NÖHÜ Müh. Bilim. Derg. / NOHU J. Eng. Sci., 2023; 12(3), 762-769

Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi
Niğde Ömer Halisdemir University Journal of Engineering Sciences

Araştırma makalesi / Research article

www.dergipark.org.tr/tr/pub/ngumuh / www.dergipark.org.tr/en/pub/ngumuh

Robotic based mask detection to prevent epidemic diseases transmitted through droplets using pre-trained deep learning models

Önceden eğitilmiş derin öğrenme modelleri kullanılarak damlacık yoluyla bulaşan salgın hastalıkları önlemek için robotik tabanlı maske tespiti

Ali Ünlütürk[1,*]

[1] *Erzurum Technical University, Electric-Electronic Engineering, 25100, Erzurum Türkiye*

Abstract

The Coronavirus disease, which emerged in Wuhan, China in December 2019 and spread rapidly all over the world, infected healthy people by being transmitted by small droplets. Medical experts have stated that the most effective fight against the Coronavirus disease is the need for people in contact to wear masks. Despite this, some people violated the obligation to wear masks. In this study, mask detection performances of pre-trained Convolutional Neural Network (CNN) models such as NasNetMobile, MobileNetV3Small, ResNet50, DenseNet121 and EfficientNetV2B0, which were previously trained, were evaluated in order to automatically detect people who violate the mask wearing obligation. At the end of this evaluation, DenseNet121 architechture has become the most successful model. This model has been tested with the image obtained from the camera on a robotic system with six Degrees of Freedom (6-DOF). The human face images taken from the camera were processed using the Jetson Xavier NX development board. As a result, this study will help the officers who carry out mask inspections in public areas and will significantly reduce the spread of new outbreaks similar to the Coronavirus.

Keywords: Mask detection, Epidemic diseases, Convolutional neural network, Transfer learning, Robotic

Öz

Aralık 2019'da Çin'in Wuhan şehrinde ortaya çıkan ve tüm dünyada hızla yayılan Koronavirüs hastalığı küçük damlacıklar ile bulaşarak sağlıklı insanları enfekte etmiştir. Tıp uzmanları Koronavirüs hastalığına karşı en etkili mücadelenin temas halindeki kişilerin maske takması gerekliliğini belirtmişlerdir. Buna rağmen bazı kişiler maske takma zorunluluğunu ihlal etmişlerdir. Bu çalışmada maske takma zorunluluğunu ihlal eden kişilerin otomatik olarak tespit edilebilmesi için önceden eğitilmiş olan NasNetMobile, MobileNetV3Small, ResNet50, DenseNet121 ve EfficientNetV2B0 gibi Evrişimli Sinir Ağı (CNN) modellerinin maske tanıma performansları değerlendirilmiştir. Bu değerlendirme sonucunda en başarılı model DenseNet121 mimarisi olmuştur. Bu model altı Serbestlik Derecesine (6-DOF) sahip robotik bir sisteminin üzerinde yer alan kameradan elde edilen görüntü ile test edilmiştir. Kameradan alınan insana ait yüz görüntüleri Jetson Xavier NX geliştirme kartı kullanılarak işlenmiştir. Sonuç olarak, bu çalışma toplu alanlarda maske denetimi gerçekleştiren görevlilere yardımcı olacak ve Koronavirüs benzeri çıkabilecek yeni salgınların yayılımı önemli ölçüde azaltacaktır.

Anahtar Kelimeler: Maske tanıma, Epidemik hastalıklar, Evrişimli sinir ağı, Transfer öğrenme, Robotik

1 Introduction

On December 31, 2019, they reported to the World Health Organization's (WHO) office in China of an unknown case of the disease detected in Wuhan City, Hubei Province, China. Later, Chinese officials reported that they detected a new type of Coronavirus by isolating this unknown disease on January 7, 2020 [1]. The scientific world has named this virus as COVID-19 (SARS-CoV-2) [2]. COVID-19 can easily be transmitted from person to person via droplets spreading to the environment due to cough and sneezing. In addition, some studies show that the COVID-19 virus tends to be transmitted by droplets or air, even three hours after aerosolization [3]. At the same time, many people with this disease usually have some symptoms such as fever and shortness of breath. However, some patients do not show any symptoms. Therefore, this virus has turned into a serious epidemic that killed many people around the world and caused the epidemic to reach serious dimensions in a short time. Therefore, due to the dramatic increase in this epidemic, many countries around the world have completely closed their borders and flights to other countries. In order to prevent the rapid spread of this disease, many states oblige people to be at certain physical distances without touching each other in social life. In addition, some serious measures have been taken in social life, such as the obligation to wear face masks. However, many people in the society do not show the necessary care to wear a mask. In order to prevent this epidemic from spreading rapidly, machine learning or Convolutional Neural Network (CNN) methods and image processing-based solutions that check whether people wear face masks have been proposed by the researchers [4-6]. In order to create a face mask detection model using classical machine learning techniques, the feature vector of the model should be extracted. In order to extract the feature vector belonging to the model, experts in the field are required on

* Sorumlu yazar / Corresponding author, e-posta / e-mail: ali.unluturk@erzurum.edu.tr (A. Ünlütürk)
Geliş / Recieved: 03.05.2023 Kabul / Accepted: 30.05.2023 Yayımlanma / Published: 15.07.2023
doi: 10.28948/ngumuh.1291781

the model examined. Unlike machine learning and classical image processing techniques, in deep learning neural networks, CNNs are designed to automatically learn and extract features from raw data. In computer vision applications including image classification, object detection, and image segmentation, CNN is therefore frequently utilized. An image or a series of images are often used as input data for CNNs. CNN has many layers such as convolution, pooling and activation functions. It is widely used for tasks where input data such as image, sound and natural language are complex and high-dimensional.

Transfer Learning, one of the machine learning techniques, is a method that enables a trained model to be reused in another related task [7]. This approach allows the model to transfer the knowledge acquired during the training process from the first task to the second task. Thus, very good performance and faster training times can be achieved for the second task. In fact, it is basically based on human learning. Because when people learn new skills or knowledge, they often benefit from their existing knowledge and experience. For example, a person who knows how to ride a bike can use this knowledge to learn how to ride a motorcycle [8]. In transfer learning, a pre-trained model, such as a neural network trained on a large dataset, is commonly used as the starting point for a new model. Also, the pre-trained model can be fine-tuned on a new dataset so that it does not adapt quickly to the new task and performs well. Another approach is to extract features from the pre-trained model. For example, a pre-trained image classification model can be used to extract features from images and these can then be imported into a new model for a different task such as object detection or image segmentation. Transfer learning has been shown to be effective in a wide variety of applications, including natural language processing, computer vision, and speech recognition [9-11].

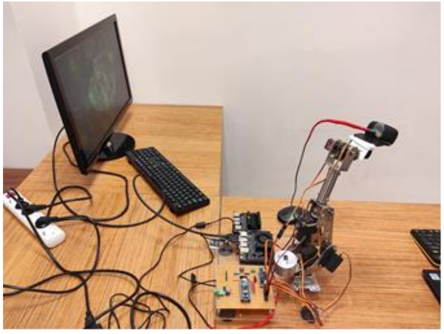
Real-time face mask images obtained with the camera should be analyzed mathematically and meaningful results should be produced. The most important point here is that the measurement capabilities, analysis capabilities and reliability of the machine vision systems are in full harmony with the software structure of the selected hardware components. Images obtained through camera systems cannot be processed precisely with a traditional Central Processing Unit (CPU) with a limited number of cores (one, two, four or eight). Because the CPUs having traditional multiple cores cannot process the huge data enough [12]. Because, on the CPU with a single core, the threads are executed by the operating system in a time-sharing manner according to their priorities and the situation. As the number of cores increases, multiple threads can be executed simultaneously. However, classical CPUs are insufficient in sensitive computer vision systems. Then the data from the layers are converted into input data of the classical neural network. In recent years, the ability to easily obtain enough data belonging to many systems or models, the increase in the processing capacity of computer or embedded development board architectures, the heterogeneous parallel processing capability of CPU and Graphics Processing Unit (GPU) have encouraged new studies on Deep Learning. In this study, Jetson Xavier NX embedded development board, which is capable of parallel computing by using different heterogeneous hardware units such as CPU and GPU, was used to quickly determine whether people are wearing a face mask. In real-time face mask detection system, the CPUs of the Jetson Xavier NX development board and the CUDA cores owned by the GPUs were actively used. As a result, a new deep learning-based algorithm has been developed in which the layers of the learning structure of the basic model are transferred, which can detect whether people are wearing a face mask to prevent the rapid spread of Covid-19 in the society.

In this paper, NasNetMobile, MobileNetV3Small, ResNet50, DenseNet121, and EfficientNetV2B0 transfer learning approaches were applied for mask detection. The results show that DenseNet121 gives high performance with %100 accuracy, against other specified architectures. As a result, the DenseNet121 model was run on the Jetson Nano Xavier development board. The studies in the literature and the face mask detection system are introduced in the first part of the article. In the second part, the hardware architecture and software structure of the Jetson Xavier NX development board used for the face mask recognition application system are discussed. In addition, the pre-trained deep learning models used in the face mask detection system are discussed. In the third chapter, the performances of the pre-trained deep learning models used in the face mask recognition detection system were evaluated. In the last part, the results from the article are evaluated.

2 Material and methods

2.1 Robotic based mask detection

In this study, a robotic arm with 6-DOF equipped with a servo motor for each joint was used. The motion of the robotic arm is controlled by a controller software in the STM32F103C8T6 embedded board, which is a development board with ARM Cortex-M3 core and running at maximum 72MHz. At the end of this robotic arm system, there is a camera that allows images to be taken. Thus, images can be taken flexibly from the robotic system. Jetson Xavier NX embedded board developed by NVIDIA company was used to process the obtained robotic images. Figure 1(a) shows the robotic based face mask detection hardware. Jetson Xavier NX development board is connected to a monitor via High Definition Multimedia Interface (HDMI) connection. For internet access, it has been connected with an Ethernet cable. In addition, the Camera Serial Interface (CSI) and camera, Universal Serial Bus (USB) and keyboard and mouse connections have been made. Jetson Xavier NX development board shown in Figure 1(b) consists of two main parts as main board and carrier card. Under the aluminum cooler there is a 6-core NVIDIA Carmel ARM®v8.2 64-bit CPU, 6MB L2 + 4MB L3, 384 NVIDIA(R) CUDA® cores and 48 Tensor cores GPU. In addition, there is 8GB 128 bit LPDDR4x 59,7GB/sn Random Access Memory (RAM) on it [13].



(a)

(b)

Figure 1. (a) Robotic face mask detection (b) NVIDIA Jetson Xavier NX development board

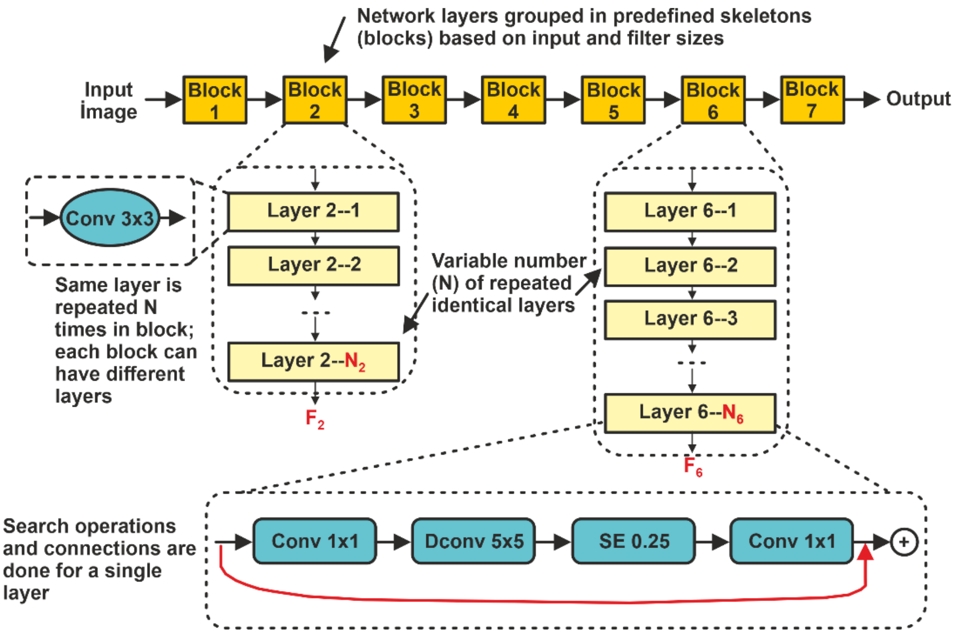


Figure 2. NasNetMobile architechture

First, the NVIDIA JetPack Software Development Kit (SDK) was installed on the Jetson Xavier NX development board. NVIDIA JetPack SDK includes operating system, libraries and Application Programming Interfaces (APIs), developer tools, samples, and documentation. As seen in Figure 2, NVIDIA JetPack SDK was installed on the MikroSD memory card to be located on the Jetson Xavier NX development board in the first stage of the installation. JetPack SDK setup file is loaded on a 64GB MicroSD card. Finally, the display is connected to the HDMI input of the development board, and the keyboard and mouse control interfaces are connected to the USB inputs. Jetson Xavier NX development board has been rebooted after all the installation steps. Jetpack, prepared by NVIDIA company; It is a software package that includes accelerated libraries such as OpenCV, Python, Tensorflow, CUDA-X, Deep Learning, Computer Vision, Accelerated Computing and Application Programming Interfaces (API) for multimedia. Once the installation process of the NVIDIA Jetson Xavier NX development board is completed, it can be used in different real-time robotic image processing applications.

2.2 Pre-Trained convolutional neural network models

One of the main features of deep learning is that it automatically learns useful patterns without the need for feature extraction from raw data. In order to extract these meaningful patterns, layers of many different structures such as Conv2D and MaxPooling2D must be superimposed. It is very important that these layer structures are selected correctly and their setup is very important. In order to create this structure correctly, the deep learning model should be trained. But there must be a lot of data to train the CNN structure. In addition, it is necessary to work on huge data sets that take days to train the network. Therefore, it is often useful to take advantage of previously worked on and resulting trained weights in models similar to each other. Big companies like Google have shared models that they have trained with big data. Structures such as NasNetMobile [14], MobileNetV3 [15], ResNet50 [16], DenseNet121 [17] and EfficientNetV2B0 [18] are some of the pre-trained models. The Google artificial intelligence team created the NasNet (Neural Architecture Search Network) family of deep neural network models utilizing an automated neural architecture search method [14]. NASNet networks are scalable Express Storage Architecture (ESA) architecture and consist of simple blocks such as separable convolution and jointing, which are improved by reinforcement learning method [19]. The aim of NasNet is to construct extremely accurate and efficient neural networks that can be trained on massive datasets with little assistance from humans. NasNet searches for the best network designs using a reinforcement learning-based method [14]. A performance benchmark is achieved by the technique after training a limited initial set of network architectures on a dataset. The program produces a new set of architectures and retrains them to assess their performance in order to replace the initial architectures based on this metric. As long as the algorithm converges on a set of high-performance architectures, this process continues to be carried out. The NasNetMobile network architecture is shown in Figure 2 [14]. For mobile and embedded devices, MobileNetV3 is a series of effective deep neural network models. Google Artificial Intelligent (AI) department introduced MobileNet, a family of TensorFlow based computer vision models, in 2017. They created the MobileNetV3 module in May 2019 [20]. The MobileNetV3 network architecture is shown in Figure 3 [15]. Then, ResNet (short for "Residual Network") is a deep neural network architecture that was introduced in 2015 by a research team from Microsoft Research [21]. ResNet is made to deal with the issue of vanishing gradients, which can slow down training in extremely deep neural networks and degrade convergence as network depth rises [22]. ResNet's fundamental concept is the use of redundant connections, which enable information to travel straight from one layer to another without passing through one or more intermediate levels [23]. Instead of learning the complete function from scratch, this method enables the network to learn a residual function that adds a layer's input to its output. The ResNet50 network architecture is shown in Figure 4 [16]. The DenseNet model was proposed by researchers from Cornwell University, Tsinghua University, and the Facebook AI Research (FAIR) group [24]. The famous ResNet architecture has been modified to include dense connections between layers in addition to residual connections [25]. This version is known as DesNet, or Dense Residual Network. Dense connections aid gradient flow and reduce the potential vanishing gradient issue in deep neural networks [26]. The DenseNet121 network architecture is shown in Figure 5 [17]. A set of deep neural network models called EfficientNet has been proposed by "Tan et al" to tackle state-of-the-art image identification tasks with minimal parameters and computational expense [27]. They have had success using a new method they call EfficienetNet to scale the depth, width and resolution of the model. Through this method, EfficientNet is able to outperform other models with comparable computational costs in terms of accuracy. The inclusion of a novel convolutional block known as the "swish" activation function is another significant innovation in EfficientNet [28]. Compared to more conventional activation functions like ReLU, Swish is a smooth, non-monotonic function that is quicker and more precise [29]. The EfficientNetV2B0 network architecture is shown in Figure 6 [18].

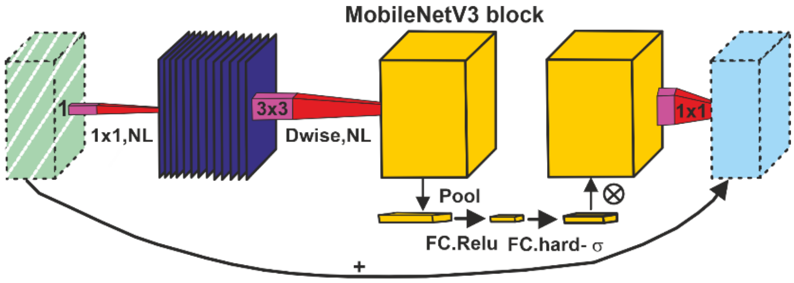


Figure 3. MobileNetV3 architechture

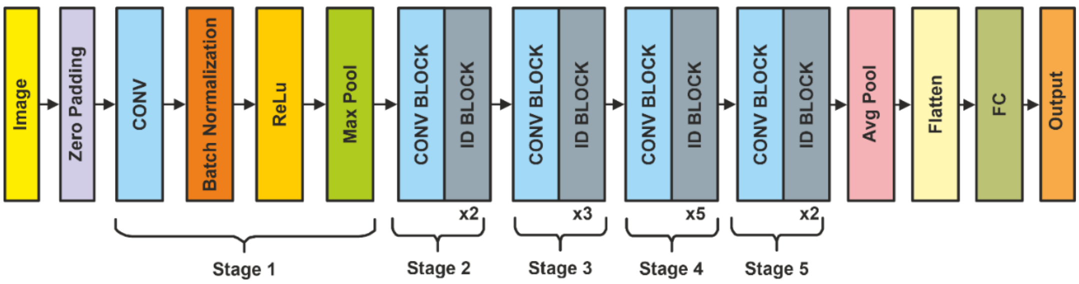


Figure 4. ResNet50 architechture

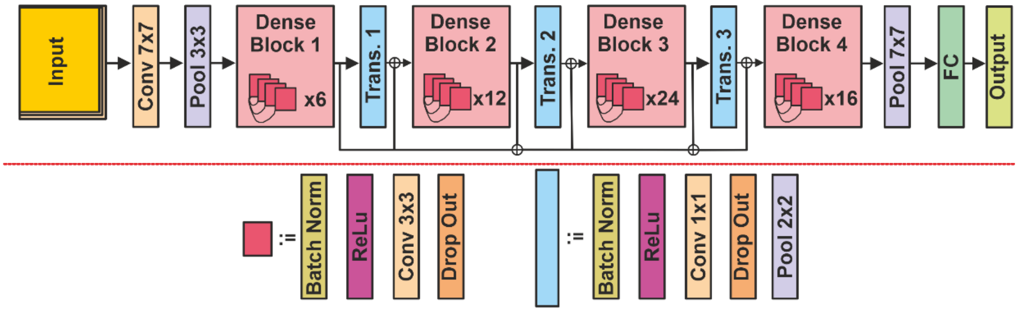


Figure 5. DesNet121 architechture

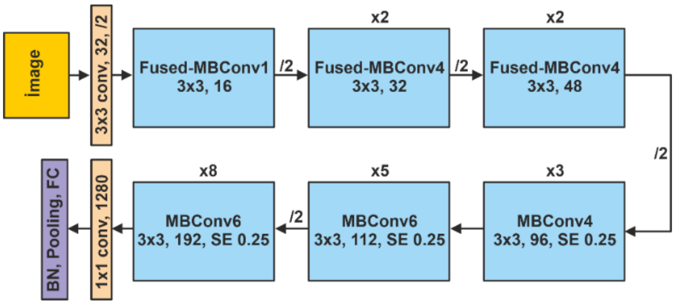


Figure 6. EfficientV2B0 architechture

3 Results and discussion

The data set in reference [30] was used in this study. It has 3833 images, where 1915 images are with masks and 1918 without masks. Train and test trials developed for the face mask detection were carried out on a laptop equipped with Intel i5-10300H processor and 8 GB RAM. This model, which was developed after the correct network structure was created, was run on NVIDIA Jetson Xavier NX board. The performance metrics of the pre-trained models are explained below Equation (1)-(4).

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{(\text{TP} + \text{FP} + \text{FN} + \text{TN})} \quad (1)$$

$$\text{Precission} = \frac{\text{TP}}{(\text{TP} + \text{FP})} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \quad (3)$$

$$\text{f1 score} = 2 * \frac{\text{Recall} * \text{Precision}}{(\text{Recall} + \text{Precision})} \quad (4)$$

Where TP=True Positive, TN=True Negative, FP=False Positive, FN=False Negative. We set the learning rate parameter to 0.0001. Due to memory limitations of this platform, the model is trained in batches of 32 images at a time. The optimization algorithm used is “Adam”. The dataset is split into two portions: training (%80) and test (%20). The performance metrics using pre-trained models are shown in Table 1. The DenseNet121 model used in face mask detection is given in Figure 7. The pre-trained different models performance during the iteration are given Figure 8.

Table 1. Pre-trained model performances

NasNetMobile	Precision	Recall	F1 Score	Support
with mask	0.99	0.99	0.99	383
without mask	0.99	0.99	0.99	384
accuracy			**0.99**	767
macro average	0.99	0.99	0.99	767
weighted average	0.99	0.99	0.99	767
MobileNetV3Small	Precision	Recall	F1 Score	Support
with mask	0.87	0.91	0.89	383
without mask	0.90	0.86	0.88	384
accuracy			**0.89**	767
macro average	0.89	0.89	0.89	767
weighted average	0.89	0.89	0.89	767
Resnet50	Precision	Recall	F1 Score	Support
with mask	0.92	0.93	0.92	383
without mask	0.93	0.92	0.92	384
accuracy			**0.92**	767
macro average	0.92	0.92	0.92	767
weighted average	0.92	0.92	0.92	767
DenseNet121	Precision	Recall	F1 Score	Support
with mask	0.99	1.00	1.00	383
without mask	1.00	0.99	1.00	384
accuracy			**1.00**	767
macro average	1.00	1.00	1.00	767
weighted average	1.00	1.00	1.00	767
EfficientNetV2B0	Precision	Recall	F1 Score	Support
with mask	0.99	0.99	0.99	383
without mask	0.99	0.99	0.99	384
accuracy			**0.99**	767
macro average	0.99	0.99	0.99	767
weighted average	0.99	0.99	0.99	767

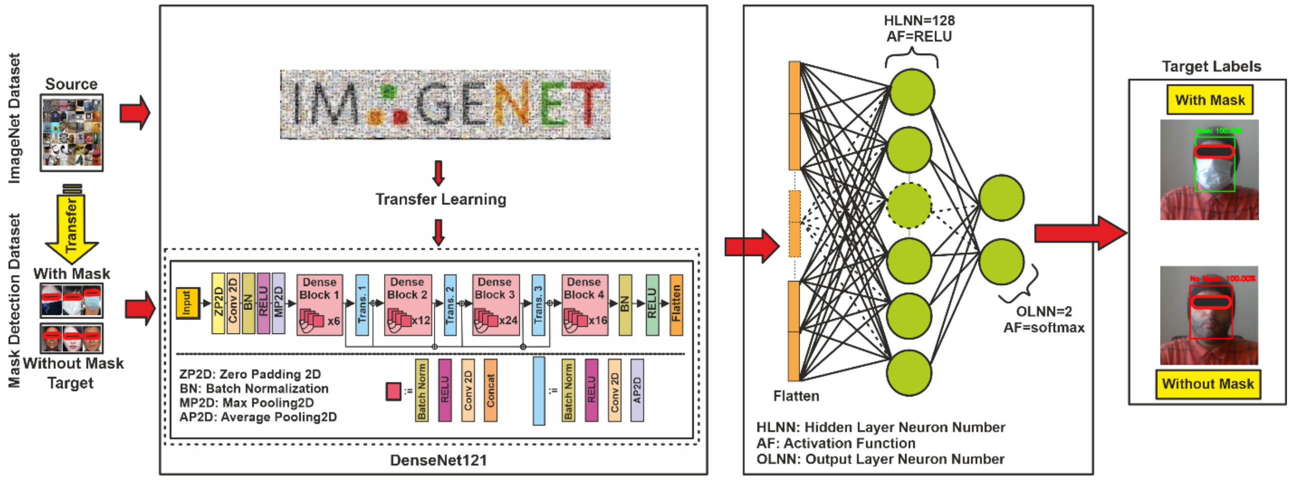
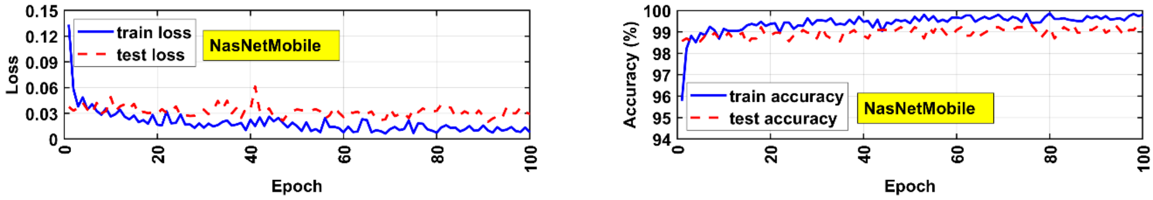
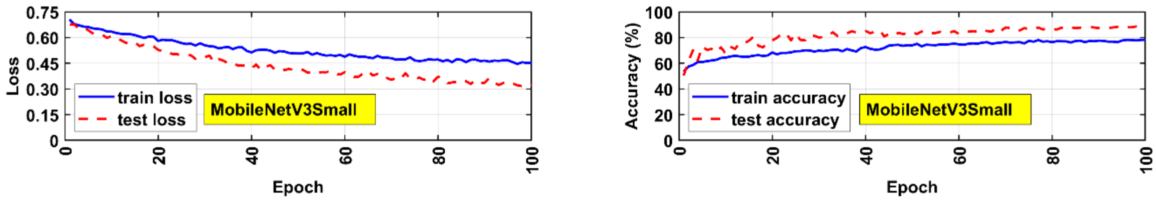


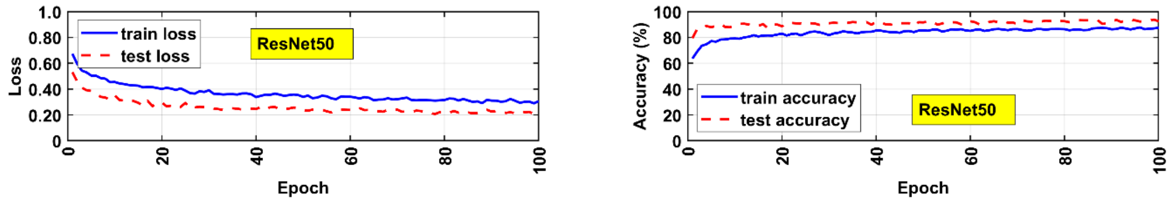
Figure 7. DenseNet121 based mask detection architechture



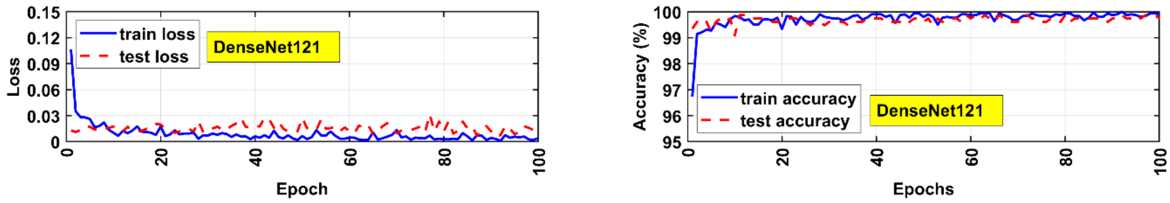
(a)



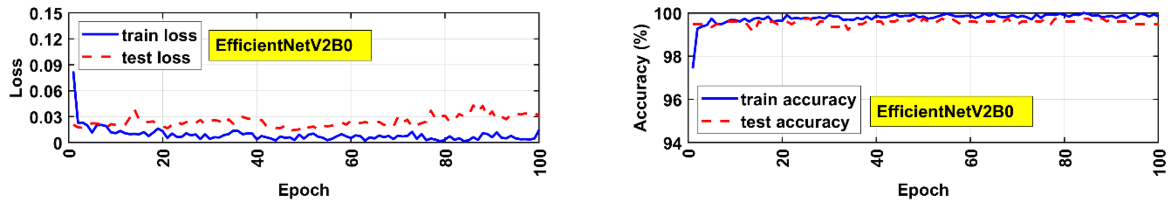
(b)



(c)



(d)



(e)

Figure 8. (a) NasNetMobile (b) MobileNetV3Small (c) ResNet50 (d) DenseNet121 (e) EfficientNetV2B0

As seen in Table 1, the performance value in the DenseNet121 model is the highest. Therefore, DenseNet121 is preferred in the real-time mask recognition basic model.

4 Conclusion

This study implements a transfer learning approach to test five pre-trained CNN models for mask detection. These models include NasNetMobile, MobileNetV3Small, ResNet50, DenseNet121, and EfficientNetV2B0. The performances of the pre-trained models have been evaluated. Obtained results show that DenseNet121 gives high performance against other specified architectures. The highest accuracy was obtained in the DenseNet121 model. The DenseNet121 model accuracy percentage value is %100. Thus, the DenseNet121 model was run on the Jetson Nano Xavier NX development board. As a result of this study, it will make an important contribution to the fight against possible future outbreaks like Covid-19, which is transmitted through droplets.

Conflicts of Interest

The author declare that there is no conflict of interest.

Similarity rate (iThenticate): %10

Reference

[1] World Health Organization, Novel coronavirus (2019-nCoV): situation report, 1, https://apps.who.int/iris/bitstream/handle/10665/330760/nCoVsitrep21Jan2020-eng.pdf?sequence=3&isAllowed=y, Accessed 20 March 2023.

[2] World Health Organization, Naming the coronavirus disease (COVID-19) and the virus that causes it, https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it, Accessed 20 March 2023.

[3] P. Bahl, C. Doolan, C. de Silva, A. A. Chughtai, L. Bourouiba and C. R. MacIntyre, Airborne or Droplet precautions for health workers treating COVID-19?. The Journal of Infectious Diseases, 225, 9, 1561-1568, 2022, https://doi.org/10.1093/infdis/jiaa189.

[4] G. Jignesh Chowdary, N. S. Punn, S. K. Sonbhadro and S. Agarwal, Face mask detection using transfer learning of InceptionV3. In: L. Bellatreche, V. Goyal, H. Fujita, A. Mondal, P. K. Reddy, Eds. Big Data Analytics. BDA 2020. Lecture Notes in Computer Science, 12581, Springer, Cham, pp. 81-90, 2020.

[5] A. Cabani, K. Hammoudi, H. Benhabiles and M. Melkemi, MaskedFace-Net -- A dataset of correctly/incorrectly masked face images in the context of COVID-19. Smart Health, 19, 1-5, 2021. https://doi.org/10.1016/j.smhl.2020.100144.

[6] X. Kong, K. Wang, S. Wang, X Wang, X. Jiang, Y. Guo, G. Shen, X. Chen and Q. Ni, Real-time mask identification for COVID-19: An edge computing-based deep learning framework. IEEE Internet of Things Journal, 8 (21), 1-10, 2021. https:// 10.1109/JIOT.2021.3051844.

[7] J. Brownlee, A Gentle introduction to transfer learning for deep learning, https://machinelearningmastery.com/transfer-learning-for-deep-learning/ Accessed 22 February 2023.

[8] JavaTPoint, Introduction to transfer learning in ML, https://www.javatpoint.com/transfer-learning-in-machine-learning. Accessed 30 February 2023.

[9] W. Wang, and J. Gang, Application of convolutional neural network in natural language processing. In 2018 International Conference on Information Systems and Computer Aided Education (ICISCAE), IEEE, pp. 64-70, Changchun, China, 2018.

[10] D. N. N. Tran, L. H. Pham, H. H. Nguyen and J. W. Jeon, A Vision-Based method for real-time traffic flow estimation on edge devices. IEEE Transactions on Intelligent Transportation Systems, 1-15, 2023. https:// 10.1109/TITS.2023.3264796.

[11] H. Avula, R. Ranjith, and A. S. Pillai, CNN based recognition of emotion and speech from gestures and facial expressions. In 2022 6th International Conference on Electronics, Communication and Aerospace Technology, IEEE, pp. 1360-1365, Coimbatore, India, 2022.

[12] T. Kalaiselvi, P. Sriramakrishnan, and K. Somasundaram, Survey of using GPU CUDA programming model in image analysis. Informatics in Medicine Unlocked, 9, 133-144, 2017. https://doi.org/10.1016/j.imu.2017.08.001.

[13] NVIDIA, Jetson Xavier NX Development Kit, Santa Clara, California, ABD, User Guide DA_09814-002, 19 May 2020.

[14] M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard, and Q. V. Le, Mnasnet: Platform-aware neural architecture search for mobile. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, (CVPR), pp. 2820-2828, Long Beach, CA, USA, 2019.

[15] A. Howard, M. Sandler, G. Chu, L. C. Chen, B. Chen, M. Tan, and H. Adam, Searching for mobilenetv3. In Proceedings of the IEEE/CVF international conference on computer vision pp. 1314-1324, Seoul, Korea (South), 2019.

[16] P. Dwivedi, Understanding and coding a ResNet in Keras, https://towardsdatascience.com/understanding-and-coding-a-resnet-in-keras-446d7ff84d33, Accessed 10 January 2023.

[17] N. Radwan, Leveraging sparse and dense features for reliable state estimation in urban environments. Ph.D. Thesis, Technische Fakultat¨ Albert-Ludwigs-Universitat Freiburg, Germany, 2019.

[18] N. H. Shabrina, R. A. Lika and S. Indarti, Deep learning models for automatic identification of plant-parasitic nematode. Artificial Intelligence in Agriculture, 7, 1-12, 2023. https://doi.org/10.1016/j.aiia.2022.12.002.

[19] F. Eryılmaz and H. A. Karacan, Akciğer X-Ray görüntülerinden COVID-19 tespitinde hafif ve geleneksel evrişimsel sinir ağ mimarilerinin karşılaştırılması. Düzce Üniversitesi Bilim ve

Teknoloji Dergisi, 9 (6), 26-39, 2021. https://doi.org/10.29130/dubited.1011829.
[20] J. Jodriguez, The evolution of Google's MobileNet architechtures to improve computer vision models, https://medium.com/dataseries/the-evolution-of-googles-mobilenet-architectures-to-improve-computer-vision-models-ffb483ffcc0a, Accessed 12 February 2023.
[21] S. Bangar, ResNet architecture expained, https://medium.com/@siddheshb008/resnet-architecture-explained-47309ea9283d, Accessed 01 January 2023.
[22] M. Shafiq and Z. Gu, Deep residual learning for image recognition: a survey. Applied Sciences, 12 (18), 8972, 1-43, 2022. https://doi.org/10.3390/app12188972.
[23] Z. Feng, An Overview of ResNet architecture and ıts variations, https://builtin.com/artificial-intelligence/resnet-architecture, Accessed 18 February 2023.
[24] S-H. Tsang, Review: DenseNet—Dense convolutional network (Image classification), https://towardsdatascience.com/review-densenet-image-classification-b6631a8ef803, Accessed 22 January 2023.
[25] B. Lodhi and J. Kang, Multipath-DenseNet: A Supervised ensemble architecture of densely connected convolutional networks. Information Sciences, 482, 63-72. 2019. https://doi.org/10.1016/j.ins.2019.01.012.
[26] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 4700-4708, Honolulu, HI, USA, 2017.
[27] M. Tan and Q. Le, Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning, pp. 6105-6114, PMLR, 2019. https://doi.org/10.48550/arXiv.1905.11946.
[28] H. Alhichri, A. S. Alswayed, Y. Bazi, N. Ammour and N. A. Alajlan, Classification of remote sensing images using EfficientNet-B3 CNN model with attention. IEEE Access, 9, 14078-14094, 2021. 10.1109/ACCESS.2021.3051085.
[29] D. Misra, Mish: A self regularized non-monotonic activation function. arXiv preprint arXiv:1908.08681, 2020, https://doi.org/10.48550/arXiv.1908.08681.
[30] S. Balaji, Face-Mask-Detection, 2021, https://github.com/balajisrinivas/Face-Mask-Detection/tree/master /dataset/ with_mask, Accessed 14 January 2023.