PAPER DETAILS

Authors   :

1) Sovon Chakroborty,   4) FM Javed Mehedi Shamrat,

2) Md.Al Jubair,   5) Md. Masum Billah,

3) Md. Alauddin,   6) Rumesh Ranjan

WHY THEY HAVE CONDUCTED THIS RESEARCH?

The researchers in this paper developed a mask detection model using deep learning methods to prevent the spread of COVID-19. The model was trained on a dataset containing masked and unmasked images obtained from a variety of sources, including images of the coauthors captured using webcams and cell phone cameras. Goal of this research is to build a model which can accurately detect whether a person is wearing a mask, which can be used in real-world applications such as public places, workplaces or hospitals.

PROPOSED SYSTEM

In this research, they proposed a system using Convolutional Neural Networks (CNNs) for detection face musk. The process involves data collection, data preprocessing and augmentation, training a CNN model using different architectures such as (Convolutional layer with Max pooling, Convolutional layer with Average pooling, and MobileNetV2), detection output, and performance evaluation.

Firstly, in the data collection stage, they gather a large dataset with total 1845 images. The images can be obtained from various sources.

Next, in the data preprocessing and augmentation stage, they clean and preprocess the data to be used for training the model. This may involve tasks such as resizing images, removing irrelevant information, and normalizing the data. They normalized images 256 x 256 pixels. Then resized images 128 x128 to pass the second layer of Convolutional2D and again 64 x 64 to pass the third Convolutional2D layer.

After preprocessing the data, they train a CNN model using different architectures such as Convolutional layer with Max Pooling, Convolutional layer with Average Pooling, and MobileNetV2. These architectures are designed to handle different types of images and tasks.

After the classification stage, the CNN model outputs a binary decision for each image as to whether a mask is present or not.

Finally, evaluate the performance of the mask detection system using various metrics such as precision, recall, F1- score, and accuracy. The evaluation can be done on a separate validation dataset or through cross-validation on the original dataset.

ARCHITECTURE

In this research they used to architecture one is Convolutional Neural Network (CNN) architecture and second is MobileNetV2 architecture. In convolutional layer architecture they used Max pooling operation with three convolution layers, as well as the average pooling operation with three convolution layers.

Firstly, Max Pooling Operation. In the proposed system, the input image is first passed through three convolutional layers with max pooling, followed by a flatten operation, a fully connected layer, and an output layer with softmax activation function.

In the first convolutional layer, the input image with a size of 256 x 256 is convolved with a set of learnable filters, and applied the RELU function to the resulting feature maps. Then, the max pooling process is performed to downsample the feature map.

In the second convolutional layer, input images resized to 128 x128, the first convolutional layer are convolved again with a different set of filters. At the same way they applied third convolutional layer. Resized images 64 x64 to pass the third Convolutional2D layer.

After the three convolutional layers with max pooling, the flatten operation is applied to transform the 3D feature maps into a 1D vector, which is then fed into the fully connected layer. The fully connected layer is used to perform classification based on the extracted features, and the output layer with softmax function used to produce the final classification results.

For function mapping again they used same architecture with the process of average pooling.

Also they used MobileNetV2 architecture for face musk detection. MobileNetV2 is a popular convolutional neural network architecture for image classification tasks. The architecture is based on depth wise separable convolutions, which allows for high accuracy with lower computational cost.

EXPERIMENTED RESULT

For detecting face mask they used a dataset of 1845 images. The proposed model achieved training accuracy 96.49% and validation accuracy 98.49% using Max pooling operation. The proposed model achieved training accuracy 95.19% and validation accuracy 96.23% using Average pooling operation. Using MobileNetV2 architecture achieved training accuracy 99.72% and validation accuracy 99.82% that is the highest accuracy.

After applying Max pooling in 1st Epoch **Training Loss-** 42.13%, **Training Accuracy-** 89.76%, **Validation Loss-** 12.32%, **Validation Accuracy-** 90.73%, in 13th Epoch **Training Loss-** 6.56%, **Training Accuracy-** 95.65%, **Validation Loss-** 6.01%, **Validation Accuracy-** 97.71%. From 1st epoch to 13th epoch there is a lot of change in results which is very excellent.

After applying Average pooling in 1st Epoch **Training Loss-** 43.54%, **Training Accuracy-** 88.92%, **Validation Loss-** 13.32%, **Validation Accuracy-** 89.95%, in 13th Epoch **Training Loss-** 5.92%, **Training Accuracy-** 95.19%, **Validation Loss-** 5.12%, **Validation Accuracy-** 96.23%.

After applying MobileNetV2 in 1st Epoch **Training Loss-** 4.43%, **Training Accuracy-** 98.67%, **Validation Loss-** 4.21%, **Validation Accuracy-** 98.71%, in 13th Epoch **Training Loss-** 3.42%, **Training Accuracy-** 99.72%, **Validation Loss-** 3.18%, **Validation Accuracy-** 99.82% After applying MobileNetV2 they achieved the highest accuracy.

In this proposed model, for training dataset best precision 99.72% and for validation dataset best precision 99.82%