POLITECNICO DI TORINO
Repository ISTITUZIONALE

Recognizing the Type of Mask or Respirator Worn Through a CNN Trained with a Novel Database

*Original*
Recognizing the Type of Mask or Respirator Worn Through a CNN Trained with a Novel Database / Marceddu, Antonio Costantino; Montrucchio, Bartolomeo. - ELETTRONICO. - (2021), pp. 1490-1495. ( 2021 IEEE 45th Annual Computers, Software, and Applications Conference (COMPSAC) Online 12-16 July 2021) [10.1109/COMPSAC51774.2021.00221].

*Availability:*
This version is available at: 11583/2914328 since: 2024-04-11T08:37:33Z

*Publisher:*
IEEE

*Published*
DOI:10.1109/COMPSAC51774.2021.00221

*Terms of use:*

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

*Publisher copyright*
IEEE postprint/Author's Accepted Manuscript

©2021 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

06 May 2026

# Recognizing the Type of Mask or Respirator Worn Through a CNN Trained with a Novel Database

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

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

***Abstract*—Since the onset of the coronavirus pandemic, researchers from all over the world have been working on projects aimed at countering its advance. The authors of this paper want to go in this direction through the study of a system capable of recognizing the type of mask or respirator worn by a person. It can be used to implement automatic entry controls in high protection areas, where people can feel comfortable and safe. It can also be used to make sure that people who work daily in contact with particles, chemicals, or other impurities wear appropriate respiratory protection. In this paper, a proof-of-concept of this system will be presented. It has been realized by using a state-of-the-art Convolutional Neural Network (CNN), EfficientNet, which was trained on a novel database, called the Facial Masks and Respirators Database (FMR-DB). Unlike other databases released so far, it has an accurate classification of the most important types of facial masks and respirators and their degree of protection. It is also at the complete disposal of the scientific community.**

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

## I. Introduction

The coronavirus pandemic has afflicted humanity for more than a year. Researchers from every area of study, as well as various companies, have used their knowledge to design and implement projects aimed at counteracting it. During this time mask detection systems, which are systems able to automatically detect whether or not a person is wearing a mask or respirator, become common. In this paper, the authors want to go further and propose a system capable of recognizing the relative typology as well. The basic idea is to implement automatic controls for entering high-security areas, where people can feel comfortable and safe. It can be used not only to counter the coronavirus advancement but also to protect people who work in contact with particles, chemicals, or other impurities: for example, it could be used in metalworking industries, in lumber mills, or in places where radioactive materials are used or studied, including nuclear power plants.

A proof-of-concept of this type of system will therefore be presented in this paper: it was implemented by training the EfficientNet-B0, a state-of-the-art Convolutional Neural Network (CNN), with a novel database, called Facial Masks and Respirators Database (FMR-DB) [1]. Unlike other databases released so far, it has an accurate classification of the most important types of facial masks and respirators and their degree of protection; having been released in open-access format, it is also at the complete disposal of the scientific community.

This paper is organized as follows:

- The next section presents some of the most important contributions related to the mask and respirator detection problem and the mask and respirator type detection problem.
- The characteristics of the FMR-DB are explained in Section III.
- Section IV reports information on the neural network and the parameters chosen to train it.
- The results of the test performed both with the dataset test and with ad-hoc images are reported in Section V.
- Finally, the conclusions and some ideas for future work are presented in Section VI.

## II. Related Work

The need to contain the advancement of the coronavirus has rapidly increased the number of databases for creating automatic mask and respirator detection systems [2] [3]. Until about two years ago, however, their number was limited and they were mainly used for the construction of automated safety systems [4] or to monitor the use of masks in highly polluted areas [5].

Unlike the simpler mask detection problem, the problem of recognizing the type of mask and respirator worn counts only a few contributions, specially made after the discovery of the coronavirus. In May 2020, Humans In The Loop released the Medical Mask Dataset [6], an open-access dataset containing 6000 images and 20 classes, labeling different accessories and also dividing between people who wear a mask on their face, who do not wear it or who wear it incorrectly. It also offers a classification for surgical, colorful, and gas masks. In the same month, an article [7] described a system capable of detecting whether a person wears a mask or not, and, in the first case, the type of mask is also indicated among the following: homemade,

XXX-X-XXXX-XXXX-X/XX/$XX.00 ©20XX IEEE

n95, and surgical. Recently, the Ways to Wear a Mask or a Respirator Database (WWMR-DB) [8] was published. It focuses on the ways in which a mask or respirator is usually worn, but also provides labels regarding the type of mask or respirator worn. The database and the system described in this paper want to offer more solutions and possibilities than the above-mentioned works: an example can be seen in Fig. 1.

Fig. 1. The proposed system must be also able to distinguish between disposable respirators with or without valve; the latter is usually small, of variable shape and its position can change both in the mask itself and due to head rotations.

## III. Facial Masks And Respirators Database

The Facial Masks and Respirators Database (FMR-DB) [1] consists of 2565 images taken from the Internet and is publicly available in an open-access format. It may be thought that the number of images is not very high, but this characteristic is not always indicative of the creation of quality systems: in fact, the classification accuracy is just as important [9]. As the coronavirus crisis required quick solutions, it was tried to best mediate between these two characteristics. The search for images was performed using common image search engines: Google Images and the one integrated into Bing and DuckDuckGo. Since they adopt different operating algorithms, this implies that the results are also different, thus maximizing the probability of finding new images. The research was conducted by ensuring that each class contained the same number of images to avoid potentially harmful imbalances in the systems implementation phase. Because of their nature, the images have different resolutions and quality from each other. They have then been carefully classified and subdivided in the manner that will be described in the next subsection.

### A. Characteristics

During the design of the database, it was thought about a classification of the images able to embrace the maximum number of possible uses. For this reason, it was first divided according to whether the person portrayed is wearing a face mask or a respirator: these cases will be seen in detail in the next subsections. Two versions of the database are available, one containing the original images as they were taken from the Internet and the other containing only the faces of the people portrayed. They have been manually cut to make them uniform and to ensure that the respiratory protection and other parts of the face were not cut during the process. This may not always be guaranteed using automatic face detection algorithms, such as Viola-Jones [10], or more recent ones such as Multi-Task Cascaded Convolutional Neural Networks (MTCNN) [11], but they can be used on the original images at the user's choice.

Most of the images were cropped in a square shape and the faces of the people portrayed were centered as much as possible. For the remaining images, this was not possible and a rectangular crop was made. Future users are therefore given the freedom to add padding to these images or to extend them to a new dimension. For the tests that were carried out it was decided not to add padding. Each image is marked by a name consisting of a series of labels that denote all the characteristics of the image itself, followed by a unique number: in Table I is possible to see the list of labels and their meaning.

TABLE I. FMR-DB: List Of Image Labels

| Label | Meaning |
|---|---|
| *MS* | With Facial Mask or Respirator |
| *NM* | Without Facial Mask or Respirator |
| *WV* | Disposable Respirators With Valve |
| *NV* | Disposable Respirators Without Valve |
| *FF* | Full-Face Respirators |
| *HF* | Half-Face Respirators |
| *NM* | Non-Medical Masks |
| *SR* | Surgical Masks |
| *EP* | With Eye Protection |
| *HP* | With Head Protection |
| *EH* | With Eye And Head Protection |
| *OC* | With Occlusions |
| *NO* | Without Occlusions |
| *HM* | Hands On Mouth |
| *HT* | Hats |
| *NW* | Neck Warmers And Bandanas |
| *SN* | Sunglasses |

### B. Facial Masks and Respirators Classification

Images that contain facial masks or respirators are classified according to their type, which falls into the following:

- Disposable Respirators With Valve.
- Disposable Respirators Without Valve.
- Full-Face Respirators.
- Half-Face Respirators.
- Non-Medical Masks.
- Surgical Masks.

These classes, containing 315 images each, also have an internal classification based on the protective clothing worn by the subject portrayed:

- They cannot wear any.
- They can wear protective glasses.
- They can wear helmets or other head protectors.

- They can wear both.

The Disposable Respirators With Valve and Without Valve have an additional classification based on their degree of protection:

- FFP1.
- FFP2 – N95 – KN95.
- FFP3.
- Other – Unknown.

This classification is intended for future use. To achieve it, the model of the mask portrayed was searched in detail or visual indicators were adopted. Some respirator manufacturers tend to visually distinguish the degree of protection of the respirator by using a different color for the strap, valve lettering, or other parts: for example, 3M Company uses yellow for the FFP1 classification, blue for the FFP2 classification, and red for FFP3 classification. It is important to notice that this classification is not viable for the Full-Face and Half-Face Respirators, since their degree of protection changes according to the filter arranged in the appropriate place. It is not feasible even for Non-Medical and Surgical Masks as there is a clear distinction between facial masks and respirators. The former have the sole purpose of reducing the likelihood of spreading infectious agents via saliva droplets, while the latter are Personal Protective Equipment (PPE): this means that they can protect the wearer from inhaling dangerous substances, the types of which depend on the degree of protection of the respirator itself.

This section of the database can be useful for implementing automatic systems for face detection and facial masks and respirators type recognition.

### C. *Occlusions Classification*

Images that do not contain facial masks or respirators are divided according to the presence or absence of occlusions, and the former also has a classification based on their type. In this regard, eyeglasses have not been considered as facial occlusion since, for many people, they are an essential tool in everyday life. A beard can also be thought of as a facial occlusion, but, like eyeglasses, it is very common to meet people who have it and the presence of such features in the database images gives, for any neural network trained with them, major resistance to variations. The following are instead classified as occlusions:

- Hands in Mouth.
- Hat.
- Neck Warmer.
- Sunglasses.

In the database, there are 90 images for each of these occlusions and 315 images of people without occlusions. This section of the database can be useful for implementing automatic systems for face and occlusion detection.

## IV. SYSTEM CREATION

The FMR-DB was used to train an Artificial Neural Network (ANN) capable of recognizing the type of mask or respirator worn. After the exploit of AlexNet [12], ANNs are increasingly used as they can solve complex problems with great precision and without knowing the most important features. They are capable of deducing them on their own at the cost of information redundancy. One of them will be used to address the aforementioned problem. In the following subsections, an orderly discussion about the following topics will be started:

- The choice of the neural network and the optimizer settings.
- The values to assign to the hyperparameters, the subdivision of the FMR-DB into the train, validation, and test datasets, and other settings.
- The configuration used for neural network training.

### A. *Neural Network Choice*

For addressing the aforementioned problem, it was sought a state-of-the-art neural network, not too complex and with a relatively low number of parameters: our choice fell into the EfficientNetB0, the smallest in the family of the EfficientNet networks [13]. The EfficientNet-B0 neural network architecture is shown in Table II.

TABLE II. EFFICIENTNET-B0 ARCHITECTURE

| **Stage** $i$ | **Operator** $\hat{F}_i$ | **Resolution** $\hat{H}_i x \hat{W}_i$ | **# Channels** $\hat{C}_i$ | **# Layers** $\hat{L}_i$ |
|---|---|---|---|---|
| 1 | Conv3x3 | 224 x 224 | 32 | 1 |
| 2 | MBConv1, k3x3 | 112 x 112 | 16 | 1 |
| 3 | MBConv6, k3x3 | 112 x 112 | 24 | 2 |
| 4 | MBConv6, k5x5 | 56 x 56 | 40 | 2 |
| 5 | MBConv6, k3x3 | 28 x 28 | 80 | 3 |
| 6 | MBConv6, k5x5 | 14 x 14 | 112 | 3 |
| 7 | MBConv6, k5x5 | 14 x 14 | 192 | 4 |
| 8 | MBConv6, k3x3 | 7 x 7 | 320 | 1 |
| 9 | Conv 1x1 & Pooling & FC | 7 x 7 | 1280 | 1 |

The choice of a neural network with complexity suitable for the task to be performed, together with the use of other techniques discussed below, helps to tackle the overfitting problem. For implementing the neural network, it was decided to recur to the high-level API Keras, using as backend Tensorflow [14] [15]. From the 2.0 version of Tensorflow, Keras has been integrated into it.

The transfer learning technique was used by initializing the model weights with the parameters previously learned by training the EficientNet-B0 with the ImageNet database. The top layers of the network were replaced with the following three:

- A GlobalAveragePooling2D layer.
- A Dropout layer with a rate of 0.5, which helps reduce overfitting.

- A Dense layer, activated by a softmax function, to obtain the final prediction.

The network was then compiled with the Adam optimizer [16] using the following parameters:

- $\beta_1 = 0.9$.
- $\beta_2 = 0.999$.
- $\varepsilon = e^{-6}$.

Finally, an input image size of 128x128 pixels was set: with these settings, the neural network requires about 4 million parameters.

### B. *Hyperparameters and Other Settings*

For all the training, the neural network hyperparameters were set as follows:

- Batch size = 32.
- Learning rate = 0.00025.
- Max number of epochs = 100.

A division of 80%-10%-10% between train, validation, and test dataset was implemented in the following way:

1) A primary division of the dataset in 90%-10% was performed to obtain the test dataset.
2) A 9-fold cross-validation was then performed on the first subset to obtain variable train and validation datasets.

Data augmentation was applied to the images with the respective parameters arranged as follows:

- Brightness range = (0.5, 1.0).
- Rotation range = +- 15 degrees.
- Width shift range = +- 0.05 (values between +- 1).
- Height shift range = +- 0.05 (values between +- 1).
- Shear range = +- 2.5 degrees.
- Zoom range = +- 0.05 (values between +- 1).
- Channel shift range = +- 50 (RGB values).
- Random horizontal flip enabled.

Some training callbacks of Keras were used. An example of these is the EarlyStopping, set to stop training before the end of the 100th epoch if the validation loss does not improve for 10 consecutive epochs. Finally, it is important to say that a Z-Score Normalization was performed on each image, since, in the same way as alternative mathematical methods such as Principal Component Analysis (PCA), it improves the learning capacity of the neural network.

### C. *Training Configuration*

The following configuration was used to train the neural network:

- Intel® Core™ i7 8750H.
- NVIDIA® GeForce® RTX 2080 8GB GDDR6 with Max-Q Design.
- 32 GB DDR4, 2666 MHz.
- 2 x 2 TB PCIe M.2 SSD.

The GPU was preferred to the CPU because of its parallelization capabilities, which boosts the training speed. To use it, there was the need to install GPU-accelerated primitives: the NVIDIA® CUDA® Toolkit [17] and the NVIDIA® CUDA® Deep Neural Network library (cuDNN) [18].

## V. Tests

### A. *Test With Some Database Images*

As previously seen, the database can be used in different contexts, but, in this paper, the authors want to focus their efforts on recognizing the type of mask or respirator worn by a person. For this reason, Fig. 2 shows the classes of this problem and the relative number of images.

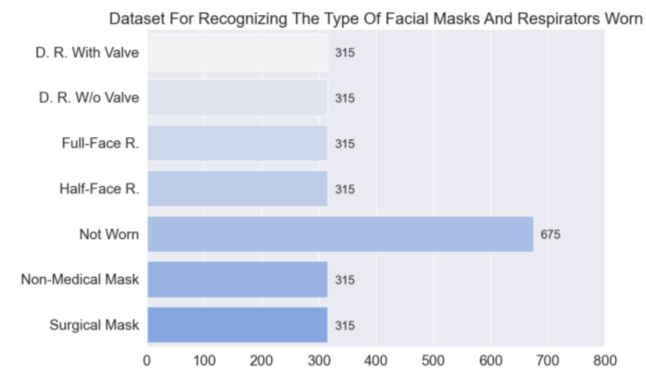


Fig. 2. The database subdivision implemented for the recognition of the type of facial mask or respirator.

Using the settings listed in the previous subsection, 9 different neural networks were trained using the 9-fold cross-validation: the results of these training are shown in Table III. The neural network that performed better was grayed out.

TABLE III. Training Results

| Fold # | Test Accuracy | Epochs Of Training | Final Validation Accuracy | Final Validation Loss |
|---|---|---|---|---|
| 0 | 0.98419 | 38 | 0.96109 | 0.11597 |
| 1 | 0.96443 | 23 | 0.97665 | 0.07127 |
| ***2*** | ***0.99209*** | ***40*** | ***0.95720*** | ***0.12897*** |
| 3 | 0.96838 | 18 | 0.93774 | 0.28144 |
| 4 | 0.96443 | 24 | 0.96498 | 0.12392 |
| 5 | 0.98024 | 24 | 0.97276 | 0.07326 |
| 6 | 0.97233 | 27 | 0.95331 | 0.20121 |
| 7 | 0.95257 | 20 | 0.96109 | 0.17073 |
| 8 | 0.97628 | 17 | 0.95703 | 0.12075 |

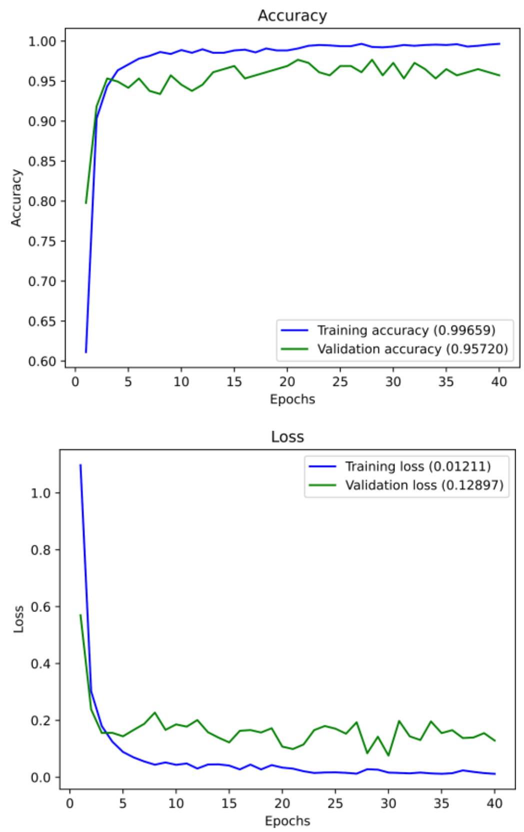


Fig. 3. Training and validation accuracy and loss achieved by the neural network trained using Fold 2 as validation dataset.

In general, all the trained neural networks obtained good results. Fig. 3 shows the evolution of the accuracy and loss during the training and validation phases of the best-performing network, obtained using Fold 2 as validation dataset. It achieved over 60% training accuracy and 80% validation accuracy only in the first epoch: in the following ones, this value has rapidly improved, reaching a stalemate in a short time.

In Fig. 4 it is possible to see the confusion matrix calculated on the test dataset. Only two mistakes were made: a Disposable Respirator Without Valve was mistaken for one With Valve and a Half-Face Respirator was mistaken for a Full-Face one. It should be noted that these details are very subtle: the first case was shown in Fig. 1. For what regards the second case, if a person is wearing a Half-Face Respirator alongside an eye protection, it could be very difficult to recognize, even for a human, whether it is a Half-Face Respirator or a Full-Face one: this problem was also encountered when classifying some database images.

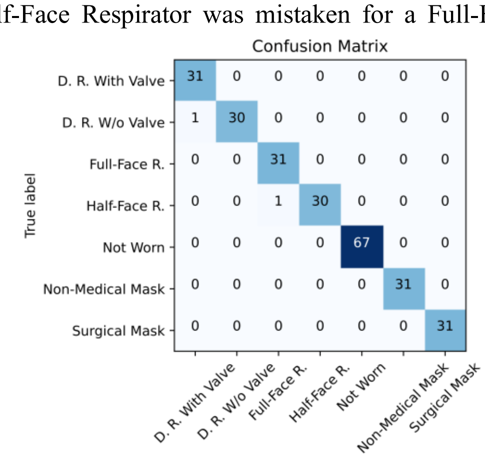


Fig. 4. Confusion matrix calculated on the test dataset of the neural network trained using Fold 2 as validation dataset.

## B. Test With Other Images

After the good results obtained with the database test images, it was decided to test the system with other images. After providing five volunteers with the necessary equipment, they were asked to take a picture of themselves wearing all types of masks and respirators supported by the database; an example of these images can be seen in Fig. 1. The test results are illustrated in the confusion matrix shown in Fig. 5 and demonstrate the good solidity of the trained neural networks.

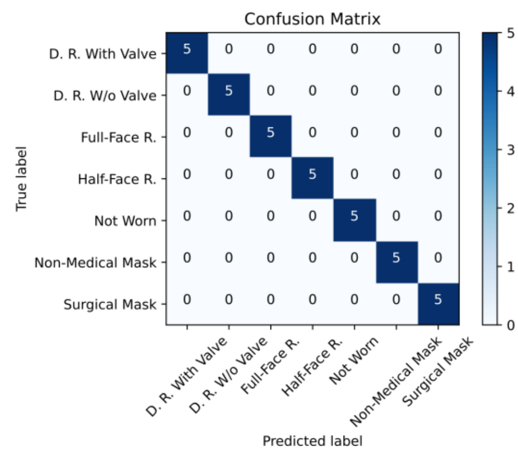


Fig. 5. Confusion matrix calculated on the dataset composed of the other images with the neural network trained using Fold 2 as validation dataset.

# VI. Conclusion

With this work, the authors hope to have created a useful database at the complete disposition of the scientific community, which can be used to create such systems. The results are encouraging, but the FMR-DB can be further improved. One possibility is to increase the database size by adding new images: in this way, it will be possible to create more robust automatic recognition systems than those achievable with the current version of the database. Furthermore, it is also possible to test other neural networks and parameters in order to improve the results obtained.

# Acknowledgment

The authors wish to thank Dr. Sandra Urpis for her help in building the database.

## References


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

[2] 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,” 2020. [Online]. Available: arXiv:2003.09093.

[3] P. Nagrath, R. Jain, A. Madan, R. Arora, P. Kataria, J. Hemanth, “SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2,” Sustainable Cities and Society, Volume 66, 2021, 102692, ISSN 2210-6707, doi: https://doi.org/10.1016/j.scs.2020.102692.

[4] S. Lin, L. Cai, X. Lin, R. Ji, “Masked face detection via a modified LeNet,” Neurocomputing, Volume 218, 2016, Pages 197-202, ISSN 0925-2312, doi: https://doi.org/10.1016/j.neucom.2016.08.056.

[5] T. Pham, B. Tran, D. Pham, L. Nguyen, “Classifying Between Masked Faces and Normal Faces with CNN and SSH”, Aug 2019. [Online]. https://github.com/aome510/Mask-Classifier/blob/master/paper/paper.pdf.

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

[7] Y. Wang, “Which Mask Are You Wearing? Face Mask Type Detection with TensorFlow and Raspberry Pi”, May 29, 2020. [Online]. Available: https://towardsdatascience.com/which-mask-are-you-wearing-face-mask-type-detection-with-tensorflow-and-raspberry-pi-1c7004641f1.

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

[9] A. C. Marceddu, J. Sini, M. Violante and B. Montrucchio, “A Novel Approach To Improve The Social Acceptance Of Autonomous Driving Vehicles By Recognizing The Emotions Of Passengers,” Proceedings of SPIE. ICMV 2020, Rome, Italy, November 2020, doi: http://dx.doi.org/10.1117/12.2586417.

[10] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Kauai, HI, USA, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517.

[11] K. Zhang, Z. Zhang, Z. Li and Y. Qiao, “Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks, ” in IEEE Signal Processing Letters, vol. 23, no. 10, pp. 1499-1503, Oct. 2016, doi: 10.1109/LSP.2016.2603342.

[12] A. Krizhevsky, I. Sutskever and G. E. Hinton, “ImageNet classification with deep convolutional neural networks,” In Proceedings of the 25th International Conference on Neural Information Processing Systems - Volume 1 (NIPS'12). Curran Associates Inc., Red Hook, NY, USA, 2012, 1097–1105, doi: https://doi.org/10.1145/3065386.

[13] M. Tan and Q. V. Le, “EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,” 2019 International Conference on Machine Learning (ICML), Long Beach, CA, 2019, arXiv:https://arxiv.org/abs/1905.11946v3.

[14] F. Chollet et al., Keras, 2015. [Online]. Available: https://keras.io.

[15] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, M. Kudlur, J. Levenberg, R. Monga, S. Moore, D. G. Murray, B. Steiner, P. Tucker, V. Vasudevan, P. Warden, M. Wicke, Y. Yu and X. Zheng, “TensorFlow: A system for large-scale machine learning”, In 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), 2016, pp. 265–283, arXiv: https://arxiv.org/abs/1605.08695

[16] D. P. Kingma and J. Ba. “Adam: A Method for Stochastic Optimization,” The 3rd International Conference for Learning Representations (ICLR), San Diego (CA), USA, 2014. doi: http://arxiv.org/abs/1412.6980

[17] NVIDIA®, NVIDIA® CUDA® Toolkit. [Online]. Available: https://developer.nvidia.com/cuda-toolkit.

[18] NVIDIA®, NVIDIA® CUDA® Deep Neural Network library (cuDNN). [Online]. Available: https://developer.nvidia.com/cudnn.