Face Mask Detection by Using AlexNet Deep CNN

**1st Khan, Md. Ifterkhar Ahmed (19-41201-1)**  
*Computer Science and Engineering*  
*American Internation University-Bangladesh.*  
Dhaka, Bangladesh

**2nd Mir, Md Ishmam (19-39925-1)**  
*Computer Science and Engineering*  
*American Internation University-Bangladesh.*  
Dhaka, Bangladesh

**3rd Shium, Sheikh Mahmudul Hasan (19-40647-2)***Computer Science and Engineering*  
*American Internation University-Bangladesh.*  
Dhaka, Bangladesh

**4nd Sumit Kanti Sarker (19-41764-3)**  
*Computer Science and Engineering*  
*American Internation University-Bangladesh.*  
Dhaka, Bangladesh

*Abstract*—COVID-19 epidemic has swiftly disrupted our day-to-day lives affecting the international trade and movements. Wearing a face mask to protect one's face has become the new normal. In the near future, many public service providers will expect the clients to wear masks appropriately to partake of their services. Therefore, face mask detection has become a critical duty to aid worldwide civilization. This paper provides a simple way to achieve this objective utilizing some fundamental [Machine Learning](https://www.sciencedirect.com/topics/computer-science/machine-learning) tools. The suggested technique successfully recognizes the face in the image or video and then determines whether or not it has a mask on it. As a surveillance job performer, it can also identify a face together with a mask in motion as well as in a video. The technique attains excellent accuracy. We investigate optimal parameter values for the Convolutional [Neural Network](https://www.sciencedirect.com/topics/neuroscience/neural-networks) model (CNN) in order to identify the existence of masks accurately without generating over-fitting.

Keywords—Machine Learning (ML), Deep Neural Learning (DL), Convolutional Neural Network (CNN), Alex Net Model.

# **Introduction**

The trend of wearing face masks in public is rising due to the COVID-19 coronavirus epidemic all over the world. Before Covid-19, People used to wear masks to protect their health from air pollution. While other people are self-conscious about their looks, they hide their emotions from the public by hiding their faces. Scientists proofed that wearing face masks works on impeding COVID-19 transmission. COVID-19 (known as coronavirus) is the latest epidemic virus that hit the human health in the last century. In 2020, the rapid spreading of COVID-19 has forced the World Health Organization to declare COVID-19 as a global pandemic. More than five million cases were infected by COVID-19 in less than 6 months across 188 countries. The virus spreads through close contact and in crowded and overcrowded areas.

The coronavirus epidemic has given rise to an extraordinary degree of worldwide scientific cooperation. Artificial Intelligence (AI) based on Machine learning and Deep Learning can help to fight Covid-19 in many ways. Machine learning allows researchers and clinicians evaluate vast quantities of data to forecast the distribution of COVID-19, to serve as an early warning mechanism for potential pandemics, and to classify vulnerable populations. The provision of healthcare needs funding for emerging technology such as artificial intelligence, [IoT](https://www.sciencedirect.com/topics/engineering/iot), big data and machine learning to tackle and predict new diseases.

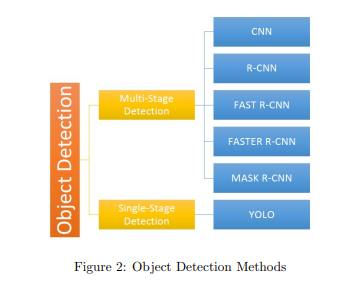
People are forced by laws to wear face masks in public in many countries. These rules and laws were developed as an action to the [exponential growth](https://www.sciencedirect.com/topics/engineering/exponential-growth) in cases and deaths in many areas. However, the process of monitoring large groups of people is becoming more difficult. The monitoring process involves the detection of anyone who is not wearing a face mask.

In this report, we introduce a mask face detection model that is based on deep transfer learning and classical machine learning classifiers. The proposed model can be integrated with surveillance cameras to impede the COVID-19 transmission by allowing the detection of people who are not wearing face masks. The model is integration between deep transfer learning and classical [machine learning algorithms](https://www.sciencedirect.com/topics/engineering/machine-learning-algorithm). We have used deep transfer leering for feature extractions and combined it with three classical machine learning algorithms. We introduced a comparison between them to find the most suitable algorithm that achieved the highest accuracy and consumed the least time in the process of training and detection.

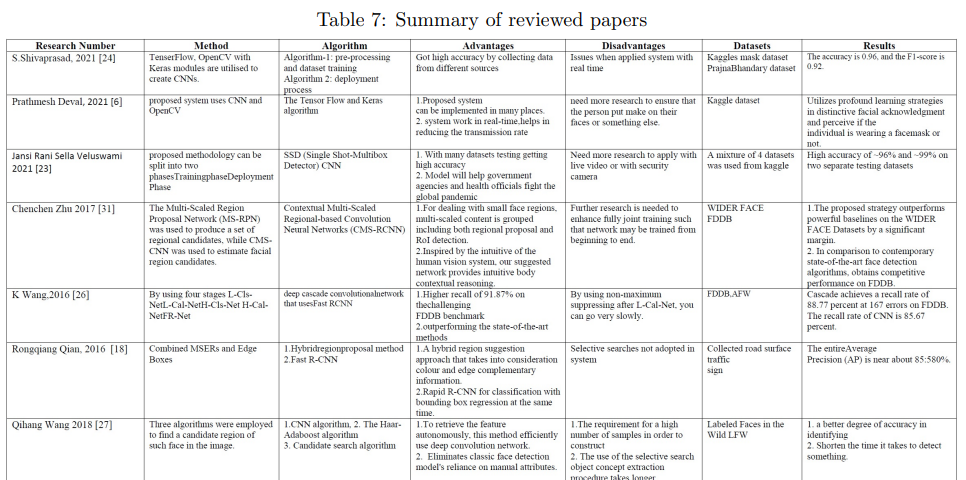
The proposed model is AlexNet method. AlexNet is the name of a convolutional neural network (CNN) architecture, designed by Alex Krizhevsky in collaboration with Ilya Sutskever and Geoffrey Hinton, who was Krizhevsky's Ph.D. advisor. AlexNet competed in the ImageNet Large Scale Visual Recognition Challenge on September 30, 2012. The network achieved a top-5 error of 15.3%, more than 10.8 percentage points lower than that of the runner up. The original paper's primary result was that the depth of the model was essential for its high performance, which was computationally expensive, but made feasible due to the utilization of graphics processing units (GPUs) during training.

# **LITERATURE REVIEW**

According to the recent literature review done by Firas Amer Mohammed Alia, Mohammed S.H. Al-Tamimi, from Department of Computer Science, College of Science, University of Baghdad, Baghdad, Iraq, there are two types of object detection method Multi-Stage Detection and Single-Stage detection [1]. There are many sub-categories that is given in Fig1



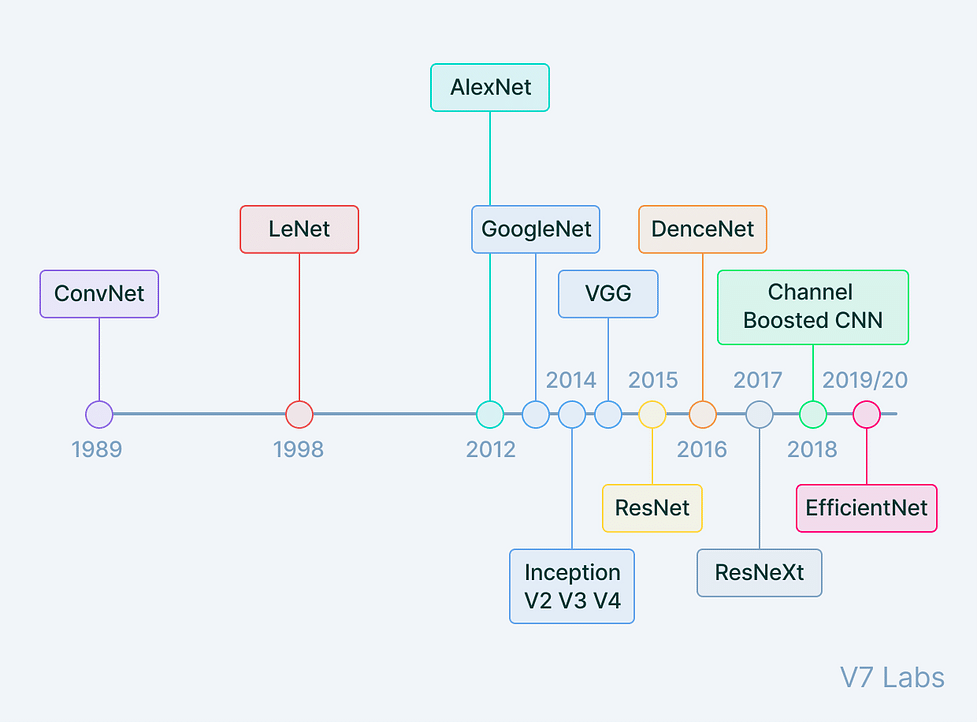
Comparison table

 Table

Description automatically generated

Graphical user interface, text, table

Description automatically generated



https://www.projectpro.io/article/introduction-to-convolutional-neural-networks-algorithm-architecture/560

Among them AlexNet is not very popular and optimized. But in order to simply understand how CNN works and practical applications AlexNet is better option with satisfactory accuracy.

# **METHODOLOGY**

The face mask detection process identifies faces in a given input image or video feed by indicating a boundary box around the face. This article will use a deep learning computer vision model to classify whether a person is wearing a mask or not. All the analysis and data preprocessing are done using the Kaggle dataset [1], containing about twelve thousand images and standard procedures. AlexNet CNN (Convolutional Neural Network) will be used for face mask detection. A flowchart of the methodology is depicted in Fig2.

Diagram

Description automatically generated

##### DATA COLLECTTION & PREPROCESSING

As explained earlier, the used dataset is collected from Kaggle [2] for research purposes. The dataset is in public domain so, the is no legal issues for using it in the project. The dataset contains about 12k images where face images with mask are about 6k in number and, all the images without the face mask are preprocessed from the “CelebFace” dataset created by Jessica Li [3].

For preprocessing all the necessary modules and settings are imported at first. After that, the dataset file is loaded that is saved in local device and split the dataset into 3 separate directories named train, valid and test. Then the image size is set 227 and there are 2 labels in training directory named ‘WithMask’ and ‘WithoutMask’. Then all 3 datasets are shuffled to reduce bias that can arise. In addition, the images are put into the X (X\_train, X\_Valid, X\_Test) and the labels in Y (Y\_train, Y\_Valid, Y\_Test). After that, pickle is used to serialize the dataset. Afterwards, all the photos are combined to get mean photo. Finally, for normalization part all images are subtracted from the mean image.

##### MODEL TRAINING AND TESTING

Alex net Architecture by transfer learning approach is used. Alexie is a pre-trained CNN (Convolutional Neural Network) model.

Our dataset is collected to train our model. The data has been spitted set into Test and Train. After that, we preprocessed our data. We have resized the image to 227 X 227 because we are using Alex net. After preprocessing, the data is imported to our base model Alex Net and this architecture is implemented by using a transfer learning approach. In this research, the model is classified by AlexNet architecture. It is a type of convolutional neural network. It is a classification algorithm that is best suited for the classification of RGB images. The architecture includes eight layers, among them, the first 5 were convolutional layers, followed by some max-pooling layers which are used in down sampling, and the last three layers were fully connected. It uses a rectified linear unit activation function named RELU. Activation in each of these layers except the output layer. That is found out using the RELU as an activation function accelerated the speed of the training process by almost six times. It also used the dropout layers, which prevented its model from overfitting. Further, the model is trained on the Imagenet dataset. The Imagenet dataset has almost 14 million images across a thousand classes. After training our model we have got our desired output with good accuracy.

Diagram

Description automatically generated

**AlexNet Architecture**

* It has 8 layers with learnable parameters.
* The input to the Model is RGB images.
* It has 5 convolution layers with a combination of max-pooling layers.
* Then it has 3 fully connected layers.
* The activation function used in all layers is Relu.
* It used two Dropout layers.
* The activation function used in the output layer is SoftMax.
* The total number of parameters in this architecture is 62.3 million.

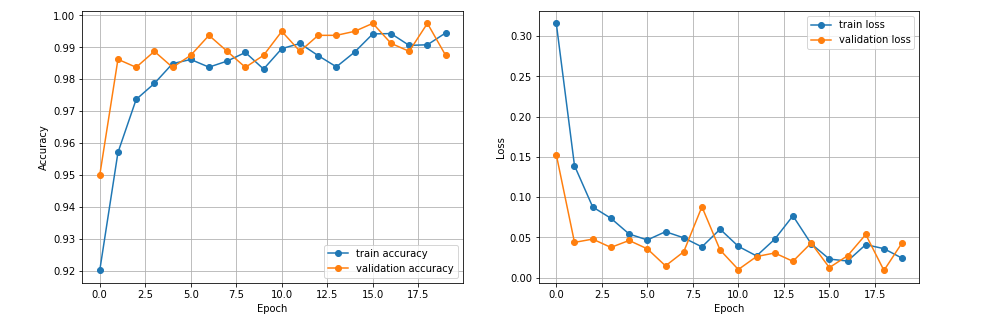
Diagram: AlexNet Architecture

Diagram, engineering drawing

Description automatically generated

##### MODEL EVALUATION

**AlexNet:** Convolutional neural networks are one of the variants of neural networks where hidden layers consists of convolutional layers, pooling layers, fully connected layers and normalization layers.



|  |  |
| --- | --- |
| Accuracy | 0.9808467626571655 |
| Loss | 0.08948305994272232 |

# **References**

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| References   |  |  | | --- | --- | | [1] | F. A. Mohammed Ali and M. S. H. Al-Tamimi, “Face mask detection methods and techniques: A review,” *International Journal of Nonlinear Analysis and Applications*, vol. 13, no. 1, pp. 3811–3823, 2022. | | [2] | A. Jangra, “Face Mask Detection ~12K Images Dataset.” 26-May-2020. | | [3] | “Jessica Li,” *Kaggle.com*. [Online]. Available: https://www.kaggle.com/jessicali9530/. [Accessed: 12-Aug-2022]. | | [4] | F. Amer, M. Ali, and M. S. H. Al-Tamimi, “Face mask detection methods and techniques: A review.” | | [5] | “Introduction to convolutional neural networks architecture,” *ProjectPro*, 20-Jun-2022. . | |  |
|  |  |