e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

EFFICIENT GENDER DETECTION USING DEEP CONVOLUTION NEURAL NETWORKS

Dr. N. Usha Rani[*1], CH Meghana Reddy[*2]

[*1,2]Computer Science & Engineering, Sri Venkateswara University College Of Engineering, India.

DOI : https://www.doi.org/10.56726/IRJMETS37871

ABSTRACT

A crucial task in intelligent applications, such as access control, human-computer cooperation, enforcement, marketing intelligence, and visual surveillance, is estimating gender from a facial image. Gender is a key facial characteristic that is important for social coordination. Human classification and identification have been used for a very long period in many different fields. The main goal of this work is to create an algorithm that accurately and efficiently determines a person's gender. In light of the expanding use of online social networking sites, automatic gender detection is now relevant to an expansion of its implementation in various software and hardware. It has been determined that using deep convolutional neural networks and learning and classification techniques. These gender classification jobs are expected to handle information at an estimated rate of growth. The method described in this work is straightforward and simple to use for human classification; it simply requires a webcam and a respectable computer system to operate.

Keywords: Gender recognition, web cam, Deep Convolution Neural Network, Accuracy.

I. INTRODUCTION

The technique of classifying people has been practised for ages and is still used today in a variety of technologies, including biometrics, forensic sciences, image processing, identification systems, etc. It has becoming much simpler to categorise humans with the advancement of Artificial Intelligence and methods like Deep Learning and Neural Networks. These new technologies make it possible to identify and categorise people without the aid of another professional or personal records. Also These technologies, which are extremely quick, can categorise millions of people far more quickly than a specialist.

Numerous hints and clues are provided by human facial image processing that are applicable to sectors like security, entertainment, etc. [1]. A person's face can reveal a great deal of information about them, including their emotional state, the tiniest nod of agreement or disagreement, irony or fury, etc. Faces have been a long-running psychological research focus because of this [2]. This information (or, in our instance, digital information) is extremely significant because it aids in the identification, selection, or recognition of individuals in accordance with the needs.

The normal multilayer perceptron (MLP) finishes with one or more fully connected layers in the convolutional neural network (CNN), which is a form of the neural network made up of a number of convolutional layers that alternate with subsampling layers. The CNN's capacity to simultaneously extract features, decrease data dimensionality, and classify in a single network structure gives it a substantial edge over traditional pattern recognition methods.

II. LITERATURE REVIEW

The support vector machine (SVM) is a well-liked classification algorithm. A gender categorization method combining a local binary pattern (LBP) and an SVM with a polynomial kernel was proposed in [7], and the CAS-PEAL face database recorded a classification rate of 94.08%. MATLAB 6.1 implementation on a 3.0 GHz CPU resulted in an average processing time of 0.12 s. The method has the drawback of requiring accurate selection of the block size for the LBP operator, which is a challenging operation, in order to obtain excellent classification performance. On the SUMS face database, the work in [8] reported a high classification accuracy of 99.30%. This study utilised K-means nearest neighbour, 2D-DCT feature extraction, and Viola and Jones face identification. One of the earliest applications of neural networks to gender classification was described in [9]. The average error rate with a fully connected MLP and many

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

image processing modules was 8.1%, which is fairly high when compared to results from state-of-the-art systems. Principal component analysis (PCA) was used to analyse the face image as part of the hybrid technique suggested in [10] in order to reduce the dimensionality. The best selection of eigenfeatures was then chosen using a genetic algorithm (GA). 11.30% mistake rate on average was reported. The key disadvantage of this method, which also has a bad error rate, is that while it is a good global random search method, the GA has a high computing complexity. The key drawbacks of the aforementioned approaches are that the feature extraction and classification modules are created and trained independently, and that they are best used in conjunction with prior application-specific knowledge.

Titive and Bouzerdoum were among the first to suggest gender categorization by combining feature extraction and classification in a single neural network in [1]. On the FERET face database, their method, which used a brand-new shunting inhibitory convolutional neural network (referred to as SICoNNets), had an average classification accuracy of 97.1%. A shunting inhibitory neuron with two activation functions and one division task was primarily responsible for the improved performance. The proposed CNN in [5] included six layers, with the output class represented by a single neuron in the last layer. On unmixed datasets using the FERET database, a classification rate of 94.7% was attained. In conclusion, the aforementioned CNN-based methods show the possibility of getting improved performance in challenges involving gender and recognition.

Through learning from data samples, the CNN performs feature extraction and classification within a single network structure [4]. By being familiar with the weights in charge of extracting features, feature selection is also incorporated into the training process [2]. Additionally, the CNN is capable of extracting topological properties from a raw input image with little to no preprocessing needed [2]. Additionally, while maintaining the spatial topology of the input data, a certain level of invariance is attained [1]. Additionally, the CNN offers some robustness and resistance to other 2D shape variations as well as geometric distortions and transformations [2]. As a result, the CNN was created expressly to address the drawbacks of the static classical feature extractor. It was created independently of the trainable classifier and does not fall under the training protocol [3]. The fact that CNNs have fewer parameters than fully connected MLP neural networks with the same number of hidden layers makes them easier to train, which is a final advantage of CNNs. As a result, the CNN has demonstrated potential performance in a variety of applications, including character recognition [4], face recognition [5], tracking of people [6], traffic sign recognition, and many more.

In 2002, Yang et al. [16] compared a number of well-known face identification techniques, however they omitted using any well-known algorithms like Haar Classifiers in their research. According to P. Viola and M. Jones [17], one of the most well-known and reliable object identification methods is the Haar Classifier. Face detection must be implemented correctly for any facial image processing or face recognition system to function successfully. A comprehensive analysis can be found in [18]. While a face is being detected in a frame, various natural (lighting, pose angle, face marks) and digital (noise, glitches) variations are imposed.

System Requirements:

Python 2.7-3.6, Open CV2, PyCharm Community Edition, and a webcam (at least 2.0 MP) are the prerequisites for the project. The project must be run on a Windows PC with the necessary hardware.

III. EXISTING SYSTEM

Gender classification:

Using Haar cascade classifiers to identify an object or its gender is a good idea. The classifier is trained using a large number of both positive and negative images in the Haar cascade machine learning method.

Positive images: These pictures include the pictures that we want our classifier to recognise.

Negative images are pictures of everything else that don't include the thing you're trying to find. To recognise faces in a picture or a live video, an object identification method called the Haar cascade is used.

• Edge detection: This function is mostly used to find objects. It identifies grey level discontinuities. We can define edge as the line dividing the regions.

• Features like Haar's The algorithm proposed by Viola and Jones is known as Haar-Classifiers for Detector with haar-like characteristics In order to find faces in images, a face detection algorithm first scans the image for patterns that could suggest the existence of a face.

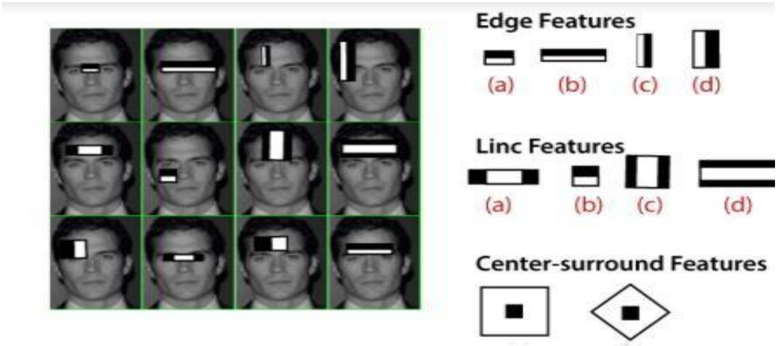

Fig 1: Working Of Haar Features

By leveraging haar-like qualities, this is accomplished. Two courtesy: (www.github.com) or three neighbouring rectangles with varying contrast values combine to form the haar-like features.

IV. PROPOSED SYSTEM

Convolution Neural Network

The convolutional layer is the foundation of CNN. A mathematical procedure called convolution is used to combine two sources of data. In our example, a convolution filter is used to apply convolution to the input data in order to create a feature map.

Deep Convolutional Neural Networks:

The LeNet-5 architecture, which is used for object recognition and detection, popularised D-CNN.

Deep Convolutional Neural Networks (D-CNN) topologies, which are neural networks with numerous layers of neurons, have a lot of promise.

Due to an unexpected increase in computational capacity with the use of Graphical Processing Units (GPU), and the volume of dataset that is readily available on the Internet or prepared by a researcher in order to do practically, they have recently become popular.

One of the most prominent applications of Deep Convolution Neural Network (D-CNN) in the real world includes image classification and recognition on a variety of facial databases containing millions of unfiltered, raw faces.

Composition of Deep Convolutional and Subsample layer:

Deep Convolutional and Subsample Layer Composition: By combining a convolutional layer and a subsample layer into one layer, Deep Convolution Neural Network (D-CNN) work is scaled down across fewer layers. Simard made this idea widespread, and Mamalet and Garcia later became familiar with it.

The equation below can be used to extract a pattern from an image:

where mef(t) is the convolutional kernel weight, pi(t-e) and p j(t) are the input and output pattern maps, respectively, and F() is function, also known as the activation function that we employed in our work. s q(t)q represents the horizontal convolution step size, sp(t)p represents the vertical convolution step size, and Rq(t) and Rp(t) are width and height, respectively. J(t) signifies bias signals the total number of input feature mapping are the convolutional kernels' relative width and height. where M (t-e) and A (t-e) and the input feature mapping's height and width.A(t) = (A(t-e) - Rp(t))/sp(t) + 1 (2)

$$M(t) = (M(t-e)- Rq(t))/sq(t) + 1 \quad (3)$$

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

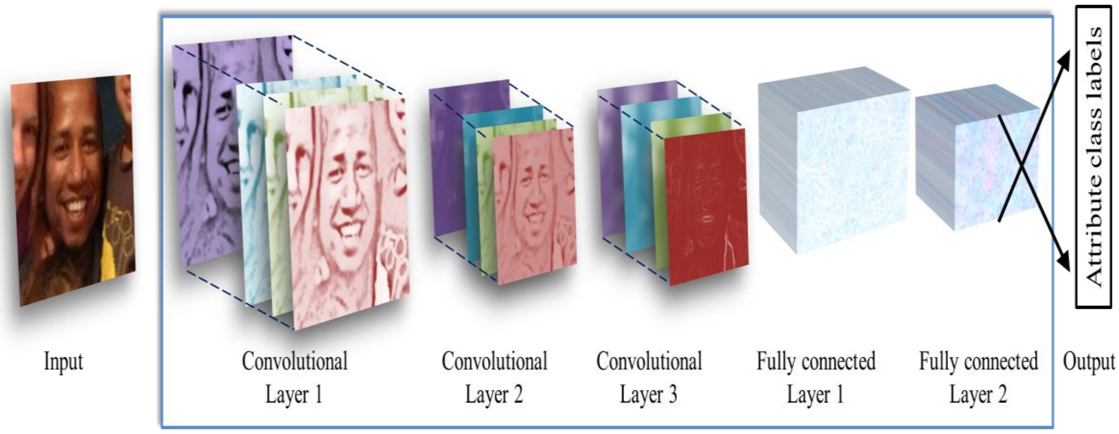

Fig. 2: CNN Architecture for Gender Recognition
courtesy: (www.github.com)

Softmax:

The softmax layer, which is the top layer in the suggested design, computes the loss term that is optimised during training as well as the category probabilities during classification. [13]

Gender Recognition through Web Cam:

Deep Learning functions as an AI system by simulating the cognitive processes of human thought. From unstructured data sets, it can recognise objects, faces, speeches, and characters.

$$f_j(z) = \frac{e^{z_j}}{\sum_k e^{z_k}}$$

Four major sections make up the developed algorithm: input, face detection, face processing (gender categorization), and output.

Flow Chart of the DCNN algorithm:

Step 1: Web Cam was launched first during algorithm implementation.

Step 2: The face in front of the webcam is captured and recognised

Step 3: It predicts the gender of the face by identifying all facial traits.

Step 4: To maximise efficiency, this procedure must continue for a while.

Step 5: Steps 2, 3, and 4 are repeated if a new image appears in front of the webcam

Step 6: If the webcam cannot detect a face, we get output as No Face Detected

Step 7: Stop

Input:

First, the user has the option of taking live data directly using the system webcam or any other webcam camera device. In addition to being quick, this allows for real-time output when using a webcam.

The following two characteristics of the human face as a pattern contribute to the difficulty of human face recognition: (1) There exists an enormous, maybe unlimited number of patterns, or faces, that can be categorised; and (2) Nearly all patterns are quite similar[5]. To address this problem and boost the algorithm's effectiveness, we employed an audience dataset that included all the various types of variations. The audience set will also serve as a benchmark for our neural network's ability to detect gender. Each image was obtained under a Creative Commons (CC) distribution licence.

Face Processing:

Face can provide enough information to study:

emotion,

ethnicity,

heritage,

biasness such as agreement/disagreement, mood, abnormality.

The face is the bodily area that biometric applications use most naturally [5]. As soon as the face is located in the frame. Its processing can begin utilising CNN, or Convolutional Neural Network. It is a form of Deep Neural Network that is primarily utilised for NLP and image processing.

In its testing and training phases, CNN will make various predictions.

The gender prediction has two options: male or female.

Three convolutional layers, each with a distinct number of nodes and kernel sizes, are used in the aforementioned architecture.

- 96 nodes with a 7-kernel configuration
- 384 nodes have kernel sizes of 3
- while 256 nodes have kernel sizes of 5.

Procedure:

We may begin evaluating the technique's accuracy now that it has been put into use. The general process to be followed is to

- Enter the data.
- Make a frame.
- Find the face.
- Classify the gender.
- Include the output as a picture.
- Display the image to the location you specify.

Key Features:

The final product had a promising output and a nearly accurate approach for classifying human gender.

There is no requirement for very precise hardware or software, which is one of this model's key characteristics. It has the ability to process images straight from cameras, including webcams.

- This method is simple to use and doesn't call for expert knowledge.
- It can process and save hundreds of faces together with the accompanying results without any lag or delay, requiring only a basic understanding of computers.

Gender Detection through Web Cam during wearing of Face Mask:

In order to determine a person's gender from a masked facial image, two advanced deep CNN networks (such as GoogleNet and ResNet50) are used in this study. On the ImageNet dataset, the two models perform similarly. We have made use of a recently released dataset known as MAFA [24], which is made up of photographs that have been pulled from the internet. After that, manually erase any pictures with a clear face.

V. SYSTEM MODEL AND METHODS

- We have implemented a pre-trained network in our proposed system that has already mastered the art of extracting strong and profound features from the massive photos from ImageNet.

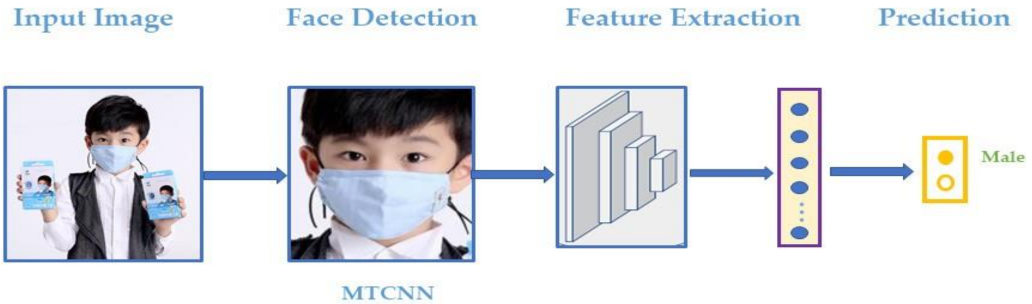

Fig 3: Pipeline of the proposed method for gender prediction from masked face image courtesy: (www.researchgate.net)

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:05/Issue:05/May-2023 Impact Factor- 7.868 www.irjmets.com

The suggested system focuses on using computer vision and deep learning to calculate how to recognise a person wearing a face veil in images and videos using OpenCV, Tensorflow, and Keras. Our main goal is to create a system that functions in crowded settings.

Method:

1. Train MobileNetV2 using transfer learning for deep learning
2. Finding ROI with MTCNN .
3. Utilise a mask detector.

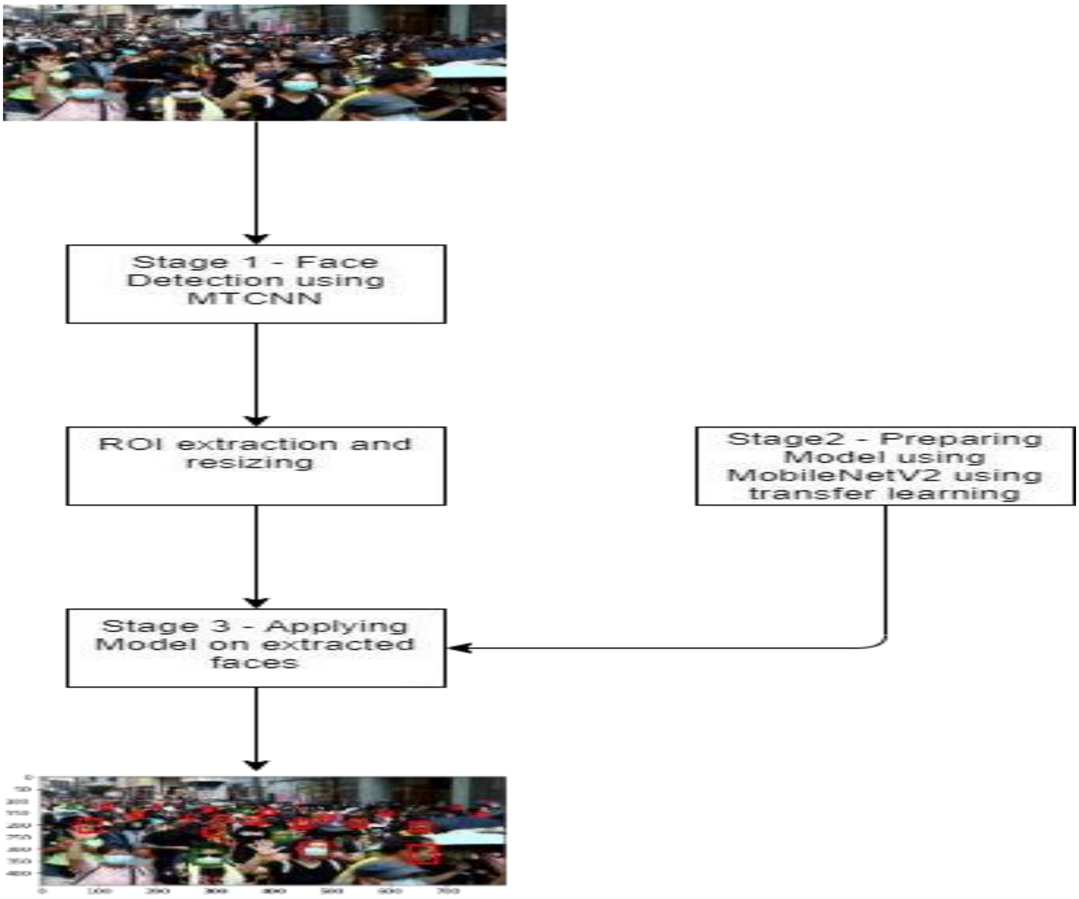

Fig 4: Flowchart of Face Mask Detection using MTCNN andMobileNetV2

Courtesy : (www.irjet.net)

A framework called Multi-task Cascaded Convolutional Networks (MTCNN) was created as a method for both face alignment and detection. Convolutional networks are used in three steps of the process to identify faces and facial landmarks such the eyes, nose, and mouth.

Use Cases:

The following are a few use scenarios for this project:

- Determining the target market within a marketing organisation.
- In the hiring process, to confirm the applicants' eligibility.
- The authenticating of applicants for government IDs.
- Mass human resource classification.
- Speech Recognition
- Computer Vision
- Finance
- Banking
- Robotics

VI. CONCLUSION

We describe a new neural network architecture in this paper named the Deep Convolution neural network. Adaptive brightness enhancement for input is used to make live videos with harsh lighting resilient. This can be acquired via the internal single-branch approach used by DCNN. Here, we use high-level haar characteristics to accurately detect facial aspects for gender detection. For image quality, a multi-branch approach is applied. Finally, DCNN successfully completes live gender recognition.

VII. REFERENCES

[1] Tivive FHC, Bouzerdoum A. A gender recognition system using shunting inhibitory convolutional neural networks. In: International Joint Conference on Neural Networks; 2006; Vancouver, Canada. New York, NY, USA: IEEE. pp.5336–5341.

[2] Khalajzadeh H, Mansouri M, Teshnehlab M. Face recognition using convolutional neural network and simple logisticclassifier. Stud Comp Intell 2013; 223: 197–207.

[3] Strigl D, Kofler K, Podlipnig S. Performance and scalability of GPU-based convolutional neural networks. In: 18th Euromicro International Conference on Parallel, Distributed and Network-Based Processing; 17–19 February 2010; Pisa, Italy. New York, NY, USA: IEEE. pp. 317–324.

[4] LeCun Y, Bottou L, Bengio Y, Haffner. P. Gradient-based quick learning applied to document recognition. P IEEE 1998;86: 2278–2324.

[5] Duffner S. Face Image Analysis with Convolutional Neural[NN] based Networks. Munich, Germany: GRIN Verlag, 2009.

[6] Fan J, Wei X, Ying W, Yihong G. Human tracking using the convolutional neural networks. IEEE T Neural Network2010; 21: 1610–1623.

[7] Lian HC, Lu BL. Multi-view gender classification using local binary patterns and support vector machines. LectNotes Comput Sc 2006; 3972: 202–209.

[8] Nazir M, Ishtiaq M, Batool A, Jaffar MA, Mirza AM. Feature selection for efficient gender classification. In: Proceedings of the 11th WSEAS International Conference; 2010. pp. 70–75.

[9] Golomb BA, Lawrence DT, Sejnowski TJ. SEXNET: A neural network identifies sex from human faces. In: Proceedings of NIPS; 1990. pp. 572–579.

[10] Sun Z, Yuan X, Bebis G, Louis SJ. Neural-network-based gender classification using genetic search for eigen- feature selection. In: Proceedings of the 2002 International Joint Conference on Neural Networks; 12–17 May 2002; Honolulu, HI, USA. New York, NY, USA: IEEE. pp. 2433–2438.

[11] Dawwd SA, Mahmood BS. Video Based Face Recognition Using Convolutional Neural Network. Rijeka, Croatia: InTech, 2011.

[12] Simard PY, Steinkraus D, Platt JC. Best Practices for Convolutional Neural Networks Applied to Visual DocumentAnalysis. Redmond, WA, USA: Microsoft Research, 2003.

[13] Mamalet F, Garcia C. Simplifying ConvNets for fast learning. In: Villa A, Duch W, Érdi P, Masulli F, Palm G, editors. Artificial Neural Networks and Machine Learning – ICANN 2012. Vol. 7553. Berlin, Germany: Springer,2012, pp. 58–65.

[14] M.H.U. Yap, "Facial image processing for facial analysis," Proceedings International Carnahan Conference on Security Technology., vol. 1, 2010.

[15] C.Y.Jiang X., Applied Pattern Recognition. Studies in Computational Intelligence, Berlin, Heidelberg: Springer, 2008.

[16] D. J. K. a. N. A. M. H. Yang, "Detecting Faces in Images: A Survey," IEEE, vol. 24, no. 1, pp. 34-58, 2002.

[17] P. V. a. M. Jones, "Rapid object detection using a boosted cascade of simple features," Proceedings of the2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, 2001.

[18] W.-Y. &. C. R. &. P. Zhao, "Face Recognition: A Literature Survey," ACM Comput. Surv.., pp. 399-458.

[19] K. K. Y. Kaya, "A BASIC STUDY ON HUMAN FACE RECOGNITION," Frontiers of Pattern Recognition, Academic Press, pp. 265-289, 1972.

[20] D. D. Zhang, "Automated Biometrics - Technologies and Systems," Kluwer Academic Publishers, vol. 1,2000.

[21] R. E. a. T. H. E. Eidinger, "Age and Gender Estimation of Unfiltered Faces," IEEE Transactions on Information Forensics and Security, vol. 9, no. 12, pp. 2170-2179, 2014

[22] G. Levi and T. Hasner, "Age and gender classification using convolutional neural networks," IEEE Conference on Computer Vision and Pattern Recognition Workshops, Boston, MA, 2015, pp. 34-42.

[23] S. Christian, et. al., "Going deeper with convolutions," In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1-9, 2015.

[24] H. Kaiming, et. al., "Deep residual learning for image recognition," In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778, 2016.