

An Enhanced Deep Learning Model for Automatic Face Mask Detection

Qazi Mudassar Ilyas¹ and Muneer Ahmad^{2,*}

¹Department of Information Systems, College of Computer Sciences and Information Technology, King Faisal University, Al-Ahsa, 31982, Saudi Arabia

²Department of Information Systems, Faculty of Computer Science & Information Technology, Universiti Malaya, 50603, Kuala Lumpur, Malaysia

*Corresponding Author: Muneer Ahmad. Email: mmalik@um.edu.my

Received: 22 February 2021; Accepted: 09 May 2021

Abstract: The recent COVID-19 pandemic has had lasting and severe impacts on social gatherings and interaction among people. Local administrative bodies enforce standard operating procedures (SOPs) to combat the spread of COVID-19, with mandatory precautionary measures including use of face masks at social assembly points. In addition, the World Health Organization (WHO) strongly recommends people wear a face mask as a shield against the virus. The manual inspection of a large number of people for face mask enforcement is a challenge for law enforcement agencies. This work proposes an automatic face mask detection solution using an enhanced lightweight deep learning model. A surveillance camera is installed in a public place to detect the faces of people. We use MobileNetV2 as a lightweight feature extraction module since the current convolution neural network (CNN) architecture contains almost 62,378,344 parameters with 729 million floating operations (FLOPs) in the classification of a single object, and thus is computationally complex and unable to process a large number of face images in real time. The proposed model outperforms existing models on larger datasets of face images for automatic detection of face masks. This research implements a variety of classifiers for face mask detection: the random forest, logistic regression, K-nearest neighbor, neural network, support vector machine, and AdaBoost. Since MobileNetV2 is the lightest model, it is a realistic choice for real-time applications requiring inexpensive computation when processing large amounts of data.

Keywords: Face mask detection; image classification; deep learning; MobileNetV2; sustainable health; COVID-19 pandemic; machine intelligence

1 Introduction

The outbreak of novel coronavirus (SARS-CoV-2), referred to here as COVID-19, was declared a pandemic by the World Health Organization (WHO) in March 2020 [1]. Health experts believe that the respiratory droplets transmitted through the nose, eyes, or mouth are among major sources of transmission of this virus [2,3]. Maintaining a high level of hygiene, minimizing touching surfaces and/or



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

use of disposable gloves, limiting social contact, and maintaining social distance in public places have been proposed by health experts as key measures to control the spread of respiratory droplets. Use of protective face masks has been another commonly suggested strategy in several pandemics in the past [4,5]. Matuschek et al. [6] provide an excellent review on use of protective face masks since the middle ages. In the early phase of the pandemic, owing to a lack of information, the WHO and several governments did not recommend use of a protective face mask for the common public. However, with improved understanding of the novel coronavirus and its spread pattern, these guidelines were revised. Today, scientists and health professionals agree that completely covering the nose and mouth with a face mask, especially in public places, is an effective strategy to limit the spread of COVID-19 [7–9]. As a result, most governments and businesses have revised their standard operating procedures (SOPs) for public conduct and have made it mandatory for individuals to wear a face mask [10–12].

Use of a protective face mask, although considered to be very effective in stopping the spread of the COVID-19 pandemic, suffers from some limitations. Several users have reported headaches and other adverse psychological effects from wearing a face mask [13]. These undesirable effects have a negative impact on certain individuals to keep wearing a face mask. In these circumstances, the authorities are left with no other option but to force wearing of a face mask. Manual inspection and enforcement are tedious processes. Additionally, it exposes the enforcing authorities to possible contraction of the virus.

The recent advancements in imaging technologies coupled with the success of machine learning applications in computer vision in several domains make a strong case for use of such technologies to assist in face mask recognition and enforcement. Deep learning methods exploit artificial neural networks and feature learning to perform complex learning tasks such as object identification [14]. Computer vision applications generally use convolutional neural networks (CNNs) for feature extraction and learning. Although CNNs are more efficient compared to fully connected multilayer perceptrons, they require huge computational resources for learning a model. When using a CNN, even a computer vision problem of modest complexity requires learning hundreds of millions of weights; these weights are learned in an iterative fashion, and may even take weeks to learn [15].

A number of enhanced network architectures have been proposed to improve the efficiency of convolutional neural networks; such enhanced networks include flattened convolutional neural networks, XNor-networks, factorized convolutional neural networks, and quantized convolutional neural networks. Howard et al. [16] introduced MobileNets, which are optimized convolutional networks for mobile vision applications. MobileNets use depthwise separable convolutions to reduce the computational resources required for learning the model without compromising performance. MobileNetV2 further improves MobileNets by using 3×3 depthwise separable convolutions, resulting in comparable performance with significant improvement in computational costs. We propose to use MobileNetV2 for feature extraction, and then feed the features into a variety of classifiers for face mask detection.

Some highlights and key observations of this study are as follows: (a) face masks are SOPs during the COVID-19 pandemic; (b) real-time manual inspection of face mask appears as a challenge to authorities; (c) we propose an automatic face mask detection method using an enhanced DL model; (d) existing DL models are computationally expensive for real-time applications; and (e) we adopt MobileNetV2 as a lightweight feature extraction model, and implement a variety of classifiers. Ultimately, the proposed model is more realistic for pattern-matching applications than existing models. The remainder of this paper is organized as follows. Section 2 critically analyzes the related works on face mask detection. Section 3 details the proposed methodology for feature extraction and classification. Results are discussed in Section 4, and the paper is concluded in Section 5.

2 Literature Review

Machine learning- and deep learning-based solutions have been widely adopted in automatic face detection and recognition for surveillance purposes. A variety of deep convolution models have significantly reduced the computational time and increased the detection accuracy, thus empowering them as excellent choices for automatic image processing applications. The recent COVID-19 outbreak has limited social interactions between people, and authorities realized the need to implement standard operating procedures to limit the virus' spread [17,18]. Among several procedures, the face mask has been declared in many countries as a mandatory measure to control the virus [19]. Law enforcement agencies and local public authorities are ensuring that people observe this face mask policy, but monitoring a large number of people at social venues is an open challenge [20].

It is evident from Fig. 1 that face detection for surveillance has widely employed machine learning and deep learning approaches in the past decade. In contrast, there is a limited number of publications addressing the automatic detection of face masks to mitigate the risk of COVID-19 spread. We present here a review of the literature that adopts machine leaning-based methods for automatic face mask detection.

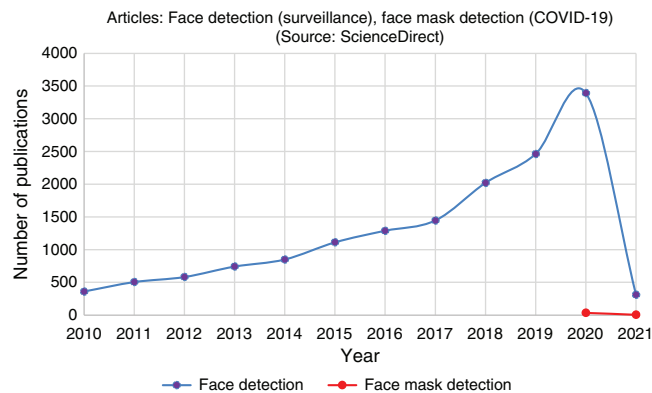


Figure 1: A trend of research publications on face detection and face mask detection

Loey et al. [21] proposed a hybrid deep learning model for face mask detection to limit the spread of COVID-19. The proposed model consisted of feature extraction employing Resnet50. The study classified face images using a variety of classifiers, including an ensemble classifier, decision tree, and support vector machine (SVM). The SVM classifier achieved 99.64%, 99.49%, and 100% testing accuracies for RMFD, SMFD, and LFW, respectively.

Similarly, Qin and Li [22] addressed the detection of face masks by proposing a face mask identification method, called SRCNet, that integrates image super-resolution networks (called SR networks) and classification networks. They quantified the three-category classification problem focusing on unconstrained 2D face images by conducting basic image preprocessing, followed by facial detection with super-resolution networks. SRCNet achieved a 98.70% detection accuracy, outperforming the conventional image classification methods that employ deep learning.

Park et al. [23] proposed an approach to remove glasses from human face images. To accomplish this, the regions surrounding the glasses were detected and observed. This method incorporated PCA reconstruction with recursive error compensation.

Li et al. [24] employed YOLOv3 for object detection to take advantage of its speed and accuracy. In this approach, face detection problems were addressed, included the detection of smaller faces. It was observed that the Softmax classifier outperformed the traditional logistic classifier in segregating the inter-class

features; this helped in decreasing the dimension of features. Meanwhile, the authors trained the procedure on the CelebA database and tested it using the Fddb database. YOLOv3 performed better than other existing techniques in the recognition of small faces while maintaining a good compromise of speed and accuracy.

In the same direction, Ud Din et al. [25] proposed a deep learning-based solution for the removal of masked objects in facial images. The study addressed the issues of face mask identification that mostly occur in facial image processing since the mask covers a large portion of the face. The proposed method segmented the process into two layers: one for mask object detection, and the other for removing the mask region. The first layer generated the binary segmentation related to the mask area. The second layer removed the mask, synthesizing with the fine details of segmenting the mask-area obtained in the first layer. The authors adopted a GAN-based network employing two discriminators. The first discriminator had a global structure, and it focused on learning relevant to the area of face not covered by the mask. They also employed the CelebA dataset for performance analysis of the proposed detection system.

Meanwhile, Khan et al. [26] emphasized the removal of certain objects from images while maintaining image quality. Microphone objects were removed from face images. The vacant areas where the microphone used to be were adjusted with the neighboring image context. This task was accomplished using an interactive method named MRGAN. However, its main limitation is that the user must specify the microphone region. Since the removal of microphone created a hole in the image, the authors then adopted an image-to-image translation method called the generative adversarial network. The procedure was comprised of two steps: inpainting, and refinement. The inpainting step performed prediction with low accuracy after filling the area occupied by the microphone earlier (it became a hole after microphone removal). The refinement step improved the prediction accuracy. Tests on the CelebA dataset showed MRGAN's superiority over other image manipulation approaches.

Hussain et al. [27] proposed a simplified three-objective face detection and recognition system with emotion classification based on deep learning. The authors employed the Open CV library with Python programming to ensure real-time efficacy. Furthermore, Wang et al. [28] identified the issues involved in the existing face mask recognition systems that employ deep learning solutions on a set of facial images. The study proposed a framework based on three masked faces datasets (MFDD, RMFRD, and SMFRD), and the proposed multigranularity masked face recognition model achieved an accuracy of 95% on RMFRD.

Among the literature related to face mask detection and mitigating the risk of COVID-19 spread, few works have addressed the optimization issues involved in face mask detection. The design and employment of a robust deep learning model can significantly enhance mask detection accuracy on publicly available datasets. Keeping pace with this current research trend, we propose an enhanced deep learning model for automatic face mask detection in the COVID-19 scenario. Since the proposed model is lightweight, it can be adopted for dedicated edge computing computer vision systems for a large number of images, especially in Industry 4.0 applications.

3 Methodology

Current machine learning models for automatic COVID-19 face mask detection require computationally inexpensive solutions for their deployment in real-time edge computing devices and other portable image processing applications. The present research intends to develop a lightweight machine learning model for automatic detection of face masks at social assembly points in the COVID-19 context. Feature extraction from a large set of images is computationally expensive in both time and space. To manage this tradeoff, MobileNetV2 is realized as a lightweight model for image feature extraction. Besides, we employ several classifiers to enhance the model's accuracy in face mask detection.

Fig. 2 describes the preprocessing and feature extraction steps. Let us assume a COVID-19 dataset that contains a collection $C = \{I_1, I_2, I_3, I_4, \dots, I_n\}$, $n \leq N$ of n images from a pool of N image space. Any image vector I of this image space C is considered as a “D” dimension vector such that $I \in \mathbb{R}^D$. Furthermore, we consider $I \in \mathbb{R}^{X \times Y \times Z}$ where each X , Y , and Z describe respective row, column, and color vectors, respectively. These vectors can be further demonstrated with particular indices a , b , and c such that $0 \leq a \leq D$, $0 \leq b \leq D$ and $0 < c \leq 3$.

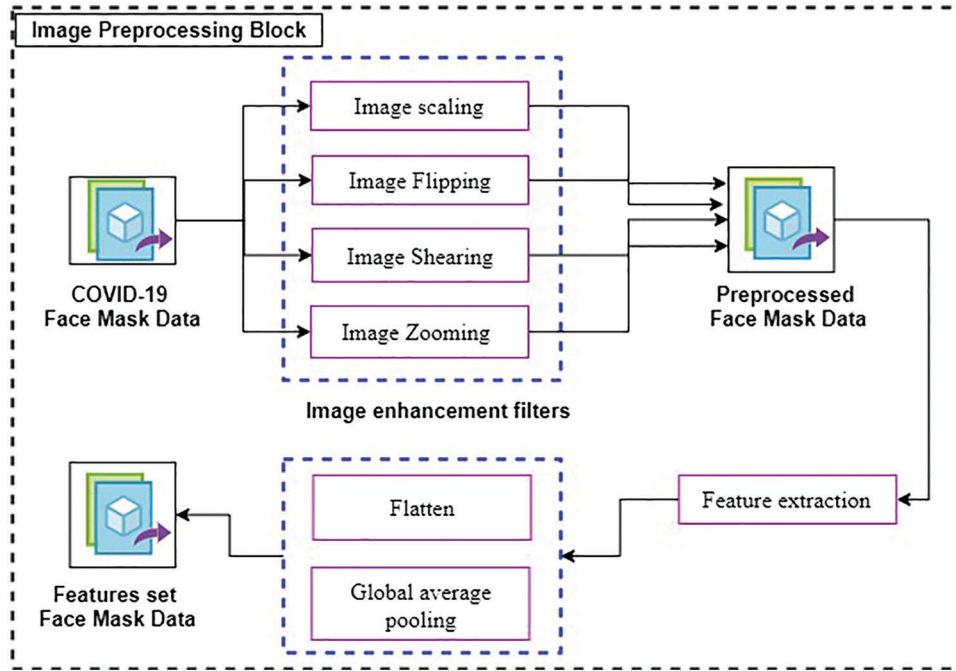


Figure 2: Preprocessing and feature extraction

The preprocessing steps are further explained below.

- Let the width and height of an image be represented with w and h , respectively. The scaling filter gives us a scaled image I' (w' , h') such that $(w', h') = S$, where S is the maximum value $(w', h') = \frac{S}{\max(w, h)}(w, h)$.
- The vertical and horizontal shearing to image vector I with coordinates p and q is described in Eqs. (1) and (2).

$$\begin{pmatrix} p' \\ q' \end{pmatrix} = \begin{pmatrix} p + mq \\ q \end{pmatrix} = \begin{pmatrix} 1 & m \\ 0 & 1 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \quad (1)$$

$$\begin{pmatrix} p' \\ q' \end{pmatrix} = \begin{pmatrix} p \\ mx + q \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ m & 1 \end{pmatrix} \begin{pmatrix} p \\ q \end{pmatrix} \quad (2)$$

- Zooming is applied to image vector I to obtain a zoomed image vector I' such that I' is approximately ($r \times 10\%$) of I . Here, r is the specific point of interest in image I with coordinates x and y . The zooming bestows us a zoomed vector with points (zx, zy) as a displacement of r . Zooming is mostly based on trial and error, we need to find the best zoomed-in image so that the required enhanced image vector best serves our purpose.
- The horizontal flip of an image vector I with coordinates x and y produces an image I' with coordinates x' and y' such that $x' = \text{width}(\text{vector } I) - x - 1$ while $y' = y$. We can demonstrate the process as follows:

Repeat (1): x in range (width of vector I)
 Repeat (2): y in range (height of vector I)
 $x' = \text{width}(\text{vector } I) - x - 1$
 End Repeat (2):
 End Repeat (1):

Fig. 3 describes the general classification model. The random forest, logistic regression, K-nearest neighbor, neural network, support vector machine, and AdaBoost are employed to enhance the face mask detection. We demonstrate here the CNN schema that applies to convolving the image vector I so that a series of convolution operations are described as

$$A^P \rightarrow B^P \rightarrow C^Q \rightarrow D^Q \rightarrow \dots \rightarrow A^Z \rightarrow B^Z \rightarrow C^T.$$

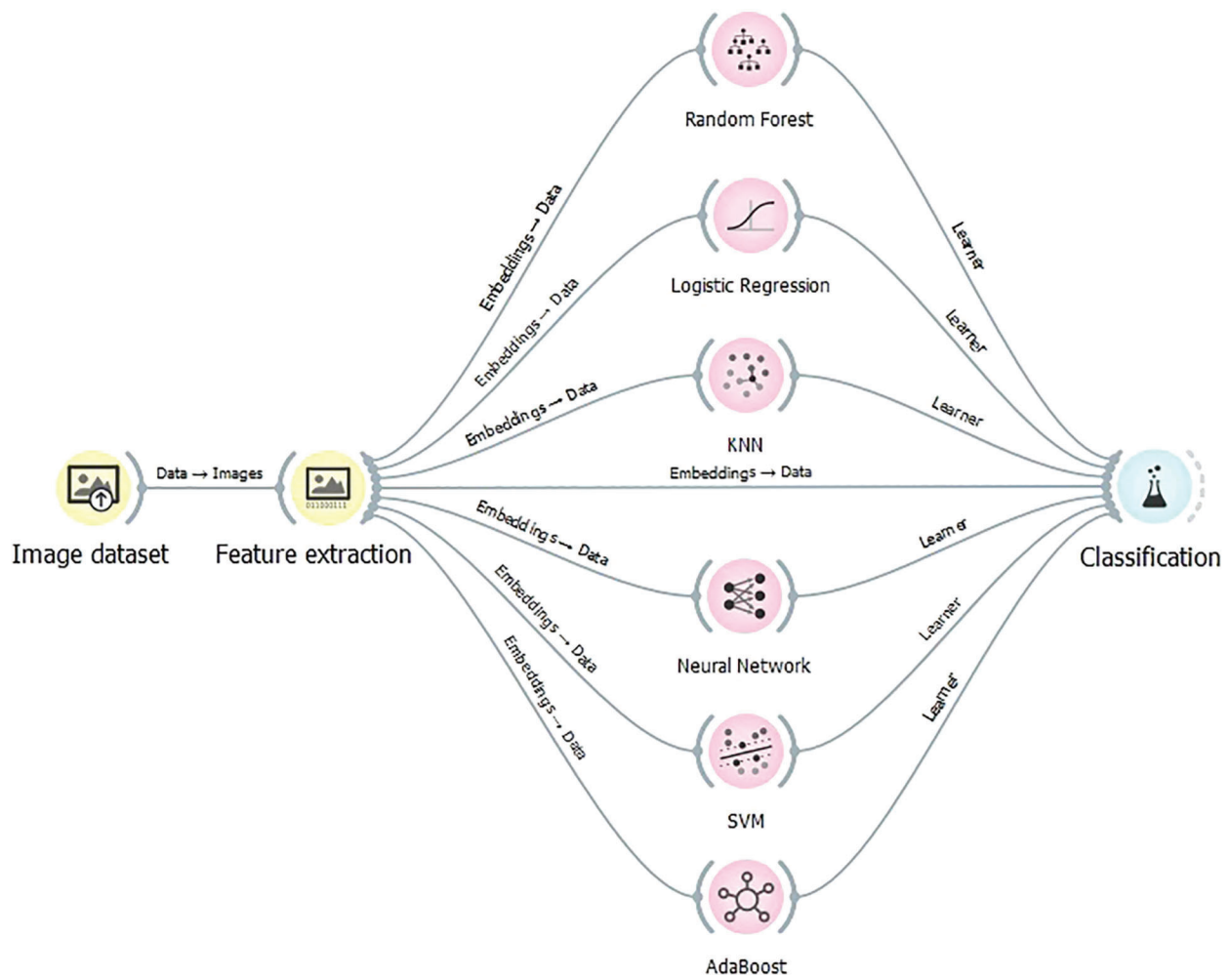


Figure 3: Classification model

Let us consider vector A as an input image. The first layer of the CNN incorporates a weight B (which is a vector numerically applied to A), and its output serves as the input for the next layer. Similarly, we take vector C^Q as the output of the first layer and D^Q as the weight of the second layer. The convolution

operation of C^Q and D^Q produces another vector. This process continues for a defined number of layers. Although, the convolution neural network is considered to produce optimal classification, we executed a variety of classification algorithms to enhance face mask detection.

Fig. 4 presents the schematic process of classifying the COVID-19 face mask images. The COVID-19 dataset contains different selected features and instances. Let us assume a collection C of data that contains N instances defined as $C = \{I_1, I_2, I_3, I_i\}$, $i \leq N$. These instances represent a collection of unprocessed images. We apply preprocessing filters, i.e., scaling, shearing, flipping, and zooming, to C_i instances for $i \leq N$ to remove inconsistencies. Next, we extract the features of I_i images for $i \leq N$. We then split C_i instances of collection C into a defined composition of training and test data using split-train-test criteria for $i \leq N$. The training data is employed for training the individual classifiers. The test data help in generating a test case scenario to find the probability of prediction, i.e., the accuracy of a certain classifier.

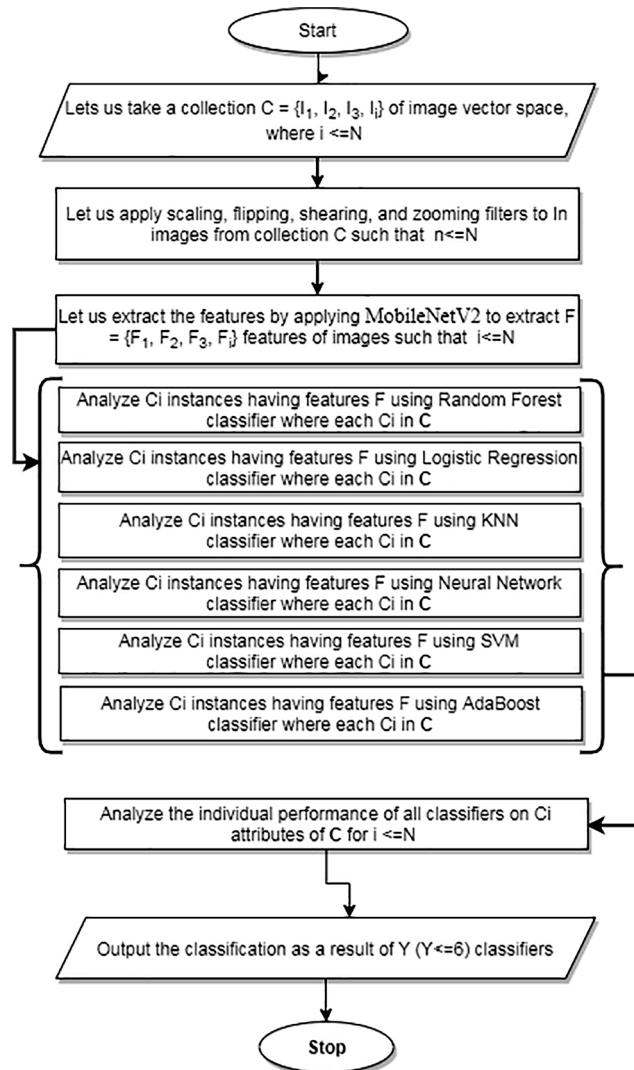


Figure 4: Schematic presentation of classification process

Next, the system analyzes C_i ($i \leq N$) instances using random forest, logistic regression, K-nearest neighbor, neural network, support vector machine, and AdaBoost to predict the accuracy of COVID-19 face mask image features.

4 Results and Discussion

The precision and accuracy of automatic face mask detection classification models require a larger, comprehensive, preprocessed set of images. Since a significant amount of time has passed since the onset of COVID-19 scenarios in most parts of the world, several publicly available datasets are readily available in different data repositories. For this study, we developed a dataset for use alongside data randomly selected from public repositories. In total, we collected 3000 images, with 1500 images each for categories “with mask” and “without mask.”

4.1 Feature Subset Selection

Overall, the dataset contains the 3000 embedded images that describe 3000 instances. Each image has 20 numeric features along with five meta-attributes (three numeric and two text attributes). The target classification (or class attribute) is represented by category. The feature extractor applied to this dataset generates 2048 features. Through experimenting with different combinations of features, we observed that the selection of the feature subset can significantly enhance detection accuracy and other metrics.

Fig. 5 illustrates the selection of the feature subset using linear projection. Twenty features that contain the strongest correlations were identified. The blue and red circles represent the data projected linearly against the 20 features for “with mask” and “without mask,” respectively. Most of the instances in this dataset can be well described using this feature subset. In addition, the disparity of data in the linear project reveals an explicit segregation between the two classes.

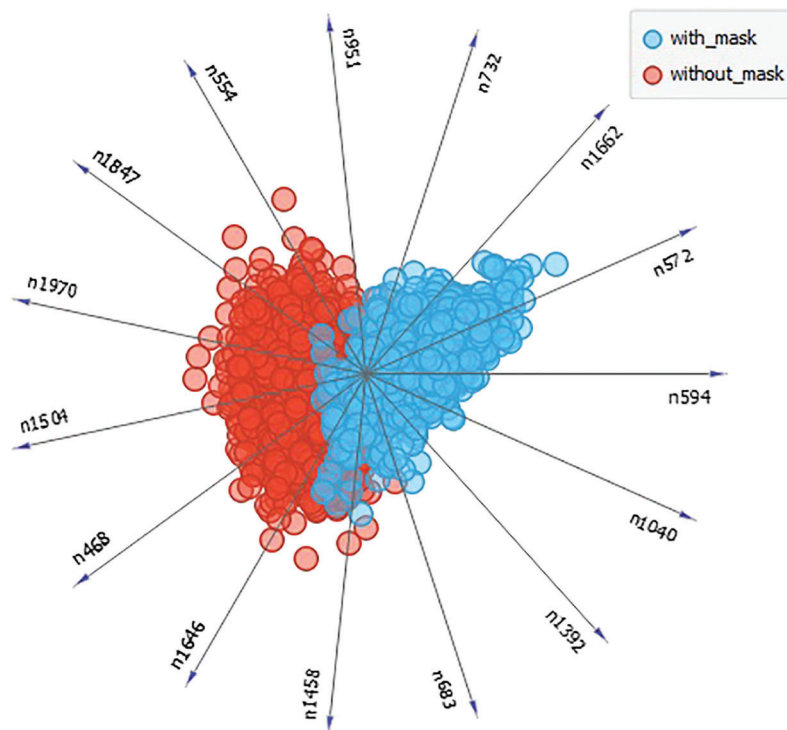


Figure 5: Feature subset selection (linear projection)

We further investigated the correlations between information gain, gain ratio, Gini index, ANOVA, and chi-square measures; the results are depicted in Fig. 6. The 20 features described in the linear projection can be validated by employing these correlation measures. The red line below the correlation value shows the relevant correlation of a feature. For instance, the first four features have higher values of information gain, gain ratio, Gini index, chi-square measures, and ANOVA. Meanwhile, features n951 to n1117 show variations in ANOVA (lower values compared to those of the first four features).

#	Info. gain	Gain ratio	Gini	ANOVA	χ^2
N n594	0.710	0.355	0.389	5967.028	1734.764
N n572	0.672	0.336	0.374	5146.399	1675.670
N n1662	0.668	0.334	0.369	5920.441	1675.670
N n732	0.666	0.333	0.368	5687.643	1673.230
N n951	0.656	0.328	0.362	3373.403	1656.200
N n554	0.644	0.322	0.357	4786.888	1634.432
N n1847	0.628	0.314	0.352	4800.213	1589.080
N n1970	0.624	0.312	0.347	3768.786	1608.022
N n1504	0.621	0.311	0.347	4884.941	1600.857
N n468	0.613	0.306	0.345	4657.140	1586.574
N n1646	0.603	0.302	0.334	4216.888	1572.356
N n1458	0.602	0.301	0.335	3516.190	1569.992
N n683	0.597	0.298	0.337	3857.991	1548.800
N n1392	0.594	0.297	0.334	4913.246	1523.094
N n1040	0.593	0.296	0.333	4210.348	1555.848
N n2033	0.591	0.295	0.331	4266.014	1553.497
N n585	0.565	0.282	0.319	3572.489	1499.912
N n569	0.564	0.282	0.323	4586.376	1499.912
N n1529	0.561	0.281	0.324	4647.072	1488.388
N n1117	0.557	0.278	0.318	3917.881	1486.088

Figure 6: Correlation among feature subsets

Fig. 7 presents the validation of the feature subset selected using the correlation measures. It is trivial to describe most of the data points amid the selected features. Features “n1662” and “n1646” give an impression of a distant-proximity relation as compared to of the other 18 features. The PCA here validates the selection of the feature subset from the primary dataset.

4.2 Performance Evaluation of Different Classifiers

We investigated the performance of different classifiers using the following evaluation metrics:

- True Negative (TN): The individual has no face mask, and the classifier correctly predicts the same.
- False Positive (FP): The individual has no face mask, but the classifier incorrectly predicts that the individual has a face mask.
- False Negative (FN): The individual has a face mask, but the classifier incorrectly predicts that the individual does not have a face mask.
- True Positive (TP): The individual has a face mask, and the classifier correctly predicts the same.

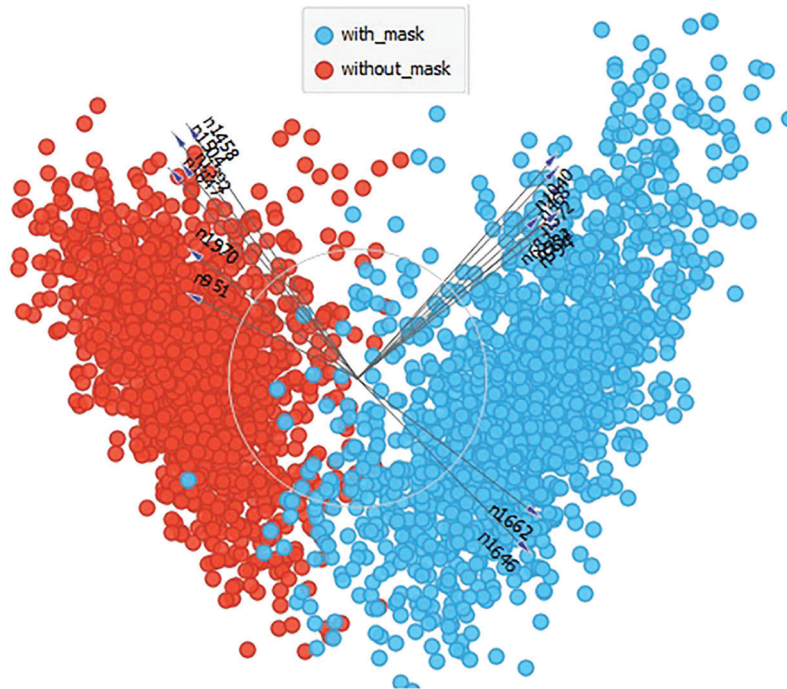


Figure 7: Validation of feature subset using principal component analysis (PCA)

To evaluate classification accuracy, we use the values from a confusion matrix. The evaluation metrics of specificity, sensitivity, and F1 score are calculated.

a) Specificity: This metric correctly identifies the individual with no face mask.

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100\% \quad (3)$$

b) Sensitivity: This metric correctly identifies the individual with a face mask.

$$\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100\% \quad (4)$$

c) F1-Score: This metric measures the overall performance of the model.

$$\text{F1 - Score} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \times 100\% \quad (5)$$

We applied these evaluations metrics to the logistic regression, neural network, KNN, SVM, random forest, and AdaBoost classifiers.

Fig. 8 presents confusion matrices depicting the prediction accuracies of the classifiers. The logistic regression performed better than the other classifiers by only 0.2%, and the neural network demonstrated comparable performance. The AdaBoost algorithm underperformed in comparison with the other classifiers, with 3.2% and 3.4% wrong prediction for [with_mask] and [without_mask], respectively.

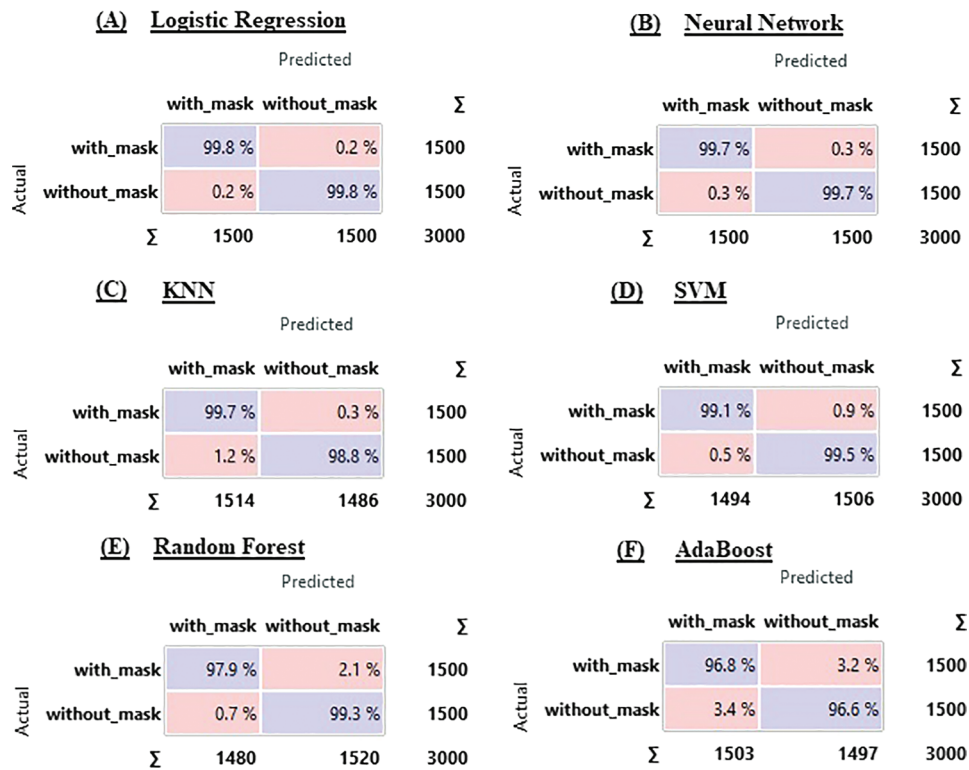


Figure 8: Confusion matrices related to prediction accuracies of different classifiers

Tab. 1 presents the cross tabulation of classifiers in terms of their performance correlation. It is worth noting that the rows present the pivot as compared to columns. Small probabilities are negligible since the difference between the classifiers’ scores is below the chosen neglecting threshold of 0.1. Logistic regression performs better than the neural network, with a significant probability of 0.836. Similarly, it performs better than the KNN, SVM, random forest, and AdaBoost, with probabilities of 0.981, 0.968, 0.997, and 1.000, respectively. For instance, there is a probability of 0.164 that the logistic regression will not perform better than the neural network classifier. Similarly, the probability that logistic regression could not perform better than AdaBoost is 0%.

Table 1: Cross-correlation (Pivot) analysis of classifier performance

Model	Logistic regression	Neural network	KNN	SVM	Random forest	AdaBoost
Logistic Regression		0.836	0.981	0.968	0.997	1.000
Neural Network	0.164		0.969	0.872	0.998	1.000
KNN	0.019	0.031		0.500	0.963	1.000
SVM	0.032	0.128	0.500		0.888	1.000
Random Forest	0.003	0.002	0.037	0.112		1.000
AdaBoost	0.000	0.000	0.000	0.000	0.000	

The performance evaluation of different classifiers is presented in Fig. 9. It can be noticed that the logistic regression classifier outperforms the other classifiers. Although the neural network classifier performs comparably for the area under the curve metric, it performed worse than the logistic regression in regards to prediction accuracy, F1-score, precision, and recall. Meanwhile, the AdaBoost classifier was inferior to all of the other classifiers. Similarly, the performance of SVM and KNN is almost the same; they have similar values for all evaluation measures.

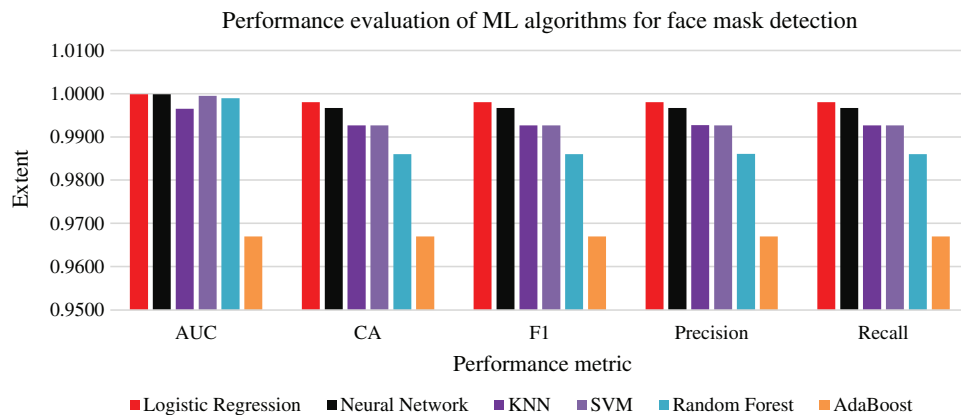


Figure 9: Performance evaluation of different classifiers

Fig. 10 presents the receiver operating characteristics curve (ROC) analysis based on the true positive rate and false positive rate. The true positive rate is the sensitivity while the false positive rate is (1-specificity). The ROC curve analysis depicts the significant achievement of logistic regression in enhancing the detection accuracy. While the AdaBoost ROC curve is far away from the desired curve (0, 1), the logistic regression curve falls very close to the ideal curve. Hence, the logistic regression outperforms the other classifiers in face mask detection.

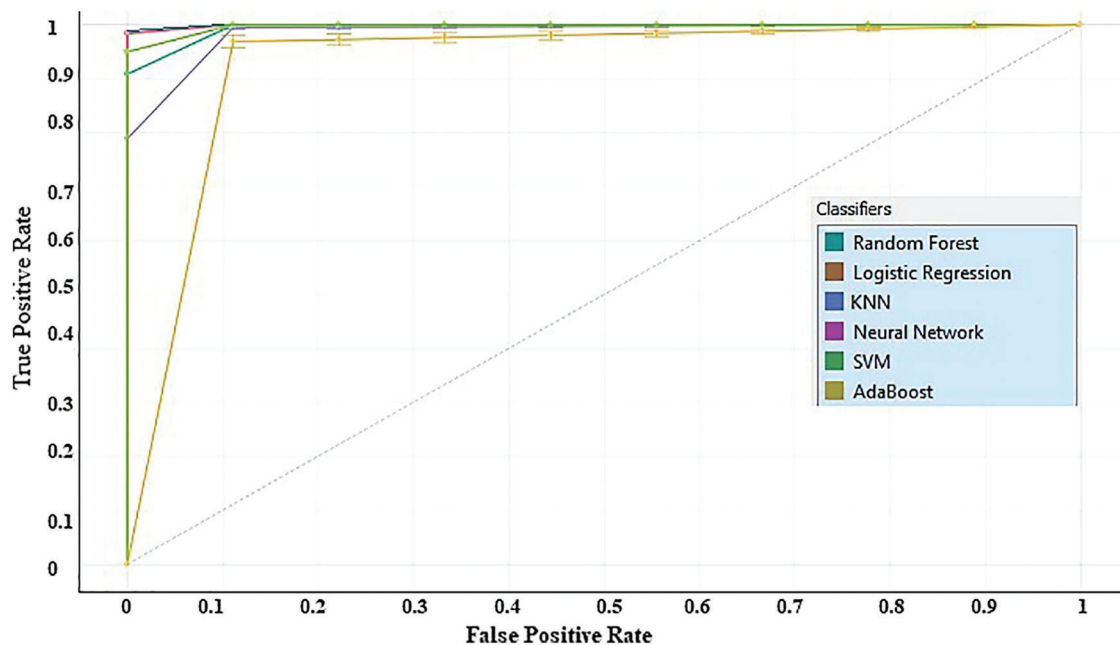


Figure 10: Receiver operating characteristics curve (ROC) analysis

5 Conclusion

COVID-19 standard operation procedures as defined by most countries require people to wear face masks, and it is quite challenging for law enforcement agencies to inspect larger crowds to ensure compliance. Such manual inspection is tedious and laborious. Thus, we presented an automatic face mask detection system to aid in prompt identification of individuals not wearing face masks. We used MobileNetV2 as a lightweight feature extraction module, and implemented a variety of classifiers for face mask detection, i.e., random forest, logistic regression, K-nearest neighbor, neural network, support vector machine, and AdaBoost. Logistic regression outperformed the other classifiers in the selection of the feature subset. This subset selection is a significant predictor for enhancing the prediction accuracy of classifiers. We also found that MobileNetV2 is a more realistic choice for real-time applications requiring inexpensive computations when processing large amounts of data.

Funding Statement: The authors extend their appreciation to the Deanship of Scientific Research at King Faisal University, Saudi Arabia for funding this research work through KFU Nasher Track 2020 with research Grant Number 206128.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] K. G. Andersen, A. Rambaut, W. I. Lipkin, E. C. Holmes and R. F. Garry, "The proximal origin of SARS-CoV-2," *Nature Medicine*, vol. 6, no. 4, pp. 450–452, 2020.
- [2] T. Galbadage, B. M. Peterson and R. S. Gunasekera, "Does COVID-19 spread through droplets alone?," *Front Public Health*, vol. 8, no. 166, pp. 1–4, 2020.
- [3] E. C. Abebe, T. A. Dejenie, M. Y. Shiferaw and T. Malik, "The newly emerged COVID-19 disease: A systemic review," *Virology Journal*, vol. 17, no. 1, pp. 1–8, 2020.
- [4] C. Lynteris, "Plague masks: The visual emergence of anti-epidemic personal protection equipment," *Medical Anthropology Cross Cultural Studies in Health and Illness*, vol. 37, no. 5, pp. 1–16, 2018.
- [5] J. T. F. Lau, J. H. Kim, H. Tsui and S. Griffiths, "Anticipated and current preventive behaviors in response to an anticipated human-to-human H5N1 epidemic in the Hong Kong Chinese general population," *BMC Infectious Diseases*, vol. 7, no. 1, pp. 1–18, 2007.
- [6] C. Matuschek, F. Moll, H. Fangerau, J. C. Fischer, K. Zänker *et al.*, "The history and value of face masks," *European Journal of Medical Research*, vol. 25, no. 1, pp. 1–6, 2020.
- [7] K. H. Chan and K. Y. Yuen, "COVID-19 epidemic: Disentangling the re-emerging controversy about medical facemasks from an epidemiological perspective," *International Journal of Epidemiology*, vol. 49, no. 4, pp. 1760, 2020.
- [8] R. V. Tso and B. J. Cowling, "Importance of face masks for COVID-19: A call for effective public education," *Clinical Infectious Diseases*, vol. 71, no. 16, pp. 2195–2198, 2020.
- [9] X. Liu and S. Zhang, "COVID-19: Face masks and human-to-human transmission," *Influenza and other Respiratory Viruses*, vol. 14, no. 4, pp. 472–473, 2020.
- [10] D. Szczesniak, M. Ciulkowicz, J. Maciaszek, B. Misiak, D. Luc *et al.*, "Psychopathological responses and face mask restrictions during the COVID-19 outbreak: Results from a nationwide survey," *Brain Behavior, and Immunity*, vol. 87, no. 1, pp. 161–162, 2020.
- [11] L. H. Chen, D. O. Freedman and L. G. Visser, "COVID-19 immunity passport to ease travel restrictions?," *Journal of Travel Medicine*, vol. 27, no. 5, pp. 1–3, 2020.
- [12] J. L. Scheid, S. P. Lupien, G. S. Ford and S. L. West, "Commentary: Physiological and psychological impact of face mask usage during the covid-19 pandemic," *International Journal of Environmental Research and Public Health*, vol. 17, no. 18, pp. 6655, 2020.
- [13] Y. Lecun, Y. Bengio and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

- [14] W. Haensch, T. Gokmen and R. Puri, "The next generation of deep learning hardware: Analog computing," *Proc. of the IEEE*, vol. 107, no. 1, pp. 108–122, 2019.
- [15] M. Rastegari, V. Ordonez, J. Redmon and A. Farhadi, "XNOR-net: Imagenet classification using binary convolutional neural networks," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, Amsterdam, The Netherlands, 2016.
- [16] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang *et al.*, "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv e-prints:1704.04861, 2017.
- [17] A. Khalid and S. Ali, "COVID-19 and its challenges for the healthcare system in Pakistan," *Asian Bioethics Review*, vol. 12, no. 3, pp. 1–14, 2020.
- [18] V. C. C. Cheng, S. C. Wong, V. W. M. Chuang, S. Y. C. So, J. H. K. Chen *et al.*, "The role of community-wide wearing of face mask for control of coronavirus disease 2019 (COVID-19) epidemic due to SARS-CoV-2," *Journal of Infection*, vol. 81, no. 1, pp. 107–114, 2020.
- [19] C. R. MacIntyre and A. A. Chughtai, "A rapid systematic review of the efficacy of face masks and respirators against coronaviruses and other respiratory transmissible viruses for the community, healthcare workers and sick patients," *International Journal of Nursing Studies*, vol. 108, no. 103629, pp. 1–6, 2020.
- [20] Y. Wang, H. Tian, L. Zhang, M. Zhang, D. Guo *et al.*, "Reduction of secondary transmission of SARS-CoV-2 in households by face mask use, disinfection and social distancing: A cohort study in Beijing," *China BMJ Global Health*, vol. 5, no. 5, pp. 1–9, 2020.
- [21] M. Loey, G. Manogaran, M. H. N. Taha and N. E. M. Khalifa, "A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic," *Measurement*, vol. 167, no. 108288, pp. 1–11, 2021.
- [22] B. Qin and D. Li, "Identifying facemask-wearing condition using image super-resolution with classification network to prevent COVID-19," *Sensors*, vol. 20, no. 18, pp. 1–23, 2020.
- [23] J. S. Park, Y. H. Oh, S. C. Ahn and S. W. Lee, "Glasses removal from facial image using recursive error compensation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 805–811, 2005.
- [24] C. Li, R. Wang, J. Li and L. Fei, "Face detection based on YOLOv3," *Advances in Intelligent Systems and Computing*, vol. 1006, no. 5, pp. 805–811, 2020.
- [25] N. Ud Din, K. Javed, S. Bae and J. Yi, "A novel GAN-based network for unmasking of masked face," *IEEE Access*, vol. 8, no. 1, pp. 44276–44287, 2020.
- [26] M. K. J. Khan, N. Ud Din, S. Bae and J. Yi, "Interactive removal of microphone object in facial images," *Electron*, vol. 8, no. 10, pp. 1–15, 2019.
- [27] S. A. Hussain and A. Salim Abdallah Al Balushi, "A real time face emotion classification and recognition using deep learning model," *Journal of Physics: Conference Series*, vol. 1432, pp. 012087, 2020.
- [28] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong *et al.*, "Masked face recognition dataset and application," 2020. [Online]. Available: <https://arxiv.org/pdf/2003.09093.pdf>.