May-June 2020
ISSN: 0193-4120 Page No. 4589-4592

Age, Gender and Emotion Detection using CNN

[1]Manasa SB, [2]Jeffy.S.Abraham, [3]Anjali Sharma, [4]Himapoornashree KS, [5]Sunil Manoli
[5]Professor
[1,2,3,4,5]Computer Science and Information Technology, Reva University, Bangalore, India
[1]manasasb17@gmail.com, [2]jeffysabraham52@gmail.com, [3]anjisharma81@gmail.com ,
[4]himapoornashree@gmail.com

Article Info
Volume 83
Page Number: 4589-4592
Publication Issue:
May-June 2020

Article History
Article Received:*19 November 2019*
Revised: *27 January 2020*
Accepted: *24 February 2020*
Publication: *12 May 2020*

Abstract

Face is one of the most dominant features in our body. We can get many information like age, gender, etc by analyzing the face of a human. In today's world, computer vision is been used to train machines to comprehend and understand the real world. Using several digital images from webcam, videos, cameras and with the help of deep learning, computers can correctly figure out and classify objects and then respond to what they "see" in real world. There are various uses of identifying age and gender from face like forensic testing, restricting access of alcohol from wending machine and adult content for young people. Emotion from a face can be used to predict human computer interaction, students/teacher's interest in class, advertisement bots etc.

Keywords: *CNN; Residual neural network; RELU -Rectified Linear Unit*

1. Introduction

Face detection is a method of recognizing if there is a face in an image or not which is also called as object detection. OpenCV is widely used for object detection. The goal of face emotion analysis is to detect and identify the different types of emotions of a person such as happy, sad, anger, confused, surprise, etc. A gender categorizing model uses face from a given image to predict the gender (male or female) based on their appearance like baldness, long hair, beard and must ache. In age classification, we classify based on wrinkles, hair color and also size of face etc. Deep learning which is a subset of machine learning that uses several layers neural network to repeatedlygain higher level of features from the given input images. The concept of deep learning was inspired by how neurons function in our brain hence it's also called deep neural network. A neural network is a sequence of process that is capable to identify hidden relationships in a set of data and the process is similar to the operation of human brain. We use a deep learning model called convolutional neural network which takes input images and allocate importance (learnable weights and biases) to different aspects/objects in an image and also capable of differentiating one from the other. When compared to other classification algorithms, the pre-processing required in a ConvNet is comparatively lesser.

2. Literature Survey

Imane et al. [1] proposed a face detection system using HAAR cascades, normalization and emotion detection using CNN on FER 2013 (KNN) for classifying and Uniform Local Gabor Binary Pattern Histogram Sequence (ULGBPHS) for pattern scanning. The model used 4 different machine learning techniques(SVM, KNN, Random Forest and Classification & Regression Trees) and good accuracy rates i.e. 70% at 106 epochs was obtained by using KNN and SVM algorithms. This model can be made better by using no machine language algorithms as well.

A better proposition is done by Rajesh et al. [2]. Here, real time emotion detection used CNN with 9 layers for training and categorizing 7 different types of emotions which gives an accuracy of around 90%.

Sepidehsadat et al. [3] proposed that the use of a Gabor filter will make it easier for the network to focus on the face as the output orientations match the face wrinkles perfectly which in turn will be an input to the CNN. The network focuses on the useful features giving a 7 percent age-accuracy and 2 percent gender accuracy.

Ari Ekmekji [4] developed a model that built-in inter-relationships linking age and gender to bind these architectures to improve overall outcome. The shortcomings were the difficulty in dividing the data into folds, train each classifier, cross-validate, and merge the resulting classifiers into a test-ready classifier.

May-June 2020
ISSN: 0193-4120 Page No. 4589-4592

3. Working

Keras library is aneural network library and also a high level API which is open source written using python and which is built on tensor flow. Keras has various functions which is mainly used for neural-network building blocks such as optimizers, objectives, layers and has in-built tools which makes processing of the image much easier and also to reduce the coding part which is required for implementing deep neural network code.

A. Face Detection:

Haar cascade XML file which is a classifier used to identify a specific object from the webcam. The haarcascasde_frontalface_default.xml provided by OpenCV used to recognize frontal face. OpenCV connects to the webcam which user can use to scan their faces for classification of age, gender and emotion.

B. Emotion detection:

FER2013 is a Kaggle dataset that contains labeled 3589 test images,28709 train images. We don't have to do data augmentation because the dataset has been built with wide range of images. The database holds grayscale pictures of human faces. We don't use transfer learning because our dataset contains grayscale images and doesn't fit in 3 channels pre-trained Models. We use 3 convolutional layers. Input [48x48x1] carries the pixel values of given image. Hence images have width equal to 48, height equal to 48, and with one color channel.

- Step 1: Normalizing the data between 0 and 1.
- Step 2: We use 3 convolutional layers. In each layer, we do Batch Normalization, RELU activation function and use MaxPooling. In fully connected layer we use RELU activation function and SOFTMAX function.
- Step 3: Calculate the loss function using Adam optimizer
- Step 4: to use the trained model later, save the weights infer.h5.

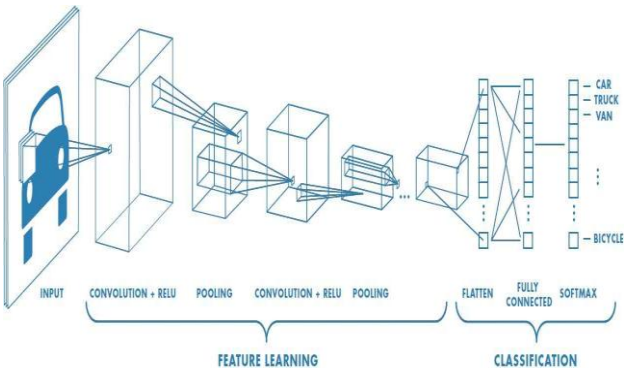


Figure 1: Workflow of convolutional neural network

C. Age and Gender Classification:

Here we use an hdf5 file as dataset which has 10,000 images.

The images have been labeled and classified into male, female and various age groups. Here we use resnet neural network architecture to train this model. ResNet is used to add a large number of layers with strong performance.

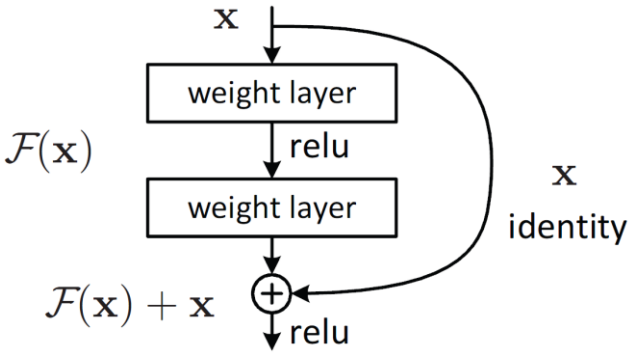
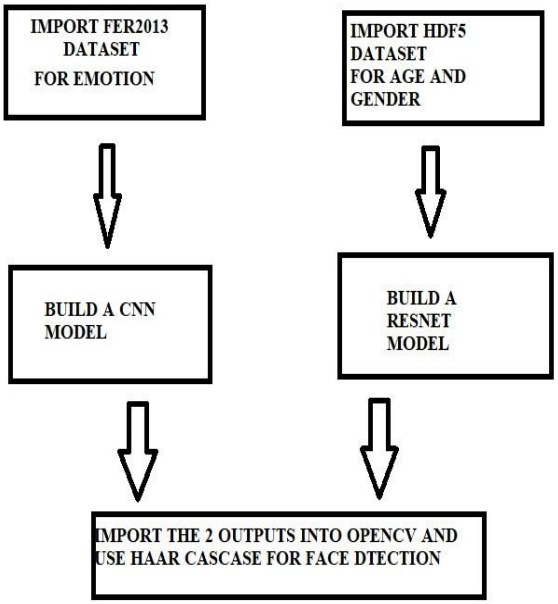


Figure 2: Single residual block

A network containing residual blocks, each layer is an input to the next layer and also inputs into the layers about 2-3 hops away. If the number of layers keeps increasing, we will notice that the accuracy will tend to saturate at certain point and after a point it will gradually decrease. This is not just because of over fitting. It indicates that shallow networks are better at learning than deeper networks. So the skip-connections skips the training of few layers and are also called residual connections as shown in above image. By including skip connections in our network architecture, we are making the network to skip training for the layers which are not useful and don't add much value in overall accuracy.

Blueprint of our model implementation



TEST Engineering & Management

May-June 2020
ISSN: 0193-4120 Page No. 4589-4592

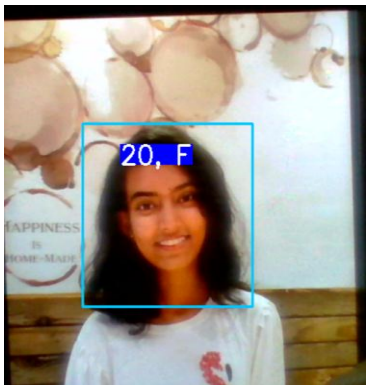
4. Conclusion

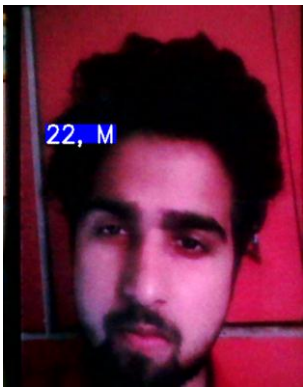
Analysis of age, gender and emotion of face was done on real time using webcam. Training the model at higher epoch value yields better result. We have used resnet architecture instead of VGG16 for age or inception v3 for gender classification. Resnet helps in handling the training error generated as the networks get deeper. We were able to achieve 80% accuracy rate. Replacing VGG-16 layers in Faster R-CNN with ResNet, we can observe a relative improvement of 28%.

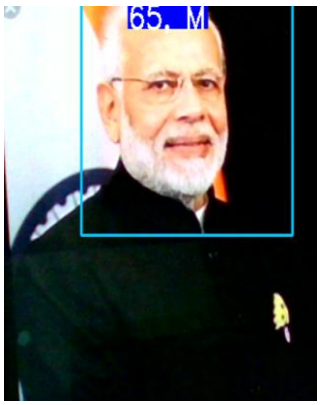
5. Result

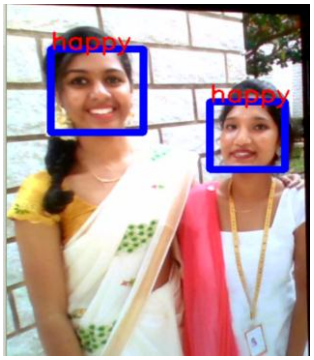
We could predict accuracy of 80% in age, gender and emotion detection.

Output:









6. Future Scope

1. Further, this model can be used to classify patients and their medicines based on age groups.
2. It can also be used for movie recommendation system in order to predict the frequency of various age

May-June 2020
ISSN: 0193-4120 Page No. 4589-4592

groups who come to watch movies the most and also to classify movies into different genres based on the viewers' emotion.

References

[1] Imane Lasri, Anouar Riad Solh & Mourad El Be lkacemi, "Facial Emotion Recognition of Students using Convolutional Neural Network", IEEE-2019

[2] Rajesh Kumar G A, Ravi Kant Kumar & Goutam Sanyal, "Facial Emotion Analysis using Deep Convolutional Neural Network", 2017 International Conference on Signal Processing and Communication (ICSPC). doi:10.1109/cspc.2017.8305872, Pg.No.-369 to 374

[3] 2018-Sepidehsadat Hosseini,Seok Hee Lee, Hyuk Jin Kwon,Hyung Il Koo &Nam Ik Cho, "Age and Gender Classification Using Wide Convolutional Neural Network and Gabor Filter",IEEE-2018

[4] Ari Ekmekji", Convolutional Neural Networks for Age and Gender Classification", Stanford University-2016

[5] Levi, G, & Hassncer, T, "Age and gender classification using convolutional neural networks", 2015 IEEE Conference on Computer Vision and PatternRecognition Workshops (CVPRW). doi:10.1109/cvprw.2015.7301352, Pg no.-34to 42

[6] Bogdan Kwolek, "Face Detection Using Convolutional Neural Networks andGabor Filters", Conference Paper- September 2005, Pg.No. - 1 to 5