Mask and respirator detection: analysis and potential solutions for a frequently ill-conditioned problem

Mask and respirator detection: analysis and potential solutions for a frequently ill-conditioned problem

Antonio Costantino Marceddu
Department of Control and Computer Engineering
Politecnico di Torino
Turin, Italy
0000-0003-4843-1267

Renato Ferrero
Department of Control and Computer Engineering
Politecnico di Torino
Turin, Italy
0000-0003-4459-4843

Bartolomeo Montrucchio
Department of Control and Computer Engineering
Politecnico di Torino
Turin, Italy
0000-0003-0065-8614

***Abstract*—During the coronavirus pandemic, the mask detection problem has become of particular interest. Usually, the goal is to create a system that can detect whether or not a person is wearing a mask or respirator. However, this tends to trivialize a problem that hides a greater complexity. In fact, people wear masks or respirators in various ways, many of which are incorrect. This makes the problem ill-conditioned and creates a bias compared to training cases, with the consequence that these systems have a considerably lower accuracy when used in practice. We claim that focusing on the ways in which a mask can be worn and classifying the problem not as binary but at least as ternary, thus adding an intermediate class containing all those ways in which a mask or respirator can be worn incorrectly, could help address this problem. For this reason, this paper describes and puts to the proof the Ways to Wear a Mask or a Respirator Database (WWMR-DB). It has a fine classification of the most common ways in which a mask or respirator is worn, which can be used to test how mask detection systems work in cases that resemble the real ones more. It was used to test a neural network, the ResNet-152, which was trained on less fine databases, like the Face-Mask Label Dataset and the MaskedFace-Net. The mixed results denote the shortcomings of these databases and the need to enhance them or resort to finer databases.**

***Keywords*—Artificial intelligence, Computer vision, Image databases, Machine learning, Neural networks**

I. Introduction

The advent of the coronavirus pandemic has radically changed both the way we live and our perception of certain behaviors. After months and months of emergency, since countries have introduced restrictions aimed at minimizing the chances of contagion, for most people across the globe it has become normal to wear masks or respirators and to avoid crowding as much as possible. Yet, many continue to assume incorrect and deleterious behaviors, which could compromise the virus containment efforts made up to now: the crowds for the race to the last console, for the fear of a new lockdown or the Black Friday and Cyber Monday offers on Christmas gifts, are clear evidence of this [1] [2] [3] [4]. Crowds are not the only problem though: people tend to wear masks or respirators in a multitude of ways, mostly incorrect. This partially or completely nullifies their actions of protection and/or reduction of the spread of respiratory droplets and considerably increases the chances of transmission of the virus. A possible solution to this problem is to design mask detection systems, capable of automatically detecting whether or not a person is wearing a mask or respirator. However, as masks and respirators are often not worn properly, this classification tends not to be exhaustive and inevitably creates a bias regarding the situations in which these systems have been trained, lowering their accuracy when used in the wild, i.e. in practice.

To put in light this problem a novel database, called Ways to Wear a Mask or a Respirator Database (WWMR-DB), was realized and released in open-access format [5]. It finely classifies how a mask or respirator could be worn, so it can be used to evaluate mask detection systems and databases. In particular, in this paper it will be used to test a neural network, the ResNet-152, which was trained on less fine databases, the Face-Mask Label Dataset and the MaskedFace-Net [6] [7], to show their limitations in scenarios similar to the real-world.

This paper is structured as follows. The next section presents some recent researches regarding the mask detection problem. The WWMR-DB features are listed in Section III. Section IV and V exploit the WWMR-DB to evaluate a neural network trained with two datasets that oversimplify the heterogeneity of masks and respirators and ways of wearing them. Finally, conclusions deducible from the obtained results and some proposals for future improvements are presented in Section VI.

II. Related Work

Since the beginning of the coronavirus pandemic, the number of databases for detecting the presence of a mask or respirator on the face has increased significantly. In March 2020 three databases were jointly presented: Masked Face Detection Dataset (MFDD), Real-world and Simulated Masked Face Recognition Datasets (respectively, RMFRD and SMFRD) [8]. In the same month, a ready-to-be-trained facial mask classifier was released on GitHub. The classifier contains a self-made dataset of 1376 images, of which 690 images of people wearing a mask [9]. It has been exploited for training and evaluating several mask detection systems, achieving validation accuracies of ~99% [10] [11]. In October 2020 two new datasets, called MFV and MFI, were presented to improve facial recognition systems in case of occlusions due to the presence of a mask [12].

Fig. 1. Example images of classes in the WWMR-DB. From top-left to right-bottom: mask or respirator not worn, mask or respirator correctly worn, mask or respirator under the nose, mask or respirator under the chin, mask or respirator hanging from an ear, mask or respirator on the tip of the nose, mask folded above the chin and mask or respirator on the forehead.

The previously mentioned databases and systems achieve great precision in detecting the presence of a mask above the mouth, but their precision decreases if the mask is worn incorrectly (under the chin, hanging from an ear, etc.). Other works have taken into account this problem in order to find a solution. In May 2020, Humans In The Loop released the Medical Mask Dataset, which contains 6K images acquired from the public domain and covering 20 classes of different accessories, as well as a classification of faces with a mask, without a mask, or with an incorrectly worn mask [13]. In November 2020, the MaskedFace-Net dataset was published [7]. It was obtained starting from the Flickr-Faces-HQ3 (FFHQ), a dataset portraying people of different ages, gender, and ethnicity, and applying a sophisticated algorithm capable of automatically generating masks, in different positions, on the faces of the people portrayed [14]. In February 2021, the Face-Mask Label Dataset (FMLD), was released. This is a more accurate relabeling of some of the images of the WIDER FACE [15] and MAFA [16] database, in which the subclass "Worn incorrectly" has been added. It was used to train a system capable of recognizing whether a mask is worn correctly (Compliant) or whether it is worn incorrectly or not worn (Non-Compliant), obtaining a maximum recognition accuracy of 98.79% [6].

III. The Ways To Wear A Mask Or A Respirator Database

This section will discuss the design, realization, and characteristics of the WWMR-DB dataset.

A. Design Phase

The design phase of the database was inspired by the following question: "what are the most common ways in which people are used to wearing masks?". After an evaluation period in which the experience of direct observation of the problem has been combined with related articles [17] [18], the following eight classes have been identified as the most representative:

- Mask or respirator not worn.
- Mask or respirator correctly worn.

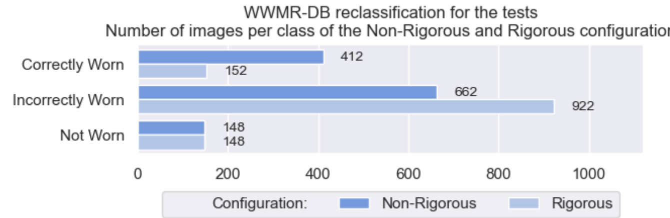


Fig. 2. Reclassification of the WWMR-DB database images into the Rigorous and Non-Rigorous configurations.

- Mask or respirator under the nose.
- Mask or respirator under the chin.
- Mask or respirator hanging from an ear.
- Mask or respirator on the tip of the nose.
- Mask folded above the chin.
- Mask or respirator on the forehead.

An example of these positions is shown in Fig. 1. A good system must work accurately regardless of the rotation of the face. By discretizing the yaw rotations of the face with respect to the camera, 3 representative values of these changes were chosen: 0, 45, and 90 degrees.

B. Realization Phase

The best way to create the database would have been to take pictures to volunteers in a specially equipped photographic laboratory. However, it was not possible to accomplish this task for three main reasons:

- Part of the research team was in a forced smart working.
- Due to the strict COVID-19 regulations, it was difficult to ensure adequate safety for all participants.
- Movement limitations were frequent.

It was therefore decided to resort to a "virtual" approach, creating Google Forms that participants use to take selfies similar to the sample photos in total safety. Each volunteer could participate more times with a different type of mask or respirator, among the following classes:

- Disposable Respirator With Valve.
- Disposable Respirator Without Valve.
- Surgical Mask.
- Non-Medical Mask.

Since 8 classes, 3 possible head yaw rotations, and 4 different types of masks or respirators were considered, the number of images required was quite high: therefore, to encourage participation, no constraints were put on the number of photos to upload, but, in case the participants had uploaded a certain number of images, they could have received one or more tickets for the drawing of some prizes. For this reason, each participant contributed with a variable number of photos. The photos contained in the database were submitted by 42 different participants in the period between September 2020 and January 2021.

Fig. 3. Some images found in the WIDER FACE and MAFA datasets, of which FMLD improved relabeling [6].

C. Database Characteristics

The database is composed of 1222 images depicting 42 people wearing a mask or a respirator in the most common positions, as defined in "Design Phase" subsection. At the time of writing, approximately 60% of participants are male, and 40% female, mostly between 20 and 40 years old. Participants were asked to take photos in three different face rotations, adequately illustrated with sample images: 0, 45 and 90 degrees. Without strict requirements, there can be small variations in image quality, position assumed, and yaw rotation of the face. However, this is beneficial because it increases the variability of the images. For each image there are two label files in PascalVOC and YOLO format, containing the positioning information of the bounding boxes relating to mask or respirator worn, person's face, and a union of previous cases.

An open-source program, LabelImg, was used for their realization [19]. Furthermore, each image was named using a format that indicates, in the name itself, subject number, position assumed, yaw rotation of the face, and type of mask or respirator worn. The presence of classification for masks and respirators opens the way for another use: the WWMR-DB could also be combined with other databases having a similar classification to create a system able to recognize the mask or respirator typology, like, for example, the Facial Mask and Respirator Dataset (FMR-DB) and the Medical Mask dataset [13] [20] [21].

D. Other Information

For the "Mask folded above the chin" case, participants were required to use only surgical or non-medical masks, as respirators are thicker and it is difficult that someone wears them this way.

The WWMR-DB database is publicly available in an open-access format [5]. Google Forms are still active, so the number of images of the database is also increasing. Anyone can participate in this improvement by sending its photos: the images will be accurately classified, labeled, and put at disposal of the research community [22] [23] [24].

E. Preface to The Test Sections

The WWMR-DB has a fine classification regarding the positions in which masks and respirators are worn, but in this paper it has not been used in full as, at the time of writing this paper, it appears that no system is capable of detecting how a mask or respirator is worn with such granularity, i.e. with such number of classes. Images were then grouped into three superclasses: Correctly Worn, Incorrectly Worn, and Not Worn.

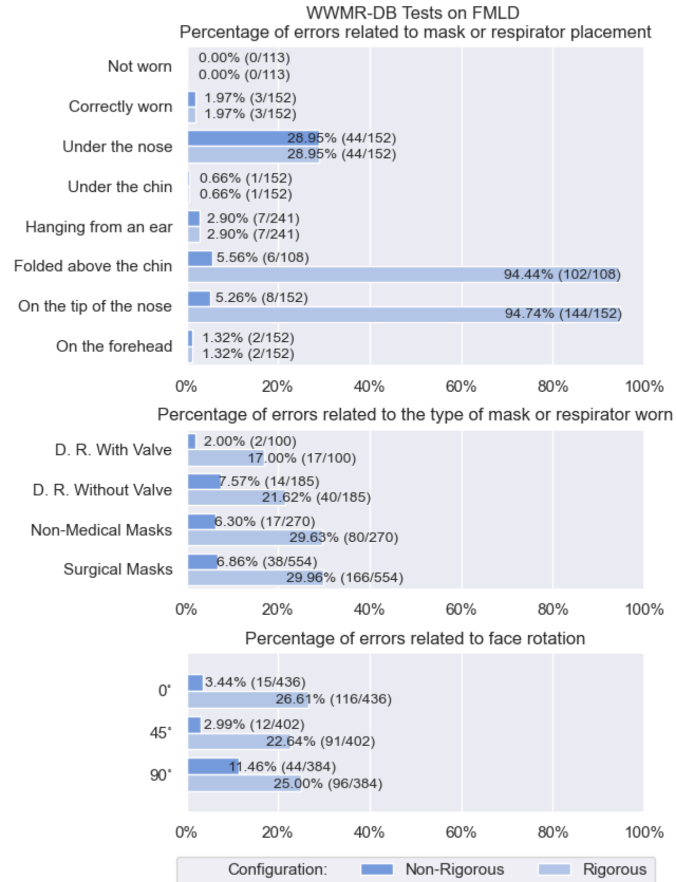


Fig. 4. WWMR-DB tests on FMLD: percentage of errors due to the positioning and type of mask or respirator worn (top and center) and the rotation of the subject's face (bottom).

The superclass referred to as "Correctly Worn" contemplates only the position defined in the WWMR-DB as "Mask or respirator correctly worn". The other positions fall into the superclass "Incorrectly Worn", except for the case where the mask or respirator is not worn. In the rest of the section, this configuration will be referred to as "Rigorous". Since this configuration is rather strict, another looser configuration, named "Non-Rigorous", has been defined. Here, a mask or a respirator is considered correctly worn even if put on the tip of the nose or if it leaves the chin exposed. These positions, however, must be considered wrong because the mask or respirator, not adhering well to the face, allows the passage of a greater quantity of air: in the first case this occurs in the spaces between the mask and the nostrils, while in the second case it occurs at the height of the cheeks, more precisely in the space created by the fold that the mask assumes. The re-subdivision of the WWMR-DB images for each class of the problem and each configuration is shown in Fig. 2.

Finally, the WWMR-DB images were cropped according to the bounding box labeled "person", which identifies people's faces in the label files. This label ensures correct cropping for each face but, for some images, it has completely removed the masks or made them barely visible. For this reason, 29 of them have been reclassified as "Not Worn". The list of reclassified images is shown in the file attached to this paper.

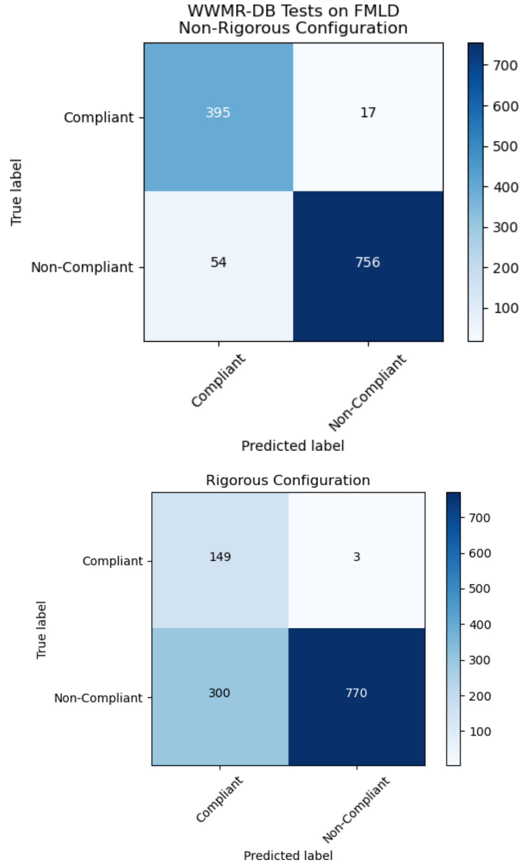


Fig. 5. WWMR-DB tests on FMLD: confusion matrices of the tests run with the pre-trained ResNet-152 neural network.

IV. Evaluation of a Low-Granularity Dataset

The Face-Mask Label Dataset (FMLD) [6] was briefly discussed in Section II. Unlike the WWMR-DB dataset, it contains a reduced number of classes and is composed of images taken from the Internet. An example is shown in Fig. 3. It was originally used to train and test different neural networks in order to design a pipeline for accurate mask detection: recurring the same settings, the neural network that performed best was ResNet-152, obtaining a 98.93% accuracy. As it was published with the paper, it was decided to use it to perform new tests. Here is a summary of the settings that have been assigned to the aforementioned neural networks:

- The outputs of the network are the “Compliant” (masks worn correctly) and “Non-Compliant” (masks worn incorrectly or not worn) cases.
- Image size was set to 224x224 pixels and a random 80:20 split of training and validation set was performed.
- Transfer learning was applied by initializing the model weights with the parameters learned pre-training the neural network with the ImageNet database [24].
- Cross-Entropy was used as learning objective.
- Stochastic Gradient Descent (SGD) with a momentum of 0.9 and a learning rate of 0.001, reduced by a factor of 0.1 every 7 epochs, was used as optimization algorithm.

Fig. 6. Example of an image of the FFHQ (left) and the corresponding image with artificial transformations in the CMFD (center) and in the IMFD (right).

- The training was repeated 10 times for a maximum of 10 epochs, shuffling images and randomly dividing them between training and validation sets.

It was therefore necessary to adapt both the Rigorous and Non-Rigorous configurations of the WWMR-DB to the aforementioned “Compliant” - “Non-Compliant” classification. So, for this test the “Incorrectly Worn” and “Not Worn” superclasses were grouped into the Non-Compliant class, while the “Correctly Worn” superclass was renamed as Compliant.

A. Results with the WWMR-DB

In the Non-Rigorous configuration, the pre-trained neural network obtained an accuracy of 94.19%. From the error analysis depicted in Fig. 4, it is possible to notice that the neural network perfectly recognizes if a person is not wearing any mask or respirator but, at the same time, it tends to make more mistakes with faces rotated at 90 degrees and with masks under the nose. In the Rigorous configuration, accuracy dropped to 75.2%, mainly because most of the images classified as “Mask folded above the chin” or “Mask or respirator on the tip of the nose” were predicted to be “Correctly Worn”. The confusion matrices shown in Fig. 5 confirm that the pre-trained system worked quite well in the first configuration, but not in the second one, which is more representative of the real case. In ~29% of cases, images classified as “Under the nose” were predicted as “Correctly Worn”. In the Rigorous classification alone, the two classes which differentiate it from the Non-Rigorous classification had an error rate of ~95%. As a final comment, it can be said that the neural network trained with FMLD performs well enough in the Non-Rigorous configuration test but, as evidenced by the Rigorous configuration test, it does not seem suitable for modeling the problem in its entirety.

V. Evaluation of a Dataset with Only One Kind of Mask

Unlike the WWMR-DB images, MaskedFace-Net [7] consists of images taken from the Internet and containing only a blue surgical mask, which is the most common type on the market. This mask was artificially superimposed on the images of the FFHQ dataset [14], so originating two datasets: the Correctly Masked Face Dataset (CMFD) and the Incorrectly Masked Face Dataset (IMFD). CMFD contains people wearing a mask correctly, while IMFD contains people wearing a mask incorrectly. An example of how an image of the FFHQ dataset was transformed to obtain the corresponding image of the CMFD and the IMFD is shown in Fig. 6. Since the “Not Worn” case is missing, the FFHQ dataset was used for it. CMFD, IMFD, and FFHQ datasets contain ~70K images each. Using the same images for the training, even if they have artificial

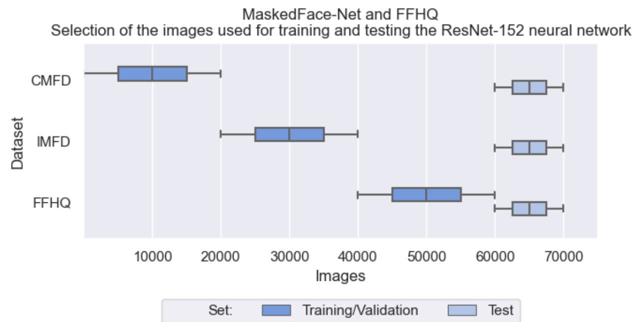


Fig. 7. Selection of the images used for training and testing the ResNet-152 neural network. This subdivision was made to reduce the bias due to the reuse, even if only partial, of the same image. To reduce the possibility of learning superfluous features, for the training and validation phases it was made sure the use of different images, while for the test phase this was not necessary.

alterations, can increase bias. For this reason, for the training and validation phases were used only a part of the databases images: in particular, ~20K images were used for each class, for a total of ~60K images. 80% of them were then used for training and the remaining 20% for validation. The bias problem does not persist in the testing phase. For this reason, the original and augmented version of each of the remaining unique images were used, for a total of ~30K images. A detailed image of the subdivision adopted is shown in Fig. 7.

To better compare the new results with those obtained previously, the same neural network (Resnet-152) and the same settings were used, except for the batch size, which was reduced to 16 because of memory constraints, and for the use of data augmentation techniques in order to reduce overfitting.

A. Training Configuration

A configuration composed of Intel® Core™ i7 8750H, NVIDIA® GeForce® RTX 2080 8GB GDDR6, 32 GB DDR4, 2666 MHz, and 2x2 TB PCIe M.2 SSD was used for the training. GPU was preferred over the CPU for its high parallelization capability, capable of increasing training performance.

B. Results

After the training, the neural network was evaluated on the test dataset, obtaining an impressive 99.96% accuracy. Only 16 misprediction errors were made, all regarding the "Incorrectly Worn" class. Due to the excellent results, it is useless to show the confusion matrix.

C. Results with the WWMR-DB

In the Non-Rigorous configuration the neural network achieved an accuracy of 24.71%, while in the Rigorous configuration the neural network achieved an accuracy of 18.33%. Those results, which may have been caused by insufficient database granularity and the lack of a variety of masks in terms of type and color, put in a different light the optimum results of the test dataset. From the error analysis and the confusion matrix of both tests, represented in Fig. 8 and 9, it is possible to notice that the neural network has only learned to perfectly recognize the people who do not wear masks. The same cannot be said for the other cases, which report a very high amount of errors. The percentage of error even reaches 100% in the case of "mask or respirator on the forehead" and, in the

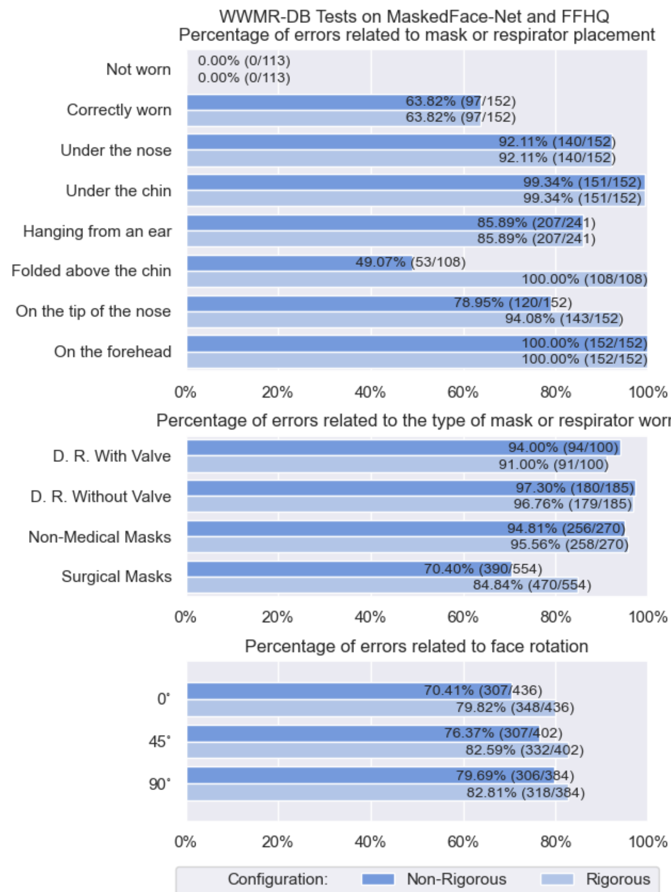


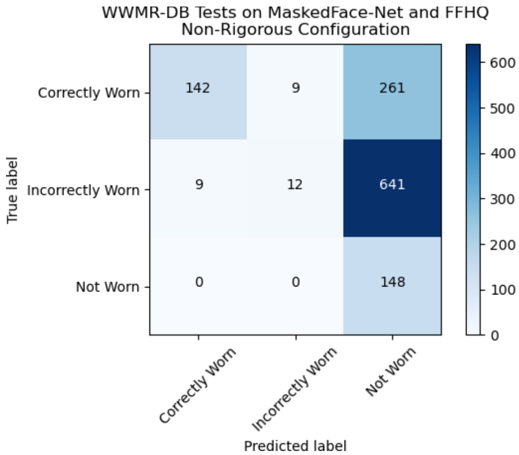
Fig. 8. WWMR-DB tests on MaskedFace-Net and FFHQ: percentage of errors, relative to the respective total, due to the positioning and type of mask or respirator worn (top and center images) and the rotation of the subject's face (bottom).

Rigorous configuration only, also in the case of "mask folded above the chin". As those results are comparable with those obtained using a random classifier, at the time of writing MaskedFace-Net may be ineffective for building systems that could work effectively in the real case.

VI. Conclusion

This paper discussed how the bad wearing of masks and respirators affects the accuracy of their detection. Since masks and respirators databases are commonly used to create systems of this type, a new database, the WWMR-DB, has been created with the intent of testing those systems in situations that resemble more the real case.

What emerges from the results is that the current mask and respirator detection systems cannot fully address this problem. Firstly, the problem should be considered at least ternary, adding at least one intermediate class between the "Correctly Worn" or "Not Worn" cases in order to take into account the "Incorrectly Worn" case. Secondly, it is necessary to focus more on the ways in which a mask or respirator is worn, which have been a major cause of test failure. The first point was already solved by some databases, while the second issue, also based on what can be seen from the results, was not satisfactorily addressed. Hence, richer databases with finer granularity are needed and WWMR-DB classes can solve this gap.



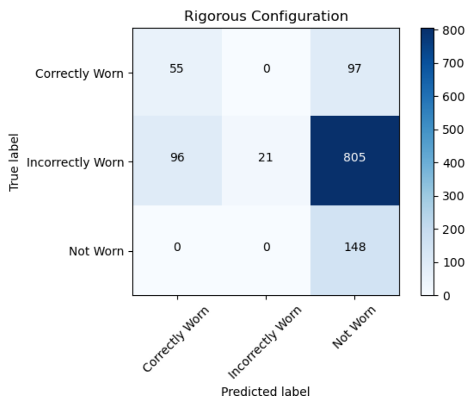


Fig. 9. WWMR-DB tests on MaskedFace-Net and FFHQ: confusion matrices of the tests run with the neural network trained on the aforementioned datasets.

References

[1] S. Best, “PlayStation 5: Huge queues form to buy console despite Sony confirming it's not available in-store,” *Mirror*, November 20, 2020. [Online]. Available: https://www.mirror.co.uk/tech/playstation-5-huge-queues-form-23034652. [Accessed Jan. 16, 2022]

[2] D. Sales, “Rationing is back! Tesco limits flour, eggs, toilet roll and other essentials as shoppers hit Oxford Street in desperate Christmas gift-buying dash before lockdown,” *Daily Mail*, November 2, 2020. [Online]. Available: https://www.dailymail.co.uk/news/article-8905187/Did-say-LOCKDOWN-Panic-buyers-form-train-trolleys-outside-Sainsburys-Costco.html. [Accessed Jan. 16, 2022]

[3] BBC Staff, “Covid-19: Crowds flee Dhaka ahead of strict Bangladesh lockdown,” BBC, June 27, 2021. [Online]. Available: https://www.bbc.com/news/world-asia-57624557. [Accessed Jan. 16, 2022]

[4] J. P. Clark, “Crowds flock to Edinburgh Christmas market despite Nicola Sturgeon's 'stay at home' warning,” Edimburgh Live, December 17, 2021. [Online]. Available: https://www.edinburghlive.co.uk/news/crowds-flock-edinburgh-christmas-market-22502541. [Accessed Jan. 16, 2022]

[5] A. C. Marceddu, R Ferrero and B. Montrucchio, “Ways To Wear a Mask or a Respirator Database (WWMR-DB),” *IEEE DataPort*, 2021, doi: 10.21227/8atn-gn55.

[6] B. Batagelj, P. Peer, V. Štruc, S. Dobrišek, “How to Correctly Detect Face-Masks for COVID-19 from Visual Information?,” *Applied Sciences*, vol. 11, no. 5, 2021, doi: 10.3390/app11052070.

[7] A. Cabani, K. Hammoudi, H. Benhabiles, M. Melkemi, “MaskedFace-Net - A dataset of correctly/incorrectly masked face images in the context of COVID-19,” Smart Health, vol. 19, 2021, doi: 10.1016/j.smhl.2020.100144.

[8] Z. Wang, G. Wang, B. Huang, Z. Xiong, Q. Hong, H. Wu, P. Yi, K. Jiang, N. Wang, Y. Pei, H. Chen, M. Yu, Z. Huang, J. Liang, “Masked Face Recognition Dataset and Application,” *ArXiv*, March 23, 2020. [Online]. Available: arXiv: 2003.09093.

[9] P. Bhandary, “Mask Classifier,” 2020, *GitHub*. [Online]. Available: https://github.com/prajnasb/observations.

[10] A. Rosebrock, “COVID-19: Face Mask Detector with OpenCV, Keras/TensorFlow, and Deep Learning,” *PyImageSearch*, May 4, 2020. [Online]. Available: https://www.pyimagesearch.com/2020/05/04/covid-19-face-mask-detector-with-opencv-keras-tensorflow-and-deep-learning/. [Accessed Jan. 16, 2022]

[11] A. Muritala, “Covid-19: Face Mask Detection,” Medium, May 18, 2020. [Online]. Available: https://medium.com/@abimbolamuritala65/covid-19-face-mask-detection-20216652ab6f. [Accessed Jan. 16, 2022]

[12] F. Ding, P. Peng, Y. Huang, M. Geng, Y. Tian, “Masked Face Recognition with Latent Part Detection,” *Proceedings of the 28th ACM International Conference on Multimedia (MM '20)*, 2020, pp. 2281–2289, doi: 10.1145/3394171.3413731.

[13] Humans In The Loop, “Medical Mask Dataset”, May 2020. [Online]. Available: https://humansintheloop.org/medical-mask-dataset/. [Accessed Jan. 16, 2022]

[14] T. Karras, S. Laine and T. Aila, "A Style-Based Generator Architecture for Generative Adversarial Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 12, pp. 4217-4228, 1 Dec. 2021, doi: 10.1109/TPAMI.2020.2970919.

[15] S. Yang, P. Luo, C. C. Loy and X. Tang, "WIDER FACE: A Face Detection Benchmark," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 5525-5533, doi: 10.1109/CVPR.2016.596.

[16] S. Ge, J. Li, Q. Ye and Z. Luo, “Detecting Masked Faces in the Wild with LLE-CNNs,” *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 426-434, doi: 10.1109/CVPR.2017.53.

[17] World Health Organization Staff, “Coronavirus disease (COVID-19) advice for the public: When and how to use masks,” *World Health Organization (WHO)*, December 1, 2020. [Online]. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks. [Accessed Jan. 16, 2022]

[18] M. Machida et al., “Incorrect Use of Face Masks during the Current COVID-19 Pandemic among the General Public in Japan,” International Journal of Environmental Research and Public Health, vol. 17, no.18, 2020, doi: 10.3390/ijerph17186484.

[19] Tzutalin et al., “LabelImg,” 2015, GitHub. [Online]. Available: https://github.com/tzutalin/labelImg.

[20] A. C. Marceddu and B. Montrucchio, “Facial Masks and Respirators Database (FMR-DB)”, *IEEE DataPort*, 2020, doi: 10.21227/wg71-v415.

[21] A. C. Marceddu and B. Montrucchio, "Recognizing the Type of Mask or Respirator Worn Through a CNN Trained with a Novel Database," *2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC)*, 2021, pp. 1490-1495, doi: 10.1109/COMPSAC51774.2021.00221.

[22] A. C. Marceddu, R. Ferrero, B. Montrucchio, “WWMR-DB. Form used to ask participants for front pictures,” Google Forms. [Online]. Available: https://forms.gle/qLBjfCVhGoaJhnSo9. [Accessed Jan. 16, 2022]

[23] A. C. Marceddu, R. Ferrero, B. Montrucchio, “WWMR-DB. Form used to ask participants for side pictures,” Google Forms. [Online]. Available: https://forms.gle/D3BuUQgjBLd6dqPj6. [Accessed Jan. 16, 2022]

[24] A. C. Marceddu, R. Ferrero, B. Montrucchio, “WWMR-DB. Form used to ask participants for profile pictures,” Google Forms. [Online]. Available: https://forms.gle/yHgCAgcGJrfC7X4YA. [Accessed Jan. 16, 2022]

[25] O. Russakovsky et al., “ImageNet Large Scale Visual Recognition Challenge,” *International Journal of Computer Vision*, vol. 115, 2015, pp. 211–252, doi: 10.1007/s11263-015-0816