Journal of Biomedical Informatics 120 (2021) 103848

Contents lists available at ScienceDirect

Journal of Biomedical Informatics

journal homepage: www.elsevier.com/locate/yjbin

Original Research

Face mask detection using deep learning: An approach to reduce risk of Coronavirus spread

Shilpa Sethi [a], Mamta Kathuria [a,*], Trilok Kaushik [b]

[a] *Department of Computer Applications, J.C. Bose University of Science and Technology, Faridabad, India*
[b] *R & D Department, Samsung India Pvt. Ltd, Noida, India*

ARTICLE INFO

Keywords:
Face mask detection
Transfer learning
COVID-19
Object deletion
One-stage detector
Two-stage detector

ABSTRACT

Effective strategies to restrain COVID-19 pandemic need high attention to mitigate negatively impacted communal health and global economy, with the brim-full horizon yet to unfold. In the absence of effective antiviral and limited medical resources, many measures are recommended by WHO to control the infection rate and avoid exhausting the limited medical resources. Wearing a mask is among the non-pharmaceutical intervention measures that can be used to cut the primary source of SARS-CoV2 droplets expelled by an infected individual. Regardless of discourse on medical resources and diversities in masks, all countries are mandating coverings over the nose and mouth in public. To contribute towards communal health, this paper aims to devise a highly accurate and real-time technique that can efficiently detect non-mask faces in public and thus, enforcing to wear mask. The proposed technique is ensemble of one-stage and two-stage detectors to achieve low inference time and high accuracy. We start with ResNet50 as a baseline and applied the concept of transfer learning to fuse high-level semantic information in multiple feature maps. In addition, we also propose a bounding box transformation to improve localization performance during mask detection. The experiment is conducted with three popular baseline models viz. ResNet50, AlexNet and MobileNet. We explored the possibility of these models to plug-in with the proposed model so that highly accurate results can be achieved in less inference time. It is observed that the proposed technique achieves high accuracy (98.2%) when implemented with ResNet50. Besides, the proposed model generates 11.07% and 6.44% higher precision and recall in mask detection when compared to the recent public baseline model published as RetinaFaceMask detector. The outstanding performance of the proposed model is highly suitable for video surveillance devices.

1. Introduction

The 209th report of the world health organization (WHO) published on 16th August 2020 reported that coronavirus disease (COVID-19) caused by acute respiratory syndrome (SARS-CoV2) has globally infected more than 6 Million people and caused over 379,941 deaths worldwide [1]. According to Carissa F. Etienne, Director, Pan American Health Organization (PAHO), the key to control COVID-19 pandemic is to maintain social distancing, improving surveillance and strengthening health systems [2]. Recently, a study on understanding measures to tackle COVID-19 pandemic carried by the researchers at the University of Edinburgh reveals that wearing a face mask or other covering over the nose and mouth cuts the risk of Coronavirus spread by avoiding forward distance travelled by a person's exhaled breath by more than 90% [3]. Steffen et al. also carried an exhaustive study to compute the community-wide impact of mask use in general public, a portion of which may be asymptomatically infectious in New York and Washington. The findings reveal that near universal adoption (80%) of even weak masks (20% effective) could prevent 17–45% of projected deaths over two months in New Work and reduces the peak daily death rate by 34–58% [4,5]. Their results strongly recommend the use of the face masks in general public to curtail the spread of Coronavirus. Further, with the reopening of countries from COVID-19 lockdown, Government and Public health agencies are recommending face mask as essential measures to keep us safe when venturing into public. To mandate the use of facemask, it becomes essential to devise some technique that enforce individuals to apply a mask before exposure to public places.

Face mask detection refers to detect whether a person is wearing a mask or not. In fact, the problem is reverse engineering of face detection where the face is detected using different machine learning algorithms

* Corresponding author.
E-mail address: mamtakathuria@jcboseust.ac.in (M. Kathuria).

https://doi.org/10.1016/j.jbi.2021.103848
Received 28 September 2020; Received in revised form 10 June 2021; Accepted 20 June 2021
Available online 24 June 2021
1532-0464/

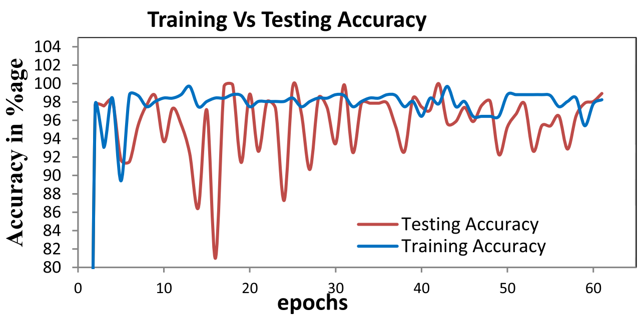

Fig. 10. Training and Testing Accuracy over 60 Epochs.

6.44% in the face and mask detection respectively. We had observed that improved results are possible due to optimized face detector discussed in Section 3.3 for dealing with complex images.

4.6. Controlling overfitting

To address RQ5 and avoid the problem of overfitting, two major steps are taken. First, we performed data augmentation as discussed in Section 3.1.2. Second, the model accuracy is critically observed over 60 epochs both for the training and testing phase. The observations are reported in Fig. 10.

It is further observed that model accuracy keeps on increasing in different epochs and get stable after epoch = 3 as depicted graphically in Fig. 10 above. To summarize the experimental results, we can say that the proposed model achieves high accuracy in face and mask detection with less inference time and less memory consumption as compared to recent techniques. Significant efforts had been put to resolve the data imbalance problem in the existing MAFA dataset, resulting in a new unbiased dataset which is highly suitable for COVID related mask detection tasks. The newly created dataset, optimal face detection approach, localizing the person identity and avoidance of overfitting resulted in an overall system that can be easily installed in an embedded device at public places to curtail the spread of Coronavirus.

5. Conclusion and future scope

In this work, a deep learning-based approach for detecting masks over faces in public places to curtail the community spread of Coronavirus is presented. The proposed technique efficiently handles occlusions in dense situations by making use of an ensemble of single and two-stage detectors at the pre-processing level. The ensemble approach not only helps in achieving high accuracy but also improves detection speed considerably. Furthermore, the application of transfer learning on pre-trained models with extensive experimentation over an unbiased dataset resulted in a highly robust and low-cost system. The identity detection of faces, violating the mask norms further, increases the utility of the system for public benefits.

Finally, the work opens interesting future directions for researchers. Firstly, the proposed technique can be integrated into any high-resolution video surveillance devices and not limited to mask detection only. Secondly, the model can be extended to detect facial landmarks with a facemask for biometric purposes.

CRediT authorship contribution statement

Shilpa Sethi: Conceptualization, Methodology, Writing - original draft. **Mamta Kathuria:** Data curation, Conceptualization, Writing - original draft. **Trilok Kaushik:** Implementation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] World Health Organization et al. Coronavirus disease 2019 (covid-19): situation report, 96. 2020. - Google Search. (n.d.). https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200816-covid-19-sitrep-209.pdf?sfvrsn=5dde1ca2_2.

[2] Social distancing, surveillance, and stronger health systems as keys to controlling COVID-19 Pandemic, PAHO Director says - PAHO/WHO | Pan American Health Organization. (n.d.). https://www.paho.org/en/news/2-6-2020-social-distancing-surveillance-and-stronger-health-systems-keys-controlling-covid-19.

[3] L.R. Garcia Godoy, et al., Facial protection for healthcare workers during pandemics: a scoping review, BMJ, Glob. Heal. 5 (5) (2020), e002553, https://doi.org/10.1136/bmjgh-2020-002553.

[4] S.E. Eikenberry, et al., To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic, Infect. Dis. Model. 5 (2020) 293–308, https://doi.org/10.1016/j.idm.2020.04.001.

[5] Wearing surgical masks in public could help slow COVID-19 pandemic's advance: Masks may limit the spread diseases including influenza, rhinoviruses and coronaviruses – ScienceDaily. (n.d.). https://www.sciencedaily.com/releases/2020/04/200403132345.htm.

[6] L. Nanni, S. Ghidoni, S. Brahnam, Handcrafted vs. non-handcrafted features for computer vision classification, Pattern Recogn. 71 (2017) 158–172, https://doi.org/10.1016/j.patcog.2017.05.025.

[7] Y. Jia et al., Caffe: Convolutional architecture for fast feature embedding, in: MM 2014 - Proceedings of the 2014 ACM Conference on Multimedia, 2014, doi: 10.1145/2647868.2654889.

[8] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. Lecun, OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks, 2014.

[9] D. Erhan, C. Szegedy, A. Toshev, D. Anguelov, Scalable Object Detection using Deep Neural Networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 2147–2154, https://doi.org/10.1109/CVPR.2014.276.

[10] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: Unified, real-time object detection, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016, vol. 2016-Decem, pp. 779–788, doi: 10.1109/CVPR.2016.91.

[11] M. Jiang, X. Fan, and H. Yan, RetinaMask: A Face Mask detector, 2020, http://arxiv.org/abs/2005.03950.

[12] M. Inamdar, N. Mehendale, Real-Time Face Mask Identification Using Facemasknet Deep Learning Network, SSRN Electron. J. (2020), https://doi.org/10.2139/ssrn.3663305.

[13] S. Qiao, C. Liu, W. Shen, A. Yuille, Few-Shot Image Recognition by Predicting Parameters from Activations, in: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2018, https://doi.org/10.1109/CVPR.2018.00755.

[14] A. Kumar, Z.J. Zhang, H. Lyu, Object detection in real time based on improved single shot multi-box detector algorithm, J. Wireless Com. Netw. 2020 (2020) 204, https://doi.org/10.1186/s13638-020-01826-x.

[15] Á. Morera, Á. Sánchez, A.B. Moreno, Á.D. Sappa, J.F. Vélez, SSD vs. YOLO for detection of outdoor urban advertising panels under multiple variabilities, Sensors (Switzerland) (2020), https://doi.org/10.3390/s20164587.

[16] R. Girshick, J. Donahue, T. Darrell, J. Malik, Region-based Convolutional Networks for Accurate Object Detection and Segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 38 (1) (2015) 142–158, https://doi.org/10.1109/TPAMI.2015.2437384.

[17] K. He, X. Zhang, S. Ren, J. Sun, Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition, IEEE Trans. Pattern Anal. Mach. Intell. (2015), https://doi.org/10.1109/TPAMI.2015.2389824.

[18] R. Girshick, Fast R-CNN, in: Proc. IEEE Int. Conf. Comput. Vis., vol. 2015 Inter, 2015, pp. 1440–1448, doi: 10.1109/ICCV.2015.169.

[19] N.D. Nguyen, T. Do, T.D. Ngo, D.D. Le, An Evaluation of Deep Learning Methods for Small Object Detection, J. Electr. Comput. Eng. 2020 (2020), https://doi.org/10.1155/2020/3189691.

[20] Z. Cai, Q. Fan, R.S. Feris, N. Vasconcelos, A unified multi-scale deep convolutional neural network for fast object detection, Lect. Notes Comput. Sci. (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (2016), https://doi.org/10.1007/978-3-319-46493-0_22.

[21] C.-Y. Fu, W. Liu, A. Ranga, A. Tyagi, A.C. Berg, DSSD : Deconvolutional Single Shot Detector, 2017, arXiv preprint arXiv:1701.06659 (2017).

[22] A. Shrivastava, R. Sukthankar, J. Malik, A. Gupta, Beyond Skip Connections: Top-Down Modulation for Object Detection, 2016, arXiv preprint arXiv:1612.06851 (2016).

[23] N. Dvornik, K. Shmelkov, J. Mairal, C. Schmid, BlitzNet: A Real-Time Deep Network for Scene Understanding, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, doi: 10.1109/ICCV.2017.447.

[24] Z. Liang, J. Shao, D. Zhang, L. Gao, Small object detection using deep feature pyramid networks, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2018, vol. 11166 LNCS, pp. 554–564, doi: 10.1007/978-3-030-00764-5_51.

[25] K. He, G. Gkioxari, P. Dollar, R. Girshick, Mask R-CNN, in: Proc. IEEE Int. Conf. Comput. Vis., vol. 2017-Octob, 2017, pp. 2980–2988, doi: 10.1109/ICCV.2017.322.