Masked-Face Recognition in Counter-Terrorism during the COVID-19 Pandemic: An Approach and Implementation

Salma Fawzy Newegy

College of Artificial Intelligence

Arab Academy for Science Technology and Maritime Transportation, Egypt

s.f.nowigy@student.aast.edu

Abstract— **Face masks have become a way of life due to the COVID-19 pandemic. People now wear them nearly everywhere we go as they reduce disease transmission by 65%. This makes conventional face recognition ineffective in facial security checks specifically at airports since certain parts of the face are hidden. Hence, this paper investigate the same problem by developing a deep learning based model capable of accurately identifying people wearing disposable masks. In order to evaluate the model, a relatively small dataset is constructed and proposed. Experimental results on this dataset integrated with ResNet-50 based architecture show good masked-faces recognition performance. This outcome could be integrated within a facial recognition system for a better intelligent surveillance analytics in combating terrorism.**

Keywords— **COVID-19, Terrorism, ResNet-50, VGG-16**

I. INTRODUCTION

Coronavirus (COVID-19) is an infectious disease caused by the SARS-CoV-2 virus. Anyone can get sick with COVID-19 and become seriously ill or die at any age. Currently, we face a critical health challenge with this disease that has infected about 287 million people and killed over 5.43 million people worldwide. The COVID-19 virus can be spread through contact and contaminated surfaces, therefore, the classical biometric systems based on passwords or finger prints are not safe anymore. The best way to protect yourself and others is wearing a properly fitted mask. Face recognition techniques, the most important means of identification, have nearly failed, which has brought huge dilemmas to authentication applications that rely on face recognition. In particular, for public surveillance analytics in places like airports, the traditional face recognition system can not effectively recognize the masked faces, but removing masks for passing authentication will increase the risk of virus infection. On the other hand, terrorism risks are increasing. The utilisation of facial recognition solutions can play a key role in improving the efficiency of police forces and intelligence agencies to prevent terrorists from fleeing the country. However, in situations where there's a need to identify terrorists with facial covering, it becomes challenging for traditional facial recognition systems to identify them. To tackle these problems, this paper handles them by using a pre-trained deep learning based model in order to generalize it in identifying faces whether covered by masks or not. The proposed solution includes: 1) Finetuning a pre-trained VGG-16 model on our masked dataset and obtaining an accuracy of 37.91% and on unmasked training dataset and obtained accuracy of 63.016%. 2) Finetuning a pre-trained ResNet-50 model on our masked dataset and obtaining an accuracy of 41.53% and on unmasked training dataset and obtained accuracy of 71.09%.

II. PROPOSED METHOD

Overview. The ultimate goal of our model is to identify terrorists while wearing face masks. Two different models of transfer learning are used on a CNN-based model, VGG-16 and ResNet-50 architectures. At first, the model's parameters are finetuned on our masked-faces small dataset. Next, we used a hybrid dataset which contains masked and unmasked faces. There does not exist lots of prior work in this area, hence, the paper presents more than one approach to get the best results possible.

A. *Transfer Learning*

Transfer learning is the reuse of a previously learned model on a new problem. It's particularly popular in deep learning right now since it can train deep neural networks with small amount of data. This is extremely valuable in our model as the dataset used is relatively small even with data augmentation.

B. VGG-16 architecture

VGG-16 is a standard deep Convolutional Neural Network (CNN) that supports 16 layers and it is one of the most popular image recognition architectures. The VGG-16 model achieves almost 92.7% top-5 test accuracy in ImageNet. ImageNet is a dataset consisting of more than 14 million images belonging to nearly 1000 classes. This model supports 16 layers and can classify images into 1000 object categories, which means that VGG16 is a pretty extensive network and has a total of around 138 million parameters. There are a 13 convolution layers followed by 5 Max Pooling layer, that reduces the height and the width, Activation layers and fully connected layers. Fig. 1 shows the architecture of the VGG-16 model.

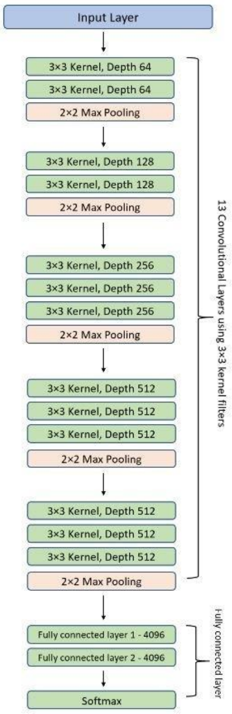

Fig.1.

C. ResNet-50 architecture

ResNet-50 has been successfully used in various pattern recognition tasks such as face detection. It is containing 50 layers trained on the ImageNet dataset. This network is a combination of Residual network integrations and Deep architecture parsing. Training with ResNet-50 is faster due to the bottleneck blocks. It is composed of five convolutional blocks with shortcuts added between layers. The last convolution layer is used to extract Deep Residual Features (DRF). Fig. 2 shows the architecture of the ResNet-50 model.

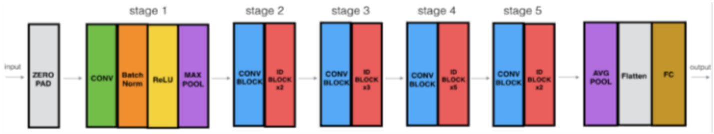

Fig.2.

III. EXPERIMENTAL SETUP

The implementation was done using Google Colab, it was evaluated at many stages.

A. Dataset

We took advantages of the existing "Real-world masked face recognition dataset" (RMFRD), which is a masked face dataset devoted mainly to improve the recognition performance of the existing face recognition technology on the masked faces during the COVID-19 pandemic. It contains 5,000 masked faces of 525 people and 90,000 unmasked faces. However, the data is not efficient as the number of images in each class various from the others. As a result, some classes are combined with our self-built Masked-Faces dataset. This dataset partitioned into two theories: a dataset full of only masked faces for training and testing, and the other one is hybrid such that the training set is the domain of unmasked faces while in testing process the masked faces are used. The dataset contains 16 classes, each has 10 images. Both masked and unmasked dataset is split into 70% training and 30% validation data. The dimension of the images are 256x256 pixels. We extracted faces from images and converted into 160x160. Next, data augmentation is applied including the random flip horizontally to our training and validation images.

A sample of the dataset is shown in Fig.3.

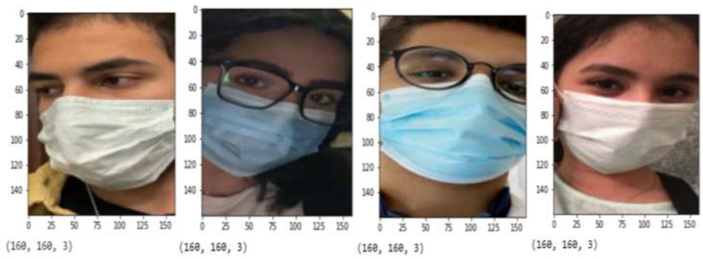

Fig.3.

B. Hyperparameters tuning

At first, we applied the best hyperparameters for traditional face recognition, then we manually tuned them to deal better with this new task. We had two different experiments on training our model, however, same hyperparameters are used as they worked the best as shown in Table I.

TABLE II
HYPERPARAMETERS USED FOR ALL MODELS

Hyperparameters tested	
Optimizer	Adam
Dropout	0.5
Epochs	20 – 25

IV. RESULTS

a- VGG-16 architecture

Masked faces

We trained our VGG-16 model on the training sets, and while testing on our masked data, we realized that it did not work well. This is because of the lower number of features due to the occlusion of the face by the mask. We used the Adam optimizer and batch size of 1. We run our algorithm for 25 epochs with cross-entropy loss, and we achieved an accuracy of 37.91%.

Unmasked faces

The VGG-16 architecture did better on the unmasked than on the masked dataset. This is as a result of having more dataset for training the unmasked than the masked dataset. Also, we could try other CNN models and compare the results to our architecture. We explain the different hyperparameters selected for our unmasked face training. We started with a Batch size of 1. Since we are using transfer learning, we froze the top layers of the pretrained network and trained the rest. We run our proposed model for 20 epochs with a cross-entropy loss. We achieved an accuracy of 63.016%.

b- ResNet-50 architecture

Masked faces

We trained our ResNet-50 model on the training set. We used the exact hyperparameters of the VGG-16 model and we achieved an accuracy of 41.53% and a loss of 2.4092.

Unmasked faces

The ResNet-50 architecture was trained on the unmasked and showing the highest performance among all the trials. Since we are using transfer learning, we froze the top layers of the pretrained network and trained the rest. We run our proposed model for 20 epochs with a cross-entropy loss. We achieved an accuracy of 71.09%.

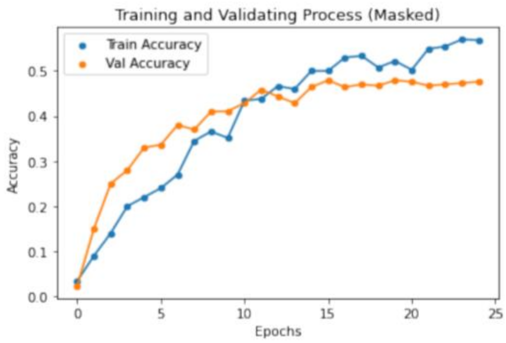

Fig.4.

V. CONCLUSIONS AND FUTURE WORK

In this approach, it is quit obvious that deep facial recognition neural network perform badly when faces are occluded during the pandemic. This paper shows many experiments that have been done with their results. The best implementation is done using ResNet-50 with a training set of unmasked-faces and testing with masked-faces dataset with accuracy 71.07%. Hence, more advanced techniques need to be implemented to improve their performance. We can use a new neural network like Efficient-Net and do some image processing on the dataset to be improved. We can also increase the dataset to perform better.

VI. REFERENCES

[1] Identifying terrorist suspects. INTERPOL. (n.d.). Retrieved from https://www.interpol.int/en/Crimes/Terrorism/Identifying-terrorist-suspects

[2] Author(s) John D. Woodward Jr., & Corporate Author Rand Corporation Address 1776 Main Street. (n.d.). Biometrics: Facing up to terrorism. Biometrics: Facing Up to Terrorism | Office of Justice Programs. Retrieved January 4, 2022, from https://www.ojp.gov/ncjrs/virtual-library/abstracts/biometrics-facing-terrorism

[3] Feature article: Identity verification during the age of covid-19. Department of Homeland Security. (2020, December 3). Retrieved January 5, 2022, from https://www.dhs.gov/science-and technology/news/2020/12/03/feature-article-identity-verification-during-age-covid-19

[4] Mandal, B., Okeukwu, A., & Theis, Y. (2021, April 19). Masked face recognition using resnet-50. arXiv.org. Retrieved from https://arxiv.org/abs/2104.08997

[5] Hariri, W. (2020). Efficient masked face recognition method during the COVID-19 pandemic. https://doi.org/10.21203/rs.3.rs-39289/v1

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015

[7] Shervin Minaee, Amirali Abdolrashidi, Hang Su, Mohammed Bennamoun, and David Zhang. Biometric Recognition Using Deep Learning: A Survey. arXiv:1912.00271 [cs], April 2020. arXiv: 1912.00271

[8] Wang, Z., Wang, G., Huang, B., Xiong, Z., Hong, Q., Wu, H., & Jiang, K. (n.d.). Masked face recognition dataset and application - arxiv.org. Retrieved January 5, 2022, from https://arxiv.org/pdf/2003.09093