



ML1

Head Pose Estimation

Name: Sarah Magdy

Date: 17-6-2022

Final Project

PREFACE

This project is used to predict the three values of angles [Pitch, Yaw, Roll].

- Rotation around the front-to-back axis is called **roll** (rotation around the Z-axis).
- Rotation around the side-to-side axis is called **pitch** (rotation around the X-axis).
- is the angle of moving the head up and down
- Rotation around the vertical axis is called **yaw** (rotation around the Y-axis).
- is the angle of moving the head left and right.

I used three models (with same type “XGBRegressor”) to get value of each angle .



CONTENTS

01 Data

02 Work flow

03 Some trials
Problems & solutions

04 References



Data

01

I used AFLW2000 data which you mentioned.

Work Flow

02

First

I used glob to read all names of images files with extension “jpg” and saved all paths in list called “imgs_pathes_list”

Second

I used opencv to read each image from paths list and pass image to mediaPipe to extract face landmarks “(468 point)” for each x,y,z but I save x,y only for first face then I split all paths to take first part and concatenate .mat to allow loading the mat files using “loadmat” to get labels values from it that are saved in the first three elements in 'Pose_Para' key and save all landmarks and labels In list and do this operation to all images and save all dataset in list
(img_landmarks.extend([pitch,yaw,roll]);all_featuresANDlabels.append(img_landmarks))

I used list comprehension to create columns names of data frame and fill data frame with data in list which contains all landmarks and labels

```
(columns_name = [['x' + str(i), 'y' + str(i)] for i in range(len(face.landmark))];columns_name=n  
p.array(columns_name);x=columns_name.reshape(1,-1)[0].tolist();x.extend(['pitch','yaw','roll'])  
;df =pd.DataFrame(all_featuresANDlabels,columns = x)
```

Head Pose Estimation



Then

I split data set into training and test set and use three models with same type “XGBRegressor” but different parameters value to overcome overfitting

First model (pitch)

```
XGBRegressor(n_estimators=400, learning_rate=0.1, max_depth=7, min_child_weight=0, gamma=0,
validate_parameters = True, subsample=0.9, colsample_bytree=0.5, nthread=4, seed=42)
,
```

Second model (yaw)

```
XGBRegressor(n_estimators=400, learning_rate=0.1, max_depth=7, min_child_weight=0, gamma=0,
validate_parameters = True, subsample=0.9, colsample_bytree=0.5, nthread=4, seed=42)
,
```

Third model (roll)

```
XGBRegressor(n_estimators=200, learning_rate=0.1, max_depth=2, min_child_weight=0, gamma=0,
validate_parameters = True, subsample=0.9, colsample_bytree=0.5, nthread=4, seed=42)
,
```

Then

I evaluate the three model using Mean Squared Error and R2 Score in training and test data .

Then

I create ‘draw_axis’ function to to draw lines on image,

and ‘fun_extract_landmarks’ function

to extract landmarks (468 3D face landmarks) for input image And save data in dictionary

Head Pose Estimation



`img_landmarks['x' + str(i)]=landmark.x;img_landmarks['y' + str(i)]=landmark.y` then

And take landmarks for first face only.

Then

I use `cv2.VideoCapture('/content/test2.mp4')` to create an object to open video then we can read video frames. I use `video.isOpened()` to check if video is opened previously or not. I use

`frame_width = int(video.get(3));frame_height = int(video.get(4))` to set resolutions

I create VideoWriter object which will create

a frame of above defined The output is stored in 'output.mp4' file.

`cv2.VideoWriter('/content/drive/MyDrive/ML1/output4.mp4',cv2.VideoWriter_fourcc(*'MJPG'),3, size)`

While true I read each frame in video using `video.read()` and pass frame to function to extract landmarks and then append this data to dataframe and uses three models to predict angles and send image and angles to function to draw axes then write frame in output4 file and When everything done I release the video capture and video write objects finally Closes all the frames.

03

Some trials (Problems & solutions)

I used pca and other models to select most important features but they cause underfitting.



I used lasso l2 to add complexity cost to cost function and delete unrelated features and overcome overfitting problem but it causes underfitting.

I spent 5 hours to extract landmarks manually but it causes underfitting too. I tried to use many models to improve accuracy like (linear regression and svm ... etc) but XGBRegressor is the best one.

04

References

- <https://techtutorialsx.com/2021/05/19/mediapipe-face-landmarks-estimation/>
- https://google.github.io/mediapipe/solutions/face_mesh
- <https://learnopencv.com/introduction-to-mediapipe/>
- https://xgboost.readthedocs.io/en/stable/tutorials/param_tuning.html
- <https://appdividend.com/2022/03/19/python-cv2-videocapture/>
- https://colab.research.google.com/github/s7s/machine_learning_1/blob/master/ML_in_practice/ProjectHelper.ipynb#scrollTo=oZVe65YyFJdF



THANKS

Head Pose Estimation