PORT SAID ENGINEERING RESEARCH JOURNAL
Faculty of Engineering - Port Said University
Volume 27 pp: 24-30

Improving Smart Infrastructure Monitoring System as a Response to Prevalent Pandemic

Nourhan Osama[*1], *Ahmed Magdy*[2], *Bassem El Hady*[3], *Khaled Abd EL Salam*[4]

[1] *Electrical Engineering Department, Faculty of Engineering, Suez Cana University, Ismailia, Egypt, email: Nor.osama@eng.suez.edu.eg .*
[2] *Electrical Engineering Department, Faculty of Engineering, Suez University, Ismailia, Egypt, email: ahmed.m.1986@ieee.org .*
[3] *Electrical Engineering Department, Faculty of Engineering, Suez Canal University, Ismailia, Egypt, email: basem_elhady@eng.suez.edu.eg.*
[4] *Electrical Engineering Department, Faculty of Engineering, Suez Canal University, Ismailia, Egypt, email: khaled.abdelsalam@eng.suez.edu.eg.*
**Corresponding author, E-mail: Nor.osama@eng.suez.edu.eg ,DOI: 10.21608/PSERJ.2023.209431.1236*

ABSTRACT

Face masks are no longer an option for protection against airborne diseases brought on by coughing, talking, or sneezing, which can spread germs into the air and infect everyone nearby, especially with the coronavirus pandemic that occurred in 2019. Also, some states and the government have made it mandatory for people to wear face masks. In recent years, Artificial Intelligence has played an important role in the medical field, as Convolution Neural Network techniques have proven to be very useful in image detection applications with different algorithms. In this paper, we propose a model using deep learning algorithms to achieve the most efficient and speedy way to detect the presence of a face mask on people in public places by using RGB cameras. The Alexnet, Googlenet, Resnet 18, and Squeezenet are trained on a dataset that consists of images of people with and without masks and is publicly available as “Mola RGB Covsurv” Mendeley Data, with 80% of the dataset being used for training and 20% for testing to get the most efficient algorithm. The proposal we recommend is Squeeznet's algorithm, which achieved an average precision of 94.1592% with a sensitivity of 91.19533% in 1700 minutes and 10 seconds.

Keywords: Coronavirus, Artificial intelligence, Deep Learning, Convolution neural network.

1 INTRODUCTION

These days, the topic of facial coverings in public is regularly raised. Masks should be used as part of an all-encompassing plan of action to stop the spread of infection and save lives because they cover the main routes of the airborne disease, which are the nose and mouth in exchange for the respiratory system; we can summaries the importance of wearing masks as we can significantly lower our chances of transmission of the virus that causes airborne diseases if I cover my face to protect you from me and you cover your face to protect me. The use of a mask by itself cannot provide an acceptable level of disease protection [1]. People also need to take other precautions besides wearing masks to stay safe, like keeping a physical distance, checking the local advice where they live or work, washing their hands, coughing into a bent elbow or tissue, keeping rooms well-ventilated, and avoiding crowded places. This, in conjunction with wearing masks, will be important in limiting the spread of the diseases. So, we could say that wearing a mask is the first step in keeping our lives safe, so we should make it a normal habit to wear it around other people [2]. Face masks are more frequently used for respiratory infections that spread through droplets, which move quickly and are transferred by coughing or sneezing, close personal contact, aerosolization of the microbe as this microbe remains in the air, and dust particles. Face masks frequently fit loosely, reducing hand-to-face contact and the dissemination of big sprays and droplets. Measles, chickenpox, and tuberculosis; the common cold, which can develop from a rhinovirus; mumps, caused by a paramyxovirus; whooping cough, a bacterial infection caused by Bordetella pertussis; and COVID-19, caused by the SARS-CoV-2 virus, are among the illnesses that call for the use of a mask as they are airborne diseases and also infectious ones [3]. Only those with respiratory infection symptoms, such as coughing, sneezing, or, in some circumstances, fever, should use face masks. Health care professionals, anyone caring for or in close proximity to those who have respiratory infections, and other people as prescribed by a doctor should also wear face masks. There is no evidence that face masks worn by healthy people help keep people healthy; thus, they should not be used by healthy people to protect themselves from respiratory infections. Masks can be scarce during times of widespread respiratory infection; thus, they should only be used by individuals who require them [4]. There are some places that are more liable for people wearing masks at them regardless of whether a physical distance of at least one meter can be maintained, in indoor environments where ventilation is known to be inadequate, cannot be measured, or the ventilation system is not adequately maintained. If a physical distance of at least one meter cannot be maintained, do it in an interior space with proper ventilation or outside if a physical distance of at least one meter cannot be maintained. Also, in hospitals by doctors, nurses, and patients; in schools; on public transportation; in workplaces and offices; in grocery stores; and in any crowded area [5]. The artificial intelligence (AI) community is facing a number of challenges. To combat this virus and enable them to receive quick feedback in real-time to stop its spread, healthcare institutions urgently require technology for decision-making. Like The use of the AI-driven chest scan has the potential to reduce the growing load for radiologists, who must keep track of and periodically assess an increasing number of patient chest scans. Also, face mask recognition has advanced significantly in the fields of image processing and computer vision since the COVID-19 pandemic's emergence. Utilizing a variety of methods and methodologies, numerous face detection models have been produced as Deep learning and convolutional neural network progression significantly make it easier to get high accuracy in image classification and object detection; besides, faces can now be detected in both static photographs and moving films, as well as in real-time inspection and supervision.

I. Background

Convolution Neural Network (CNN)

CNNs have been put to work in many ways, including image classification, localization, detection, segmentation, and registration. In recent years, CNN has undergone rapid evolution, resulting in a wide range of architectures that are available today. CNNs function better with images than other neural networks, and they feature three primary types of layers:

Convolution layer

Pooling layer

Fully-connected layer (FC)

The convolution layer is the top of a convolution network that performs feature extraction and it can be followed by additional convolution layers or a pooling layer where feature maps are pooled for translational invariance. The spatial size of the convolved feature has been decreased by the pooling layer. By reducing the dimensionality of the data, less processing power will be needed to process it. Additionally, it helps in extracting dominant characteristics that are rotational and positional invariant, preserving the effectiveness of the model training process. Then the fully connected layer is the final convolution or pooling layer's output feature maps are typically flattened, or converted into a one-dimensional (1D) array of numbers (or vector), and then connected to one or more fully connected layers, also known as dense layers, in which every input is connected to every output by an adaptable weight. Once the features extracted by the convolution layers and down sampled by the pooling layers are created, they are mapped by a subset of fully connected layers to the final outputs of the network, such as the probabilities for each class in classification tasks. The final, fully connected layer typically has the same number of output nodes as the number of classes [6].

And since there is a lot of variety in CNN's architecture, here are some examples of the CNN architecture we used in this paper:

A) Alexnet (2012)

Alexnet (paper) was the first winner of the ImageNet challenge, it has 8 layers in total (5 convolutional layers plus 3 fully connected layers). There are 96 to 384 filters in each convolutional layer, with filter size varying from 3×3 to 11×11, and it has about 60 M parameters .

B) Googlenet (2014)

In 2014, Google introduced Googlenet, also known as Inception-v1, which is a deep convolutional neural network consisting of 27 layers, including 22 layers with trainable parameters and 5 pooling layers. It includes 1 convolutional layer, 9 inception modules, and 1 fully connected layer. Each inception module contains multiple branches with different types of layers, such as 1x1 convolutions, 3x3 convolutions, and pooling layers. Googlenet is designed to reduce the number of parameters, memory utilization, and calculation while maintaining high accuracy.

C) Resnet 18 (2015)

The residual network can have variable sizes, depending on how big each of the layers of the model are and how many layers it has. Resnet 18 contains 18 deep layers, it consists of conv layers with filter 3*3, and uses a single pooling layer at the beginning of the network and no pooling layers at the end.

D) SqueezeNet (2016)

It was introduced in 2016 as a response to the increasing demand for deep learning models that can run on resource-constrained devices such as mobile phones or embedded systems. SqueezeNet achieves high accuracy by using a combination of fire modules and squeeze modules to reduce the number of parameters in the network while preserving the expressive power of deeper models. This makes it possible to run SqueezeNet on devices with limited computational resources while still achieving high performance [7].

2 LITERATURE REVIEW

There are several studies and researches focused on mask detection, Face construction and identity recognition while wearing face masks are the main topics of this paper. Our goal in this study is to identify those who are not donning face masks in order to reduce COVID-19 transmission and dissemination. Face masks have been shown to reduce the rate of COVID-19 dissemination by researchers and scientists.

In A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic paper, developed a hybrid approach integrating deep learning and machine learning for the detection of a two-component face mask. The first component, Resnet50, is used for feature extraction. In the second component, decision trees, ensemble algorithms, and support vector machines are employed for classification [8].

In Fighting against COVID-19: A novel deep learning model based on YOLO-v2 with ResNet-50 for medical face mask detection paper Two parts make up the suggested model. The first part is created for the ResNet-50 deep transfer learning model's feature extraction process. YOLO v2 is used in the second component, which is intended for the identification of medical face masks [9].

In YOLO-face: a real-time face detector, to enhance face identification performance, they suggest YOLO-face, a face detector based on YOLOv3. The current method uses a more accurate regression loss function and anchor boxes that are more suited for face detection. The enhanced detector maintained a constant level of accuracy while rapidity of detection Our enhanced method performs better than YOLO and its variants, according to tests on the WIDER FACE and FDDB dataset [10].

and the most recently research is AI Based Monitoring of different Risk Levels in covid 19 context, that proposed a real-time system that focuses on the application of deep learning algorithms to detect the presence of masks on people in public spaces (using RGB cameras), as well as the detection of the caruncle in the human eye area to make an accurate measurement of body temperature (using thermal cameras). For RGB mask detection, the Yolo V5 four models are used to get accurate results with fast detection [11].

3 SYSTEM MODEL

In this research, we suggested a smart system for screening people without face masks. All public areas in the smart city are monitored by CCTV cameras. The cameras are employed to take pictures in public areas, and these pictures are then fed into a system that determines if somebody is visible in the picture without a face mask. If a person without a face mask is found, the appropriate authorities are informed, and the necessary steps are then taken. Then we use deep learning algorithms to determine whether a face mask is present or not. And for this, we divided the dataset into 20% for testing and 80% for model training, then applied the four deep learning algorithms (Alexnet, Googlenet, Resnet18, and Squeeznet) to the dataset, which first made feature extraction, then trained and tested the model, and classified at the end. The model has been implemented using the MATLAB R2021a environment, and the test was performed with a Core i7-6820HQ laptop with 8 GB of RAM and 512 GB of SSD system storage and with input images of different sizes.

4 DATASET COLLECTION

The samples utilised with this Model system are drawn from the pre-existing datasets shown in Table 1. In addition to increasing the sample size, using various datasets improves the algorithms that will be trained because there are samples with various levels of quality, occlusion, luminosity, background, and population. And this dataset was made publicly available as “Mola RGB Covsurv”, Mendeley Data, v1 [12]. In this research We use images contains front face pose with single face in the frame 18251 with mask and 10961 without mask. For training purposes, 80% images of each class are used and the rest of the images are utilized for testing purposes. Fig. 3 shows some of the images of two different classes

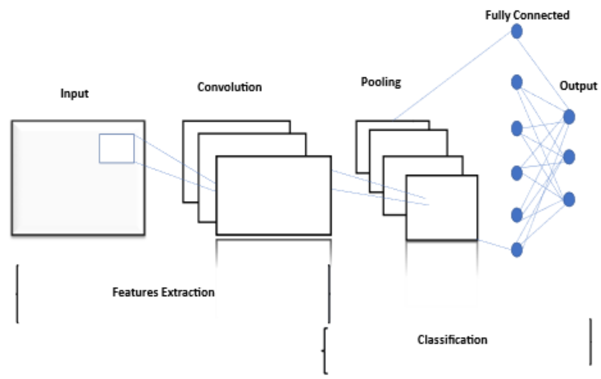

Fig.1 Convolution Neural Network architecture.

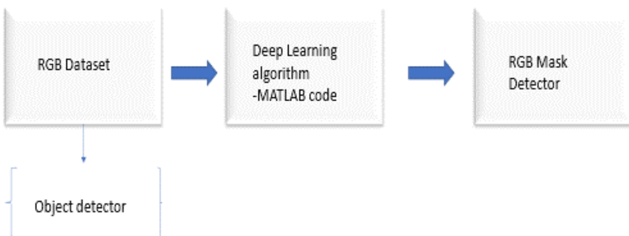

Fig.2 system model

Table 1. Datasets where proposed tool was applied.

Dataset	Description
Celeba [13]	More than 200000 images of faces. Images of just one person.
Coco [14]	More than 320000 images and more than 91 different objects, widely used for object detection tasks. We only used images where people are present in the most varied environment.
Helen [15]	It consists of 2330 images of one or several people.
IMM [16]	Consists of 240 single face images of 40 different people
Wider [17]	Over 32000 images with different levels of scale and occlusion

(a) (b)

Fig.3 sample of dataset; (a)people with mask, (b) people without mask.

5 RESULT ANALYSIS

In this model, we use deep learning algorithms to determine whether a face mask is present or absent, selecting CNN as one of the most well-known deep learning architectures that learns directly from data without the need for feature extraction procedures. Tens or even hundreds of layers can be present in a CNN, and each layer can be trained to detect various aspects of an image. Each training image is subjected to filters at various resolutions, and the result of each convolved image is used as the input to the following layer. Beginning with relatively basic properties like brightness and borders, the filters can get more complicated until they reach characteristics that specifically identify the object The design of a CNN changes to categorization after learning data through several layers, the next-to-last layer, which is a fully connected layer, outputs a vector of K dimensions, where K is the number of classes that may be predicted, and contains the probabilities for each class that a classification algorithm can assign to an image. The final classification output is provided by a classification layer in the CNN architecture's final layer. [18]. And in this research Employing various CNN architectures (Alexnet, Googlenet, SqueezeNet, and Reset18).

20% of the dataset was used for testing, and 80% for training the model. The dataset contains a total of 29212 images, of which 23369 are utilized to train the classification module and the remaining 5843 to test its performance. With a mini batch size of 64, we used the mini-batch stochastic gradient-descent technique, which produced 1825 iterations per epoch for 5 epochs. The models were configured with an initial learning rate (μ) of $1e^{-4}$.

To evaluate the performance of the different classifiers, performance matrices are needed to be investigated through this research.

And the most common performance measures to be calculated are Accuracy, Precision, Recall, and F1 Score [19], and they are presented from Eqs. (1) to Eqs. (4).

1. **Accuracy**

Accuracy is the ratio of the number of correctly predicted instances to the total number of instances. It measures how often the classifier correctly predicts the class of an instance.

$$Accuracy = \frac{Tp+Tn}{Tp+Tn+Fp+Fn} \quad (1)$$

where Tp is the number of true positives, Tn is the number of true negatives, Fp is the number of false positives, and Fn is the number of false negatives.

2. **Precision**

Precision is the ratio of true positives to the total number of predicted positives. It measures the accuracy of positive predictions.

$$Precision = \frac{Tp}{Tp+Fp} \quad (2)$$

Tp for the true positive, and Fp representing the false positive.

3. **Sensitivity (Recall)**

Recall is the ratio of true positives to the total number of actual positives. It measures the ability of the classifier to find all positive instances.

$$Sensitivity(Recall) = \frac{Tp}{Tp+Tn} \quad (3)$$

Where Tp for the true positive, and Fn stands for false-negative.

4. **Specificity**

Specificity is the ratio of true negatives to the total number of actual negatives. It measures the ability of the classifier to identify negative instances.

$$specificity = \frac{Tn}{Tn+Fp} \quad (4)$$

Where Tn for the true Negative, and Fp stands for false-Positive.

Utilizing four DL algorithms (Alexnet, Googlenet, Squeezenet, and Resnet 18), applying the parameters that have been kept in the system model section, training the algorithms on 80% of the dataset, and assessing the results on the remaining 20% of the dataset, we have achieved an accuracy of (94.745, 95.156 ,94.574, 95.036) respectively. Googlenet has achieved best Accuracy between the four algorithms. Table 2 illustrates the performance of the four classifier that had been used. As seen in table (2), Googlenet achieved the highest accuracy with 91.925 percent, while on the other side, Squeeznet achieved the highest precision with 94.159 percent.

The result of simulation in MATLAB for 4 algorithms shown in fig. (4).

Table 2. Performance of the four algorithms.

Classifier / Algorithm	Squeez-net	Google-net	Alexnet	Resnet 18
Accuracy	94.574%	95.156%	94.745%	95.0359%
Sensitivity	91.196%	93.430%	91.9252%	92.8376%
Specificity	96.603%	96.192%	96.4384%	96.3562%
True positive Rate (Recall)	1999	2048	2015	2035
False Positive Rate	124	139	130	133
False Negative Rate	193	144	177	157
True Negative Rate	3526	3511	3520	3517
Precision	94.159%	93.640%	93.939%	93.865%

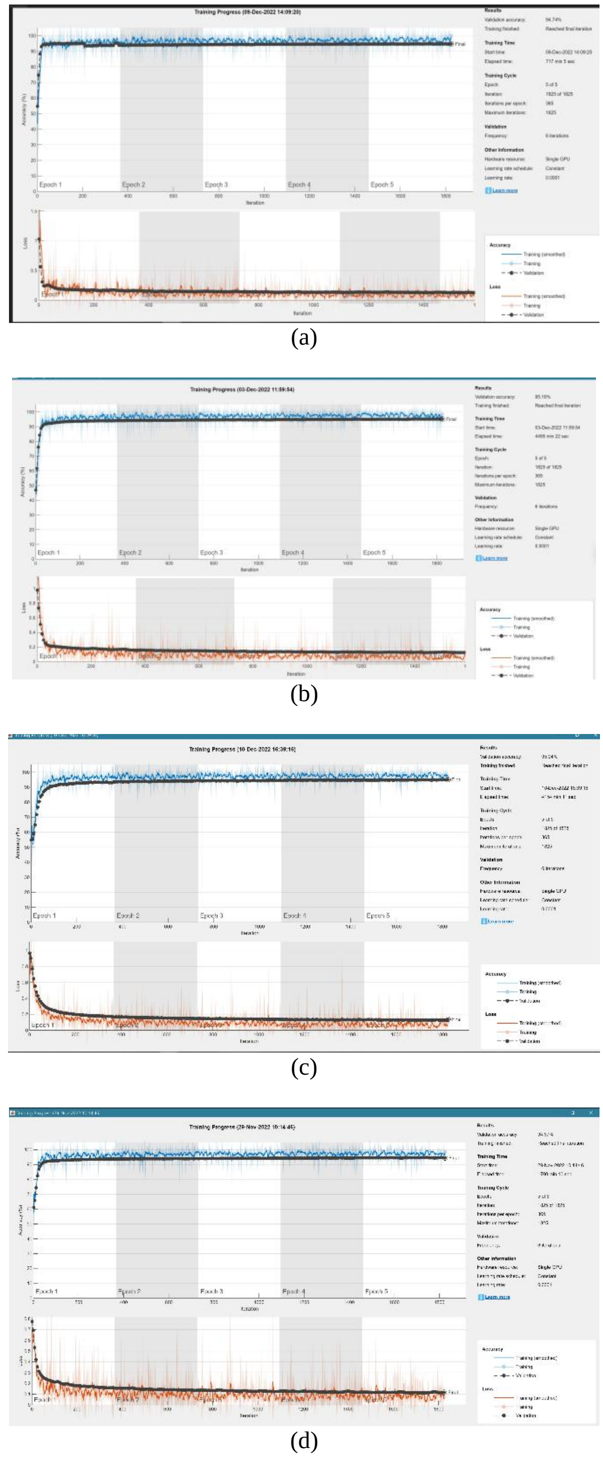

Fig.4 Graphs of the results for the four algorithms; a) Alexnet Algorithm, b) Googlenet Algorithm, c) Resnet18 Algorithm, d) Squeeznet Algorithm.

6 CONCLUSION

In conclusion, the use of face masks is essential in protecting individuals from airborne diseases, particularly in public places, and has been made mandatory in some areas. Artificial intelligence has played a significant role in the medical field, particularly in image detection applications using convolution neural network techniques. In this paper, a model was proposed that uses deep learning algorithms to efficiently detect the presence of face masks on people in public places using RGB cameras. The proposed algorithm, Squeezenet, achieved an average precision of 94.1592% with a sensitivity of 91.19533% in 1700 minutes and 10 seconds. This model's implementation can help curb the spread of airborne diseases and could be beneficial in places like hospitals, schools, public transportation, and other crowded areas.

7 REFERENCE

[1] "Why Wearing a Mask Is Important: Opelika & Auburn, Alabama (AL), East Alabama Health." East Alabama Medical Center, https://www.eastalabamahealth.org/news-and-media/why-is-wearing-a-mask-important (accessed on 14 February 2023)

[2] "Coronavirus Disease (Covid-19): Masks." World Health Organization, World Health Organization, https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19-masks (Accessed on 14 February 2023)

[3] The Global Infectious Disease Threat and Its Implications for the United States, https://irp.fas.org/threat/nie99-17d.htm (Accessed on 15 February 2023)

[4] "Use Masks to Slow the Spread of Covid-19." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/masks.html (Accessed on 15 February 2023)

[5] "Wearing Masks in Travel and Public Transportation Settings." Centers for Disease Control and Prevention, Centers for Disease Control and Prevention, https://www.cdc.gov/coronavirus/2019-ncov/travelers/masks-public-transportation.html (Accessed on 16 February 2023)

[6] Yamashita, R., Nishio, M., Do, R.K.G. et al. "Convolutional neural networks: an overview and application in radiology". Insights Imaging 9, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9 (Accessed on 11 February 2023)

[7] Tapomoy Adhikari. "Designing a Convolutional Neural Network for Image Recognition: A ..." Senior Researcher, Bengaluru, Karnataka, India, February 2023. http://dx.doi.org/10.13140/RG.2.2.17991.16804. (Accessed on 15 February 2023)

[8] Loey M, Manogaran G, Taha MHN, Khalifa NEM." A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic". Measurement (Lond). 2021 Jan 1; 167:108288. doi: 10.1016/j.measurement.2020.108288. Epub 2020 Jul 28. PMID: 32834324; PMCID: PMC7386450. (Accessed on 9 April 2022)

[9] NEM;, Loey M;Manogaran G;Taha MHN;Khalifa. "Fighting against Covid-19: A Novel Deep Learning Model Based on Yolo-V2 with Resnet-50 for Medical Face Mask Detection." Sustainable Cities and Society, U.S. National Library of Medicine, https://pubmed.ncbi.nlm.nih.gov/33200063/ (Accessed on 10 April 2022)

[10] Chen, Weijun, et al. "Yolo-Face: A Real-Time Face Detector - the Visual Computer." SpringerLink, Springer Berlin Heidelberg, 12 Mar. 2020, https://link.springer.com/article/10.1007/s00371-020-01831-7 (Accessed on 10 April 2022)

[11] Melo C;Dixe S;Fonseca JC;Moreira AHJ;Borges J; "AI Based Monitoring of Different Risk Levels in COVID-19 Context." Sensors (Basel, Switzerland), U.S. National Library of Medicine, https://pubmed.ncbi.nlm.nih.gov/35009846/ (Accessed on 10 April 2022)

[12] Melo, César. "Mola RGB COVSURV." *Mendeley Data*, Mendeley Data, 20 Dec. 2021, https://data.mendeley.com/datasets/vzf939jbxy/1 (Accessed on 1 may 2022)

[13] Liu, Z., Luo, P., Wang, X. and Tang, X. (2015) Deep Learning Face Attributes in the Wild. - References - Scientific Research Publishing, https://www.scirp.org/reference/ReferencesPapers.aspx?ReferenceID=2321305 (Accessed on 1 May 2022)

[14] Lin, Tsung-Yi, et al. "Microsoft Coco: Common Objects in Context." ArXiv.org, 21 Feb. 2015, https://arxiv.org/abs/1405.0312 (Accessed on 1 May 2022.)

[15] Le, Vuong, et al. "Interactive Facial Feature Localization." University of Illinois Urbana-Champaign, 30 Oct. 2012, https://experts.illinois.edu/en/publications/interactive-facial-feature-localization (Accessed on 1 May 2022)

[16] Nordstrøm, Michael M., et al. "The Imm Face Database - an Annotated Dataset of 240 Face Images." , 1 Jan. 1970, https://orbit.dtu.dk/en/publications/the-imm-face-database-an-annotated-dataset-of-240-face-images (Accessed on 1 May 2022)

[17] Yang, Shuo, et al. "Wider Face: A Face Detection Benchmark: Semantic Scholar." 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1 Jan. 1970, https://www.semanticscholar.org/paper/WIDER-FACE%3A-A-Face-Detection-Benchmark-Yang-Luo/52d7eb0fbc3522434c13cc247549f74bb9609c5d (Accessed on 2 May 2022)

[18] "What Is a Convolutional Neural Network? 3 Things You Need to Know." What Is a Convolutional Neural Network? | 3 Things You Need to Know - MATLAB & Simulink, https://www.mathworks.com/discovery/convolutional-neural-network-matlab.html (Accessed on 10 February 2023)

[19] Goutte, Cyril, and Éric Gaussier. "a Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation: Semantic Scholar." A Probabilistic Interpretation of Precision, Recall and F-Score, with Implication for Evaluation | Semantic Scholar, 1 Jan. 1970, https://www.semanticscholar.org/paper/A-Probabilistic-Interpretation-of-Precision%2C-Recall-Goutte-Gaussier/e3999a46cb8aaf71131a77670da5c5c113aad01d (Accessed on 10 February 2023)