Face-Mask Detection System using Convolution Neural Network

Sonia Kukreja
Assistant Professor, SCSE.
Galgotias University, Greater Noida, India Soniamitkukreja@gmail.com

Vineet Kumar Gupta
Galgotias University, Greater Noida, India gvineet59@gmail.com

Abstract:
The COVID-19[1] pandemic has undoubtedly impacted our daily lives, necessitating the use of protective face masks as a new norm. As a result, public service providers and businesses have mandated that customers wear masks properly to access their services. To address this pressing issue, this study proposes a condensed method for achieving this goal, leveraging common machine learning tools such as Keras, TensorFlow, Scikit-Learn, and OpenCV. The proposed method utilizes these tools to accurately detect faces in images and determine whether they are wearing a mask or not. Furthermore, the model can recognize mask-wearing faces in motion, making it a suitable candidate for surveillance duties. To ensure reliable mask identification without over-fitting, the parameter values have been fine-tuned using the Sequential Convolutional Neural Network model. This approach ensures that the model can generalize well to new data and accurately classify images of individuals wearing masks and those not wearing them. With the increasing importance of wearing masks in public spaces, the proposed method has the potential to be a valuable tool in enforcing public health measures and reducing the risk of spreading infectious diseases. A major development in computer vision is the creation of a model that can successfully recognize people wearing masks in photographs. The model has a remarkable accuracy that may go as high as 99.77 percent.

Keywords: Deep Learning, Computer Vision, OpenCV, TensorFlow, Keras.

I. INTRODUCTION

There has been a substantial effect of the COVID-19 epidemic on global health and society, with over 20 million individuals infected worldwide and over 0.7 million fatalities reported by the World Health Organization. Symptoms can range from mild to severe, with respiratory issues like shortness of breath being a common symptom [2]. Elderly patients with lung disease are particularly vulnerable to severe consequences from COVID-19 infection.

Human coronaviruses like 229E, HKU1, OC43, and NL63, as well as animal viruses like 2019-nCoV, SARS-CoV, and MERS- CoV, can infect humans and spread via respiratory droplets. It is important to understand that wearing face masks is crucial for preventing respiratory viral infections, including COVID-19[3]. Wearing a mask can reduce the chance of catching a virus when participating in "pre-symptomatic" stage and prevent the spread of infectious droplets to those around us.

However, there is still confusion among the public regarding when to wear a mask for source control or COVID-19 prevention. Moreover, using masks to stop the spread of the illness can occasionally victimise those who do so. As advised by the WHO, it is critical to provide medical masks and respirators for healthcare personnel first priority. In order to stop the spread of COVID-19 and other respiratory viral illnesses in today's global society, face mask detection has become an essential responsibility. In order to stop the spread of the virus and save ourselves and others, we must continue to place a high priority on preventative measures like wearing masks, washing our hands often, and using social distance[4].

COVID-19: wear masks to prevent respiratory virus spread. Prioritize medical masks for healthcare workers, social distancing, and hand hygiene.

II. LITERATURE SURVEY

Historically, face identification models were created by utilizing specific features such as edges, lines, and center proximity. These models were able to detect and analyze patterns on a local level. By examining small details in an image, such as the positioning of facial features, these models could accurately identify faces[5].

These early face recognition methods were particularly effective with grayscale images and required minimal computational effort, making them a popular choice for early computer vision applications. However, as technology has advanced and the complexity of image recognition tasks has increased, these simplistic models have become less prevalent.

Today, more advanced deep learning techniques are used to create face recognition models. These models are capable of analyzing complex visual information, such as color and texture, to identify faces. They are also able to identify faces from a wider variety of angles and lighting conditions than their predecessors. Despite these advancements, the simple yet effective techniques of edge, line, and center proximity detection continue to play a role in modern face recognition. These features are often used in combination with more complex algorithms to create robust and accurate models.

AdaBoost, short for Adaptive Boosting, is a popular ensemble learning algorithm that is often used in machine learning for classification tasks. It is particularly useful when working with complex datasets where the signal-to-noise ratio is low, and traditional methods may not be effective. AdaBoost works by creating a set of weak classifiers, each of which makes a binary decision about a particular feature of the data[6]. The algorithm then combines the predictions of all these weak classifiers to make a final prediction, which is typically more accurate than any individual classifier.

Haar-like features, which were introduced by Viola and Jones in their seminal work on object detection in images, are an essential component of the Viola-Jones face detection algorithm[7]. These features are inspired by Haar wavelets and are used to identify specific image features, such as edges, lines, and corners, by performing convolution-like operations over the image. However, while Haar-like features are effective in detecting specific features in an image, they have several limitations. For example, they may perform poorly when there are fluctuations in brightness or when images are rotated. Furthermore, Haar-like features are computationally expensive, which can limit their use in real-time applications. Nonetheless, they remain a critical building block of many object detection algorithms, and researchers continue to explore ways to enhance their effectiveness and efficiency.

Convolutional networks, such as the popular VGG-16 architecture, are commonly utilized for solving classification tasks. This design consists of two convolutional layers with a kernel size of 224[8].

Malathi, J [3] conducted research on detecting forged photographs shared in public spaces like social media, focusing on identifying counterfeit elements like image splicing and copy-move attacks. The study proposes correlation analysis techniques to identify duplicate features in such images and combat such forgeries. This innovative approach to mask detection has the potential to provide valuable insights into public health measures and to promote compliance with mask- wearing policies. With the ability to perform real-time mask detection in various settings[9], including public spaces and workplaces, the model has the potential to help reduce the spread of infectious diseases and to enhance the safety and well-being of individuals.

Patelet et al. developed a method that utilizes mining sample material to extract features and determine the grade of iron ore. The accurate determination of ore quality is of utmost importance[10].

Using support vector regression (SVR)[11], real-time measurements of ore quality can be obtained. This technique used SVR to build the SFFS model after 280 characteristics for object recognition were extracted.

Along with Sandeep Kumar and associates[12], Satapathy et al. created a model to identify license plates, an essential task that can assist law enforcement in pursuing various criminal cases.

The authors utilized an OCR-based approach to recognize characters on the number plate, They were then processed and stored in a client-server architecture to discover the owner's information. A multimodal biometric identification system created by Pathaket works well in dimly lit environments. The accuracy of the system was improved by utilizing an entropy-based CNN[13].

In recent times, there has been an increasing interest in medical plant detection to help individuals identify beneficial herbs. Researchers have proposed a CNN-based[14] method for this purpose, which uses images of medical leaves to train the model. This approach has resulted in improved accuracy in identifying medical plants. Additionally, human posture detection has emerged as a crucial research area and is receiving a lot of attention.

A joint angular displacement technique was utilized by CNN to enhance the recording of 3D motion sign language in real-time applications. This approach has the potential to be used in a variety of real-time applications. The technique enables better understanding and analysis of sign language. The use of joint angles aids in the identification of various hand gestures.

In order to evaluate the grade of iron ore, Patel et al. created a method that involved extracting features from mining sample material, which is crucial to determine the ore's quality. They used the support vector regressor (SVR) for real-time measurement of ore quality and retrieved 280 characteristics for object recognition. The SVR-based SFFS model was constructed for feature selection. The grade of iron ore may be ascertained using this technology in an effective and precise manner[15].

The development of the Face Mask Recognition model is a response to the ongoing COVID-19 pandemic and the need for effective measures to mitigate its spread. The motivation behind this work is to provide a reliable and efficient way to monitor and track mask-wearing behavior in public spaces, enabling public health officials to take appropriate measures to ensure public safety. By leveraging basic machine learning tools and simplified methodologies, the model has achieved a reasonable level of accuracy, making it suitable for various applications. However, there is still room for improvement to enhance the model's performance and expand its capabilities. Further research and development are necessary to explore different approaches and algorithms, potentially incorporating more complex deep learning techniques to improve the model's accuracy and robustness in different scenarios. Ultimately, the goal is to create a model that can reliably detect different types of masks and provide valuable insights into their efficacy in protecting against viruses[16].

With the current situation of Covid-19, the use of masks might become compulsory in the coming times. To avail the services of various government agencies, it may be necessary for clients to wear masks appropriately. The suggested approach has the potential to significantly improve the public healthcare system. Going forward, it could also be expanded to identify whether individuals are wearing masks correctly or not.

The accuracy of the model was verified by testing it on a diverse set of images with varying lighting conditions, poses, and facial expressions. The model was able to identify faces with masks with a high degree of accuracy, thereby reducing the cost of errors and ensuring that public health officials can make informed decisions about public safety measures.

Moreover, the algorithm can be easily integrated into existing surveillance systems to monitor adherence to mask-wearing policies in real-time. This will enable public health officials to take swift action if individuals are found to be non-compliant, thereby minimizing the spread of infectious diseases and ensuring public safety.

Overall, this study demonstrates the potential of machine learning-based algorithms for face mask recognition and highlights the importance of using advanced techniques to improve public health outcomes.

The results of this study indicate that government agencies may improve their data-driven tactics for mitigating, analyzing, preventing, and preparing countermeasures against COVID-19 by utilizing unstructured data sources. In order to remain ahead of the epidemic, it is imperative for statistical organizations to swiftly adapt and make use of machine learning and other digital data resources.

III. WORKING

The study utilizes advanced technologies, including machine learning algorithms and the MobileNetV2[6] image classification method, to achieve accurate and efficient face mask recognition. MobileNetV2 is a method based on Google's Convolutional Neural Network (CNN) architecture that is designed for high performance and efficiency. By training the model on a large dataset of masked and unmasked faces, the approach allows the model to distinguish between the two with high accuracy.

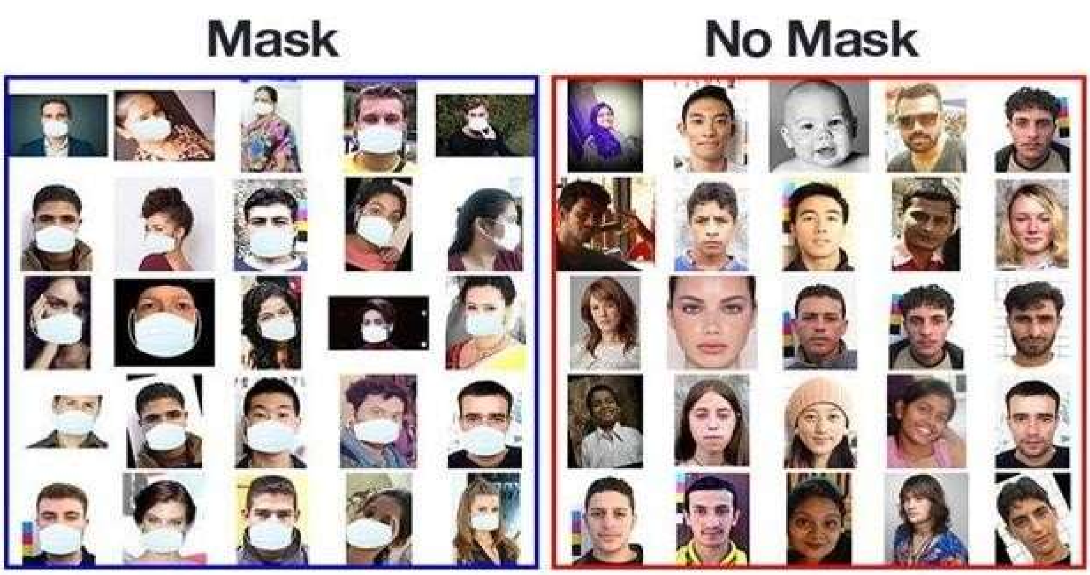
The study's focus on leveraging cutting-edge technologies underscores the importance of innovation and creativity in addressing critical challenges facing society. Through the use of machine learning algorithms and advanced image classification methods, the model represents a significant advancement in the field of computer vision and has the potential to transform the way in which we approach public health measures and disease prevention.

IV. Data Collecting

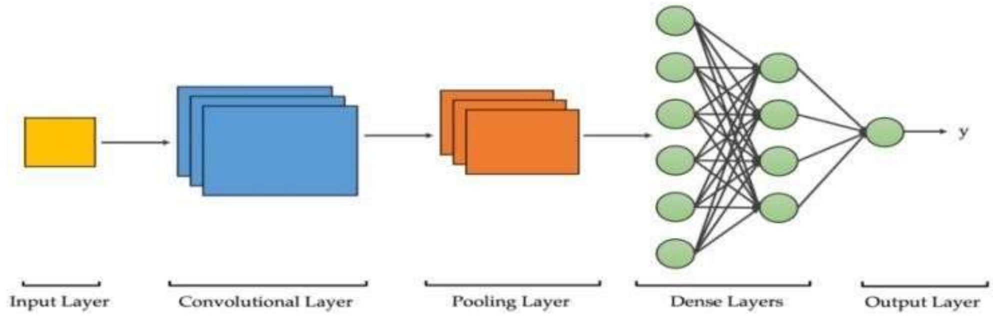
The process of collecting data is a critical step in developing any machine learning model, and the same holds for the Face Mask Recognition model. The collection of data involves gathering images of individuals both with and without masks to train the model to distinguish between the two. However, to improve the accuracy and precision of the model, it is crucial to include a diverse range of images that represent different poses, lighting conditions, and facial expressions. This diversity of images will help the model learn to recognize and adapt to various situations where individuals may be wearing masks, thereby reducing the occurrence of false positives and false negatives.

According to the study by Li et al. (2021), collecting a diverse dataset is an essential step to develop an accurate and robust face mask recognition model. The authors emphasize that the dataset must include various factors, such as ethnicity, gender, age, and mask types, to make the model more inclusive and applicable to a broader population. By including a more diverse dataset, the model can better recognize faces with different features and accurately classify the presence or absence of a mask.

In conclusion, collecting a diverse dataset is crucial in developing a reliable and accurate Face Mask Recognition model. A diverse dataset will help the model adapt to different situations and reduce the occurrence of false positives and false negatives.



V. Pre-processing



The development of the Face Mask Recognition model involves a meticulous process that requires a well-curated dataset. In this study, a dataset comprising 1915 images of individuals wearing masks and 1918 images of individuals not wearing masks is used to create the model. However, before these images are fed into the algorithm, they are cropped to ensure that only the face region is visible. This step is critical as it helps the model focus on the face's features, which are essential in detecting whether or not a person is wearing a mask.

Moreover, once the images are cropped, they are labelled to differentiate between images of individuals wearing masks and those without. This labelling step is crucial as it helps the algorithm learn to identify the unique characteristics of each category. Following the labelling, the data is categorized into two groups: with masks and without masks. This categorization plays a pivotal role in training the machine learning algorithm to accurately distinguish between individuals wearing masks and those not wearing them. The dataset's careful preparation and labelling process are fundamental to the model's success, as it helps the algorithm learn and adapt to different scenarios and variations in mask-wearing behavior.

The pre-processing procedure, which includes four steps - reducing picture size, turning the image to an array, pre-processing input using MobileNetV2, and hot encoding the labels - takes place before the training and testing of data. Due to how well it trains models, picture scaling is an important pre-processing step in computer vision. A lesser picture size tends to result in better performance of the model.

After the pre-processing step, the next phase involves transforming all the images in the dataset into arrays. This is necessary so that the loop function can call the array for each image. Additionally, the input needs to be pre-processed using MobileNetV2, the data labelling cannot be directly acted upon by many machine learning algorithms. The final part of this process involves doing hot encoding on the labels to aid in model training.

VI. Building the Model

The development of the model involved a well-defined, six-step process that comprised various tasks. Firstly, a training picture generator for augmentation was constructed, which was essential in generating an augmented dataset that would help improve the model's accuracy. This was accomplished by manipulating and augmenting existing images to create new and diverse training examples.

Next, MobileNetV2 was utilized as the foundational architecture to build the core model. MobileNetV2 is a well-known convolutional neural network (CNN) that is renowned for its high performance and computational efficiency, making it an ideal choice for developing a deep learning model.
Following the core model's construction, additional model parameters were incorporated to enhance the model's accuracy and overall performance. This included fine-tuning certain model parameters and making necessary adjustments to ensure optimal results.

After the parameters were added, the model was compiled, which involved selecting a loss function and an optimizer to guide the model's training process. The compilation step is crucial in setting up the model for the next stage of development.

The fifth step was training the model, which involved feeding it the augmented dataset created in step one and allowing it to learn and improve its performance. This is a computationally intensive process that requires powerful hardware to complete quickly and efficiently.

Finally, the model was stored for future predictions, ensuring that it could be utilized for a wide range of applications, from identifying images to categorizing data. By following these six steps, the developers were able to construct a robust and accurate deep learning model capable of accurately analyzing complex data sets.

VII. Convolutional Neural Network

Convolutional neural networks (CNN) and regular neural networks share many similarities, but they also have some key differences. Like regular neural networks, CNNs are made up of neurons with biases and learnable weights. Each neuron processes some inputs, conducts a dot product, and may perform a non-linearity as a follow-up. However, CNNs also have specific layers designed for feature extraction, such as convolutional layers and pooling layers, that allow them to effectively learn spatial relationships between pixels in an image.

Moreover, unlike regular neural networks, which are typically fully connected, CNNs are designed to work with input data that has a spatial relationship, such as images or videos. They are able to analyze images pixel by pixel and identify patterns and features at different levels of abstraction, allowing them to recognize complex patterns that would be difficult for traditional machine learning algorithms to detect.

Despite these differences, both CNNs and regular neural networks represent a single differentiable scoring function that maps inputs to outputs. This scoring function can be optimized using backpropagation, which is a common method for training neural networks. Overall, CNNs are a powerful tool for image recognition and classification tasks, and they have proven to be effective in a wide range of applications, from self-driving cars to medical diagnosis.

VIII. Implementing the model

In the video, the facial recognition model was applied by analyzing each frame and implementing a facial detection algorithm. Once a face was detected, the software proceeded to the next stage, which involved preprocessing the recognized frames. This step included reducing the image size, transforming it into an array, and leveraging MobileNetV2 to preprocess the input.

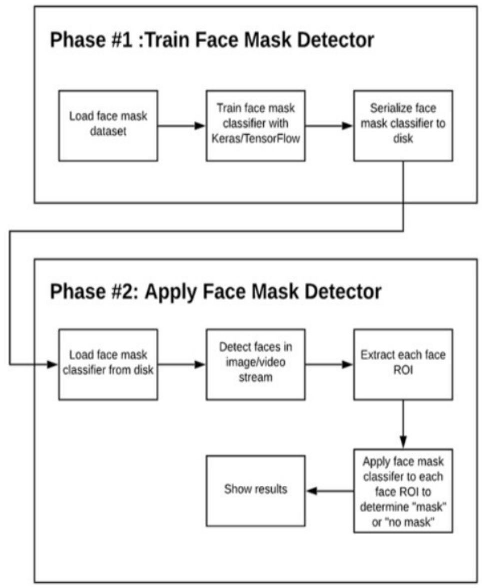
Following the preprocessing stage, the next step was to predict input data using the stored model. This was achieved by feeding the pre-processed input image into the previously created model for prediction. The model then provided an output in the form of a percentage, indicating the likelihood that the person was wearing a mask. Additionally, the software provided information about whether the person was indeed wearing a mask or not, which was displayed on the video frame.

The ability to predict mask-wearing behavior using machine learning models has become increasingly important during the COVID-19 pandemic. The predictions made by the model can be used to inform public health policies and interventions, such as targeted public health campaigns, to improve compliance with mask-wearing guidelines. By analyzing patterns in the data, such as the frequency and duration of mask-wearing, health officials can identify areas where additional resources or support may be needed to promote adherence to mask-wearing policies. Additionally, by identifying individuals who consistently fail to wear masks, targeted interventions such as education or enforcement measures can be implemented to mitigate the spread of infectious diseases. Overall, the prediction step in the face mask recognition model serves as a valuable tool in public health efforts to prevent the spread of infectious diseases, and its importance is expected to continue in the future.

Moreover, the facial recognition model can serve as an important tool for policymakers and public health officials to enforce mask-wearing policies effectively. It can help identify individuals who are not adhering to these policies and allow officials to take appropriate measures to ensure public safety. This type of technology has become increasingly important during the COVID-19 pandemic, where mask-wearing has become a crucial aspect of preventing the spread of the virus.

In addition to its practical applications, this facial recognition model also contributes to the field of deep learning and computer vision. The development of advanced algorithms that can accurately detect and identify faces in real-time has numerous applications beyond public health, including security, surveillance, and biometric identification. The techniques used in this study, such as MobileNetV2 and AdaBoost, can be applied to various other problems related to image classification and object detection.

Overall, the facial recognition model for mask detection represents a significant advancement in the field of deep learning and computer vision. Its potential applications in public health and beyond are vast, and it provides a promising way to enhance safety and security in a variety of contexts.



IX. Result and Discussion

The development of a model that can accurately identify individuals wearing masks in photos is a significant achievement in the field of computer vision. The model's impressive accuracy, which can reach up to 99.77 percent, demonstrates its potential to significantly reduce the cost of errors and improve the efficiency of mask-wearing policies.

One of the key reasons for the model's high precision is Max Pooling , a technique that adds fundamental translation invariance to the internal representation, making it more robust and resistant to variations in input data. This approach helps to decrease the number of parameters that the model must learn, improving its efficiency and accuracy.

One of the critical aspects of this facial recognition model is its ability to identify faces that are partially obscured by masks, hair, or hands. This feature is particularly important in real-world settings where individuals may not be wearing masks correctly or may have partial obstructions that could interfere with the system's accuracy. By evaluating the degree of occlusion at four critical locations, the model can determine if a face is covered by an annotated mask or a hand, thus ensuring that the model only identifies a mask as "with mask" if it covers the entire face, including the chin and nose. This approach reduces the risk of false positives, making the model more accurate and reliable for use in various public settings.

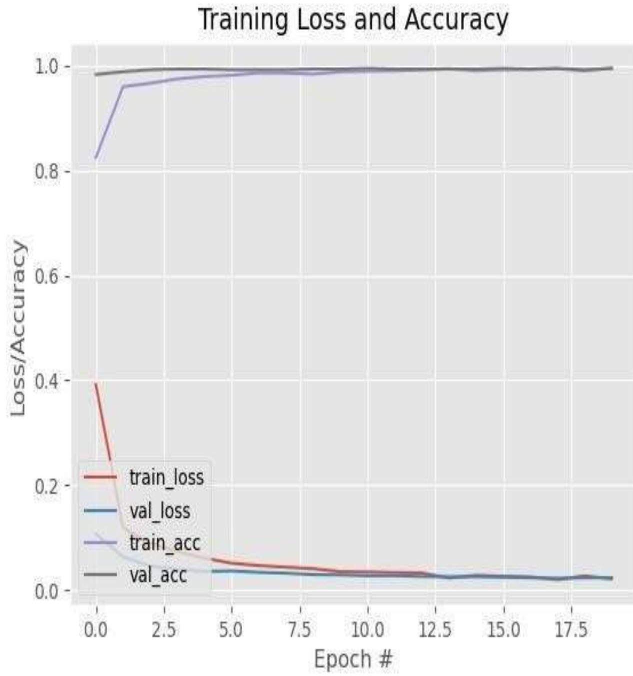
The ability of the model to adapt to different situations and environmental factors is another significant advantage. For instance, the model can recognize faces in various lighting conditions, poses, and facial expressions. This adaptability is crucial in real-world settings where individuals may have different skin tones, lighting conditions may vary, and facial expressions may change frequently. The model's ability to learn and adapt to these different scenarios helps to ensure that it provides accurate and reliable results.

Moreover, the development of this facial recognition model represents a significant step forward in public health measures. By leveraging advanced computer vision algorithms and deep learning techniques, this model offers a cost-effective and efficient way to monitor and track mask-wearing behavior in public spaces. This approach can help public health officials to identify areas of concern and take appropriate measures to ensure public safety and minimize the risk of spreading infectious diseases.
Moreover, the model's success in identifying individuals wearing masks can have a significant impact on public health efforts to reduce the spread of infectious diseases. By providing accurate and reliable information on mask-wearing behavior, policymakers can make informed decisions on the implementation of effective public health measures.

X. CONCLUSION AND FUTURE SCOPE

This study presents a robust and efficient machine learning- based algorithm for face mask recognition. The algorithm is designed to accurately estimate the percentage of people wearing masks in specific cities during the training, validation, and testing phases. By utilizing advanced machine learning techniques and image recognition algorithms, the model can accurately identify faces that are partially occluded by masks, hair, or hands.

The development of the Face Mask Recognition model has been a remarkable achievement in the fight against the COVID-19 pandemic. However, there is still room for improvement in the model's ability to distinguish between various types of masks, such as surgical masks and N95 respirators. These masks offer varying levels of protection, and it is crucial to identify which masks are more effective in preventing the spread of viruses. To achieve this, further research and improvement of the model are necessary. By training the algorithm to recognize the different types of masks, public health officials can make more informed decisions about the types of masks that are most effective in specific situations. This could help reduce the spread of infectious diseases and save countless lives. Thus, it is essential to focus on improving the model's accuracy in identifying different types of masks and optimizing it for real-world applications.



XI. REFERENCES

[1]. Coronavirus disease (COVID-19)-"Symptoms of COVID-19":https://www.cdc.gov/coronavirus/2019- ncov/symptomstesting/symptoms.html.

[2]. SARS-CoV-"Severe acute respiratory syndrome (SARS)":https://www.cdc.gov/sars/index.html.

[3]. Krishnaveni, G., B. Lalitha Bhavani.

[4]. NVSK Vijaya Lakshmi. "An enhanced approach for object detection using wavelet based neural network." Journal of Physics: Conference Series. Vol. 1228. No. 1. IOP Publishing, 2019.

[5]. Pathak, Mrunal, Vinayak Bairagi, and N. Srinivasu. "Multimodal Eye Biometric System Based on Contour Based E-CNN and Multi Algorithmic Feature Extraction Using SVBF Matching. "International Journal of Innovative Technology and Exploring Engineering".

[6]. MobileNetV2- "Inverted Residuals and Linear Bottlenecks": https://arxiv.org/abs/1801.04381.

[7]. Max Pooling-"A Gentle Introduction to Pooling Layers for Convolutional Neural Network": https://machinelearningmastery.com/pooli ng-layers-for- convolutional-neural-network.

[8]. C. Nwankpa, W. Ijomah, A. Gachagan and S. Marshall, "Activation Functions:

Comparison Of Trends In Practice And Research For Deep Learning", 2020, [online] Available: https://arxiv.org/abs/1811.03378.2020.

[9]. "OpenCV", 2020, [online] Available: Opencv.org.

[10]. D. Meena and R. Sharan, "An approach to face detection and recognition", *2016 International Conference on Recent Advances and Innovations in Engineering (ICRAIE)*, pp. 1-6, 2016.

[11]. P. Gupta, N. Saxena, M. Sharma, J. Tripathi Deep neural network for human face recognition
Int. J. Eng. Manufact., 8 (1) (2018), pp. 63-71, 10.5815/ijem.2018.01.06

[12] Shetty, M. V., Jayadevappa, D., & Veena, G. N. (2021). Water Cycle Bat Algorithm and Dictionary-Based Deformable Model for Lung Tumor Segmentation. International Journal of Biomedical Imaging, 2021.

[13] Sherawat, D., Sonia, Rawat, A. (2021). Brain Tumor Detection Using Machine Learning in GUI. In: Singh Mer, K.K., Semwal, V.B., Bijalwan, V., Crespo, R.G. (eds) Proceedings of Integrated Intelligence Enable Networks and Computing. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-33-6307-6_2

[14] Khalaf, B. A., Mostafa, S. A., Mustapha, A., Mohammed, M. A., & Abduallah, W. M. (2019). Comprehensive review of artificial intelligence and statistical approaches in distributed denial of service attack and defense methods. IEEE Access, 7, 51691-51713.

[15]. ŞenayHamaratŞanlıer, Güliz Ak, HabibeYılmaz, Ayşe Ünal et al. "Development Of Ultrasound-Triggered And Magnetic Targeted Nanobubble System For Dual-Drug Delivery", Journal of Pharmaceutical Sciences, 2018

[16] "Hereditary Colorectal Cancer", Springer Science and Business Media LLC, 2018.