Real Time Face Mask Detection

Benson Baby Thomas
SID: 2190563

Ajin Joseph
SID: 2238709

Ansila Thalathil Poolakkamannil
SID: 2238707

Abstract: In this Project, we use deep learning models to propose a real-time face mask detection system. On a dataset of 23,345 images of faces with and without masks, we trained three distinct models, named MobilenetV2, Inceptionv3, and DenseNet121. The task of detecting faces was accomplished using a pre-trained SSD model. Images from many different sources and the well-known MaskedFace-Net dataset were combined to create a more diverse data set. For testing, we used two distinct unseen datasets, one with 10,000 images and the other with 5733 images. Through the use of both individual and ensemble modelling methodologies, we assessed the performance of the models. The models' mean accuracy varied from 0.9348 to 0.9878. The best three models based on our evaluation were DenseNet121 + Inceptionv3, MobilenetV2 + DenseNet121, and DenseNet121 respectively. Considering the real-time application, we suggest using Densenet121 as the best model, even though it is the third-best in terms of accuracy. We found that ensemble modelling techniques only provided a slight improvement in accuracy. Overall, our system can detect face masks in real-time and can be used in various settings to promote public health and safety.

Keywords-Deep Learning, MobileNetV2, InceptionV3, DenseNet121, SSD, Ensembling.

I.INTRODUCTION

The COVID-19 pandemic has forced the world to change its way of living, and one of the most important measures to prevent the spread of the virus is to wear a face mask. With the requirement for face masks in public settings, it is crucial to have a system that can tell whether or not someone is wearing one. This technique can be helpful in a variety of environments, including hospitals, public transit, and other open spaces where maintaining social distance is challenging. In this project, we want to create a real-time face mask detection system that can determine whether or not someone is wearing one

Our research questions include: Which deep learning model is the best for real time face mask identification in terms of accuracy and efficiency? Can ensemble modelling techniques improve the accuracy of the face mask detection system? To answer these questions, we will train and test three different deep learning models (Mobilenetv2, Inceptionv3, and Densenet) on a dataset of images with and without masks.

The main objectives of this project is to develop and evaluate a real-time face mask detection system using deep learning models. First, we will build a dataset of images with and without masks to train and validate the deep learning models. Second, we will develop three different models (MobilenetV2, InceptionV3, and DenseNet121) to detect whether a person is wearing a face mask or not. Third, we will evaluate the performance of each model using two different unseen datasets. Fourth, we will compare the performance of each model and determine the best-performing model for real-time face mask detection. Fifth, we will explore the use of ensemble modelling techniques to improve the accuracy of the face mask detection system. Ultimately, the goal of this project is to develop a real-time face mask detection system that can accurately and efficiently detect whether an individual is wearing a face mask, which can be useful in various settings where social distancing is difficult to maintain.

II. Literature Review.

Real-time face mask detection has been a significant topic of research. Different machine learning and deep learning-based algorithms to determine whether a person is wearing a face mask or not have been proposed by a number of researchers. Based on the articles offered, we will explore some of the most current efforts in this field in this literature review.

A real-time face mask identification system utilising Single Shot identification (SSD) and MobileNetV2 architecture was proposed by K. Balaji and S. Gowri[1]. The suggested approach recognised face masks in real-time video streams and shown high accuracy on the validation dataset. They used transfer learning to fine-tune the MobileNetV2 architecture on their dataset.

[3] G. Lakshmi Durga et al. also proposed a face mask detection system using MobileNetV2 architecture. The proposed system can be integrated with surveillance cameras to monitor compliance with mask-wearing guidelines. Using the MobileNetV2 and InceptionV3 models, N. [2] A. Nainan et al. suggested a real-time face mask identification system. The authors obtained great accuracy by fine-tuning both models using transfer learning on their dataset. The suggested technology could recognise face masks in real-time video feeds and could be linked with CCTV cameras to track compliance with mask wearing. A face mask identification system employing machine learning approaches was proposed by A. Velip and A. Dessai [6]. Face masks were discovered in real-time video streams using SVM and KNN classifiers, according to the authors. The proposed system may be applied in low-resource environments where deep learning-based methods might not be practical.

[8] Santoso, A. J., & Saragih, R. E. (2022) present a system for automatic face mask detection using deep learning models. They train MobileNetV2 and DenseNet121 models to analyse images of human faces and check whether or not the person is wearing a face mask. The authors also evaluate the system's performance and discuss the training process of the models. This article is of interest to researchers and practitioners in computer vision, deep learning, and public health during the COVID-19 pandemic. Overall, the article provides valuable insights into the development of an automatic face mask detection system using deep learning models.

III. Technical Aspects.

a) Dataset and Task

The data used for this project consists of approximately 60% image data from the MaskedFace-Net[7]dataset and the remaining data obtained from various internet sources to ensure diversity in the dataset. The data was only used for training and validation testing, with an 80:20 split ratio. To evaluate the performance of the models created, a separate unseen dataset consisting of 5,733 and 10,000 images was used for testing. Pre-processing of the data involved resizing all images to a fixed size of 224 x 224 pixels using bilinear interpolation. Additionally, data augmentation techniques such as random rotation, horizontal and vertical flipping were used to increase the size of the dataset and improve the model's ability to generalize to new data. In our mixed dataset, we collected 11,756 images of faces with masks and 11,589 images without masks. The images were captured from various angles, under different lighting conditions, and included individuals from different ethnicities (Figure 1). Additionally, the without mask dataset included images of individuals wearing masks incorrectly. (Figure 2) Furthermore, we included a few images of faces with illustrated masks in both classes image. A sample dataset.

Figure 1: Mixed Dataset with masks

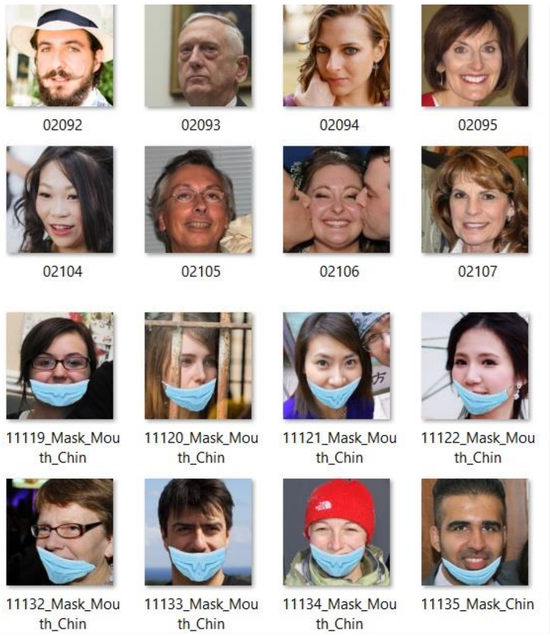

Figure 2: Mixed Dataset without masks

The primary objective of this project is binary image classification, which involves categorizing images into two classes: masked and unmasked. To evaluate the performance of the model, metrics such as precision and recall were used to measure aspects such as the percentage of accurate positive predictions and the percentage of correctly identified actual positive instances. The overall aim of the project was to develop a model that can efficiently and accurately classify masked and unmasked faces.

b) Methodology.

In this project, three popular pre-trained deep learning models have been used for face mask classification: DenseNet121, MobileNetV2 and InceptionV3. Additionally, we have also employed a pre-trained face detection model SSD (Single shot detector).

DenseNet121.

DenseNet121 is a CNN architecture that uses dense connections between layers to improve the flow of information and reduce the number of parameters in the model. In this project, we used DenseNet121 as one of the base models for our image classification task because of its high accuracy and ability to capture complex features in images. The input shape for this model is (224, 224, 3), which is the size of the input images. We set the "include top" parameter to "False" to exclude the fully connected layer at the end of the model. We added our custom layers on top of the base model for classification. The custom layers include a Flatten layer to flatten the output from the base model, a dense layer with 128 units and ReLU activation function, a Dropout layer to prevent overfitting, and a dense layer with 2 units and softmax activation function for the output. We used the Categorical Crossentropy loss function and Adam optimizer for compiling the model.

MobileNetV2.

MobileNetV2 is a CNN architecture designed for efficient mobile devices. It has shown to provide excellent accuracy for image classification tasks while keeping the model size small. Custom layers are added in the same manner as with DenseNet121 for image classification

InceptionV3.

InceptionV3 is a convolutional neural network architecture with deep convolutional layers, which has been shown to be effective for image classification tasks. This model is designed to be both computationally efficient and accurate. This model is fine-tuned same as DenseNet121 for face mask classification.

Single Shot Detector (SSD).

Single Shot Detector (SSD) is a popular object detection algorithm that uses a single deep neural network to predict object classes and bounding boxes. SSD is designed to be fast and accurate, making it suitable for real-time object detection on mobile devices and other embedded systems. In this project, we used SSD to perform face detection in images and video feed for mask detection in real time. The Single shot detector(SSD) model is trained on a large dataset of labelled images, and was able to accurately detect faces in our test images with high precision and recall. The face detection results were used to pre-process the images before feeding them into our image classification models.

Ensemble Modelling.

In addition to these three pre-trained deep learning models, we also attempted an ensemble modelling approach by combining the predictions of two models using the averaging technique to further improve the classification accuracy.

IV. Experiments.

a) Experiment 1

In Experiment 1, we tested our three trained models on two large and diverse unseen datasets, and the results are shown in the below table.

Test Data 1(5773 images belonging to 2 classes)

Model	F1 Score	Precision	Recall	Accuracy
MobileNetV2	0.9797	0.9928	0.9670	0.9800
InceptionV3	0.9847	0.9943	0.9753	0.9849
DenseNet121	0.9933	0.9978	0.988	0.9934

Test Data 2(10000 images belonging to 2 classes)

Model	F1 Score	Precision	Recall	Accuracy
MobileNetV2	0.8932	0.8643	0.9242	0.8896
InceptionV3	0.9698	0.9437	0.9974	0.969
DenseNet121	0.9758	0.9588	0.9934	0.9754

The results presented in the table demonstrate that all three models perform well on the two diverse and unseen datasets. However, the DenseNet121 model stands out as the top-performing model with highest accuracy.

b) Experiment 2

In Experiment 2, we are testing our trained ensemble models on two large and diverse unseen datasets, and the observations are shown in the below table.

Test Data 1(5773 images belonging to 2 classes)

Model	F1 Score	Precision	Recall	Accuracy
MobileNetV2 + InceptionV3	0.9866	0.9978	0.97575	0.9868
MobileNetV2 + DenseNet121	0.9942	0.9989	0.9895	0.9942
DenseNet121 + InceptionV3	0.9921	0.9982	0.9861	0.9922

Test Data 2(10000 images belonging to 2 classes)

Model	F1 Score	Precision	Recall	Accuracy
MobileNetV2 + InceptionV3	0.9769	0.9579	0.9966	0.9764
MobileNetV2 + DenseNet121	0.9775	0.9622	0.9934	0.9772
DenseNet121 + InceptionV3	0.9837	0.9709	0.9969	0.9835

The results shows all three Ensemble models perform well on the two diverse and unseen datasets.

DenseNet121+InceptionV3 model stands out as the top-performing model with the highest accuracy and F1 score.

V. CONCLUSION AND FUTURE WORK.

The DenseNet121 model was found to be the best option for real-time face mask detection, based on its performance in terms of accuracy and speed. Ensemble modeling techniques were also explored, but their effectiveness was limited. To improve the model's accuracy, we suggest using the confusion matrix to identify false positives and patterns that may be causing the model to misclassify images. By doing so, we can enhance the model's accuracy and reduce the number of false positives in future predictions. In future studies, further investigation into other deep learning models and techniques may be conducted to improve the system's accuracy and efficiency even further.

VI. REFERENCE.

1. Balaji, K. and Gowri, S., 2021. A Real-Time Face Mask Detection Using SSD and MobileNetV2. In 2021 4th International Conference on Computing and Communications Technologies (ICCCT) (pp. 1-5). IEEE.
2. Nainan, N.A., J, J., Jalan, R., N, R.S., C, K.V. and Shankar, S.P., 2021. Real Time Face Mask Detection Using MobileNetV2 and InceptionV3 Models. In 2021 IEEE Mysore Sub Section International Conference (pp. 1-6). IEEE.
3. Durga, G.L., Potluri, H., Vinnakota, A., Prativada, N.P. and Yelavarti, K., 2022. Face Mask Detection using MobileNetV2. In 2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS) (pp. 123-127). IEEE.
4. Rokhana, R., Herulambang, W. and Indraswari, R., 2021. Multi-Class Image Classification Based on MobileNetV2 for Detecting the Proper Use of Face Mask. In 2021 International Electronics Symposium (IES) (pp. 24-27). IEEE.
5. Chowdary, P. V., Gutta, S. S., Punn, N. S., Sonbhadra, S. K., & Agarwal, S. (2020). Face Mask Detection Using Transfer Learning of InceptionV3. In Advances in Intelligent Systems and Computing (pp. 63-73). Springer.
6. Velip, A., & Dessai, A. (2022). Face Mask Detection Using Machine Learning Techniques. In 2022 2nd Asian Conference on Innovation in Technology (ASIANCON) (pp. 1-6). IEEE.
7. Cabani, A., Hammoudi, K., Benhabiles, H., & Melkemi, M. (2020). MaskedFace-Net: A dataset of correctly/incorrectly masked face images in the context of COVID-19. Smart Health, 19, 100144.
8. Santoso, A. J., & Saragih, R. E. (2022). Automatic face mask detection based on MobileNet V2 and DenseNet 121 models. Journal of Physics: Conference Series, 1993(1), 012019.
9. Tang, J., Peng, X., Chen, X., & Luo, B. (2021). An Improved Mobilenet-SSD Approach For Face Detection. In 2021 40th Chinese Control Conference (CCC) (pp. 14160-14165). IEEE.