Report of the work : Calib-challenge

NOTE: Several versions are available, please read the TEST chapter of this document for more information.

In this document I will share with you my thoughts on the problem as well as the methodology I have put in place in order to propose a coherent solution with its different steps. I will also bring a critical glance on my solution and I will evoke very quickly various perspectives of evolution.

• Approach and method:

Here are the steps I followed to predict the pitch and yaw angles:

1- Load the labeled videos and extract the pitch and yaw angles for each image and extract the images.
2- Preprocess the images by resizing and normalizing them.
3- Build a neural network model that takes an image and provides the estimated pitch and yaw angles (CNN).
4- Training the model using the labeled video data and evaluating its performance (using metrics such as mean square error (MSE) or mean absolute error (MAE) to evaluate the performance of the model).
5- Load the unlabeled videos and extract the images.
6- Use the trained model to predict the pitch and yaw angles for each frame in the unlabeled videos.
7- Save the predicted pitch and yaw angles in 2D tables and write them to files 5.txt to 9.txt.

During this project I was able to create, test and use different neural networks to compare the results and choose the best solution for the given problem.

Here are some of them :

- Normal CNN
- Custom VGG16
- Keras VGG16 with transfer learning
- ResNet50

I also used data generation techniques by modifying for example the contrast or the luminosity.

• TEST :

I studied the different results obtained with different values of hyperparameters such as the optimization algorithm, the learning rate, the number of epochs, the batch size... with the neural network models presented above.

In order to be able to estimate the level of accuracy I performed several tests by training my algorithm on 4 labeled videos and predicting then generating a file for the 5th labeled video. Thanks to this methodology I could know if the results were relevant or not.

- **Version 1** was trained on 4 videos (0, 1, 2, 3) and I predicted the angles for the 4.hevc video, as well as for the unlabeled videos.
- **Version** 2 was trained on 4 videos (0, 1, 2, 4) and I predicted the angles for the 3.hevc video, as well as for the unlabeled videos.
- **Version 3** was trained on the 5 labeled videos and I predicted the angles for the unlabeled videos.

These three versions were trained with the same model and hyperparameters only the training data varied. The predictions of version 3 should be more accurate since it was trained on all the labeled videos.

You will find 3 directories corresponding to the 3 versions with a folder containing the labels of the unlabeled videos and for version 1 and 2 a second directory containing the prediction of the labeled file that was not used for training (for example for version 1: 4.txt and 4_predict.txt).

Thank you for reading and I hope my work will allow me to join comma.ai. I believe that companies like comma.ai are at the forefront of innovation and problem solving, and I would like to be an active member of this community. I am motivated by challenges and enjoy working in an environment where things are constantly changing. I believe that working at comma.ai will allow me to grow professionally, giving me the opportunity to learn new skills and hone the ones I already have.