Convolutional Neural Network Based Gender Classification From Facial Images

Miss. Vaishnavi Y. Mali[1], Dr. Babasaheb G. Patil[2]
[1, 2] Dept of Electronics
[1, 2] Walchand College of Engineering,Sangli

Abstract- *Gender classification has become applicable to a huge amount of demand in the area of human and computer interactions. Human gender is become the interested field in the research area. For each intelligent applications, machine learning become crucial part in today's world. This paper explores a deep learning based solutions for gender classification from face images. For this we have created a dataset which contains frontal, left and aligned, smiling, non-smiling as well as expression images. Method includes transfer learning framework where information is reused out of deep learning model. Our aim is to inspect inceptionv3 model, the high resolution classifier which is based on convolutional neural network.*

Keywords- Convolutional neural network, Face recognition, Gender Classification, Machine Learning

I. INTRODUCTION

Gender recognition is the gateway device for any approach to improve their functionality by doing targeted interaction with the gender of its preference. Automatic age and gender classification has become applicable to rise in amount of applications, like social platforms and social media.

Gender classification can be beneficial in the area of business intelligence. For advertising or better customer service, it is essential to understand in addition about the customers, in terms of their age, gender, ethnicity, etc. So gender attributes can be merged with age to group customers into distinct groups. The dispute is how to boost the gender classification accuracy when the image quality is not fine and the image resolution is insignificant. Gender classification probably gainful for online image filtering or database organization of images. To well assemble the enormous images and video data, gender might play a vital role by splitting the images and videos into two classes, male-centered or female- centered. In Japan, vending machines that use age and gender information of customers to recommend drinks have spotted increased sales.

Supportive information from gender classification to upgrade the user experience in mobile applications and video games. In mobile apps, some analyzers utilize this method to promote the use of the mobile Internet by designing apps as per gender. In video games, males and females usually have unique favors, which would allow the use of gender knowledge to provide their preferred content.Earlier knowledge of the application domain is necessary to grasp the finest feature extractor to design.

Furthermore, the performance of the recognition system is highly dependent on the kind of classifier taken, which is based on the feature extraction method applied. It is hard to find a classifier that merges the best with the chosen feature extractor such that an optimal classification performance is achieved. Any small change to the problem domain need a complete redesign of the system.

After speedy development of artificial intelligence, face feature has grandstand due to its non-intrusive identity and since it is major method of identification for human when it is compared with other variety of biometric techniques. As compared to existing machine learning techniques, deep learning based approaches have shown better results in terms of accuracy and processing speed in image recognition.

II. LITURATURE REVIEW

Various research papers that are connected to gender classification for distinct types of databases have been reviewed to recognize numerous ways for gender classification problems. Furthermore, the varying classification analyses that are explained in research paper have been viewed to get ideas on how the classification is served.

Sandeep Kumar et al. [1] proposed recognition algorithm for face images. In the algorithm, SIFT (Scale Invariant Fourier Transform) and SVM classifier (Support Vector Machine) were utilized for feature extraction and classification. The SVM classifier for a textual description of facial characteristic attributes such as K-means clustering totally based on the normalized face. Live images, FEI and SCIEN databases were compared for better performance results.

KyoungsonJhang et al. [2] provided a face photo based gender and age prediction method using camera. It processed the RGB colour images to predict the gender because of good results with RGB colour images and testing was done with the help of camera instead of image files. Googlenet was employed for age and gender classification which contains good prediction accuracy. Though, CNN's trained with grayscale images exhibited better prediction accuracy in camera based testing and model learned with RGB colour images displayed well classification accuracy in file based testing.

The paper [3] represented gender recognition from facial image with the help of CNN. The method explored the model of transfer learning for training using deep learning on a small length dataset of facial images. The method investigated the reusability of VGG16 paradigm which was pre-trained upon natural images dataset. However fine-tuned network should yet to be limited.

Lin Yuan et al. [4] employed a gender identification at individual level using human brain images with a novel 3D descriptor and 3DWHGO. They used MRI scans of a large sample of healthy adults which distinguish the gender of person. Presented way gave well accuracy as compared to 3D PCA and SIFT method. However in case of damaged brain, the method was not efficient.

Philip Smith et al. [5] proposed deep learning network with pre-trained weights were used for image based gender recognition and age estimation. Transfer learning was traversed using both VGG19 and VGGface pre-trained model by testing transforms in diverse design schemes and trained parameters to upgrade prediction accuracy. For training, they were used the input standardization, augmentation and label distribution encoding. Proposed model obtained accuracy rate of 98.7%.

Shuo Yang et al. [6] proposed a deep CNN (Convolutional Neural Network) for face detection on facial attributes. The result stimulated a method for finding a faces via scoring of facial parts replied by spatial arrangements and its structure. However, the results were compared by using FDDB, WIDER face, AFW and PASCAL faces.

Sergiu et al. [7] introduced an automatic gender recognition from facial images based on convolutional neural network. Author used several face databases and annotated approximately 70,000 facial images. The results was achieved with the help of InceptionV4 architecture. For recognition of gender, InceptionV4 gave 98.2% accuracy on their own dataset and 84% accuracy on Adience dataset.

Christian Szegedy et al. [8] presented a deep CNN architecture named Inception attained for classification and detection in ImageNet. Googlenet networks were trained for gender classification. The setup and result details of ILSVRC 2014 classification challenge involved in the paper.

In the paper [9], the analysis of different machine learning algorithms were discussed. To select the best algorithm, author conducted experimental study on several algorithms like KNN, decision tree, logistic regression etc. for classification of gender. The performance was analyzed using voice dataset. Parameter tuning gave accuracy of 98.6% with SVM and 99.8% with ANN.

III. EXPERIMENTATION

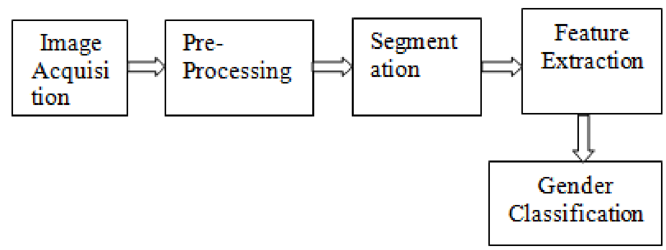

Fig.1 Block Diagram of human gender classification

The Fig.1 shows block diagram of human gender classification which includes Acquisition, pre-processing, feature extraction, classification. The steps involved in the human gender classification are as follows:

Step 1: START
Step 2: Collect dataset i.e. importing the image dataset for training
Step 3: Splitting Dataset into train and test data.
Step 4: Pre-processing of the dataset.
Step 5: Define methods for training and create the model using CNN.
Step 6: Compilation of the model with selected parameters.
Step 7: Model fitting.
Step 8: Classify the gender and evaluate it.
Step 9: STOP.

A. Image Acquisition:

ML depends mainly on data, without data, it is impossible for an "AI" to learn. A data set is a collection of data. It is the actual dataset used to train the model for performing various actions. The simplest way to describe any machine learning project is that it is a simple program that when given unknown pieces of data, it will process them based on previous experience and gives output something that user did not already known.

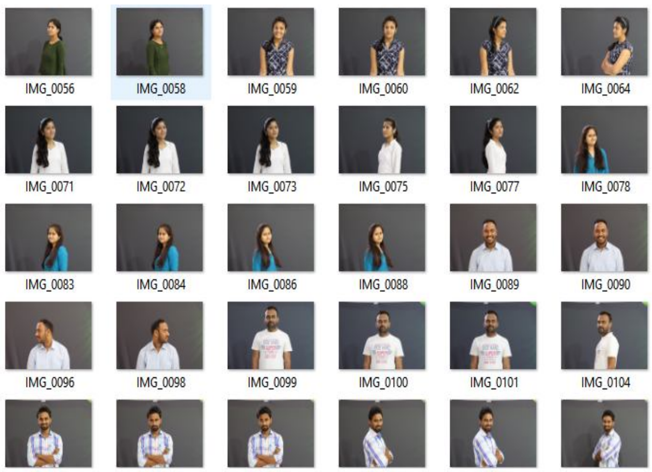

Fig.2 Gender classification dataset

Human gender classification is one of the most interested and critical area of research. We are using 360 severalcombinations of face images. Images consist of frontal, aligned, smiling, non-smiling as well as expression images. In particular, three data sets are commonly used in different stages of the creation of the model.

The model is trained on the training dataset using a supervised learning method. This data set is used to adjust the weights on the neural network. The training data is a dataset of examples used for learning. If we are using large number of data i.e. more features for training, it leads to more accuracy and causes overfitting. Validation set is the set of examples used only to tune the parameters (i.e., architecture, not weights) of a classifier, for example to choose the number of hidden units in a neural network. This data set is used to minimize overfitting. Test dataset is independent of training dataset and also called as holdout dataset. The test set is a set of observations used to evaluate the performance of the model using some performance metric. It is important that no observations from the training set are included in the test set.

B. Image Pre-processing:

Pre-processing is the essential procedure to upgrade the quality of raw data where not only training but also test images will be normalized to reduce noise from images. Pre-processing refers to the transformations applied to our data before feeding it to the algorithm.Pre-processing is not a single step. It includes number of operations as mentioned below:

1) First step is usually importing the libraries that will be needed in the program. A library is a collection of modules that can be called and used.NumPy is probably the most fundamental package for scientific computing in Python. It provides a highly efficient interface to create and interact with multi-dimensional arrays.Matplotlib is a library to produce high-quality and interactive two-dimensional plots. Google. It is called Tensorflow because it takes input as a multi-dimensional array, also known as **tensors.** Tensor flow library incorporates different API to build at scale deep learning architecture like CNN or RNN.

2)Create an image dataset for the purposes of gender classification. Each folder in the dataset, one for testing, training, and validation, has images that are organized by class labels.

- Create a folder named male_female.
- Create three subfolders named Train_Data, Test_Data, Validation_Data which are used for training, testing and validation respectively. Since we have used male, female create folders to indicate the types of documents which will be the classes that machine learning model learns and classifies images.
- Provide the path of created folder and import the dataset.

3) Generally dataset is divided into 80% of training set and 20% of test set respectively. Percentage of data entirely depends on user.

4) When our data is comprised of attributes with varying scales, many machine learning algorithms can benefit from rescaling the attributes to all have the same scale. This is useful for optimization algorithms in used in the core of machine learning algorithms like gradient descent. The ImageDataGenerator class can be used to rescale pixel values from the range of 0-255 to the range 0-1 preferred for neural network models.

5) flow_from_directory() generates augmented images from directory with arbitrary collection of images. So there is need of parameter target_size() to make all images of same shape. The target_size can be either (width, height) or (height, width).

6) In machine learning (ML), we are often presented with a dataset that will be further split into training, testing & validation datasets. It is very important that dataset is shuffle well to avoid any element of bias/patterns in the split datasets before training the ML model.

C. Image Extraction and Classification using CNN:

Transfer learning is a machine learning method which utilizes a pre-trained neural network. The image recognition model called Inceptionv3 consists of two parts: Feature extraction part with a convolutional neural network and Classification part with fully-connected and softmax layers. Transfer learning allows to retrain the final layer of an existing model, resulting in a significant decrease in not only training time, but also the size of the dataset required. One of the most famous models that can be used for transfer learning is Inception V3. Main difference between CNN and Inception-v3 is Inception-v3 has a 48 layer convolutional network, while our CNN has only two convolutional layer.

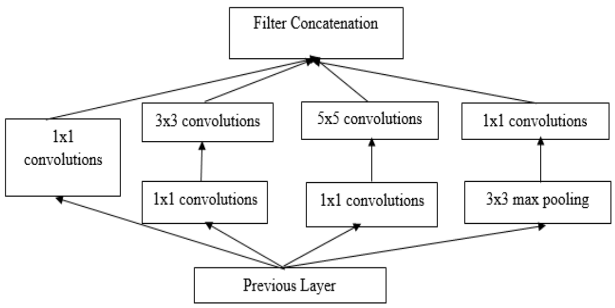

Fig.3 Inceptionv3 for dimensionality reduction

Inception architecture is used to account how an optimal local sparse structure of a convolutional vision network can be approximated and covered by readily available dense components.Inception-v3 is Deep Neural Network architecture that uses inception blocks. In order to avoid patch-alignment issues, current incarnations of the Inception architecture are restricted to filter sizes 1×1, 3×3and 5×5.i.e. architecture is a combination of all those layers with their output filter banks concatenated into a single output vector forming the input of the next stage. In na¨ıve architecture, modest number of 5×5 convolutions can be prohibitively expensive on top of a convolutional layer with a large number of filters. This problem introduced once pooling units are added to the mix the number of output filters equals to the number of filters in the previous stage. The merging of output of the pooling layer with outputs of the convolutional layers would lead to an inevitable increase in the number of outputs from stage to stage. In dimensionality reduction architecture, 1×1 convolutions are used to compute reductions before the expensive 3×3 and 5×5 convolutions. Besides being used as reductions, they also include the use of rectified linear activation. Inception network is a network consisting of modules of the above type stacked upon each other, with occasional max-pooling layers with stride 2 to halve the resolution of the grid. The main characteristics is it allows for increasing the number of units at each stage significantly without an uncontrolled blow-up in computational complexity at later stages and attained by the ubiquitous utility of dimensionality reduction prior to expensive convolutions with larger patch sizes.

Convolutional Neural Networks (CNNs) is the most popular neural network model being used for image classification problem. Having fewer parameters greatly improves the time it takes to learn as well as reduces the amount of data required to train the model. Instead of a fully connected network of weights from each pixel, a CNN has just enough weights to look at a small patch of the image. The purpose of the convolution is to extract the features of the object on the image locally. It means the network will learn specific patterns within the picture and will be able to recognize it everywhere in the picture. Convolution is an element-wise multiplication.

At the end of the convolution operation, the output is subject to an activation function to allow non-linearity. The usual activation function for convnet is the Relu. All the pixel with a negative value will be replaced by zero. The next step after the convolution is to down sample the feature max. The purpose is to reduce the dimensionality of the feature map to prevent overfitting and improve the computation speed. Max pooling is the conventional technique, which divides the feature maps into sub regions (usually with a 2x2 size) and keeps only the maximum values. The last step consists of building a traditional artificial neural network.Connect all neurons from the previous layer to the next layer. Then use a softmax activation function to classify the number on the input image. All neurons from the previous layers are connected to the next layers. The CNN will classify the label according to the features from the convolutional layers and reduced with the pooling layer. Final pooling layer feeding the features to a classifier useCsSoftmax activation function. The sum of output probabilities from the Fully Connected Layer is 1. This is ensured by using the Softmax as the activation function. The Softmax function takes a vector of arbitrary real-valued scores and squashes it to a vector of values between zero and one that sum to one.

IV. RESULTS

Model Evaluation is an integral part of the model development process. It helps to find the best model that represents our data and how well the chosen model will work in the future. Evaluating model performance with the data used for training is not acceptable in data science because it can easily generate overoptimistic and overfitted models. There are two methods of evaluating models, Hold-Out and Cross-Validation. Normalization tries to improve the contrast by stretching the intensity values of an image to fill the entire dynamic range. Scaling data to the range of 0-1 is traditionally referred to as normalization. Min-max normalization is one of the most common ways to normalize data. For every feature, the minimum value of that feature gets transformed into a 0, the maximum value gets transformed into a 1, and every other value gets transformed into a decimal between 0 and 1. A class prediction is given the finalized model and one or more data instances, predict the class for the data instances. Model predict the class for new data instances using our finalized classification model in Keras using the predict classes() function. In two-class (binary) classification problem, the sigmoid activation function is often used in the output layer. The predicted probability is taken as the likelihood of the observation belonging to class 1, or inverted (1 – probability) to give the probability for class 0.

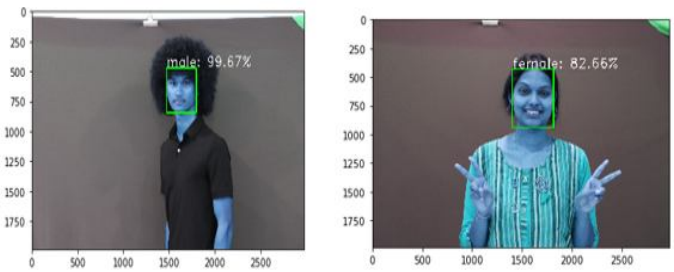

Fig.4 Gender recognition for known dataset using bounding box

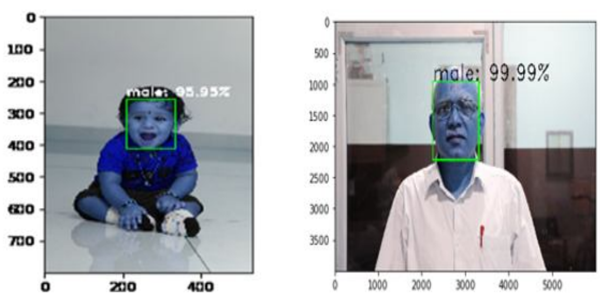

Fig.5 Gender recognition for unknown dataset using bounding box

V. CONCLUSION

This project is intended to build the system that recognizes gender of individual. This method is used to test the model for different images which consists of different styles. Preprocessing is a technique that is used to convert the raw data into a clean data set. Convolutional Neural Networks (CNNs) is the most popular neural network model being used for image classification problem. Having fewer parameters greatly improes the time it takes to learn as well as reduces the amount of data required to train the model. For achieving better results from the applied model in Machine Learning projects the format of the data has to be in a proper manner.During the convolutional part, the network keeps the essential features of the image and excludes irrelevant noise. During ReLU, All the pixel with a negative value will be replaced by zero.The purpose of pooling is to reduce the dimensionality of the feature map to prevent overfitting and improve the computation speed. Softmax activation function to classify the number on the input image. All of the models are pre-trained on ImageNet. According to the experiments, all of the models perform well. Image based gender and age group prediction, we compared CNNs in two settings, i.e. file-based testing and camera-based testing.

REFERENCES

[1] Sandeep Kumar, Sukhwinder Singh and JagdishKumar, “Gender Classification Using Machine Learning with Multi-Feature Method”,978-1-7281-0554-3/19/ @ 2019 IEEE.

[2] KyoungsonJhang, Junsoo Cho, “CNN Training for Face Photo based Gender and Age Group Prediction with Camera”, 978-1-5386-7822-0/19/$31.00 ©2019 IEEE.

[3] Shubham Mittal, Shiva Mittal,”Gender Recognition from Facial Images using Convolutional Neural Network”, 978-1-7281-0899-5/19/$31.00 ©2019 IEEE.

[4] Lin Yuan, Fanglin Chen, Ling-Li Zeng, Lubin Wang, and DewenHu,“Gender Identification of Human Brain Image with A Novel 3D Descriptor”,IEEE/ACM Transactions On Computational Biology And Bioinformatics, VOL. 15, NO. 2, MARCH/APRIL 2018.

[5] Philip Smith, Cuixian Chen, “Transfer Learning with Deep CNNs for Gender Recognition and Age Estimation”, 2018 IEEE International Conference on Big Data.

[6] Shuo Yang, Ping Luo, Chen Change Loy, “Faceness-Net: Face Detection through Deep Facial Part Responses”, IEEE Transactions on Pattern Analysis and Machine Intelligence, DOI 10.1109/TPAMI.2017.2738644.

[7] Sergiu Cosmin Nistor, Alexandra-Cristina Marina, Adrian Sergiu Darabant, Diana Borza, “Automatic gender recognition for “in the wild” facial images using convolutional neural networks”, 978-1-5386-3368-7/17/$31.00 © 2017 IEEE.

[8] Christian Szegedy1, Wei Liu2, Yangqing Jia1, Pierre Sermanet1, Scott Reed3, “Going Deeper with Convolutions”, 978-1-4673-6964-0/15/$31.00 ©2015 IEEE.

[9] Laxmi Narayana Pondhu, GovardhaniKummari, “Performance Analysis of Machine Learning Algorithms for Gender Classification”, 2nd International Conference on Inventive Communication and Computational Technologies (ICICCT 2018)

[10] Jinhyeok Jang, Hyunjoong Cho, Jaehong Kim, Jaeyeon Lee, andSeungjoon Yang, “Facial Attribute Recognition by RecurrentLearning With Visual Fixation”,IEEE

TRANSACTIONS ON CYBERNETICS, VOL. 49, NO. 2, FEBRUARY 2019.

[11] Juan E. Tapia And Claudio A. Perez, “Clusters of Features Using Complementary Information Applied to Gender Classification From Face Images”,IEEE Access, VOLUME 7, 2019.

[12] George Azzopardi, Antonio Greco, AlessiaSaggese and Mario Vento, “Fusion of Domain Specific and Trainable Features for Gender Recognition from Face Images”,IEEE Access, VOLUME 6, 2018.

[13] Ahmad Jamal KHDR and Cihan VAROL,“Age and Gender Identification by SMS Text Messages”,978-1-5386-68788/18/©2018 IEEE.

[14] Juan E. Tapia, Claudio A. Perez and Kevin W. Bowyer, “Gender Classification from the Same Iris Code Used for Recognition”, IEEE Transactions on Information Forensics and Security, Vol. 11, No. 8, August 2016.

[15] Jiale Chen, Sen Liu, Zhibo Chen, “Gender Classification in Live Videos”,978-1-5090-2175-8/17/@2017 IEEE.

[16] Shannon Silessi, CihanVarol, Murat Karabatak, “Identifying Gender from SMS Text Messages”, 15th IEEE International Conference on Machine Learning and Applications, 2016.

[17] LahoucineBallihi, Boulbaba Ben Amor, Mohamed Daoudi,“Boosting 3-D-Geometric Features for Efficient Face Recognition and Gender Classification”, IEEE Transactions on Information Forensics And Security, VOL. 7, NO. 6, DECEMBER 2012.

[18] Huaining Cheng, and Kathleen M. Robinette, “Gender Recognition Using 3-D Human Body Shapes”, IEEE Transactions on Systems, Man, and Cybernetics—Part C: Applications and Reviews, VOL. 41, NO. 6, NOVEMBER 2011.

AUTHORS PROFILE

Miss. Vaishnavi Yashwant Mali: She is M-Tech student of Walchand College of Engineering, Sangli, Maharashtra, India in Electronics. She completed her graduation in Electronics and Telecommunications from PVPIT, Budhgoan. She has keen interest in the field of Image Processing.

Dr. Babasaheb G. Patil: He received his M.E. Electronics degree in 1990 and B.E. Electronics in 1988 and Ph.D from Shivaji University, Kolhapur in 2015.He is currently working as Associate Professor in department of Electronics, in Walchand College of Engineering, Sangli, Maharashtra, India.He has guided numerous students for graduation and Post-graduation. He is having keen interest in image processing and Communication.