International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

AI TO DESIGN A MASK INSENSIBLE TO THE DISTANCE FROM CAMERA TO THE SENSE OBJECTS

[1]Sharanabasavaraj H Angadi[2]Manjula Mashyal

Associate Professor Tech Fortune Technology Bengaluru

Rural Engineering College,Hulkoti_KarnatakaEmail:**Manjushreemashal922@gmail.com**

Email:shangadicse@gmail.com

ABSTRACT

Face Detection has evolved as a very popular problem in Image processing and Computer Vision.. The convolutional architectures have made it possible to extract even the pixel details. In this paper,a binary face classifier is designed to detect any face present in the frame irrespective of its alignment. To do so, a method is proposed to generate accurate face segmentation masks from any arbitrary size input image. Beginning from the RGB image of any size, the method uses Predefined Training Weights of VGG – 16 Architecture for feature extraction. Training is performed through Convolutional Neural Networks to semantically segment out the faces present in that image. Gradient Descent is used for training while Binomial Cross Entropy is used as a loss function. Further the output image from the CNN is processed to remove the unwanted noise and avoid the false predictions if any and make bounding box around the faces. Furthermore, proposed model has also shown great results in recognizing non frontal faces. Along with this it is also able to detect multiple facial masks in a single frame. Experiments were performed on Multi Parsing Human Dataset obtaining mean pixel level accuracy of 93.884 % for the segmented face masks.

Index Terms—Convolutional Neural Network, Semantic Segmentation, Face Segmentation and Detection.

INTRODUCTION

Face detection has emerged as a very interesting problem in image processing and computer vision. It has a range of applications from facial motion capture to face recognition. Face detection is more relevant today because it not only used on images but also in video applications like real time surveillance and face detection in videos. High accuracy image classification is possible now with the advancements of Convolutional networks. Pixel level information is often required after face detection which most face detection methods fail to provide. Obtaining pixel level details has been a challenging part in semantic segmentation. Semantic segmentation is the process of assigning a label to each pixel of the image. In our case the labels are either face or non-face. Semantic segmentation is thus used to separate out

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

the face by classifying each pixel of the image as face or background. Also most of the widely used face detection algorithms tend to focus on the detection of frontalfaces.
This Paper proposes a model for face detection using semantic segmentation in an image by classifying each pixel as face and non-face i.e. effectively creating a binary classifier and then detecting that segmented area. The model works very well not only for images having frontal faces but also for non-frontal faces. The paper also focuses on removing the erroneous predictions which are bound to occur. Semantic segmentation of human face is performed with the help of a Convolutional Neural Network. The next section discusses the related work done in the domain of face detection. In section III we describe the method followed for face segmentation and detection using semantic segmentation on any arbitrary RGB image. Finally, the generated facial masks are demonstrated in experimental results in section IV. Post processing on the predicted images has also been discussed at length which also entails the removal of erroneous predictions.

LITERATURE SURVEY

Sanjay Kumar, AshishNegi, J.N Singh; HimanshuVerma[1] Attractive significance image base medicinal picture study and how to improve detection of brain tumours MRI be achievement thought inside current period outstanding towards augmented require of competent with accurate reports using semantic segmentation

Xiaomeng Fu; HuimingQu, proposed a segmentation method based on full convolutional neural network (CNN) based on the semantic segmentation of high-resolution remote sensing images. The method improves the traditional convolutional neural network (CNN) replaces the final fully connected layer of the CNN network with a convolutional layer.

In biomedical image processing, typical application of deep learning semantic segmentation. However, classical deep learning methods require hardware consumption, computational costs. In order to resolve problem, Kaiyue Li; Guangtai Ding; Haitao Wang propose a new lightweight convolutional Neural network.

Xuelei Li devoted to accelerate this promising framework in the inference application. FPGA (Field Programmable Gate Array) using OpenCL programming language. Firstly, a supplemented deep learning accelerator is constructed to realize the residual. function in ResNet. Secondly, our construction reschedules three on-chip buffers in order to store the feature data and to stream it to processor elements alternately. In addition, we also implement data parallel and pipeline execution such that the filter parameters can be synchronously processed with the image data on FPGA

Christian Szegedy; Wei Liu; YangqingJia; Pierre Sermanet, Scott Red proposed deep convolutional neural network architecture codenamed Inception that achieves the new state of the art for classification and detection in the ImageNet Large- Scale Visual Recognition Challenge 2014 (ILSVRC14). The main hallmark of this architecture is the improved utilization of the computing resources inside the network. By a carefully crafted design, we increased the depth and width of the network while keeping the computational budget

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

constant. To optimize quality, the architectural decisions were based on the Hebbian principle and the intuition of multi-scale processing. One particular incarnation used in our submission for ILSVRC14 is called GoogLeNet, a 22 layers deep network, the quality of which is assessed in the context of classification and detection.

In the workof K Andrew ZisSer man investigated the effect of convolutional network depth on its accuracy in large-scale image recognition setting. Their contribution is a thorough evaluation of networks of increasing depth using an architecture with very small (3x3) convolution filters, which shows that a significant improvement on the prior-art configurations can be achieved by pushing the depth to 16-19 weight layers. These findings were the basis of our ImageNet Challenge 2014 submission, where our team secured the first and the second places in the localisation and classification tracks respectively. We also show that our representations generalise well to other datasets, where they achieve state-of-the-art results. We have made our two best-performing ConvNet models publicly available to facilitate further research on the use of deep visual representations in computer vision.
P.Viola and M.J.Jonesdescribes and discusses the algorithms required to perform face detection and face recognition in real-time. Simple features, similar to Haar basis functions, are used for detection and the eigenfaces technique is used for recognition. Further to the above, a novel method of increasing face recognition rates is presented for situations where a database containing multiple images of the same subject is being used. It is shown that these well- known, existing techniques for both detection and recognition can be combined in a manner that runs in real-time, but still preserves the original success rates mentioned in literature.

P. Viola describes a machine learning approach for visual object detection which is capable of processing images extremely rapidly and achieving high detection rates. This work is distinguished by three key contributions. The first is the introduction of a new image representation called the "integral image" which allows the features used by our detector to be computed very quickly. The second is a learning algorithm, based on AdaBoost, which selects a small number of critical visual features from a larger set and yields extremely efficient classifiers. The third contribution is a method for combining increasingly more complex classifiers in a "cascade" which allows background regions of the image to be quickly discarded while spending more computation on promising object- likeregions.

Human parsing is attracting increasing research attention. In this work, Alex Krizhevsky, IlYaSutskeveraim to push the frontier of human parsing by introducing the problem of multi-human parsing in the wild. Existing works on human parsing mainly tackle single-person scenarios, which deviates from real-world applications where multiple persons are present simultaneously with interaction and occlusion. To address the multi-human parsing problem, we introduce a new multi- human parsing (MHP) dataset and a novel multi-human parsing model named MH-Parser. The MHP dataset contains multiple persons captured in real-world scenes with pixel-level fine-grained semantic annotations in an instance-aware setting.

Alex Krizhevsky, IlYaSutskevertrained a large, deep convolutional neural network to classify the 1.2 million high- resolution images in the ImageNet LSVRC-2010 contest into the 1000 different classes. On the test data, we achieved top-1 and top-5 error rates of 37.5% and

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

17.0% which is considerably better than the previous state-of-the-art. The neural network, which has 60 million parameters and 650,000 neurons, consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully-connected layers with a final 1000-way softmax. To make training faster, we used non-saturating neurons and a very efficient GPU implementation of the convolution operation.

A multi-class classifier-based AdaBoost algorithm for the efficient classification of multi-class data is proposed by T.-H. Kim, D.-C.Park, D.-M.Woo,T. Jeong, and S.-Y. Min.The traditional AdaBoost algorithm is basically a binarIn order to overcome the problems of the AdaBoost algorithm for multi-class classification problems, devisedaAdaBoost architecture with its training algorithm that uses multi-class classifiers for its weak classifiers instead of series of binary classifiers. Typical multi-class AdaBoost architecture based on binary weak classifiers.

T Ojala, M Pietikainen, T Maenpaaproposed a method based on recognizing that certain local binary patterns, termed "uniform," are fundamental properties of local image texture and their occurrence histogram is proven to be a very powerful texture feature. We derive a generalized gray-scale and rotation invariant operator presentation that allows for detecting the "uniform" patterns for any quantization of the angular space and for any spatial resolution and presents a method for combining multiple operators for multiresolution analysis.

Face Recognition (FR) systems are increasingly gaining more importance. Face detection and tracking in a complex scene forms the first step in building a practical FR system. In this paper Amit Pal proposed a hybrid algorithm to detect face(s) in color images is introduced. The algorithm uses color histogram for skin (in the HSV color space) in conjunction with shape information and facial feature detection to quickly locate faces in a given image.

In this paper, DarijanMarčetić, TomislavHrkać, Slobodan Ribarićpropose a two-stage model for unconstrained face detection. The first stage is based on the normalized pixel difference (NPD) method, and the second stage uses the deformable part model (DPM) method. The NPD method applied to in the wild image datasets outputs the unbalanced ratio of false positive to false negative face detection when the main goal is to achieve minimal false negative face detection.

Human face detection and face organ feature location are both important steps in automatic visual interpretation and human face recognition systems. In this paper, a robust system is proposed by Moshu Wu, Guangda Su, Jun Zhou, which focuses on human face detection and face organ location is introduced. We improve the cascade-structured classifier of the human face detector, and integrate human face superresolution algorithm into our system to help processing low-resolution faces. Template matching, feature space analysis method (PCA) combined with AdaBoost algorithm are used to get the precise location of eye feature, and the locations of other face organs are incorporated in the similar framework. The experimental results demonstrate that the proposed system has good performance on actual images.

The objective of the paper proposed by P Shanmugavadivu, Ashish Kumar is to implement the methodology for the of strategic plans in organizations through the prevention and, in its

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

case, the definition and solution of the problems that frequently affect the implementation processes with many negative manifestations and harmful consequences. By elaborating the concept of implementation under the systems approach and cybernetic paradigm, two types of these problems have been identified: the organizational and the functional ones.

Many challenges on face detectors like extreme pose, illumination, low resolution and small scales are studied in the survey paper of by Yuqian Zhou, Din Liu,ThomasHumg. However, previous proposed models are mostly trained and tested on good-quality images which are not always the case for practical applications like surveillance systems. In this paper, we first review the current state-of-the- art face detectors and their performance on benchmark dataset FDDB, and compare the design protocols of the algorithms. Secondly, we investigate their performance degradation while testing on low-quality images with different levels of blur, noise, and contrast. Our results demonstrate that both hand-crafted and deep-learning based face detectors are not robust enough for low-quality images. It inspires researchers to produce more robust design for face detection in the wild.

EXISTING SYSTEM

1. Initially researchers focus on edge and grey value of face image. It was based on patternrecognition model.

2. The face detection technology of famous Viola Jones Detector, which greatly improved real time face detection.

3. Viola Jones detector failed to tackle the real world problems and was influenced by various factors like face brightness and face orientation.

4. Viola Jones could only detect frontal well lit faces and failed to work well in dark condition and with non-frontal images .

5. These issues have made the independent researchers work on developing new face detection models based on deep learning, to have better results for the different facial conditions.

PROPOSED SYSTEM

1. Propose to work with twin objective of creating a Binary face classifier.

2. To detect faces in any orientation irrespective of alignment and train it in an appropriate neural network to get accurate results.

3. The model requires inputting an RGB image of any arbitrary size to the model (feature extraction and class prediction).

4. Aim to generate accurate face masks for human objects from RGB channel images containing localized objects.

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

5. Also the problem of erroneous predictions has been solved and a proper bounding box has been drawn around the segmented region.

6 Developing our face detection model using Multi Human Parsing Dataset, based on convolutional neural networks, such that it can detect the face in any geometric condition frontal or non-frontal for that matter.

7. CNN have always been used for image classification tasks.

IMPLEMENTATION

Machine learning is the concept that a computer program can learn and adapt to new data without human intervention. Machine learning is a field of artificial intelligence (AI) that keeps a computer's built-in algorithms current regardless of changes in the worldwide economy.

DEEP LEARNING

Deep learning is a subset of machine learning where artificial neural networks, algorithms inspired by the human brain, learn from large amounts of data. Deep learning allows machines to solve complex problems even when using a data set that is very diverse, unstructured and inter-connected.

OPEN CV

Open CV (Open Source Computer Vision) is a popular computer vision library started by Intel in 1999. It shows you how to perform face recognition with Face Recognizer in Open CV (with full source code listings) and gives you an introduction into the algorithms behind. To build our face recognition system, we'll first perform face detection, extract face embedding's from each face using deep learning, train a face recognition model on the embedding's, and then finally recognize faces in both images and video streams with Open CV. OpenCV's deployed uses span the range from stitching street view images together, detecting intrusions in surveillance video in Israel, monitoring mine equipment in China, helping robots navigate and pick up objects.

TENSOR FLOW

Tensor Flow is an open-source library developed by Google primarily for deep learning applications. It also supports traditional machine learning. Tensor Flow accepts data in the form of multi-dimensional arrays of higher dimensions called tensors. Multi-dimensional arrays are very handy in handling large amounts of data. It is a symbolic math library, and is also used for machine learning applications such as neural networks. Tensor Flow can run on multiple CPUs and GPUs (with optional CUDA and SYCL extensions for general-purpose computing on graphics processing units).

Keras

Keras is an API designed for human beings, not machines. Keras follows best practices for reducing cognitive load: it offers consistent & simple APIs, it minimizes the number of user actions required for common use cases, and it provides clear & actionable error messages. It also has extensive documentation and developer guides. Keras contains numerous implementations of commonly used neural network building blocks such as layers,

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

objectives, activation functions, optimizers, and a host of tools to make working with image and text data easier to simplify the coding necessary for writing deep neural network code.

METHODOLOGY

In this paper,proposed to createtwin objective of a Binary face classifier which can detect faces any orientation irrespective of the alignment and train it in an appropriate neural network to get accurate results. The model requires inputting an RGB image of any arbitrary size to the model. The model's basic function is feature extraction and class prediction. The output of the model is a feature vector which is Optimized using Gradient descent and the loss function used is Binomial Cross Entropy.

PROPOSED WORK FLOW

The authors proposed a method of obtaining segmentation masks directly from the images containing one or more faces in different orientation. The input image of any arbitrary size is resized to 224 × 224 × 3 and fed to the CNN network for feature extraction and prediction. The output of the network is then subjected to post processing. Initially the pixel values of the face and background are subjected to global thresholding. After that it's passed through median filter to remove the high frequency noise and then subjected to closing operation to fill the gaps in the segmented area. After this bounding box is drawn around the segmented area.

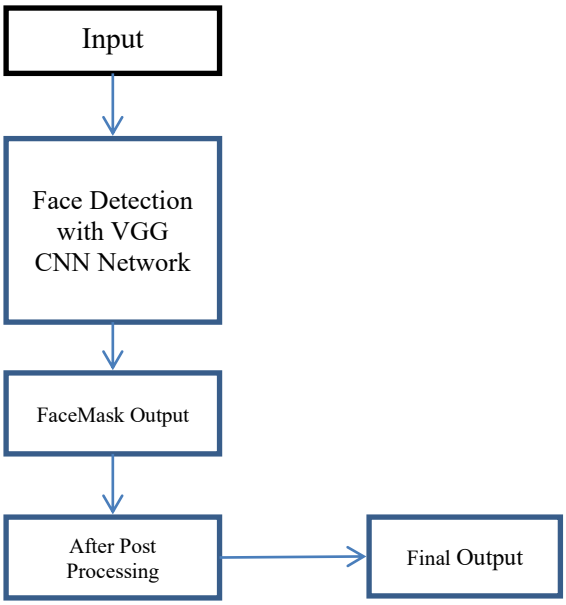

Figure 1:Proposed Flow Diagram

ARCHITECTURE

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

The feature extraction and prediction is performed using pre-defined training weights of VGG 16 architecture. The basic VGG-16 architecture is depicted fig.5.1. Our proposed model consists of a total of 17 convolutional layers and 5 Max pooling layers. The initial image size which is fed to the model is 224×224×3. As the image is processed through the layers for feature extraction it's passed through convolutional layers and max pooling layers. Convolutional layer convolutes the input image with another window while the max pooling operation ensures that the size of the feature vector being produced in every layer is halved so as to reduce the number of parameters. This is a very crucial step in feature extraction, if the number of parameters are not reduced then it will become very difficult to predict the classes of each pixel in a convolutional neural network. The initial layers extract the lower level features while as the subsequent layers extract the mid-level and higher level features. The segmentation task requires that the spatial information be stored in a pixel wise classification, this we have achieved by converting the VGG layers to convolutional layers. After the final max pooling layer. This is further up sampled to bring the image to standard size i.e. 224 × 224 × 2 since it's a binary classifier – hence creates two channels for both the classes, face and background

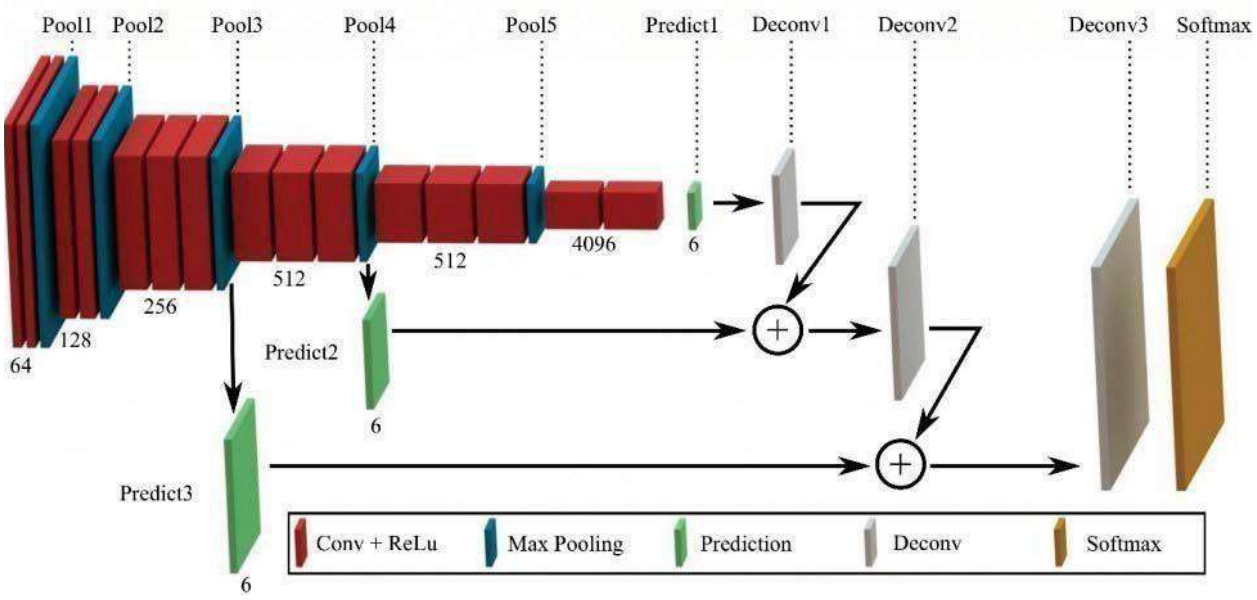

Figure2 :CNN Architecture

FACE DETECTION AND ERRONUOUS DETECTION

Post processing on the predicted mask obtained is performed so that the irregularities in the region can be filled and to remove the unwanted errors (which may have crept during the processing). This we perform by first passing the mask through Median filter and then performing the Closing Operation. This ensures that the gaps in the segmented region are filled and most of the unwanted false erroneous prediction removed. In spite of this there is a possibility that some large error may not have been removed. We have designed the model

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

such that all those erroneous predictions are not considered while showing the final detected faces. We find out the following parameters in each region – Centroid, Major Axis Length and Minor Axis Length. These values are depicted in Table for all the facial (segmented) regions detected (including false predictions). Even after post processing through median filter and dilation, the unwanted erroneous predictions have not completely gone. This results in false face detection .

TRAINING MODEL

The proposed system focuses on how to identify the person on image/video stream wearing face mask with the help of computer vision and deep learning algorithm by using the OpenCV, Tensor flow, Keras.
Approach:
1 Train Deep learning model (MobileNetV2)
2. Apply mask detector over images / live video stream

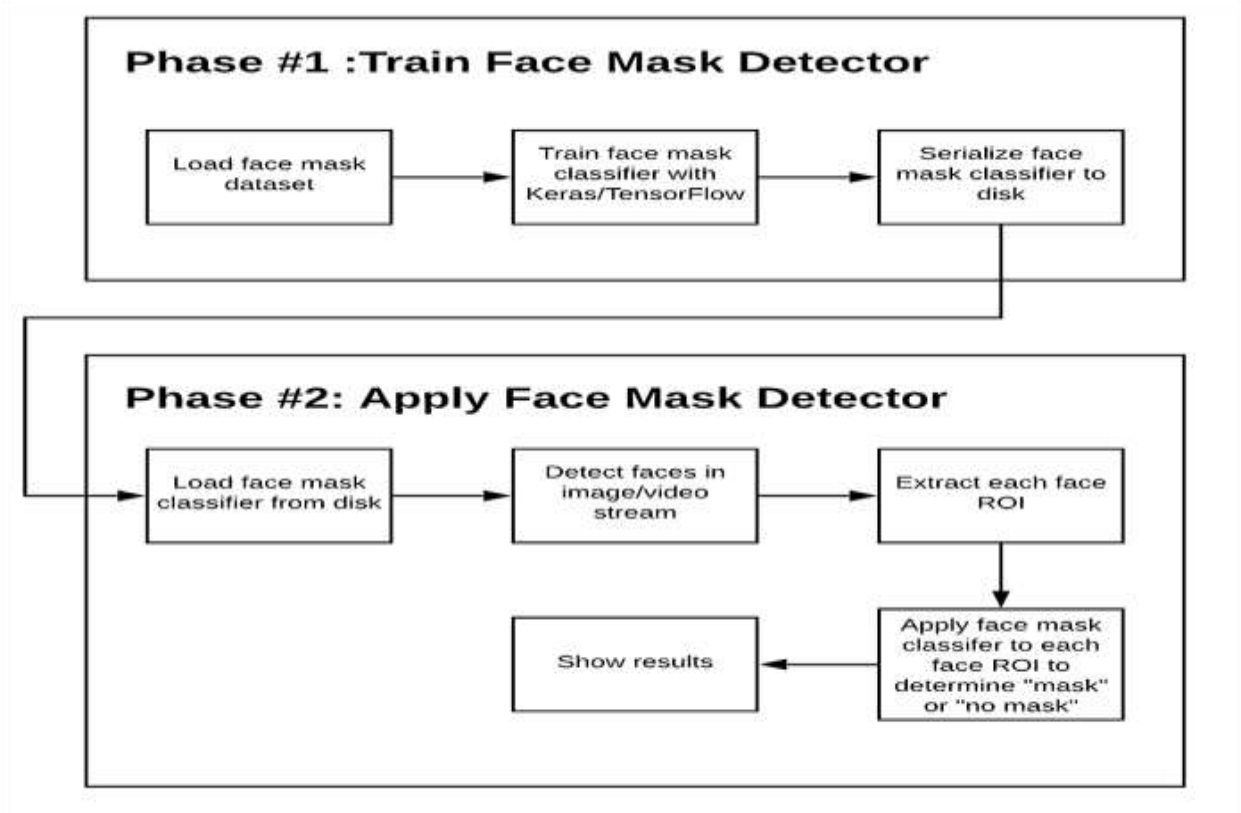

Fig. 3: Training Model

DATA AT SOURCE

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

The majority of the images were augmented by Open CV. The set of images were already labelled —mask and —no mask. The images that were present were of different sizes andresolutions, probably extracted from different sources or from machines (cameras) of different resolutions.

DATA PREPROCESSING

Preprocessing steps as mentioned below was applied to all the raw input images to convert them into clean versions, which could be fed to a neural network machine learning model

1. Resizing the input image (256 x 256)

2. Applying the color filtering (RGB) over the channels (Our model MobileNetV2 supports 2D 3 channel image)

3. Scaling / Normalizing images using the standard mean to build in weights

4. Center cropping the image with the pixel value of 224x224x3.

5. Finally converting them into tensors (Similar to NumPy array).

MobileNetV2

MobileNetV2 builds upon the ideas from MobileNetV1, using depth wise separable convolution as efficient building blocks. However, V2 introduces two new features to the architecture:

1) Linear bottlenecks between the layers, and

2) Shortcut connections between the bottlenecks.

The typical MobilenetV2 architecture has as many layers listed below, In Pytorch we can use the models library in Torch Vision to create the MobileNetV2 model instead of defining/building our own model. The weights of each layer in the model are predefined based on the ImageNet dataset. The weights indicate the padding, strides, kernel size, input channels and output channels.MobileNetV2 was chosen as an algorithm to build a model that could be deployed on a mobile device. A customized fully connected layer which contains four sequential layers on top of the MobileNetV2 model was developed. The layers are

1. Average Pooling layer with 7×7 weights

2. Linear layer with ReLu activation function

3. Dropout Layer

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

4. Linear layer with Softmax activation function with the result of 2 values.

The final layer softmax function gives the result of two probabilities each one represents theclassification of mask or not mask.
Train Mask Detector
Train the model to detect the mask by converting all the images into image arrays and append them into lists. Preprocessing steps are applied to all the raw input images to convert them into clean versions, which could be fed to a neural network machine learning modelFace Mask Detection in webcam stream
The flow to identify the person in the webcam wearing the face mask or not.
The process is two-fold.

1. To identify the faces in the webcam

2. Classify the faces based on the mask

Identify the Face in the Webcam
To identify the faces a pre-trained model provided by the OpenCV framework was used. The model was trained using web images. OpenCV provides 2 models for this face detector:

1. Floating-point 16 version of the original Cmake implementation.

2. 8 bit quantized version using Tensor flow The Cmake model in this face mask detector. There has been a lot of discussion around deep learning based approaches for person detection. This encouraged us to come up with our own algorithm to solve this problem. Our work on face mask detection comprises of data collection to tackle the variance in kinds of face masks worn by the workers. The face mask detection model is a combination of face detection model to identify the existing faces from camera feeds and then running those faces through a mask detection model. Detect Mask Video
Applying model in camera by using face detector model to detect the face and by using mask detector model to detect the face with mask or without mask. Along with that for camera operations we use OpenCV.

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

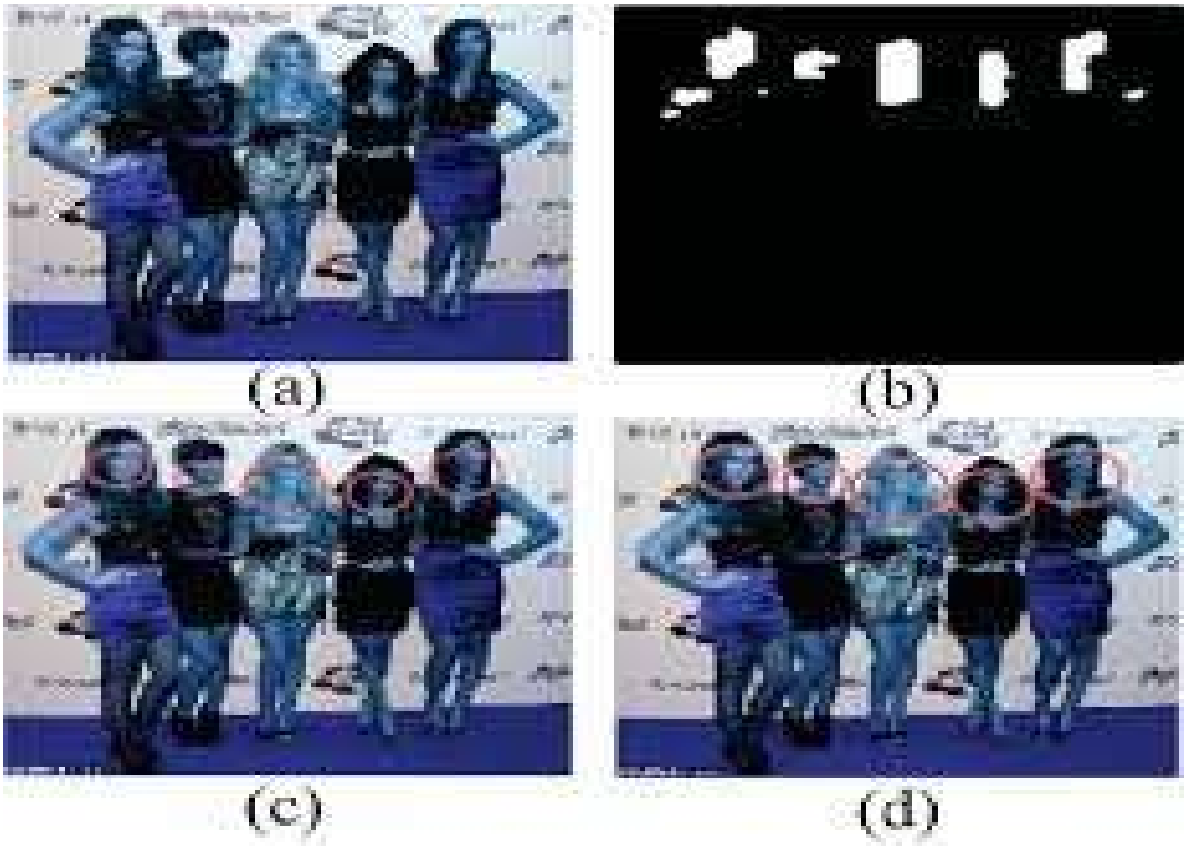

Figure 4: a)Actual Image (b) Erroneous Prediction (c) False Face Detection (d) Correct Face Detection

International Research Journal in Global Engineering and Sciences. (IRJGES)
ISSN : 2456-172X | Vol. 3, No. 2, June – August 2018
Pages 120-132 | Cosmos Impact Factor (Germany): 5.195
Received: 02.07.2018 Published: 28.08.2018

REFERENCE

1. T. Ojala, M. Pietikainen, and T. Maenpaa, "**Multiresolution gray-scale and rotation invariant texture classification with local binary patterns**," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971–987, July2002.

2. T.-H. Kim, D.-C.Park, D.-M.Woo, T. Jeong, and S.-Y. Min, "**Multi-class classifier-based adaboost algorithm**", in Proceedings of the Second Sinoforeign-interchange Conference on Intelligent Science and Intelligent Data Engineering, ser. IScIDE'11. Berlin, Heidelberg: Springer-Verlag, pp.122–127, 2012.

3. P. Viola and M. J. Jones, "**Robust real-time face detection**," Int. J. Compute.Vision, vol. 57no.2, pp.137–154, May2004.

4. P. Viola and M. Jones, "**Rapid object detection using a boosted cascade of simple features**", in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 1, pp. I–I, Dec 2001.

5. J. Li, J. Zhao, Y. Wei, C. Lang, Y. Li, and J. Feng, "**Towards real world human parsing: Multiple-human parsing in the wild**",CoRR, vol. abs/1705.07206.

6. A. Krizhevsky, I. Sutskever, and G. E. Hinton, "**Imagenet classification with deep convolutional neural networks**" ,in Advances in Neural Information Processing Systems 25, Eds. Curran Associates, Inc., 20, pp. 1097–1105,2012

7. K. Simonyan and A. Zisserman, "**Very deep convolutional networks for large-scale image recognition**," CoRR, vol. abs/1409.1556, 2014.

8. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "**Going deeper with convolutions**," 2015.

9. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778,2016.

10. K. Li, G. Ding, and H. Wang, "**L-cnn: A lightweight convolutional neural network for biomedical semantic segmentation**," in 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2363–2367, Dec 2018.

11. X. Fu and H. Qu, "**Research on semantic segmentation of high-resolution remote sensing image based on full convolutional neural network**," in 2018 12th International Symposium on Antennas, Propagation and EM Theory (ISAPE), pp. 1–4, Dec 2018.

12. S. Kumar, A. Negi, J. N. Singh, and H. Verma, "**A deep learning for brain tumor mri images semantic segmentation using cnn**," in 2018 4th International Conference on Computing Communication and Automation (ICCCA), , pp. 1–4, Dec 2018.