Gender Classification Based on Image Processing using Deep Learning

VIDHYADHARAN K
Dept. of Computer Science Engineering
Rajalakshmi Engineering College
Chennai, India.
Vidhyadharan.k.2018.cse@rajalakshmi.edu.in

VIKNESH A
Dept. of Computer Science Engineering,
Rajalakshmi Engineering College
Chennai, India
viknesh.a.2018.cse@rajalakshmi.edu.in

Dr ANAND SIVA PRAKASHAM
Dept. of Computer Science Engineering
Rajalakshmi Engineering College,
Chennai, India
AnanthaSivapraskasam.s@rajalakshmi.edu.in

***Abstract*—Gender records using facial patterns serve a crucial use in various person interplay applications. This venture proposes a technique primarily based on Convolutional Neural Networks to identify gender from faces and be counted the overall quantity of men and women from the given image or video. By using the image validation approach to test whether it's a celebrity or not. The network is trained using back-propagation and Adam optimization. The usage of convolution operations and the overall performance of proposed CNN communities were evaluated on a publicly available gender detection dataset in Kaggle. The models we will use have been trained on this dataset. In this method, we use Convolutional Neural Network to recognize multiple images to predict males and females which is better than the SVM method to classify. The CNN algorithm used in this project is CNN VGG-16. The CNN VGG-16 provides more accuracy compared to other CNN algorithms. The CNN is outperformed the SVM and the accuracy attained by the Convolutional Neural Network is 98.5% whereas the accuracy attained by SVM is 94.81%.**

***Keywords*— facial patterns, gender, CNN, deep learning**

I. INTRODUCTION:

Conventional approaches to face recognition involve image pre-processing, feature extraction, and further classification. Thus performance is highly dependent on the type of classifier used and the number of features extracted[1]. The advantage of using CNNs is that these multilayer perceptron achieve feature extraction, dimensionality reduction and classification are all in one network. This is done by using convolutional layers, Max pooling layers, and finally fully connected layers. Thus accuracy is improved to a large extent, with the use of CNN for facial recognition and analysis. CNN is easier to train as compared to traditional classifiers. Also, they have fewer parameters than fully connected multi-layer perceptron networks and are robust to shape and transformation[2]. SVM classifier achieves an accuracy of 94.81% whereas the accuracy achieved on the CNN model is 98.5 percent. So it is more efficient. The first step is to capture the video stream and In the video Stream using python code, we get some images. These images reduce the areas of the unwanted image by Max Pooling in CNN. Compare the remaining areas of the image with an already predefined dataset that is written binary format. Create a bounding box using python code to recognize the image, extract the pixels, and then compare it with predetermined datasets. And then it detects whether the image is male or female with the most probability that will occur in the SoftMax layer[3]. And it finally counts the total no of males and females in a given image. And then it also checks whether the given image is a celebrity or not. The dataset used in this project serves as a benchmark for face photos and is inclusive of various real-world imaging conditions like noise, lighting, pose, and appearance. The images have been collected from Flickr albums and distributed under the Creative Commons (CC) license in Kaggle. It has a total of 26,580 photos of 2,284 subjects and is about 1GB in size. The models we will use have been trained on this dataset. The convolutional neural network for this python project has 3 convolutional layers:

- Convolutional layer; 96 nodes, kernel size 7
- Convolutional layer; 256 nodes, kernel size 5
- Convolutional layer; 384 nodes, kernel size 3

It has 2 fully connected layers, each with 512 nodes, and a final output layer of SoftMax type. For the

image Validation process, we use an image array to append all images and check whether it is a celebrity or not.

II. RELATED WORKS:

In previous methodologies, they use the SVM algorithm for gender Detection. But SVM for multiple object detection is not accurate because we aim to detect multiple objects .so SVM failed to prove it. gender classification approach by exploring the inherent properties of multi-spectral imaging sensor by SVM is 95% accuracy is published in 2017[13]. Another classification method based on the FERET dataset by CNN is done With 95.7% accuracy[12]. Another classification method based on the LFW dataset is done with an accuracy of 97.31% dataset by using minimalistic CNN and it provides a 10 times lesser number of training images[14]. In this paper, we use a pre-trained dataset that was done by CNN VGG-16. The approach for gender classification using single neural network was given by Titive and Bouzerdoum [20] in which an average classification accuracy of 97% was achieved. Afifi [21] has worked on solving gender recognition tasks using a large dataset of hands, as they give important gender related information. so this method provides better accuracy for multiple object detection. Biometric identification is increasingly being used to ensure the security and privacy of individuals in a variety of applications. Various characteristics such as the face, fingerprints, and iris are utilized to identify people. a person's identity In facial recognition applications, a single image can provide us with crucial information such as gender, age, ethnicity, and so on. The task of gender recognition is crucial because as well as difficult due to the nature of the situation, it is possible that faces will not be identified Pose, lighting, and expressions are all different. It is conceivable that a biometric system might make a mistake in identifying a person in high-risk situations. Using soft biometrics (which aren't as accurate as hard biometrics include characteristics like age, gender, and so on) can assist in obtaining a valuable piece of knowledge about a person and a solution to difficulty with identification Apart from resolving issues like Gender recognition, as well as shadows and lighting, can be used.

III . CNN ARCHITECTURE

The CNN architecture described below of some information are taken from the paper of robust approach for gender recognition . The working of the pretrained dataset model is described as follows. First, we get the dataset from Kaggle that consists of two parts namely the training dataset and the validation dataset. The training dataset consists of male images in one folder and female images in another folder the validation dataset is also similar to the training dataset. In training, the dataset consists of a large no of images compared to the validation Dataset because the training process plays a major role to classify the gender and train the model. Train data directory we should copy the path of the training dataset and Validation data directory we should copy the path of a validation dataset. Image Data Generator is used to get a different types of images and features from the trained model single image. There are four layers in the convolutional neural network they are 1)convolutional layer 2) ReLU 3) Pooling Layer 4) Fully Connected Layer are shown in fig 1

The convolutional layer is used to extract features from the images. ReLU is the Activation layer. The Activation is mainly used to remove the Negative part. Max pooling is mainly used to reduce the size of the images. A fully connected layer provides the output with the help of SoftMax. These are the process done by the convolutional neural network. In compiling the CNN model we use optimizer is ResProp optimizer help us to understand unknown information and the fit of the model in CNN occurs in it. After that, we get the accuracy and loss of the model in the graph. These processes are done in a pre-trained model

VGG16 is a Convolution Neural Network(CNN) architecture which was used to win ILSVR(Image net) competition in 2014. It is considered to be one of the excellent vision model architecture till date. Most unique thing about VGG16 is that instead of having a large number of hyper-parameter they focused on having convolution layers of 3x3 filter with a stride 1 and always used same padding and Max Pool layer of 2x2 filter of stride 2. It follows this arrangement of convolution and max pool layers consistently throughout the whole architecture. In the end it has 2 fully connected layers followed by a SoftMax for output. The 16 in VGG16 refers to it has 16 layers that have weights. This network is a pretty large network and it has about 138 million parameters.

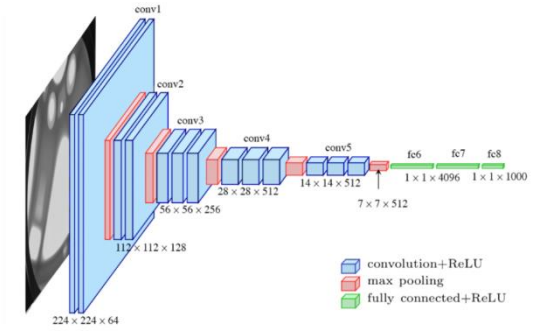

Fig 1

A. PROPOSED ARCHITECTURE

In previous method they use the SVM to predict multiple images but its accuracy is low. So we use CNN algorithm to detect multiple images which is better than SVM. The dimension of the image is 512 X512. and the number of training images used is 26400 images and the number of validation images used is 15000 images and the number of epochs is 100 and Optimizer Adam is Loss and Categorical is cross-entropy and Augmentation parameters are Flip left, right The dimension of the image input to the network is of dimensions 512x512x1. This 1 represents the number of colour channels. The input image is in form of the coloured image have been resized into a grayscale image. In CNN it consists of 10 convolution layers and max-pooling layers that are followed by a fully connected layer. The filter is in size 5x5.In CNN the filters are in size 5x5 and are convolved with no padding, strides that result in a reduced Space size of an image. In CNN number of filters is 32. Then next rectified unit function and Max pooling reduces the sizes to smaller and it negotiates any negative values that are present in the image matrix. Then it helps to extract features from the images of the previous layers. Each layer is connected with max-pooling layers and the activation function is ReLU. ReLU is the best activation function compared to others. The output from the previous layer is fed into another two convolutional layers with a filter size of 5. Then fully connected layer is used at end of the network with a dropout factor of 0.8 and the fully connected layer is used as a Soft Max layer. The Soft Max layer help to find probability with each image i.e. to predict whether a person is male or female. These are done by a pre-trained model. By using this model we detect multiple objects and detect whether the person is male or female and then it counts the total no of males and females. By image validation process we append all images in an image array and check whether the person is a celebrity or not as shown in Fig 2. If it is a celebrity it will display yes otherwise it will display no

First we import a pretrained dataset and we design a code for creating bounding boxes and recognize the face that are created from bounding boxes. And we take video using python code or we can capture images from video using python code .And then with help of video or images the code create a bounding boxes by using threshold value and recognize the face and it detects the gender with the help of SoftMax layers by using CNN VGG-16 algorithm. The Bounding Boxes will be Created when the threshold value should be greater than 0.7 . The threshold value lies between 0 and 1.And then we count total number of males and females .Then we check whether the given image is celebrity or not by using the Image Validation process. In Image Validation process we append some celebrity images in image array and compare whether the given image is celebrity or not with the Image array. If it is a celebrity then it displays yes otherwise it displays no. These process are described in below given Fig 2.

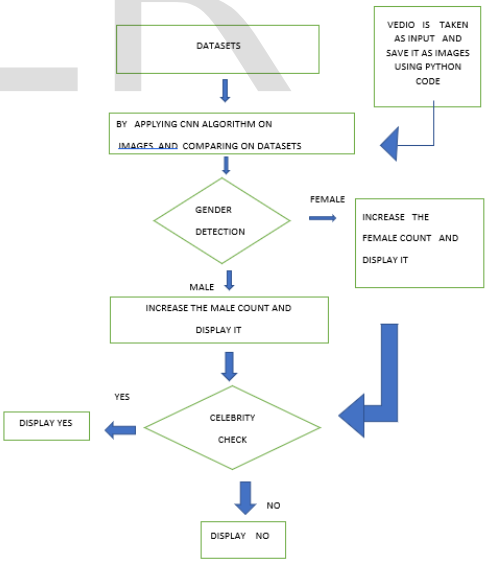

Fig 2

B. ADIENCE BENCHMARK

The Image Database consists of 26580 facial images of a variety of people. These faces have been

captured using a Camera. The people involved are mostly employees, students, etc.

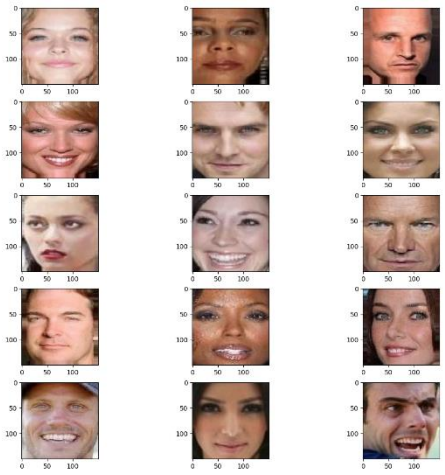

Fig 3

The resolution of images is 512x512. In this paper, we have opted for the cropped version of the images. These images have been converted from coloured images to grayscale images using a library called OpenCV. While 15000 images out of the available 26400 images have been considered to be a part of the training set and 12000 are a part of the validation set. Next Followed training and validation process has to be done. Based on Training it accurately classifies males and females and it has been done by the ADIENCE benchmark.

C. TRAINING AND TESTING

The dataset is divided into training and validation sets. The training dataset only provides training and the validation dataset provides validation. Usually, we provide more numbers of Images for training because training plays a significant role to improve efficiency. optimization used in CNN is Adam. In this model, we use 100 epochs to provide more efficiency which is better than the existing method. Image augmentation is used to flip the image left and right

1 epoch = (Number of training images/ Batch size)

D. IMAGE VALIDATION:

Image validation we used to append celebrity images inside the image arrays in it. so we validate the images by the given image and predefined images in the image array and compare with it. If it is the same image and then it displays yes otherwise it displays no. These processes are done under image validation are shown in Fig 7 and Fig 8.

IV. EXPERIMENTAL RESULTS

The accuracy obtained by ADIENCE Benchmark is 98.5 and it can be shown in the graph and it was done by CNN VGG-16.In a previous paper, it has been identified that multiple Object detection CNN is good when compared to SVM as shown below in table 1. In this, they use the image classification method for medical diagnosis. Fig 4 describes the model loss that is equal to 10% which is less. Fig 5 indicates model accuracy which is equal to 98.5%. Fig 6 describes the output of counting total number of males and females.

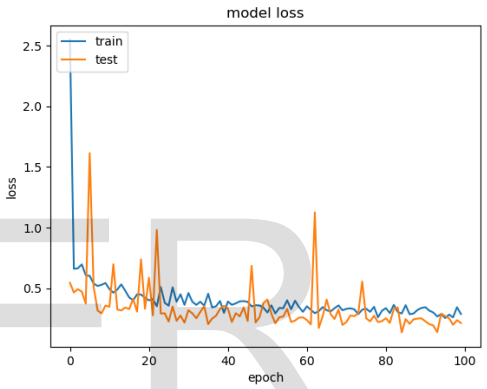

Fig 4

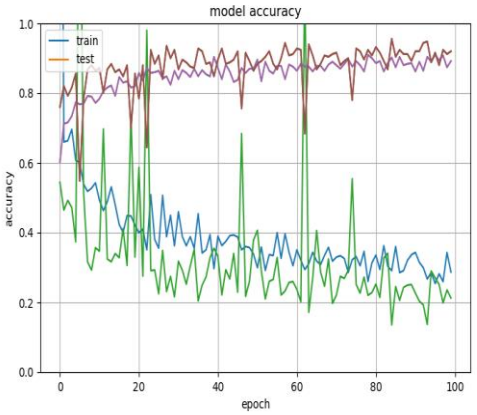

Fig 5

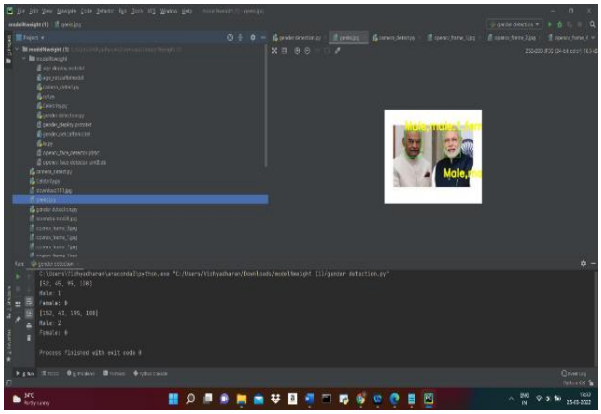

Fig 6

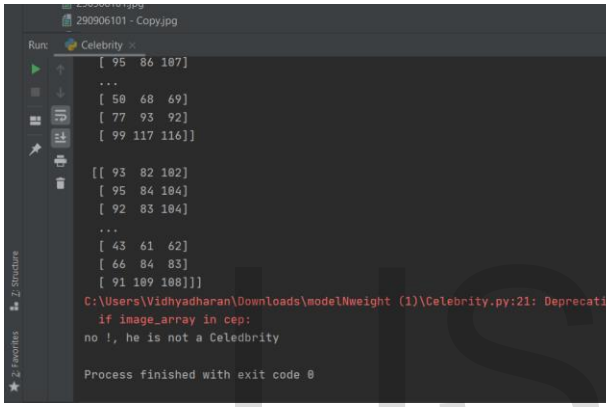

Fig 7

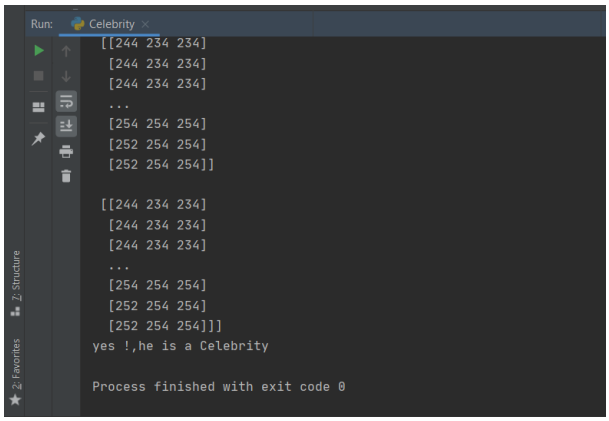

Fig 8

Classification Method	Accuracy %
Support Vector Machine Binary Classification	80.95%
Support Vector Machine Multiple Class Classification	50%
Convolution Neural Network Binary Classification	75%
Convolution Neural Network Multiple Class Classification	78.4%

Table 1: A Comparison Between Support Vector Machine (SVM) and Convolutional Neural Network (CNN) Models For Image Classification Based on the accuracy

V. CONCLUSION AND FUTURE WORK

In this paper, we use the Audience Benchmark dataset to classify gender. Here we use CNN algorithm to classify and predict the gender. CNN algorithm contains many number of hidden layers. The accuracy of CNN is 98.4% when compared to SVM. Here the image used is coloured image and coloured is taken as the input and fed into CNN and Input layer. Input layer send the input image into hidden layer and hidden layer process all kind of input to provide the output. For the classification, process SVM is better than CNN and For prediction CNN is better than SVM . For identification process , the accuracy of CNN is better than SVM and also SVM takes more amount of time than CNN. so in future SVM algorithm for classification and CNN algorithm for prediction can be used to improve the accuracy. For image validation, we use an image array to append all the images. But it is not possible to store 1GB of images for validation because the ram will corrupt it. So in future work, we will find a better method for image validation. In Image recognition combining these two algorithms for face, detection will provide a more accurate to predict this model.

VI REFERENCE

[1] H. Khalajzadeh, M. Mansouri and M. Teshnehlab, "Face Recognition Using Convolutional Neural Network and Simple Logistic Classifier", Advances in Intelligent Systems and Computing, vol. 223, pp. 197-207, 2013.
[2] S. Xie and H. Hu, "Facial expression recognition with FRRCNN", Electronics Letters, vol. 53, no. 4, pp. 235-237, 2017.
[3] O. Ayodeji Arigbabu, S. Mumtazah Syed Ahmad, W. Azizun, W. , S. Yussof and S. Mahmood, "Soft Biometrics: Gender

Recognition from Unconstrained Face Images using Local Feature The descriptor", Computer Vision and Pattern Recognition, pp. 1-9, 2017.
[4] Chengjun Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," in IEEE Transactions on Image Processing, vol. 11, no. 4, pp. 467-476, Apr 2002.
[5] E. Gonzalez-Sosa, J. Fierrez, R. Vera-Rodriguez and F. Alonso Fernandez, "Facial Soft Biometrics for Recognition in the Wild: Recent Works, Annotation, and COTS Evaluation," in IEEE Transactions on Information Forensics and Security, vol. 13, no. 8, pp. 2001-2014, Aug.2018.
[6] G. Levi and T. Hassncer, "Age and gender classification Using convolutional neural networks," 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Boston, MA, pp. 34-42,2015.
[7] A. Toshev and C. Szegedy, "DeepPose: Human Pose Estimation via Deep Neural Networks", in IEEE Conference on Computer Vision and Pattern Recognition, pp. 1653-1660.,2014.
[8] D. Reid, S. Samangooei, C. Chen, M. Nixon, and A. Ross, "Soft biometrics for surveillance: an overview", in Handbook of Statistics, vol. 31, pp. 327-352,2013.
[9] G. Antipov, M. Baccouche, S. Berrani and J. Dugelay, "Effective training of convolutional neural networks for face-based gender and age prediction", Pattern Recognition, vol. 72, pp. 15-26, 2017.
[10] C. Shan, "Learning local binary patterns for gender classification on real-world face images", Pattern Recognition Letters, vol. 33, no. 4, pp. 431-437, 2012.
[11] V. Mirjalili and A. Ross, "Soft Biometric Privacy: Retaining Biometric Utility of Face Images while Perturbing Gender", in Proc. of International Joint Conference on Biometrics, 2017, pp. 1-10.
[12] P. Phillips, H. Wechsler, J. Huang, and P. Rauss, "The FERET database and evaluation procedure for face-recognition algorithms", Image and Vision Computing, vol. 16, no. 5, pp. 295-306, 1998.
[13] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet classification with deep convolutional neural networks", Communications of the ACM, vol. 60, no. 6, pp. 84-90, 2017.
[14] G. Antipov, S. Berrani, and J. Dugelay, "Minimalistic CNN-based ensemble model for gender prediction from face images", Pattern Recognition Letters, vol. 70, pp. 59-65, 2016.
[15] G. Huang, M. Ramesh, T. Berg, and E. Learned-Miller, "Labeled faces in the wild: A database for studying face recognition in unconstrained environments", University of Massachusetts, Amherst, 2007.
[16] M. D. Coco, P. Carcagnì, M. Leo, P. L. Mazzeo, P. Spagnolo and C. Distance, "Assessment of deep learning for gender classification on traditional datasets," 2016 13th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), Colorado Springs, CO, pp. 271-277,2016.
[17] C. Szegedy et al., "Going deeper with convolutions," 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, pp. 1-9,2015.
[18] D. Jaswal, S. V and K. Soman, "Image Classification Using Convolutional Neural Networks", International Journal of Scientific and Engineering Research, vol. 5, no. 6, pp. 1661-1668, 2014.