Comparative Analysis of Machine Learning Algorithms for Face Mask Detection and Alerting System.

Snehit Vaddi[1], Vishnu Vipul Maddi[2]

[1]*Student, Department of Computer Science and Engineering, GITAM Deemed to be University, Visakhapatnam, Andhra Pradesh, India.*
[2]*Student, Department of Computer Science and Engineering, GITAM Deemed to be University, Visakhapatnam, Andhra Pradesh, India.*

---***---

Abstract - *According to the World Health Organization, the coronavirus COVID-19 pandemic is producing a worldwide health catastrophe, and the most effective protective strategy is wearing a face mask in public places (WHO). According to reports, wearing a facemask in general settings minimizes the chance of transmission. This paper describe and construct an ensemble of Deep Learning and Machine Learning algorithms to recognize persons without face masks and send an email alert when the count threshold crosses. The facemask dataset consists of images of people with and without masks. Our objective is to take advantage of existing cameras and detect persons wearing masks and avoid virus transmission.*

***Key Words*: DeepLearning, SVM, RandomForest, CNN, InceptionResNetV2, Caffe, OpenCV, Keras.**

1. INTRODUCTION

The COVID-19 coronavirus outbreak all over the world has mandated the use of face masks. Scientists demonstrated that wearing facemasks reduces COVID-19 transmission. COVID-19's fast spread has compelled the World Health Organization to designate it as a worldwide pandemic in 2020. The rapid development of the COVID-19 pandemic has resulted in the world's most serious public health disaster. There is no indication of a clinically approved antiviral treatment or immunization against the virus until early 2021. People were forced by laws to wear face masks in public in many countries [18]. However, the process of monitoring large groups of people is becoming more complex in public places. So, we create an automation process for detecting the faces.

The coronavirus outbreak has led to substantial levels of international scientific cooperation. Artificial Intelligence (A.I.) based on Machine Learning and Deep Learning can aid in the fight against Covid-19 in various ways. Machine learning enables researchers and physicians to analyze massive amounts of data to estimate the dissemination of COVID-19, act as an early warning system for possible pandemics, and categorize susceptible groups. We can tackle and predict new diseases with the help of new technologies such as artificial intelligence, IoT, Big data, and Machine learning.

Here we introduce a facemask detection model based on computer vision and deep learning. The proposed model integrates with Surveillance Cameras to impede the COVID-19 transmission by detecting people wearing masks, no face masks. The model is an integration between deep learning and classical machine learning. We have implemented multiple Machine Learning and Deep Learning algorithms to achieve the highest accuracy and consume the least time in training and detection.

2. LITERATURE REVIEW

2.1 "A Face-Mask Detection Approach based on YOLO."

An object detection model is used in the first technique to discover and categorize masked and unmasked faces. In the second technique, a YOLO face detector detects faces (whether masked or not), and then a CNN architecture classifies the faces into masked and unmasked categories. We used the Caffe model, a lightweight and computationally efficient model that makes working in real-time more manageable. For optimum performance at high speeds, we examined five different models. Our proposed approach solves the problem of camera specification.

2.2 "Rapid object detection using a boosted cascade of simple features"[3]

Viola-Jones is an early face detection algorithm. It was designed to identify frontal faces, which it does better than faces looking laterally, above, or downwards. This algorithm recognizes and locates human faces regardless of their size, location, or surroundings. Viola-Jones fails when the face is covered, which makes it not a suitable approach. Our proposed system uses a Caffe model capable of detecting faces in distorted angles, low brightness, and masks.

2.3 "Face Mask Detection using MTCNN and MobileNetV2"[13]

This solution consists of MobileNet and MTCNN as the backbone. The system was used for high and low computation scenarios. MobileNet is a lightweight CNN model built with the aim of modularity and efficiency. Though MTCNN is robust and efficient, it is slow in the detection process and complex in methodology. The proposed solution can overcome this problem and is easy to implement using the OpenCV DNN module.

2.4 "*FACE MASK DETECTION.*"

The proposed system is based on a deep learning algorithm for image segmentation [19]. The technique of isolating the item of interest from the rest of the actual material using input video processing is called segmentation. Face landmarks are used to identify segmentation. Detecting, mapping, and analyzing landmarks necessitates a high level of computer capacity. The current proposed solution would bypass the need for face-landmark detection and additional steps.

3. METHODOLOGY

The proposed system aims to find human faces using a pre-trained Caffe Face detection model and then classify if the face has a mask[17] or not using a machine learning model. Caffe model is capable of localizing face even with slight distortion in face angle. We compare multiple Deep Learning and Machine learning models for picking the best performing Face mask detection model. The Neural Network architectures implemented using Keras and Scikit-learn modules in Python. The real-time implementation made possible using OpenCV.

Approach

We follow a two-phased approach.

A. Training Phase: Training Dataset on multiple Machine learning and Deep Learning models, do comparative analysis on various performance metrics.

B. Deployment Phase: Apply Face detection, face mask classification on the real-time video feed from the webcam. Send an email alert to the admin when the "without a mask" count crosses a threshold.

A. *Training Phase*

The dataset is adopted from an open-sourced repository by Prajnasb. The dataset is categorized into two folders, namely "with_mask" and "without_mask." Each folder consists of people's faces with or without masks respective to their category. The images in the dataset underwent preprocessing steps of noise reduction and image resizing to maintain consistency and robustness. Data augmentation was done to make the model robust and avoid overfitting.

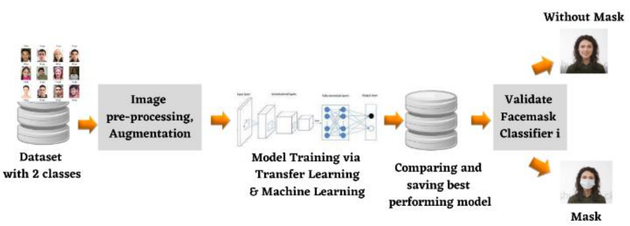

Fig 1: Training of the dataset with Conventional ML and CNN Algorithms.

To implement a best-performing model, we utilize transfer learning to adapt weights from a similar task face detection trained on an extensive dataset. Comparison among four conventional Machine learning models: SVM, Random Forest, Decision Tree, KNN (K Nearest Neighbors), and two popular DNN models (InceptionV3, InceptionResNetV2).

	Precision		Recall		F1 Score		Accuracy
	Mask	Without Mask	Mask	Without Mask	Mask	Without Mask	
SVM	0.66	0.70	0.73	0.61	0.69	0.65	67%
Random Forest	0.91	0.93	0.93	0.91	0.92	0.92	92%
Decision Tree	0.87	0.91	0.91	0.87	0.89	0.89	89%
KNN	0.88	0.75	0.70	0.91	0.78	0.82	80%
InceptionV3	0.78	0.96	0.97	0.73	0.86	0.83	85%
Inception-ResNet-v2	0.95	0.81	0.77	0.96	0.85	0.88	87%

Fig 2: Comparison table of Machine Learning & Deep Learning Algorithms for Facemask classification.

In comparison, Random Forest has mis-proven the conventional thought of Neural Network performance over traditional machine learning models. Out of all the algorithms, SVM has shown mere accuracy. Support Vector Machine, or SVM, is a well-known Supervised Learning technique used for both classification and regression issues. However, it is primarily utilized in Machine Learning for Classification difficulties.

The SVM algorithm's purpose is to find the optimum line or decision boundary for categorizing n-dimensional space to place new data points in the proper category in the future. A hyperplane is the optimal choice boundary. Out of the other three ML algorithms, SVM has poor accuracy of 67%. Random Forest is a classifier that combines several decision trees on different subsets of a dataset and averages the results to increase the dataset's predicted accuracy. Instead, depending on a single decision tree, the random forest collects the forecasts from each tree and indicates the output with the highest votes of predictions. Random Forest has proven to be the best performing model compared to other ML and Neural Network algorithms with an accuracy of 92%.

Transfer Learning in layman terns is a methodology that uses previously acquired knowledge in new problems. It is a widely popular technique in Deep Learning. Transfer Learning involves taking the pre-trained weights in the first layers, which are generally universal to many datasets, and randomly initializing and training the remaining layers for classification purposes is a standard transfer learning method. On the ImageNet dataset, Inception v3 is a commonly used image recognition model demonstrated to achieve higher than 78.1 percent accuracy. The Inception-ResNet-v2 model is a variant of the Inception V3 model, and it is far more accurate than the preceding Inception V3. The Inception-ResNet-v2 architecture outperforms earlier state-of-the-art models in terms of accuracy.

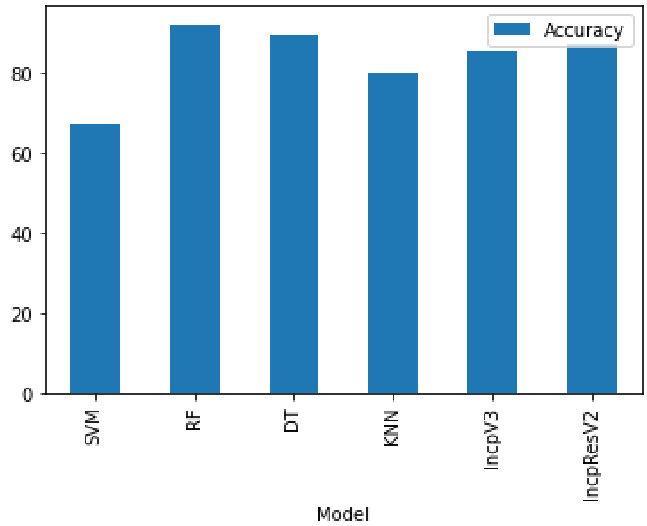

Fig 3: Bar plot of Model Vs. Accuracy.

B. *Deployment Phase*

The procedure for determining whether or not the individual on the camera is wearing a mask is a two-step process.

1. Recognize the faces in the webcams frame.

2. Categorize the faces into any of the two classes (With and Without masks).

The Caffe model is used to detect faces from the input frame. Caffe model employs the ResNet-10 architecture and is based on the Single Shot-Multibox Detector (SSD). It was included in OpenCV's deep neural network module after version OpenCV 3.3. Although a quantized Tensorflow version is available, we utilize the Caffe Model.

You'll need two sets of files to use OpenCV's deep neural network module with Caffe models:

- The model architecture is defined by the .prototxt file(s) (i.e., the layers themselves)

- The weights for the actual layers are stored in the .caffemodel file.

According to internal Facebook benchmarks, Caffe outperforms TensorFlow by 1.2 to 5 times. TensorFlow is the most popular deep learning library for pictures and sequences, whereas Caffe is suitable for images but not for sequences or recurrent neural networks.

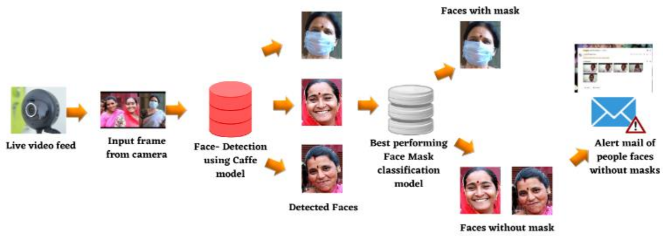

Fig 4: Facemask detection and classification on live feed during the deployment phase.

Classification of facemask from detected individual faces is done using the best performing Machine learning model. When the number of "Without mask" face count crosses a specific count, the system alerts the admin with images of individuals without masks.

Fig 4.1: Webcam input frame of People With and Without the mask

Fig 4.2: Output of System after Face detection and Facemask classification

Input frames having individuals without the mask would be saved, and an email is sent to the admin email after some saved frames cross a certain count threshold. The alert email has a subject and images saved without mask frames as specified in Fig 5.

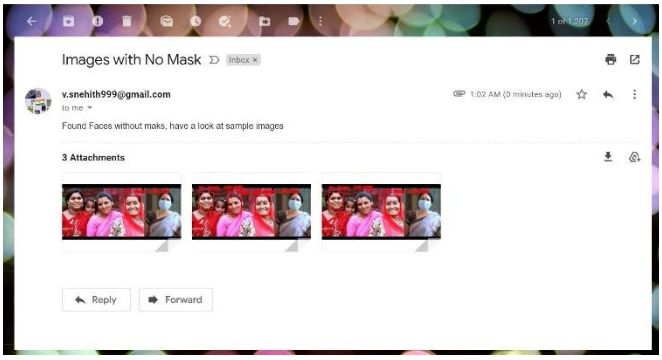

Fig 6: Email alert of individuals in the frame without masks at the admin's end.

4. BENEFITS

1. We are checking whether people are wearing masks or not is challenging for cops using manual monitoring. As a result, we use a webcam to recognize people's faces and prevent virus transmission with our method.
2. It is quick and accurate.
3. We can keep people secure from our approach by implementing this system at ATMs, banks, and other places.
4. It provides an automated way to monitor and save data for subsequent studies.

5. CONCLUSION

Our proposed system can identify and recognize human faces in a real-time context. Face detection and identification using the Caffe and Machine learning models provide faster detection and more reliable recognition. Compared to traditional face detection and recognition systems, this technique may reduce miss and error rates. When several models are compared, the best model with the highest accuracy and prediction emerges. The findings indicate that current face detection and identification technology are slower and inaccurate, and the suggested method might replace it.

REFERENCES

[1] S. Ge, J. Li, Q. Ye, and Z. Luo, "Detection of Masked Faces in the Wild with LLE-CNN's" June 2016.

[2] Guangchengwang, yumiao "Masked face recognition data sets and application" National natural science foundation of china 2020

[3] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517

[4] A. S. Joshi, S. S. Joshi, G. Kanahasabai, R. Kapil and S. Gupta, "Deep Learning Framework to Detect Face Masks from Video Footage," 2020 12th International Conference on Computational Intelligence and Communication Networks (CICN), Bhimtal, India, 2020, pp. 435-440, doi: 10.1109/CICN49253.2020.9242625.

[5] Amit Chavda, Jason Souza, SumeetBadgujar "Multi-Stage CNN Architecture for Face Mask Detection," September 2020

[6] Roshan M Thomas, Motty Sabu, Tintu Samson, Shihana Mol B, Tinu Thomas, "Real Time Face Mask Detection and Recognition using Python" June 2021

[7] Sammy v. militante, Nanettev.dionisioReal time face mask recognition with alarm system using deep learning 2020 11th IEEE control and system graduate research colloquium

[8] "Real Time Face Mask Detection and Recognition using Python," Roshan M Thomas, Motty Sabu, Tintu Samson, Shihana Mol B, Tinu Thomas

[9] "Effectiveness of Mask Wearing to Control Community Spread of SARS-CoV-2", John T. Brooks, MD1; Jay C. Butler, MD1

[10] Wenyunsun, Yusong, Changsheng, Face spoofing detection based on local ternary label supervision in fully convolutional networks IEE transactions on information forensis and security 2020.

[11] "FACE MASK DETECTION," A.P. Nithya Priya, V.Mavitha Sri, S. Monisha, P.Prathibha, K.Rekha, April 2021.

[12] S. Abbasi, H. Abdi, and A. Ahmadi, "A Face-Mask Detection Approach based on YOLO Applied for a New Collected Dataset," 2021 26th International Computer Conference, Computer Society of Iran (CSICC), 2021, pp. 1-6, doi: 10.1109/CSICC52343.2021.9420599.

[13] "Face Mask Detection using MTCNN and MobileNetV2", Vansh Gupta1, Rajeev Rajput2

[14] CiteSeerX "Rapid Object Detection Using a Boosted Cascade" http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.415.8118

[15] "Rethinking the Inception Architecture for Computer Vision" by Szegedy et. al.

[16] "Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning", Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, Alex Alemi

[17] 翻訳結果: Real-time Face Mask Detection in Video Data.

[18] R. Suganthalakshmi, A. Hafeeza, P. Abinaya, A.Ganga Devi, 2021, Covid-19 Facemask Detection with Deep Learning

and Computer Vision, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) ICRADL – 2021 (Volume 09 – Issue05)

[19] arXiv:1804.03447v2 [cs.CV] 18 Apr 2018. https://arxiv.org/pdf/1804.03447.pdf

[20] © 2021 JETIR July 2021, Volume 8, Issue 7 www.jetir.org

BIOGRAPHIES

Snehit Vaddi,
B.Tech Final year,
Dept. of Computer Science and Engineering,
GITAM deemed to be University

Vishnu Vipul Maddi,
B.Tech Final year,
Dept. of Computer Science and Engineering,
GITAM deemed to be University