CS574

Computer vision and machine learning

(Assignment #3)

Supervised By - Arijit Sur

Created by:-

(Group 07)

Aditya(160101010)

Avinash(160101018)

Phool Chandra(160101051)

Savinay(160101062)

Kevin(160101063)

Deep Neural Decision Forest

Age estimation has been tried by different approaches, out of which CNN has been now widely used for age estimation in a different form. Like using DenseNet, OR-CNN, Fusion network, etc. But till now deep neural decision tree is not tried with CNN. Earlier with decision tree hand made splitting was done which is referred has hard but with the use of NDF here it has been tried to have a soft splitting function. CNN has been evolving and with it, residual learning has shown much improvement and as been recognized as one of the best approaches for making problems optimally solvable. So here new approach is tried in which NDF is tried with CNN and residual learning is used for making more optimal. To have a better understanding of results i.e., for making observations network visualisation technique were used. It uses the divide-and-conquer method. It reduces the main problem into the subproblem by dividing the main problem.

Implementation Architecture:-

Firstly facial land masking is used to detect a face in given image. This part is important as it should give region which contains all features required for age estimation. The result obtained is then flipped and rotated and cropped to get 256 *256 pixel image. If it fails then it is done manually. Then normalization is applied and it is further flipped horizontally finally giving 224*224 pixel image. Then it is given to our model which consists of initially convolution layer then max-pooling layer and then three convolution layer and finally fully connected layer. This with the help of sigmoid function gives final output for the tree. There are 5 trees used here and output at leaf node is predicted after 20 iterations. Here SGD optimizer is used for optimizing.

Limitations:-

1. This model ignores the hair part of the face. During ignoring the hair part of the face it also ignores nearby regions of the hairy part. By ignoring it means a feature is dropping of facial characteristics. It should take the hair part also because it will add one more feature and also help in age estimation.
2. In this model binary tree of deep neural decision, forest take input only from fully connected layers. This feature makes this model inflexible.
3. Here image segregation is used in the form to convert image into something meaningful for which image pixels are separated into segments. But this approach can be improved by using a different way to view the network.
4. Here model architecture used can be improved to give better accuracy and also increase performance.

Improvements:-

1. Facial landmarking should be selected such that it can use hair also like one feature for age estimation.
2. This model can be made better if the tree can take input from other layers also, not only just fully connected layer.

Ordinal Regression

Earlier, it implements using multiclass classification and metric regression methods. All class labels are independent of one another in the multiclass classification method. But there is a strong relationship in age labels. In metric regression non-stationary kernel is used which is has often lead to overfitting.

Here, they changed ordinal regression into a series of subproblem of simpler binary classification. the advantage of this transformation is,

already known bounds for binary classification can be used to derive bounds for ordinal regression. Binary classification subproblems are solved using CNN. In CNN, All binary classification subproblem gives its own output i.e., multiple output layer. Earlier ordinal regression approach, Extracting features and learning a regression model are optimized independently and also separated independently.

It is difficult to design good features for age estimation because there is no specific way in which the human face changes as with the age.

Here CNN model is built as implementation where the regression method and feature extraction is used.

Implementation Architecture -

Preprocessing is required before training or testing of images hence file name preprocessing.py does preprocess like get frontal face and remove the unwanted noise and objects in the image and only left face image.

In implementation we made training script for each data set and loss combination. Name of the file given in the format <dataset>-<loss>.py where < dataset > can be AFAD (afad), MORPH-2 (morph), UTKFace (utk), or CACD (cacd).

Basically we are training on each database and using all loss combinations all have different scripts.

The training data has been saved in different directory name training-results each respective directory of training script contain a tensor file and two txt file containing training data.

Limitations & Improvements

1. The proposed algorithm uses the facial feature to process images. So, it requires a nose position to be aligned in the center but if the image is clicked from side angle then the position of the nose will not be in the center so proposing will not be able to aligned image correctly and hence will result in wrong results.

2. The given model performs worse on the UTKFace database then AFAD this is because the quality of images is not good in this database. To improve this image quality must be very good.
3. Imbalance datasets will also lead to worse performance as CACD and AFAD databases have the same size but have a different range of labels CACD has 14-62 while AFAD has 15-40. To handle that range of the labels and data sample must be chosen appropriately hence must be balanced.