SHIV SHAKTI
International Journal in Multidisciplinary and Academic Research (SSIJMAR)
Vol. 7, No. 2, April 2018 (ISSN 2278 – 5973)

Facial Recognition Using Deep Learning

Rajeshwar M, Sanjit Singh Chouhan, Pattan Asif Khan, Sampath K & Pavan

* HITAM, JNTUH, Hyderabad, rajeshwarm.cse@hitam.org, +918374151675
HITAM, JNTUH, Hyderabad, sanjitschouhan@gmail.com, +917382978847
HITAM, JNTUH, Hyderabad, aisf.khan.399@gmail.com, +919177026529
HITAM, JNTUH, Hyderabad, sampath.varanasi1310@gmail.com, +918019845239
HITAM, JNTUH, Hyderabad, pavan48chowdary@gmail.com, +919652187419

Abstract

Facial Recognition is an emerging technology and is considered to be a authentication system that is secure enough and is accessible to everyone. The methodology looks easy to follow but, there exist a lot of loopholes that are to be considered, the existing face recognition methodologies follow or need an excellent or at least a good image quality to perform the recognition operation. But, these algorithms fail to address the need in case of low quality images. We here describe the methodologies that could be implemented to overcome the hurdle of recognition of low quality images using few algorithms like super resolution.

Keywords: Facial Recognition, Super Resolution

Introduction

Convolutional Neural Networks (CNNs) have been majorly concentrated in the areas of research. The Major reason noted is that the availability of the training data. The Faces data available gives us a major advantage to improve the existing facial recognition methods using CNNs. In [Table 1] as you can see there are various kinds of face datasets available. But, We will here mostly concentrate on the [[CelebFaces[25]] dataset from all the datasets we have. The paper will also elucidate on improving the existing quality of the images using the super resolution. The company Giants like Google, Amazon, and Apple are investing a huge amount of time and money to improve the facial recognition technique. The problem even faced by these giants is they are not able to apply facial recognition when the image quality is not up to the mark or if the image resolution is not proper. This problem arises because there is a non-availability of a public dataset of low quality on which the facial recognition algorithms should be trained. Gathering real time data from CCTV's has become a problem because it would lead to personal integrity and privacy problems of the people.

This is why even giants fail to address the low quality facial recognition problem. So, our paper would majorly concentrate on two areas i.e. getting the data from various sources and building a deep learning on those images. Getting data includes finding low resolution image

data from heterogeneous sources or to generate a data set using volunteers and making the data freely available to the open source community. The second goal is to generate a model using Deep Learning module to improve existing face recognition algorithm.

Few of the existing works include Timo Ahonen, Abdenour Hadid, and Matti Pietika¨inen(2005), proposed a methodology to divide the image into sub images and each sub image is further divided into 3*3 matrix and the center pixel is compared with other pixel, if other pixels value is greater than the center pixel value then the value in 3*3 matrix is noted to be 1 else 0 thus generating a binary number, these numbers are compared locally in the sub image as well as globally with the total image and prediction is done. Taigman, Yaniv(2014) extracted Compact features, training them on nine layers deep neural network on one of the largest dataset LFW . They have applied 3d – alignment and frontalization . Sun, Y., Liang, D., Wang, X. and Tang, X.(2015) used Google Net and VGG the top ranked deep neural networks in the general image classification ILSVRC2014 for face recognition and they were able to add few inception layers to the existing facial recognition techniques.

Dataset	Identities	Images
LFW	5,749	13,233
WDRef [4]	2,995	99,773
CelebFaces [25]	10,177	202,599

Table 1: Datasets

Research Methodology

The Existing methodologies follow the methodology in which existing data sets are loaded into the model and the data sets are further divided into two parts i.e. for training 75% of the images are given and the rest 25% are used to test the model. The Algorithm stands out to be taking the images and performing the division of images into sub images and getting the details of the pixels in the form of binary representations, all details are stored, later the same is followed with other images and training is done. Later, the testing is done on the 25% test data and accuracy is noted. We here use a set of CNN'S followed by a fully connected layer. The layers perform the operations like breaking of images into parts, getting the finer details from the images and storing the details.

Data Analysis and Interpretation

The dataset used in the model is CelebData available publicly on net. The dataset consist of over 2 Lakh images of 50 famous celebrities from all over the world. The dataset is organised in folders with names of celebrity containing all pictures of them from different sources cropped to view only the faces of celebrities. The dataset is preprocessed manually to remove any corrupted image files and are cropped to show only faces of the celebrities. The Dataset is then loaded to model and preprocessed to match the input size of the model which in our case is 128x128x3. The dataset is divided into three parts, (1) Training set - which is used to train the model to recognise the faces, (2) Validation set - which is used to validate the performance of model while it is being trained and (3) Testing Set - which is used to test the accuracy of model after it has been trained completely. We keep testing set separate to avoid overfitting of dataset on model when training and validation as neural networks learn even when predicting the results.

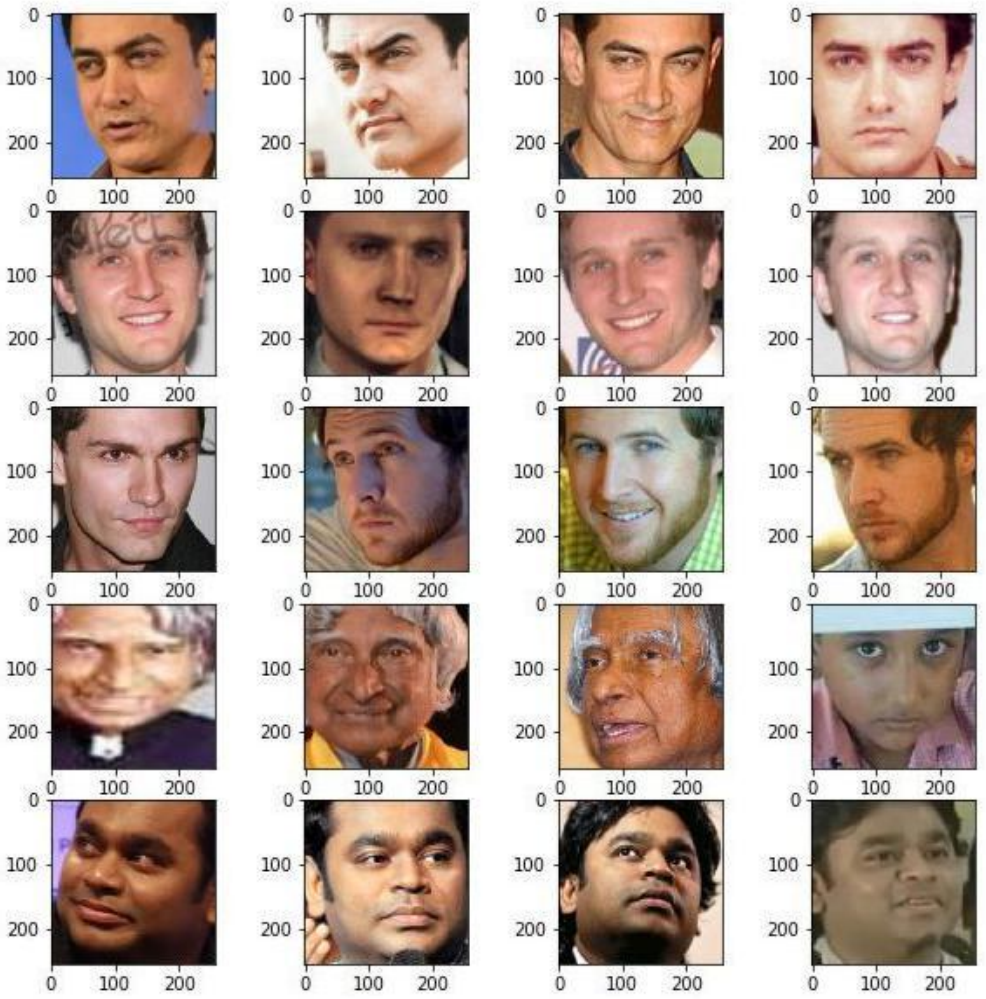

Figure 1: Dataset - CelebData

The model is a Convolutional Neural Network(CNN) which takes input tensor as array of images of size 128x128 with 3 channels and outputs tensors of array of one class probabilities for each input image. The images are passed in batch of 32 images per batch and run for 50 epochs. One epoch is one complete cycle of all the available training images after which they are shuffled to form new batches with different combinations. The model has 5 layers of convolutions each consisting of a convolution 2d layer of string [1,1,1,1] and 'SAME' padding, and a max pool layer of size 2x2 with kernel size of 3x3 and 'SAME' padding. The model is optimised with an AdamOptimizer with a learning rate of 0.001 to minimize the loss in the model and fine tune the Weights and bias of model. The training graph of model is as shown in Figure 2

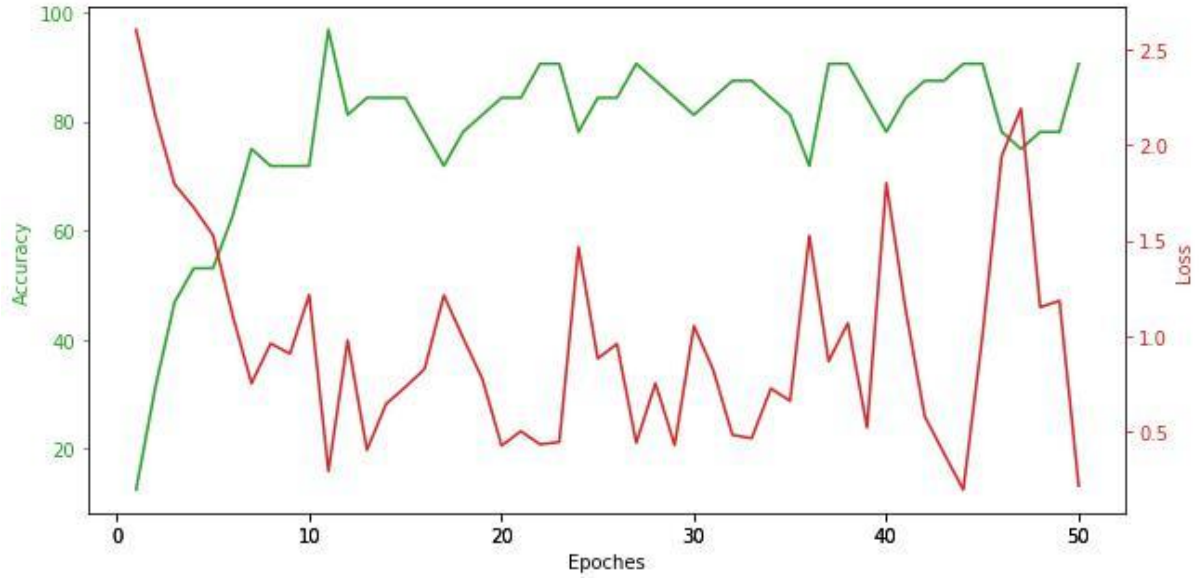

Fig 2: Training Graph

The model trained with images of 10 celebrities, gave an accuracy of 84.46% on the testing set which consist of 15% of all the available images. The accuracy is less as it is on very less data and low computation power, but is pretty good as compared to earlier models which required large data to train the model.

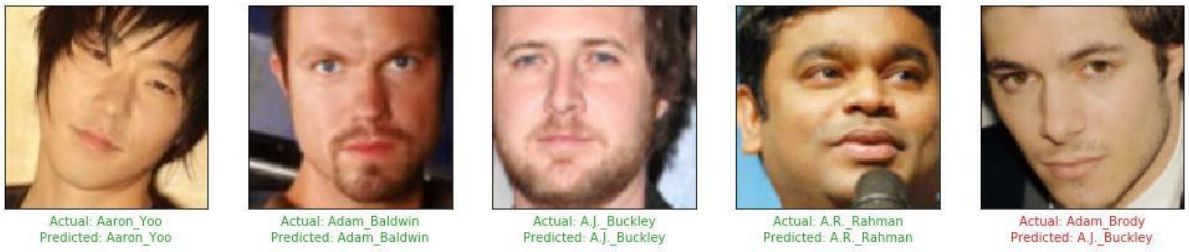

Fig 3: Prediction Results

Conclusion

Facial Recognition has been trend that has been adapted rapidly by the people. It has reached into the area of authentication and is being considered to be most acceptable way of authentication after the biometric authentication. The usage of CNN'S for image recognition would improve the existing methodology as described, but, there needs to be a lot of work required to achieve the accuracy where it could replace the existing authentication methodologies. The work proposed in the paper concentrates on improving the accuracy of existing models. The work combined with few other techniques would lead to make a robust model and would bring significant change in the areas of authentication, recognition of people.

Scope for Future Research

The future research includes the application of the technique called super resolution in which the enhancement of the low quality and small size images then facial recognition would be applied. Application of this model to blurred images, application of the model to the real time CCTV data i.e prediction while video recording. These can be considered to be the areas where model lacks and could be improved.

References

Sun, Y., Liang, D., Wang, X., & Tang, X. (2015). Deepid3: Face recognition with very deep neural networks. *arXiv preprint arXiv:1502.00873*.

Taigman, Y., Yang, M., Ranzato, M. A., & Wolf, L. (2014). *Deepface: Closing the gap to human-level performance in face verification*. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1701-1708).

Yu, H., & Yang, J. (2001*). A direct LDA algorithm for high-dimensional data—with application to face recognition. Pattern recognition*,34(10), 2067-2070.

Ahonen, T., Hadid, A., & Pietikainen, M. (2006). Face description with local binary patterns: Application to face recognition. IEEE transactions on pattern analysis and machine intelligence, 28(12), 2037-2041.

Wiskott, L., Krüger, N., Kuiger, N., & Von Der Malsburg, C. (1997). Face recognition by elastic bunch graph matching. IEEE Transactions on pattern analysis and machine intelligence, 19(7), 775-779.