ISSN: 2347-6540

International Journal of Latest Research In Engineering and Computing (IJLREC)
Volume 12, Issue 1: Page No. 45-48
www.ijlrec.com

EFFICIENT FACE MASK DETECTION ON EDGE DEVICES USING MOBILENetV2

Rajkumar Verma[1], Kuldeep Chauhan[2], Varun Bansal[3]
Research Scholar, Department of Computer Science & Engineering, Shobhit University, Gangoh,
raj.verma513@gmail.com[1]
Assistant Professor, Department of Computer Science & Engineering, Shobhit University, Gangoh,
kuldeep.chauhan@shobhituniversity.ac.in[2]
Associate Professor, Department of Computer Science & Engineering Shobhit University, Gangoh,
varun.bansal@shobhituniversity.ac.in[3]

Abstract :- The COVID-19 pandemic has created a need for automatic face mask detection to ensure public safety. This study examines the identification of individuals wearing face masks in practical settings utilizing MobileNetV2, a lightweight Convolutional Neural Network (CNN) model. The model achieves a 99% accuracy rate in distinguishing between masked and unmasked faces, utilizing datasets like RMFD and Kaggle. The research supports MobileNetV2's efficiency in processing and deployment, making it an ideal choice for embedded systems and smart surveillance applications.

Keywords: MobileNetV2, Face Mask Detection, CNN, Deep Learning, COVID-19, Image Segmentation, Real-Time Surveillance.

I. INTRODUCTION

The global COVID-19 pandemic has highlighted the importance of face masks in preventing the airborne transmission of viruses. However, monitoring compliance is often inconsistent and requires significant labor. Using deep learning for real-time surveillance of face mask detection proves to be an efficient and scalable solution. This study focuses on MobileNetV2, an optimized convolutional neural network (CNN) model that is particularly suitable for edge devices due to its minimal computing requirements. In crowded areas, automated face mask detection systems can be extremely beneficial for maintaining public health standards. These systems not only enhance scalability in places such as hospitals, schools, train stations, and airports but also reduce the need for human oversight. By integrating deep learning with computer vision, these systems can operate continuously at a low cost.

II. LITERATURE REVIEW

A. For face mask detection, a variety of methods have been put forth. These consist of ensemble learning [5], ResNet [4], Faster R-CNN [3], and YOLO [2]. MobileNetV2, on the other hand, balances computational efficiency with accuracy.

Illustration, Nagrath et al.(6) introduced SSD- MobileNetV2, which achieved an delicacy of 93. A multi-model armature was developed by Sethi et al.(5), performing in an emotional delicacy of 98.2. Loey et al.(7) attained 81 perfection using YOLOv2 combined with ResNet-50. By exercising YOLOv4, Abbas et al.(8) created a new dataset and reached an delicacy of 99.5. Xu et al. [10] and Benini et al. [9] investigated semantic segmentation techniques. For precise classification, Rusli et al. [11] used the LeNet method. In light of this, MobileNetV2's lightweight architecture makes it an excellent choice for real-time applications [12].

Apart from the above, research has also considered hybrid approaches that integrate deep learning and conventional image processing techniques. These hybrid approaches aim to enhance detection quality in hard-to-reach scenarios such as low light, partial occlusion, or when there is headgear and sunglasses. Furthermore, databases such as as FMD and MAFA are being utilized for enhanced generalization over various types of masks and populations

III. METHODOLOGY

Data Set

The first stage in developing a Face Mask Recognition model is the data collection process. The collection contains training data for both mask users and non-users. The model will distinguish between individuals who are wearing masks and those who are not.

The 4095 photos in this collection are divided into two classes:

* ith_mask: 2165 images
* without_mask: 1930 images

The following sources were used to gather the pictures:

* **Kaggle datasets** - Kaggle is an online platform for data scientists and machine learning enthusiasts to connect. In addition to using notebooks with GPU integration and working collaboratively, Kaggle users can access and share datasets and compete with other data scientists to solve data science issues.
* **RMFD dataset** - The Real-World Masked Face dataset is a large dataset for masked face detection. Dataset (RMFD).

- **Data Pre-processing**

There is a pre-processing step before the data are trained and tested. Four processes make up the pre-processing: reducing the image's size, converting it to an array, pre-processing the input with MobileNetV2, and then hot encoding the labels.

In computer vision, scaling the image is an essential pre-processing step since training models perform so well. A smaller image will work better for the model. A 224 x 224 pixel image has been scaled for the study's purposes.

In the next step, the dataset's images will be transformed into an array. Before calling the loop method, the image is converted to an array. The input will then be pre-processed using the image using MobileNetV2.

Since many machine learning algorithms cannot operate directly on data labelling, the last step in this phase is hot encoding labels. They require that all input and output variables, as well as this method, be of the numerical variety. In order for the algorithm to understand and use the tagged data, it will be transformed into a numerical label.

- Split the Data
- Building the Model
- Testing the Model

Architecture

MobileNetV2 utilizes depthwise separable convolutions and inverted residuals to maintain accuracy while reducing complexity [1].

Training

Framework: TensorFlow + Keras
Optimizer: Adam
Epochs: 20
Loss Function: Binary Cross-Entropy

Implementation Environment

Google Colab with GPU acceleration was used for the implementation. Keras was used to train the model with a TensorFlow backend. We evaluated the model on a Raspberry Pi 4 with a camera module to assess deployment practicality, and we were able to detect in real time at 10 frames per second.

IV. Result and Environment

- **Metrics**

Epoch	Accuracy	Validation Accuracy
1	85.04%	97.19%
20	99.04%	98.78%

- **Classification Report**

Class	**Precision**	**Reca ll**	**F1-Score**
With Mask	0.99	0.99	0.99
Class	**Precision**	**Reca ll**	**F1-Score**
Without Mask	0.99	0.99	0.99
Accuracy	—	—	**99%**

MobileNetV2 performs well under variable lighting, head poses, and mask types. Detection accuracy decreases if key landmarks are occluded.
Confusion matrix analysis was also used to comprehend instances of misclassification. With only slight uncertainty in photos where the face was partially occluded or turned, the model demonstrated great specificity and sensitivity. The classifier's robustness was further validated via ROC curve analysis.

V. Discussion

The proposed system can be deployed in:
- Surveillance cameras
- IoT-based smart doors
- Public transport systems
- Educational institutions

Deployment on low-power devices such as the NVIDIA Jetson Nano or Raspberry Pi is made possible by its architecture.

In addition to promoting public health safety, this study shows how feasible it is to implement cutting-edge deep learning methods in settings with limited resources. The method is adaptable and can be used for additional safety compliance duties like hand sanitiser use or helmet detection.

VI. Conclusion

The MobileNetV2-based model is computationally efficient and achieves state-of-the-art performance in face mask identification. For real-time embedded applications, it is ideal. By providing a deployable solution, this work closes the gap between scholarly research and practical implementation. MobileNetV2's lightweight design allows it to be used in a variety of embedded devices without sacrificing functionality.

VII. Future Scope

Future improvements can include:
- Detecting improperly worn masks
- Classifying different mask types
- Integrating with thermal screening
- Adding social distancing detection

In order to facilitate privacy-preserving training across dispersed camera networks, future versions of this model might integrate federated learning approaches. Additionally, authorities can detect non-compliance zones in real time by combining the technology with crowd analytics.

VIII. References

[1] Sandler, M., et al. (2018). 'MobileNetV2: Inverted Residuals and Linear Bottlenecks,' CVPR.
[2] Redmon, J., et al. (2016). 'You Only Look Once: Unified, Real-Time Object Detection,' CVPR.
[3] Ren, S., et al. (2017). 'Faster R-CNN: Towards Real-Time Object Detection,' IEEE TPAMI.
[4] He, K., et al. (2016). 'Deep Residual Learning for Image Recognition,' CVPR.
[5] Sethi, S., et al. (2021). 'Face Mask Detection Using Deep Learning,' J. Biomedical Informatics.
[6] Nagrath, P., et al. (2021). 'SSDMNV2: Real-time DNN-based Face Mask Detection,' Sustainable Cities and Society.
[7] Loey, M., et al. (2021). 'Hybrid Deep Learning for Medical Face Mask Detection,' Measurement.
[8] Abbasi, S., et al. (2021). 'Face-Mask Detection with YOLO for a New Dataset,' IEEE CSICC.
[9] Benini, S., et al. (2019). 'Face Analysis through Semantic Face Segmentation,' Signal Processing: Image Communication.

[10] Xu, M., et al. (2020). 'Mask Wearing Detection Using SSD-Mask,' IEEE ICCSMT.
[11] Rusli, M. H., et al. (2021). 'Evaluating Masked and Unmasked Faces with LeNet,' IEEE CSPA.
[12] Mbunge, E., et al. (2021). 'Application of Deep Learning for COVID-19 Face Masks,' Sustainable Operations and Computers.
[13] Kumar, A., et al. (2021). 'Scaling up Face Masks Detection with YOLO,' Optik.
[14] Wang, K., et al. (2016). 'Using Fast R-CNN for Face Detection,' IEEE VCIP.
[15] Singh, S., et al. (2021). 'YOLOv3 and Faster R-CNN Models for COVID-19 Mask Detection,' Multimedia Tools and Applications.