A Low-Cost Embedded Security System for UAV-Based Face Mask Detector Using IoT and Deep Learning to Reduce COVID-19

Nashwan Adnan OTHMAN
Department of Computer Engineering,
College of Engineering
Knowledge University
Erbil 44001, Iraq
nashwan.adnan@knu.edu.iq

Ilhan AYDIN
Department of Computer Engineering
College of Engineering
Firat University
Elazig 23200, Turkey
iaydin@firat.edu.tr

2022 International Conference on Decision Aid Sciences and Applications (DASA) | 978-1-6654-9501-1/22/$31.00 ©2022 IEEE | DOI: 10.1109/DASA54658.2022.9765223

***Abstract*— Nowadays, the most effective method against the virus is wearing a mask. Hence, it is fundamental to wear a mask appropriately at open places like general stores and shopping malls. This paper proposes a novel human face mask detection method from UAV-captured frame sequences to solve the aforementioned problem. The proposed approach involves an offline stage and an inference stage. The offline stage generates the mask or no-mask by utilizing a convolutional neural network. We trained our model on a face mask dataset, and this enhancement allows the suggested system to obtain high accuracy in detecting unmasked people. The inference stage uses the already generated model to detect no mask humans and sends the alert to the smartphone-based Internet of Things. At this stage, Jetson nano was used to implement an embedded powerful real-time application for UAV-based face mask detection that runs at high frames per second. The proposed system monitors and detects people who have not worn a mask. Also, we used IoT techniques to send the pictures and notifications to the nearest police station to apply forfeit when it detects unmasked people. The main contributions in this paper lie in adjusting the deep learning, embedded platforms, IoT techniques, and Tello drone, generally dedicated to detecting unmasked people at a low cost. On average, detection accuracy is 99% based on the experimental evaluation of the proposed deep learning model for UAV-based face mask detection on the provided dataset. Overall, the proposed method can help decrease the spread of COVID-19 and other transmissible diseases.**

Keywords—UAV, COVID-19 virus, face mask detection, Internet of Things, Jetson nano

I. INTRODUCTION

COVID-19 pandemic has emerged in China, in December 2019. From there, the virus spread over the world, reaching nearly every country. World Health Organization (WHO) describes the two main factors that contribute to the spread of this infection as respiratory droplets and physical contact across people groups [1,2]. In the event that a tainted individual sneezes or coughs, respiratory droplets from other people in the vicinity (within 1 meter) may flow through the air and reach different surfaces that are closer. This illness can be found on almost any surface, which could lead to contact transmission.

Everyone is encouraged to use face masks in public areas during the COVID-19 outbreak. Agreeing to the WHO, masks can be utilized for controlling this painful disease [2].

Deep learning is a creative branch of machine learning techniques that uses organic neural networks to tackle bioinformatics, computer vision, and other fields [3]. Applications of deep learning in diagnostics and monitoring can always help reduce COVID-19's prevalence [4]. The proposed face mask detection systems use deep learning and computer vision techniques to detect unmasked people in a public area. A mask detector system depends on features attained by the face part. The unmask detector system has received a lot of attention recently because of its potential to reduce the transmission of diseases [5].

A low-cost embedded system proposed in this paper detects whether or not a human is wearing a mask using the Tello drone and Jetson nano platform. It is similar to an object detection system, which detects a specific type of thing. The proposed system could assist in ensuring people's safety in public spaces. This method can be utilized in various settings, including supermarkets and shopping malls, schools, colleges, and train stations. As a whole, the proposed system comprises of two core stages. The first stage is the offline stage, which includes preprocessing the dataset and training a deep learning model using the MobileNetV2 architecture to detect unmasked people. The second stage is the inference stage, which includes capturing Unmanned Aerial Vehicles (UAV) frames, face detection, face mask detection, and sending notifications to the smartphone-based IoT when unmasked people are detected.

The IoT has incorporated a modern measurement of living beings through smart items. Using this technology makes the connection between any media and anything at any time and place possible [6–8].

The significant contributions of the proposed system are concise as follows:

- We developed a low-cost, innovative embedded system to assist healthcare organizations during pandemics like the COVID-19 virus using the IoT and deep learning.
- We trained a novel Convolutional Neural Network (CNN) with approximately 100% accuracy to solve the problem of detecting unmasked people based on face parts.
- The proposed system is connected to the internet, detecting unmasked people and directly sending the

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

notification and pictures to the smartphone by using IoTs.

The rest of the paper is broken up like the following: In section 2, similar works on face mask detection systems are explained. The proposed methodology is given in section 3. Section 4 explores the findings and accuracy. In the end, the paper concludes with a suggestion for future work improvement in Section 5.

II. Related Works

Numerous studies have been conducted on the face mask detection system. Arjya Das et al. [9] describe a disentangled strategy to discover this explanation by combining TensorFlow, Keras, OpenCV, and Scikit-Learn with a few fundamental machine learning tools. The suggested approach accurately recognizes the face from the image and then determines whether or not a mask covers it. It can also detect a face and a mask in motion while performing an observation task. The technique achieves up to 95.77% and 94.58% accuracy on two different datasets, respectively. The Sequential Convolutional Neural Network is used to examine optimum parameter values to detect the nearness of masks without generating over-fitting accurately.

Aniruddha Srinivas Joshi et al. [10] offer a deep learning-based method for detecting facial masks in recordings. The proposed method uses MTCNN face detection to demonstrate that it can distinguish faces and their associated facial landmarks within a video frame.

Adnane Cabania et al. [11] provide a method for picture editing in addition to three kinds of masked face detection datasets: a combination of the datasets with successfully and wrongly masked faces is used.

III. Proposed Methodology

In the context of the human face mask detector from UAV-frames, we suggested a new method that tends to reduce the cost and improve the performance. This method comprises two stages; an offline phase and an inference phase. The offline phase generates the mask/no-mask model utilizing CNN architecture. The inference phase uses the pre-existing model to detect the faces and the proposed model to detect unmasked people. Fig. 1 shows the proposed approach of these two stages explained in the following sections.

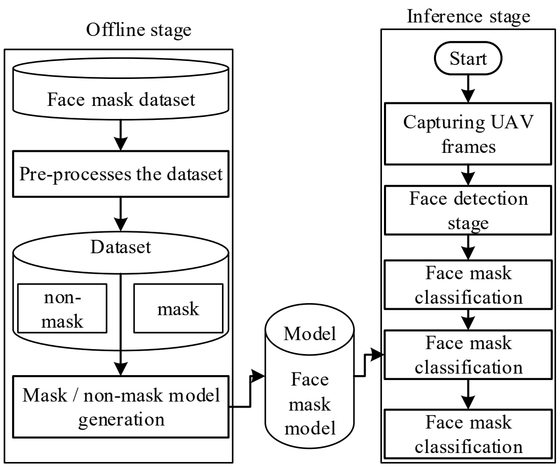

Fig. 1. The proposed approach for face mask detection from UAV-captured video sequences.

The above figure shows that the proposed UAV-based face mask detector is a two-stage system. The dataset will be utilized to train the unique face mask detector model by utilizing the Tensorflow and Keras libraries [12]. Then, we will use the generated model to detect unmasked people in real-time UAV video frames. Finally, the unmasked people with notifications will be sent to the smartphone by using IoT technology.

A. *Offline Stage*

The offline stage integrates two steps: preparing the face mask dataset and generating a mask/no-mask model. This section will load the proposed face mask dataset to train a model utilizing TensorFlow and Keras.

1) Dataset: The first step in the offline stage is to prepare the proposed dataset [13] to generate a face mask detector model. Normal images of faces were used to construct this dataset. Then, a new computer vision python software was built to add face masks to these photographs, resulting in an artificial yet real-world usable dataset. Face detection is utilized to identify the face's bounding box location in the image. The Region of Interest (ROI) is extracted once we know where the face is in the image. Then, facial landmarks locate the eyes, nose, mouth, and other facial features [14]. his method is resistant to rigid and nonrigid facial deformations caused by natural motions, poses, and expressions. It is used to make this approach easier. A mask with a translucent background is photographed. This mask will be artificially suited to the face part utilizing facial landmarks extracted along with the nose and chin. Then, the mask will be resized and rotated before being applied to the face. The artificial face mask dataset is constructed by repeating the procedure for all input photographs. The proposed dataset comprises of (1,376 pictures) divided into two groups, with masks (690 pictures) and without masks (686 pictures). Fig. 2 shows the sample of the proposed dataset.

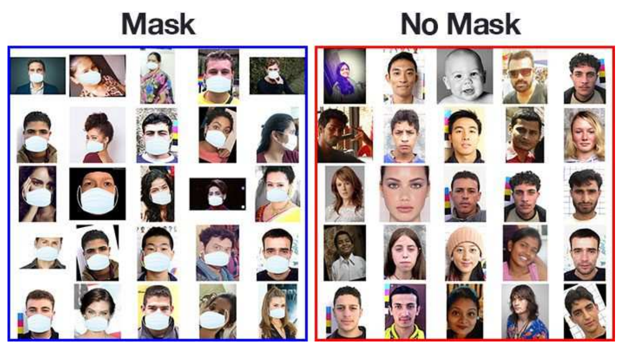

Fig. 2. The proposed dataset [13].

2) Training Deep Learning Model for Mask/No-Mask: In this stage, the loaded dataset will build a deep learning model that can identify if someone is wearing a mask or not. The MobileNetV2 architecture was utilized to generate the face mask detector model [15]. Although the architecture is very efficient, it was chosen because it can be used on embedded devices with little computational resources, like the Jetson nano or Raspberry pi. The cost of making face mask detection systems could be reduced by deploying trained MobileNet models to embedded devices. Firstly, we will set the hyperparameter constant values: learning rate, batch size, and number of training epochs. We can load and preprocess the training dataset when done with this step. Processing starts with resizing the frames and changing them to an array format. Then, we will split our dataset into two portions: one for training and another for testing. Then, we will define the structure of MobileNetV2 for training. The trained model will be saved to disk at that point. Small bottlenecks exist at both the input and output of the MobileNetV2 residual block, which is built on an inverted residual architecture. In contrast to standard residual models, MobileNetV2 filters feature lightweight depth-wise convolutions in the intermediate expansion layer instead of expanded representations in the input. Furthermore, we found that removing non-linearities in the thin layers is necessary to sustain representational power. Because our method separates the input and output domains, it provides a more precise analytical context for the transformation's expressiveness. In addition to a wide range of activities and benchmarks, MobileNetV2 improves the performance of mobile models across a variety of model sizes. The MobileNetV2 structure is displayed in Fig. 3.

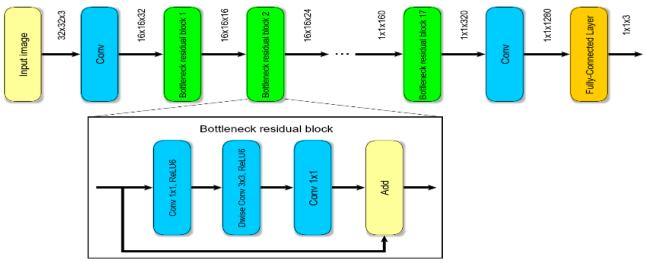

Fig. 3. The MobileNetV2 architecture.

B. Inference Stage

In the inference stage, we take a UAV frame and apply a face detection model to detect faces. To recognize a human wearing or not wearing a mask class, we will use the trained face mask detector model to categorize each face as with or without a mask. The proposed system may be implemented in a few stages. Firstly, Tello drone takes aerial frames. Secondly, the face detection method detects all faces in UAV frames. Then, the face mask detection technique is utilized to determine the face mask, followed by the face mask detector model. Next, the face mask is estimated based on the stages mentioned above. Finally, the warning message will be transmitted over the IoT system if a person is not wearing a mask based on the IoT Telegram API. The flowchart for the real-time proposed approach is depicted in the complete steps in Figure 4.

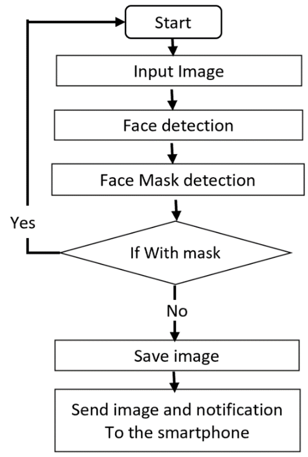

Fig. 4. Flowchart for real-time proposed system

1) Capturing UAV frames: We live in the twenty-first century, and technology has an impact on almost everything. Nowadays, we can use intelligence technology like UAV to operate a smart system in creating smart cities. We used the Tello drone to capture frames and analyse them using deep learning techniques to detect unmasked people in the proposed system.

2) Face Detection: Face detection is a crucial stage of the proposed system. In computer vision and machine learning, several methods can be utilized to detect human faces. However, with the rapid advancement of deep learning, most work in the field of face identification is expected to utilize it is algorithms. A pre-trained deep learning model utilized to find faces in this study. The backbone of this model is the ResNet-10 Architecture, which is based on the Single Shot Detector (SSD) [16]. Fig. 5 depicts the SSD architecture.

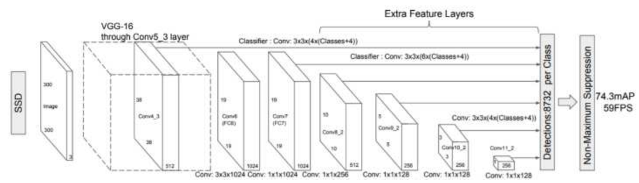

Fig. 5. SSD architecture [16].

3) Face Mask Classification: After the face detection stage, the face mask detector model generated in the offline stage is used to distinguish whether someone is wearing a mask or not. Our system's goal is to identify those who aren't wearing a face mask. The learning architecture produces a response based on the input frames, categorizing it as a mask or no mask. An alarm will ring if someone is seen without a mask. Furthermore, if everyone puts on a mask, they will resist the disease. Using the softmax function, the final layer of the convolutional neural network produces a vector of n real numbers that can be used to perform the classification. These numbers, which add up to 1, represent the probability of activity.

4) IoT API: In this paper, the Telegram bot API is utilized to announce the smartphone when unmasked people are detected. The TeleBot API [17] provides an HTTP-based interface for bot developers. TeleBots are accounts that do not require a phone number to set up. The Newbot command creates a new bot. The Bot Father will ask for a username and a name and create a mandate token for the new bot. The token is a string of characters required to authorize the bot. Fig. 6 displays the steps to creating a Telebot for the proposed system.

Fig. 6. Bot creation for unmasked people detector system.

IV. Experimental Results

Deep convolutional neural networks have lately been proven to be more effective than humans at detecting and recognizing objects. This study employs deep learning algorithms to determine whether the person is wearing a facemask. In the suggested method, we trained the deep learning model to detect face masks, and it is capable of achieving approximately 99% accuracy on the test set. The evaluated results for the trained model are shown in Table I. Fig. 7 illustrates the accuracy and training loss for the trained face mask detector model. It has a high degree of accuracy and only a small amount of overfitting, with the validation loss less than the training loss.

TABLE I. Accuracy of Face Mask Detector Model

	Precision	Recall	F1-score
Mask class	0.99	1.00	0.99
No-mask class	1.00	0.99	0.99
Accuracy			0.99

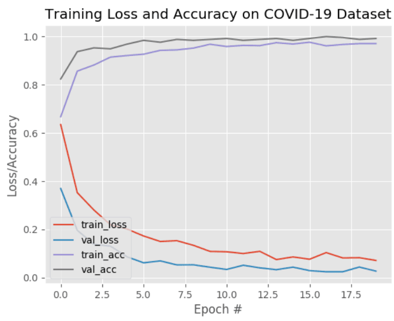

Fig. 7. Training loss and accuracy

After training stage, we will be ready to deploy the model based the proposed pipeline. We developed an embedded intelligent security system that uses deep learning and IoT to reduce COVID-19 disease in real-time UAV-video frames. When an unmasked person is discovered, the system will raise alarms. As a programming language, Python was utilized to develop this brilliant application. Therefore, a low-cost, high-performance embedded system can be created using the Jetson nano platform. The Jetson nano board, Tello drone, and power bank are included in the low-cost experimental setup represented in figure 8. All the necessary software has been installed, apart from the hardware connections.

Fig. 8. Low-cost experimental setup.

The real-time UAV-based face mask detector system is prepared by utilizing a deep learning method for the human face. When a person's face is detected, the system next determines whether or not the person is wearing a mask. Fig. 9 displays the result when the human is wearing a mask. The IoT Telegram application is used to integrate the suggested system. As shown in Figure 10, after unmasked people are detected, the system will notify the smartphone by sending an image of the unmasked person with a notification. Conversely, when a human is wearing a mask, the system will not alert the smartphone.

Fig. 9. Result when people wear a mask.

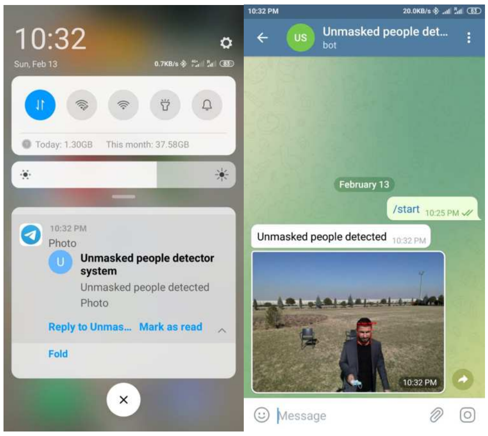

Fig. 10. Smartphone notification for unmasked people.

From there, the owner of the system that get this alert can oblige him/her to wear a mask. The proposed system, as seen above, is quite effective in identifying unmasked humans in UAV images. The suggested system utilizes the Jetson nano platform. The Jetson nano's NVIDIA GPU and 4 GB of RAM make it the fastest embedded artificial intelligence platform at a low price. Table II compares the Jetson nano and the Raspberry Pi4 for their performance. The Jetson nano-based system appears to outperform the Raspberry Pi4. There are 9 frames per second (FPS) approximately on the Jetson nano, while only 3 FPS approximately are achievable on the Raspberry Pi 4. Due to its NVIDIA GPU, the Jetson nano is three times faster than the Raspberry pi4. The execution time includes taking frames, loading a deep learning model for face detection, loading a model for face mask detection, and sending notifications with unmasked pictures to the smartphone through IoT.

TABLE II. EXECUTION TIME FOR THE PROPOSED SYSTEM

Embedded Platform	FPS	GPU
Jetson nano	~ 8.72	Yes
Raspberry pi4 – 4GB	~ 2.6	No

The suggested method is robust and figures out the fast result. This face mask detector model relies on the pre-trained face detection model for accuracy. Using aerial photos to train a face detector model improves the proposed system's accuracy. Moreover, given the above results, we can recognize unmasked people accurately on UAV frames by using a trained face mask detector model.

V. CONCLUSIONS

This paper proposed a low-cost embedded security system to reduce the COVID-19-based UAV face mask detector by utilizing deep learning and IoT. The proposed intelligent security system may prevent people from removing their masks while doing their daily activities in public spaces. Mainly, humans without a mask in crowded places cause danger to all those in the area, especially since the coronavirus has been infected with people, and it is a dangerous disease. To solve this crisis, the custom deep learning model trained using the MobileNetV2 architecture to classify whether a human is wearing a mask or not. The trained model attained a classifier that is approximately 100% accurate. Then, we applied the mask detection model to real-time UAV videos by identifying faces in each frame. Additionally, this system is suitably associated with the internet to alert the system's owner based on IoT when an unmasked person is detected. Using MobileNetV2's architecture, the suggested face mask detector is computationally efficient and easy to execute on embedded devices like the Jetson nano. COVID-19 will be reduced by using this low-cost intelligent system, which also works fast, and approximately 9 FPS can be achieved. Nevertheless, the performance of the suggested method is compared with a Raspberry Pi4-4GB. Using a face mask in this manner has the drawback of obscuring the face. The face mask detector will not work if enough of the face is obscured. To avoid this, training a two-class object detector with and without masks enhances the accuracy of the proposed system.

REFERENCES

[1] E. Mahase, Coronavirus covid-19 has killed more people than SARS and MERS combined, despite lower case fatality rate, BMJ. 368 (2020) m641. https://doi.org/10.1136/bmj.m641.

[2] Advice for the public: Coronavirus disease (COVID-19), (n.d.). https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public (accessed October 1, 2021).

[3] A. Abbasian, A. Rajabzadeh, U.R. Acharya, N. Khadem, A. Mohammadi, Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images : Results of 10 convolutional neural networks, Comput. Biol. Med. 121 (2020) 103795. https://doi.org/10.1016/j.compbiomed.2020.103795.

[4] N.A. Othman, M.Z.N. Al-Dabagh, I. Aydin, A New Embedded Surveillance System for Reducing COVID-19 Outbreak in Elderly Based on Deep Learning and IoT, 2020 Int. Conf. Data Anal. Bus. Ind. W. Towar. a Sustain. Econ. ICDABI 2020. (2020). https://doi.org/10.1109/ICDABI51230.2020.9325651.

[5] A.A.R. Alsaeedy, E.K.P. Chong, Detecting Regions At Risk for Spreading COVID-19 Using Existing Cellular Wireless Network Functionalities, (n.d.). https://doi.org/10.1109/OJEMB.2020.3002447.

[6] Y. Perwej, K. Haq, F. Parwej, M. M., The Internet of Things (IoT) and its Application Domains, Int. J. Comput. Appl. 182 (2019) 36–49. https://doi.org/10.5120/ijca2019918763.

[7] N.A. Othman, I. Aydin, A face recognition method in the Internet of Things for security applications in smart homes and cities, Proc. - 2018 6th Int. Istanbul Smart Grids Cities Congr. Fair, ICSG 2018. (2018) 20–24. https://doi.org/10.1109/SGCF.2018.8408934.

[8] N.A. Othman, I. Aydin, A new IoT combined body detection of people by using computer vision for security application, Proc. - 9th Int. Conf. Comput. Intell. Commun. Networks, CICN 2017. 2018-Janua (2018) 108–112. https://doi.org/10.1109/CICN.2017.8319366.

[9] A. Das, M. Wasif Ansari, R. Basak, Covid-19 Face Mask Detection Using TensorFlow, Keras and OpenCV, 2020 IEEE 17th India Counc. Int. Conf. INDICON 2020. (2020). https://doi.org/10.1109/INDICON49873.2020.9342585.

[10] A.S. Joshi, S.S. Joshi, G. Kanahasabai, R. Kapil, S. Gupta, Deep Learning Framework to Detect Face Masks from Video Footage, Proc. - 2020 12th Int. Conf. Comput. Intell. Commun. Networks, CICN 2020. (2020) 435–440. https://doi.org/10.1109/CICN49253.2020.9242625.

[11] 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 Heal. 19 (2021) 100144. https://doi.org/10.1016/j.smhl.2020.100144.

[12] Google, An end-to-end open source machine learning platform, (n.d.).

[13] P. Bhandary, Face Mask Dataset, (n.d.). https://github.com/prajnasb/observations/tree/master/experiements/data (accessed March 26, 2020).

[14] K.S. Khabarlak, L.S. Koriashkina, Fast Facial Landmark Detection and Applications: A Survey, n.d.

[15] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.C. Chen, MobileNetV2: Inverted Residuals and Linear Bottlenecks, Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (2018) 4510–4520. https://doi.org/10.1109/CVPR.2018.00474.

[16] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.Y. Fu, A.C. Berg, SSD: Single shot multibox detector, Lect. Notes Comput. Sci. (Including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). 9905 LNCS (2016) 21–37. https://doi.org/10.1007/978-3-319-46448-0_2.

[17] Telegram Bot API, (n.d.). https://core.telegram.org/bots/api (accessed December 30, 2021).