FACE POSITION DETECTION

MACHINE LEARNING MINI PROJECT

ABSTRACT

- Evaluation of driving performance is of atmost importance in order to reduce road accident rate. Since driving ability includes visual-spatial and operational attention, among others, head pose estimation of the driver is a crucial indicator of driving performance.
- We rely on a set of geometric features computed from just three representative facial keypoints, namely the center of the eyes and the nose tip.
- Despite the very few facial keypoints required, the results are comparable to the state-of-the-art techniques.
- Face plays a core part in distinguishing and identifying a person and hence face detection is much sought after.
- Accurate face position estimation is a challenging problem in itself due to the variability introduced by multiple factors such as illumination, identity and expression, to name a few

PROBLEM DESCRIPTION

- Face Position detection is the problem of detecting the head position of driver while driving in real scenarios.
- These positions are as looking-right,frontal, looking left.
- It is composed of 606 samples acquired over different days from 4 drivers (2 women and 2 men) wit several facial features like glasses and beard.

PROBLEM STATEMENT

- Modeled as a multi classification problem.

- There are 3 classes in this dataset
- Class 1 - looking-right
- Class 2 – frontal
- Class 3 - looking-left

- Associated tasks:Classification, Regression, Clustering

- Evaluation metrics - Accuracy.

DATASET DESCRIPTION

- A set of labels assigning each image into 3 possible gaze direction classes are given.
  The first class is the looking-right class and contains the head angles between -45Â° and -30Â°.
  The second one is the frontal class and contains the head angles between -15Â° and 15Â°.
  The last one is the looking-left class and contains the head angles between 30Â° and 45Â°.
- Along with the facial key point positions(eyes,nose and mouth).
- Face position [xF yF wF hF] and also head pose angle.
- No of features: 18, No of patterns: 606.
- The Given Dataset Is Highly Imbalanced,The Provided Dataset Contains Class Imbalance Problem.The No of Patterns For Each of The Given 3 Classes-{1,2,3} are:
- Class 1=27 Patterns, Class 2=546 Patterns, Class 3=33 Patterns.

Why use ML?

- Here The Input-Output Relation Is Not Straight Forward.
- So, Here We Can Design a Model Using Algorithms In ML Such That, By Passing The Different Training Samples from The Dataset and Make The Model Learn From The Experience.
- The Model Should Correctly Detect The Respective Class By Using Features Of the Face From The Dataset and Classify The Images Into Corresponding Classes.
- We Use Some ML Algorithms Like “LOGISTIC REGRESSION”,”SINGLE LAYER PERCEPTRON(SLP)”,”MULTI LAYER PERCEPTRON(MLP)”.

LITERATURE SURVEY

- The classification methods learned a mapping between images and a discretized space of poses. Given a new image, the classifiers assign it to a discrete class .
- Since the majority of such methods have discretized outputs, only allowing coarse head pose estimation, it is difficult to derive a reliable continuous estimation from the results.
- Different from classification methods, regression methods estimate Face Position by learning a functional mapping from the image space to one or more pose directions .
- The allure of these approaches is that with a set of labeled training data, a model can be built to provide a precise Face Position estimation for any new data samples. Due to the breakthrough results achieved by Machine learning 90 technologies in many research field.

LITERATURE SURVEY

- Zavan et al. proposed an automatic pipeline based on convolutional neural networks for detecting different facial regions, processing them, and combining the results generated from each, resulting in a robust head pose estimation and gender recognition. And some recent work can estimate head pose with high accuracy and perform in real time.
- Accurate face position estimation is a challenging problem in itself due to the variability introduced by multiple factors such as illumination, identity and expression, to name a few. During the last decade there has been an increasing interest in developing head pose estimation methods for different applications such as security and surveillance systems , human-robot interaction , meeting rooms , intelligent wheelchair systems , and driving monitoring .

Result Analysis

- The Applied Algorithms are:
- Logistic Regression
- SLP
- MLP
- For each algorithm applied, we will calculate the class wise accuracy oversampling and undersampling under the case of logistic regression.

Result Analysis

- Logistic Regression:
- By Using MSE Cost-Function:
- Hyper-Parameter Tuning On The Validation Set,Best Hyper Paramaeters are
- Alpha=0.1,Rho=0.0001,Epochs=10.
- After Training The Model, Using 30% of The Samples
- Correctly Predicted : 166
- Total Test Samples : 181
- confusion Matrix= [[0 5 0]
  [0 166 0]
  [0 10 0]]
- Train Accuracy: 91.71270718232044

Result Analysis

- After The Training Is Completed, Testing The Model By Using The Test Set:20% of The Samples
- Correctly Predicted : 110
- Total Test Samples : 119
- confusion Matrix= [[0 6 0]
- [0 110 0]
- [0 3 0]]
- Test Accuracy: 92.43697478991596
- As The Dataset Is Class Imbalanced, Predicton of The Model Is Biased Towards The Class 2
- Applying Kfold- Cross Validation by Using K=5
- Accuracies Fold Wise:
- Fold 1: Accuracy : 90.98360655737704
- Fold 2: Accuracy : 92.56198347107438
- Fold 3: Accuracy : 89.25619834710744
- Fold 4: Accuracy : 90.9090909090909
- Fold 5:Accuracy : 86.77685950413223
- Average Accuracy: 90.0975477577564

Result Analysis

- Using LoggLoss(Convex)-Cost Function:
- By Hyper-Parameter Tuning On The Validation Set, Best Hyper Parameters are Alpha=0.001,Rho=0.0005,Epochs=80.
- After Training The Model:
- Train Accuracy : 90.60773480662984
- After The Training Is Completed, Testing The Model By Using The Test Set:20% of The Samples
- Test Accuracy : 91.52542372881356

Result Analysis

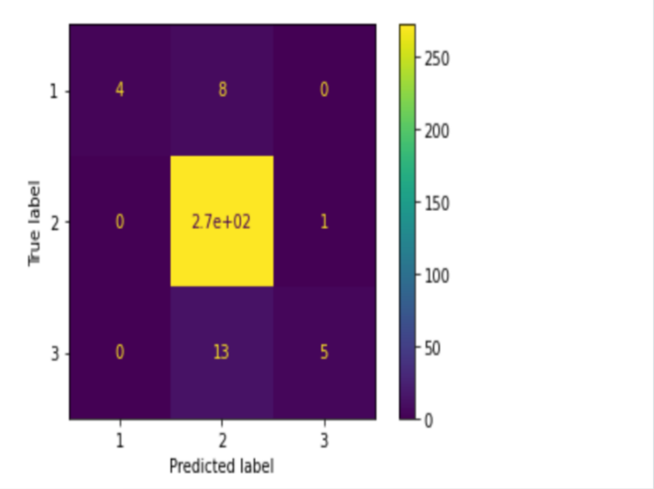
- SLP:
- Uses The Concept Of ANN and Performs Classification Using One Architecture By Performing One-Hot Encoding On The Dataset. Used Inbulit From SKlearn
- Accuracy = 90.0990099009901
- Confusion Matrix= [[0 12 0]
  [0 273 0]
  [0 18 0]]
- class_accuracy_one= nan
- class_accuracy_two= 0.900990099009901
- class_accuracy_three= nan
- precision for 1 class : 0.0 , recall for 1 class : nan
- precision for 2 class : 1.0 , recall for 2 class : 0.900990099009901
- precision for 3 class : 0.0 , recall for 3 class : nan

Result Analysis

- MLP:
- Accuracy For The Classification :92.73927392739274
- Confusion_Matrix=[[4 8 0]

  [0 272 1]

  [0 13 5]]



REFERENCES

- https://www.researchgate.net/publication/323863444_Continuous_Head_Pose_Estimation_Using_Manifold_Subspace_Embedding_and_Multivariate_Regression
- https://hal.archives-ouvertes.fr/hal-03344132/document
- https://fairyonice.github.io/Driver-facial-keypoint-detection.html