ISSN 2278-3091
Volume 9, No.4, July – August 2020
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse212942020.pdf
https://doi.org/10.30534/ijatcse/2020/212942020

Coronamask: A Face Mask Detector for Real-Time Data

Chhaya Gupta[1], Nasib Singh Gill[2]
[1, 2]Department of Computer science & Applications
Maharishi Dayanand University, Rohtak
Haryana, India
Email: chhayagupta.spm@gmail.com , nasibsgill@gmail.com

ABSTRACT

COVID – 19 (2019 novel coronavirus) which started in China had spread all over the world rapidly. It is the worst health crisis the whole world has suffered after World War II. Many precautionary measures have been indicated by the World Health Organisation (WHO) like to maintain social distancing, wear masks, wash hands with soap for 20 seconds and many more. Wearing masks in public places is quite an effective measure to stay protected from this pandemic. There is very few research done for detecting face masks. This paper contributes to the welfare of human beings and proposes CoronaMask, a highly effective face mask detector. The proposed model uses the deep learning convolutional neural network (CNN) algorithm as a base for detecting faces. In this study, the dataset has been created which consists of 1238 images which are divided into two classes as “mask” and “no_mask”. This model also takes live streaming videos as input and detects faces which are wearing masks and which are not wearing a mask. The convolutional neural network is trained on the dataset and it gives 95% of accuracy. CoronaMask, a two-phase face mask detector works in identifying masks in images and also in real-time video streams.

Key words : Convolutional Neural Network (CNN), Deep Learning, Face Detection, Face Mask Detection, Real-time video streaming.

1. INTRODUCTION

COVID-19 pandemic first appeared in Wuhan, China in December 2019. It is a severe respiratory syndrome coronavirus which affected people worldwide. This virus has been presumed to be transferred to humans from bats. Many researchers have used the chest CT scan images to identify the corona cases amongst normal persons chest CT scan images. Symptoms of this pandemic are cough, fever, and respiratory disorders. People are more concerned about their health. WHO has recommended washing hands often, sneeze in sleeves, wear masks and to take care of them from this pandemic. Services are available for people only if they are following the safety measures properly and are wearing their masks. Therefore, face mask detection has become an important part of computer vision to help society. Unfortunately, research is very limited in this field as of now.

The idea behind face mask detection is to determine whether a person is wearing a mask or not. It is an object detection problem as here face is referred to as an object which needs to be detected [1],[2]. There are varied applications of object detection like self-driving vehicles [3], the education system [4], surveillance systems and many more. Unlike traditional object detectors, this model is using Haar Cascade features. Many of the other facial detection algorithms use different techniques like HOG (Histogram of Oriented Gradients), SIFT (Scale-Invariant Feature Transform) and others [5]. Deep Learning is a science that works with neural networks. Various neural networks are present like Artificial Neural Network (ANN), Convolutional Neural Network (CNN), and many more are there [6].

This paper proposes a two-phase face mask detector, CoronaMask, which detects face masks in images as well as in real-time video streams and ensure the safety of people. CoronaMask works in two phases where the first phase is of training the face mask detector and in the second phase, the trained face mask detector is applied to the real-time images and videos. In simple words, an image with a face is taken and is analysed through a cascade classifier. Classifier gives the region of interest (ROI) and in this sense, the classifier is trained on a dataset of 1238 images shown in figure 1. The dataset consists of various masked and unmasked images. CoronaMask uses the convolutional neural network as the classifier algorithm and after 50 epochs it gives an accuracy of 95%.

Figure 1 : Dataset images with mask and no-mask images

The paper is organised as follows. In section II, Literature Survey on object detection and neural networks. Section III consists of the proposed methodology. Section IV describes neural network used. Section V briefs the results and finally, section VI gives conclusion and future scope.

2. LITERATURE SURVEY

Various traditional object detectors use multiple-step processes [7]. Viola Jons detector extracts features by Haar Cascade feature descriptor and can achieve real-time detection [8]. The author in [9] a new feature extractor called Histogram of Oriented Gradients (HOG) is introduced that works with magnitudes of gradients over the image. A DPM (Deformable Part-base Model) is also proposed that detects different parts of objects [5].

Authors in [10], introduced that face detection can be categorized into 4 different categories like knowledge-based method, template-based method, appearance-based method and feature invariant method. However, in [11],[12] authors proposed a distributed based system object detection method in which positive and negative examples are used by the object class for learning. In [13] retinal connected bootstrap neural network is used and works on positive face examples and negative face examples. In [14], [15], [16], AdaBoost classifier is used with LBP (Local Binary Pattern) and Haar Cascade features to detect objects and faces. Authors in [17], [18] have proposed Support Vector machine classifier with HOG features for face detection. In [19], plant infection diseases have been identified by user with the help of Convolutional Neural Network by using the images of leaves. In [20], various pre-trained CNN models have been connected together to identify pneumonia in various patients by using their X-Ray images. They provided an accuracy of 93% for training data and 82% for validation data.

In [21], authors perform facial detection using a fully convolutional network (FCN)that uses predefined training weights (VGG – 16) and also uses binomial cross-entropy as a loss function. The research was focused on greyscale images in the field of face detection.

In [22], object detection was based on prior information of the face model. Different architectures namely AlexNet [23] and VGGNet [24] consists of stacked convolutional layers, former consist of five convolutional layers with three fully connected layers but later was an enhancement over AlexNet. A decision tree classifier is trained for images for face detection yielding good accuracy [25]. Authors in [26] proposed a hierarchical knowledge-based model foe object detection that is an improved version of the edge detection method.

Recent research works with two-stage detectors for facial detection as well as single-stage detectors like Faster R-CNN, SSD and YOLO respectively [27], [28]. Various machine learning classifiers like Support Vector Machine and K-Nearest Neighbours have been compared in this field. SVM used with Principal Component Analysis (PCA) provides a 92% face recognition accuracy [29].

In [30] authors proposed Locally Nonlinear Feature Fusion-based Network (LNFF-Net) that fuses heat map and saliency map and shows better accuracy. The objective of this network is to identify the single masked face regions. Authors proposed a new Deep Pyramid Single Shot Face Detector (DPSSD) with a new loss function namely crystal loss for better accuracy and better performance for object detection [31]. In [32] authors proposed a deep convolutional neural network that uses Discriminative Complete features (DCF) that extracts multi-scale features. A study was conducted to filter selfie face images based on hashtags, Haar cascade method filter images with 71.48% accuracy [33].

Convolutional Neural Network is an important neural network in computer vision related pattern recognition. Inception neural network has been proposed that makes networks to learn about best kernels to be used [34]. Residual network (ResNet) that is capable to learn and identify mapping from previous layers was proposed in [35] and is a deep residual neural network. These are foundations for segmentations.

3. METHODOLOGY

The dataset used in this study consists of a total of 1238 images with masked and no-masked people. The platform used id Anaconda and spyder is used for python programming. CoronaMask architecture is shown in figure 2. In CoronaMask, CNN is the backbone that uses Haar cascade features for face detection. CoronaMask is a two-phase face detector that works on images as well as live video streams to detect whether a person is wearing a mask or not. This model ensures the safety of people completely in this COVID-19 pandemic.

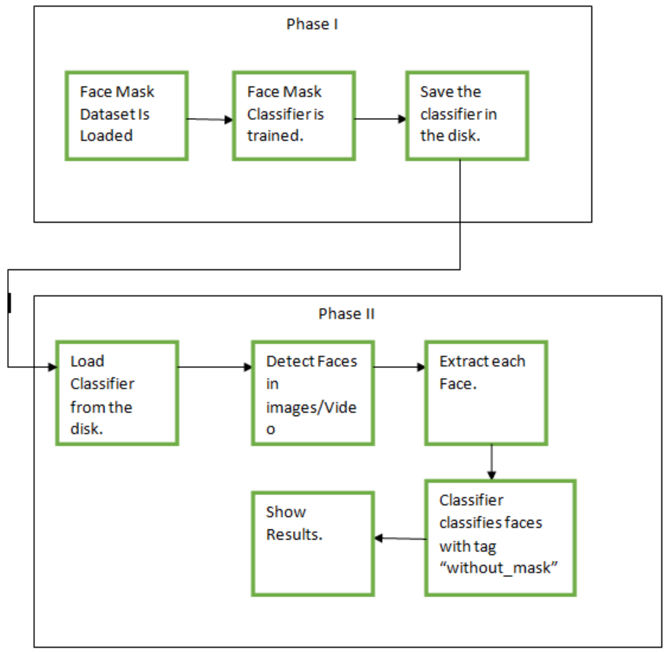

Figure 2 : CoronaMask Architecture

The whole model is divided into two main steps:

- Training Phase: In this step, the dataset is loaded and the model is trained on the dataset and is saved back to the disk.
- Release Phase: In this step, the trained model is loaded back to work as a face mask detector and classify the images and real-time video streams into two classes of “mask” and “no_mask”.

Face detectors allow inferring the locations of facial features like eyes, nose, eyebrows, mouth and jawline. The detector helps to identify the region of interest (ROI) from facial images and the ROI is extracted. Then facial features are identified. The model uses various steps which are described as follows:

Step 1: DATA PREPROCESSING – The dataset used in this study consists of 1238 images of the various mask and no_mask people. All the images are coloured images and therefore they all are needed to be converted into grayscale. All the images are resized into (100x100) to be of the same size pixel values. Dataset is splitted into training and testing sets in 75:25 ratio with 825 images training set and 207 images in the testing set.

Step 2: TRAINING CNN – The convolutional model used in this research study is a two-layer neural network. To avoid overfitting problem, Dropout method has been incorporated. As there are two categories of images “mask” and “no_mask” binary_crossentropy is being used as a loss function [36]. While training the CNN, 50 epochs have been implemented. RelU activation function has been used as a function for internal layers and softmax is used as an activation function for the output layer.

Step 3: DETECTING REAL-TIME VIDEO STREAMS WITH NO_MASK – The trained model is loaded from disk and camera is set up. The camera can be either webcam camera or mobile phone camera. For this, a label is created with “without_mask”. A bounding rectangle with “Red “ colour is used. The camera is used to view live stream videos and the streaming video is read frame by frame and is converted into grayscale and faces are detected. The result shows probability result of no_mask and mask. Figure 3 and figure 4 show the output from live video stream captured images with label “without_mask” when person is not wearing the mask.

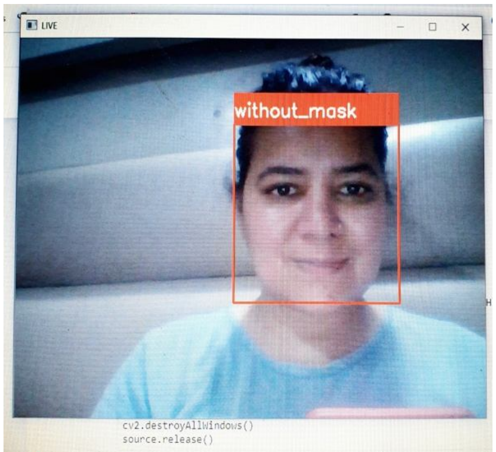

Figure 3 : Person not wearing mask

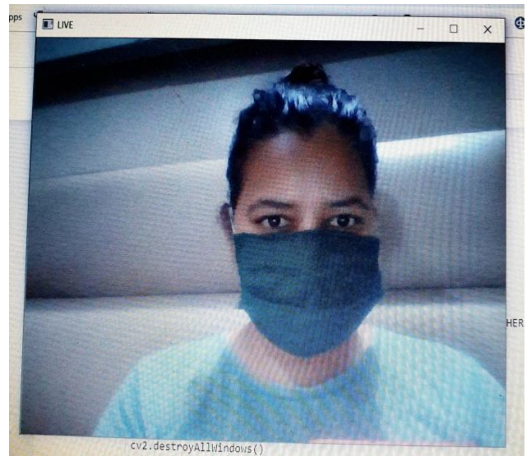

Figure 4 : Person wearing mask

4. CONVOLUTIONAL NEURAL NETWORK (CNN)

Face detection is always being a challenging problem. It is an important problem in computer vision. Computer vision always needs the same perception between humans and machines. It works on making a machine as smart as a human in perceiving this world. Computer vision is responsible for various tasks such as image analysis, image and video recognition, image classification etc. Deep learning embedded with computer vision had made it possible by giving the convolutional neural network (CNN) [37].

CNN is a deep learning algorithm that takes image as input, analyse the image with weights and biases and returns the output. In traditional methods, handcrafted engineered filters [38] were used but this is not the case with CNN. CNN has the capability to capture spatial as well as temporal dependencies in image via relevant filters. It also helps in avoiding the overfitting problem with images. A traditional CNN is shown in figure 5 [39] in which it has layers having different planes. Each plane receives input from the planes of the previous layer. Each plane has its weights. A plane can be considered as a feature map. There are multiple planes in each layer which help detect multiple features. These layers are known as convolutional layers. This neural network is trained by backpropagation gradient descent method [40]. As there are convolutional layers, pooling layers are also present on CNN. There are two types of pooling layers:

- MaxPooling: It returns the maximum value from the image covered by the kernel. It is useful in removing noise from the images along with the dimensionality reduction.
- AveragePooling: It returns average of all values from protion of the image covered by the kernel. It simply performs dimensionality reduction.

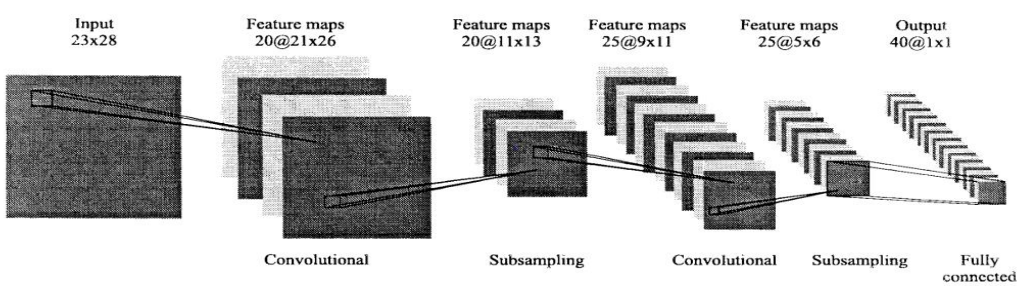

Figure 5 : Traditional Convolutional Neural Network [39]

After Pooling layers there is ReLu layer. ReLu (Rectified Linear Unit) is an activation function that removes negative values from the activation map by setting them to zero. Finally, there is the fully connected layer which is responsible for high-level reasoning in the neural network. Nodes in this fully connected layer have a connection to all activations done in previous layers. After going through all the epochs the model can distinguish low-level and dominating features in the image.

CNN also consists of different loss functions which are :

- Softmax – used for predicting a single class of L mutually exclusive classes.
- Sigmoid cross-entropy – used for predicting M probability values.
- Euclidean – used for predicting the real-valued labels.

5. RESULTS

The dataset used in this research study consists of 1238 images with masked and unmasked faces as shown in figure 1. The model after 50 epochs provides an accuracy of 95% which is far better than many other neural networks used for face detection. The paper also provides a graphical representation of Validation Loss and Training Loss as well as Validation Accuracy and Training Accuracy which are shown in figure 6 and 7 below and are helpful for better validation decisions. The model is able to analyse the live streaming videos and is detetcting people wearing a mask and not wearing a mask and figure 8 below is an output from the live streaming video which was analysed by this model.

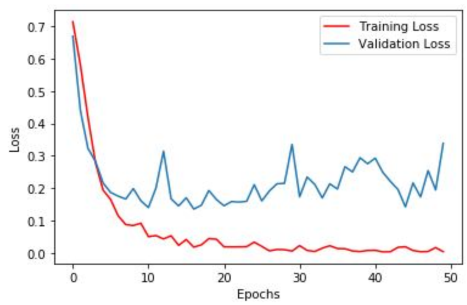

Figure 6 : Training Loss and Validation Loss VS epochs

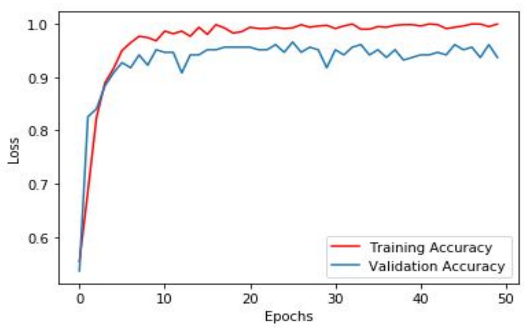

Figure 7 : Training Accuracy and Validation Accuracy VS epochs

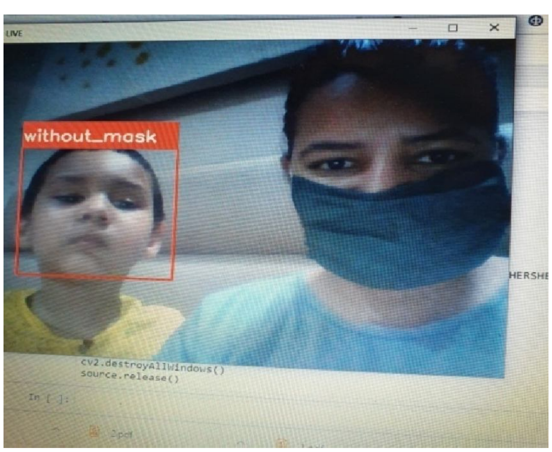

Figure 8 : Live Streaming Video captured no_mask and mask person

6. CONCLUSION AND FUTURE WORK

The CoronaMask provides a real-time safety measure for human beings by detecting whether a person is wearing the mask or not as wearing a mask is an essential need of the hour in this COVID-19 pandemic. The accuracy given by the neural network on image dataset is 95% which can be increased in future by using MobileNet neural network as a backbone model. This model can be implemented in various government bodies in different market places, airports, railway stations, other crowded areas to ensure people follow the safety measures. In future, this model can be embedded with an alarm or buzzer and IoT can be implemented with deep learning CNN which will buzz whenever a person is detected not wearing a mask. If this model is deployed correctly it will be helpful to ensure the safety of human beings.

REFERENCES

[1] Z. Q. Zhao, P. Zheng, S. T. Xu, and X. Wu, "Object Detection with Deep Learning: A Review," *IEEE Trans. Neural Networks Learn. Syst.*, vol. 30, no. 11, pp. 3212–3232, 2019 doi: 10.1109/TNNLS.2018.2876865.

[2] D. Meena and R. Sharan, "An approach to face detection and recognition," *2016 Int. Conf. Recent Adv. Innov. Eng. ICRAIE 2016*, pp. 0–5, 2016 doi: 10.1109/ICRAIE.2016.7939462.

[3] D.-H. Lee, K.-L. Chen, K.-H. Liou, C.-L. Liu, and J.-L. Liu, "Deep Learning and Control Algorithms of Direct Perception for Autonomous Driving," pp. 2–7, 2019, [Online]. Available: http://arxiv.org/abs/1910.12031.

[4] K. S. Savita, N. A. Hasbullah, S. M. Taib, A. I. Z. Abidin, and M. Muniandy, "How's the Turnout to the Class? A Face Detection System for Universities," *2018 IEEE Conf. e-Learning, e-Management e-Services, IC3e 2018*, pp. 179–184, 2019, doi: 10.1109/IC3e.2018.8632630.

[5] J. Redmon *et al.*, "2008 Dpm," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 330, no. 6, pp. 1299–1305, 2014, doi: 10.1109/CVPR.2018.00889.

[6] K. Lin *et al.*, "Face Detection and Segmentation Based on Improved Mask R-CNN," *Discret. Dyn. Nat. Soc.*, vol. 2020, pp. 1–11, 2020, doi: 10.1155/2020/9242917.

[7] Z. Zou, Z. Shi, Y. Guo, and J. Ye, "Object Detection in 20 Years: A Survey," pp. 1–39, 2019, [Online]. Available: http://arxiv.org/abs/1905.05055.

[8] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 1, no. July 2014, 2001, doi: 10.1109/cvpr.2001.990517.

[9] B. R. de Forges and P. K. L. Ng, "New records and new species of Homolidae De Haan, 1839, from the Philippines and French Polynesia

(Crustacea : Decapoda : Brachyura)," *Raffles Bull. Zool.*, no. 16, pp. 29–45, 2007
doi: 10.1109/CVPR.2005.177.

[10] S. M. Lajevardi and H. R. Wu, "Facial expression recognition in perceptual color space," *IEEE Trans. Image Process.*, vol. 21, no. 8, pp. 3721–3733, 2012, doi: 10.1109/TIP.2012.2197628.

[11] M. H. Yang, N. Ahuja, and D. Kriegman, "Face detection using mixtures of linear subspaces," *Proc. - 4th IEEE Int. Conf. Autom. Face Gesture Recognition, FG 2000*, no. February, pp. 70–76, 2000, doi: 10.1109/AFGR.2000.840614.

[12] G. Moazzam, M. Rubayat, and M. Al-Amin, "Human Face Detection under Complex Lighting Conditions," *Int. J. Adv. Comput. Sci. Appl.*, vol. 1, no. 1, pp. 85–90, 2011
doi: 10.14569/specialissue.2011.010112.

[13] A. F. Wolfschoon, "Análise do Leite por Espectroscopia Infravermelha - 2a parte," *Rev. do ILCT*, vol. 20, no. 1, pp. 3–9, 1977.

[14] Y. Heydarzadeh and A. T. Haghighat, "An efficient face detection method using adaboost and facial parts," *Int. J. Simul. Syst. Sci. Technol.*, vol. 12, no. 4, pp. 1–9, 2011
doi: 10.5013/IJSSST.a.12.04.01.

[15] T. Mita, T. Kaneko, and O. Hori, "Joint Haar-like features for face detection," *Proc. IEEE Int. Conf. Comput. Vis.*, vol. II, pp. 1619–1626, 2005
doi: 10.1109/ICCV.2005.129.

[16] L. Chen, Y. H. Wang, Y. D. Wang, and D. Huang, "Face recognition with statistical local binary patterns," *Proc. 2009 Int. Conf. Mach. Learn. Cybern.*, vol. 4, no. February, pp. 2433–2439, 2009, doi: 10.1109/ICMLC.2009.5212189.

[17] I. Kukenys and B. McCane, "Support vector machines for human face detection," *New Zeal. Comput. Sci. Res. Student Conf. NZCSRSC 2008 - Proc.*, no. January 2008, pp. 226–229, 2008.

[18] M. M. Abdelwahab, S. A. Aly, and I. Yousry, "Efficient Web-based Facial Recognition System Employing 2DHOG," no. February, 2012, [Online]. Available:
http://arxiv.org/abs/1202.2449.

[19] V. Prashanthi and K. Srinivas, "Plant disease detection using convolutional neural networks," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 3, pp. 2632–2637, 2020
doi: 10.30534/ijatcse/2020/21932020.

[20] J. C. Young and A. Suryadibrata, "Applicability of various pre-trained deep convolutional neural networks for pneumonia classification based on x-ray images," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 3, pp. 2649–2654, 2020
doi: 10.30534/ijatcse/2020/24932020.

[21] T. Meenpal, A. Balakrishnan, and A. Verma, "Facial Mask Detection using Semantic Segmentation," *2019 4th Int. Conf. Comput. Commun. Secur. ICCCS 2019*, no. October, pp. 1–5, 2019, doi: 10.1109/CCCS.2019.8888092.

[22] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, 2002
doi: 10.1109/TPAMI.2002.1017623.

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017, doi: 10.1145/3065386.

[24] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*, pp. 1–14, 2015.

[25] M. J. Jones and M. Jones, "Fast multi-view face detection Fast Multi-view Face Detection," no. December 2003, 2017.

[26] G. Yang and T. S. Huang, "Human face detection in a complex background," *Pattern Recognit.*, vol. 27, no. 1, pp. 53–63, 1994, doi: 10.1016/0031-3203(94)90017-5.

[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, 2017, doi: 10.1109/TPAMI.2016.2577031.

[28] J. Zhang, X. Wu, S. C. H. Hoi, and J. Zhu, "Feature agglomeration networks for single stage face detection," *Neurocomputing*, vol. 380, pp. 180–189, 2020
doi: 10.1016/j.neucom.2019.10.087.

[29] P. Rakshit, R. Basu, S. Paul, S. Bhattacharyya, J. Mistri, and I. Nath, "Face Detection using Support Vector Mechine with PCA," *SSRN Electron. J.*, pp. 1–5, 2020, doi: 10.2139/ssrn.3515989.

[30] X.-Y. Peng, J. Cao, and F.-Y. Zhang, "Masked Face Detection Based on Locally Nonlinear Feature Fusion," pp. 114–118, 2020, doi: 10.1145/3384544.3384568.

[31] R. Ranjan *et al.*, "A Fast and Accurate System for Face Detection, Identification, and Verification," *IEEE Trans. Biometrics, Behav. Identity Sci.*, vol. 1, no. 2, pp. 82–96, 2019
doi: 10.1109/tbiom.2019.2908436.

[32] G. Guo, H. Wang, Y. Yan, J. Zheng, and B. Li, "A fast face detection method via convolutional neural network," *Neurocomputing*, no. xxxx, 2019, doi: 10.1016/j.neucom.2018.02.110.

[33] A. Priadana and M. Habibi, "Face detection using haar cascades to filter selfie face image on instagram," *Proceeding - 2019 Int. Conf. Artif. Intell. Inf. Technol. ICAIIT 2019*, pp. 6–9, 2019, doi: 10.1109/ICAIIT.2019.8834526.

[34] G. Zeng, Y. He, Z. Yu, X. Yang, R. Yang, and L. Zhang, "Preparation of novel high copper ions removal membranes by embedding organosilane-functionalized multi-walled carbon nanotube," *J. Chem. Technol. Biotechnol.*, vol. 91, no. 8, pp. 2322–2330, 2016, doi: 10.1002/jctb.4820.

[35] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[36] Y. Ho and S. Wookey, "The Real-World-Weight Cross-Entropy Loss Function: Modeling the Costs of Mislabeling," *IEEE Access*, vol. 8, pp. 4806–

4813, 2020, doi: 10.1109/ACCESS.2019.2962617.

[37] L. Haoxiang and Lin, "A convolutional neural network approach for face identification," *Proc. CVPR*, pp. 5325–5334, 2015, [Online]. Available: http://www.cs.nyu.edu/~xd283/face2012.pdf.

[38] S. S. Farfade, "Multi-view Face Detection Using Deep Convolutional Neural Networks Categories and Subject Descriptors," pp. 643–650.

[39] G. Rostami, M. Hamid, and H. Jalaeikhoo, "Impact of the BCR-ABL1 fusion transcripts on different responses to Imatinib and disease recurrence in Iranian patients with Chronic Myeloid Leukemia," *Gene*, vol. 627, no. 1, pp. 202–206, 2017, doi: 10.1016/j.gene.2017.06.018.

[40] R. Lippmann, "Book Review: 'Neural Networks, A Comprehensive Foundation', by Simon Haykin," *International Journal of Neural Systems*, vol. 05, no. 04. pp. 363–364, 1994 doi: 10.1142/s0129065794000372.