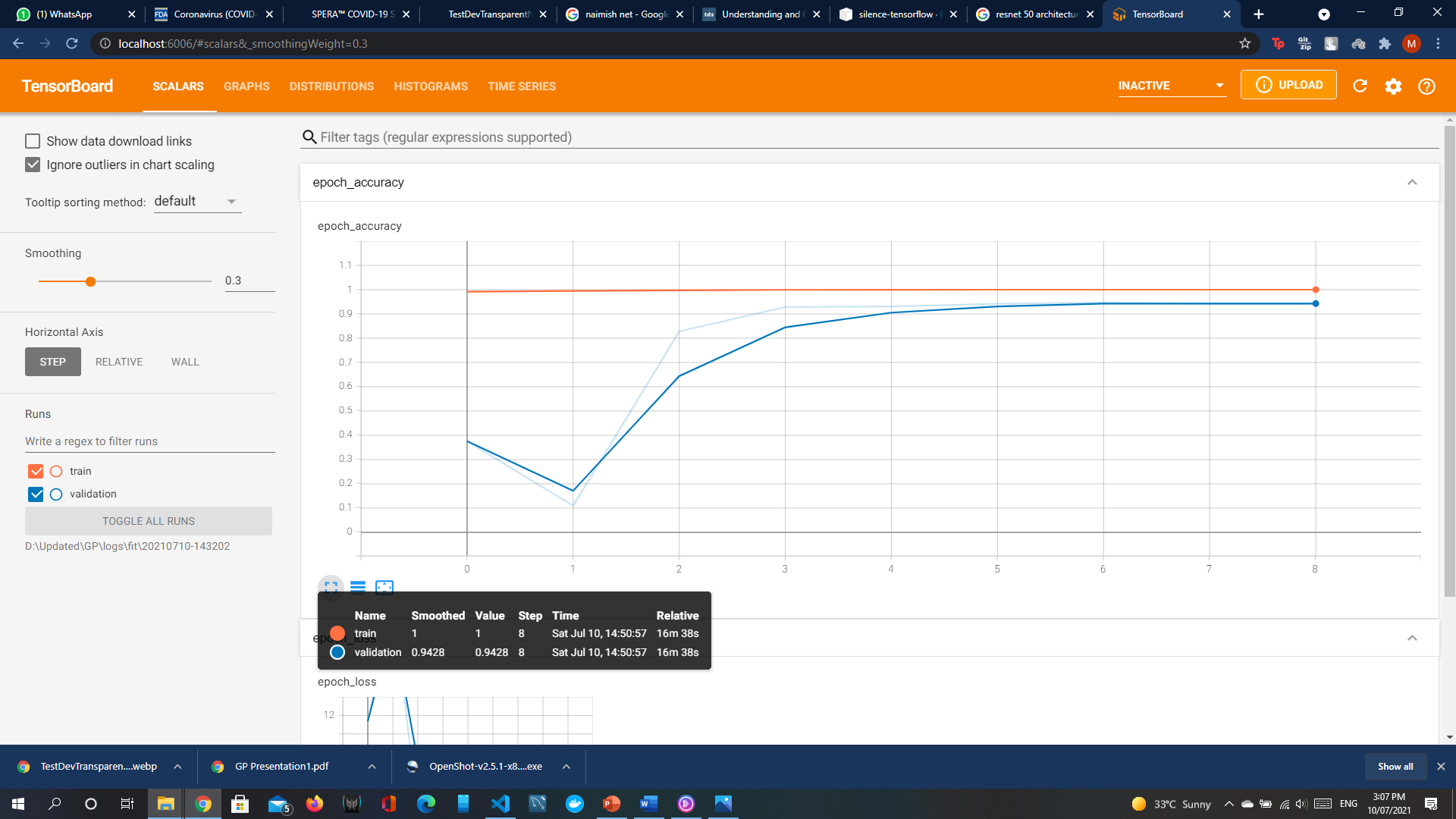


Figure 5 Epoch Accuracy Graph