International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 01 | Jan 2022 www.irjet.net p-ISSN: 2395-0072

Designing of Application for Detection of Face Mask and Social Distancing During Covid-19 using CNN and Yolo v3

Gaurav Garje[1], Rasika Lande[2], Shardul Shinde[3], Nishant Inamdar[4], Prof. A.S.Deokar[5]

[1,2,3,4]Undergraduate Students, Dept. of Computer Engineering, AISSMS COE, Pune, Maharashtra, India
[5]Professor, Dept. of Computer Engineering, AISSMS COE, Pune, Maharashtra, India

---

Abstract - *The aftermath of the Covid-19 pandemic has ushered in a new normal, forcing numerous unavoidable changes in social life around the world. Wearing a face mask and maintaining social distance, especially in crowded settings, was made compulsory by the government, some disobeyed the orders, resulting in a rise in covid cases.*

Current work presents the study of machine learning algorithms for a face mask and social distance detection at crowded places during the pandemic. It suggests a system architecture that supports functions like detecting whether people are wearing masks (partially/completely) and identifying whether they are following social distancing.

Study comprises of 3 modules – face detection, mask detection, and social distance detection. YOLO (You Only Look Once) algorithm is used for Object Detection as well as Object Tracking, because of its highest accuracy and precision achieved to date amongst prevalent techniques. It detects the people and their faces in the frame for counting the objects and keeps a record of those objects in the next frame using Object Tracking.

The custom MAFA dataset is used to understand facemasks. The CNN (Convolutional neural network) Algorithm will be used to train the model on those datasets for classification and detection. Implementation of the application will be done in python and efficiency and accuracy are calculated. A notification is sent if the separation between the two entities is below the standard specifications. Faces, both masked and unmasked, are labelled appropriately.

Key Words: **Covid-19, Object Detection, Object Tracking, Distance Estimation, MAFA, CNN, YOLO v3, Open CV**

1.INTRODUCTION

The pandemic has brought us a lot of challenges, especially in the working patterns.[1] Along with Covid-19, several other large-scale, serious respiratory diseases, such as SARS and MERS, emerged, adding to the pandemic's chaos.

In 2020, the World Health Organization recommended that people must wear masks to avoid infection by a novel coronavirus and maintain a social distance of at least 3 meters between themselves to prevent the virus's proliferation and spread.

People are more conscious about their wellbeing, and governments are placing a greater emphasis on public health. The virus's next wave has begun, but people have continued to congregate in crowded places without wearing a mask. The government has been unable to maintain control over the situation, triggering the need for artificial intelligence and computer technology to solve this problem. Distancing oneself from others and wearing masks help to slow the disease transmission. The researchers believe in developing a surveillance system that makes use of existing IP and CCTV cameras, as well as Computer Vision and Object Detection algorithms, to identify people who are violating Covid norms.[5]

Current research work comprises of 2 stages, namely designing a framework for detection of face mask and social distancing with the help of CNN and Yolo v3 and in the next part application will be developed.

This research understands various aspects of the image or videos based on frames those would be made available as inputs to the algorithms. [3] The Object Detection algorithm YOLO v3 will be used to find people in a frame.

The distance between people in the image will be calculated by the system using the Euclidean distance between the centroids of the detected boxes around the faces in the images. The six most important characteristics of masked faces are face locations, face orientation, eyes, masks, occlusion degree, and mask type. These will be considered while creating the model.

Using Deep learning and OpenCV, a CNN model will be generated for mask detection[5]. Face masks will be detected using predictive analysis on both static images and videos.

As a result, this system will automate the process of checking up on citizens, assisting the government in enforcing COVID-19 rules and regulations which will significantly reduce number of Covid cases.

1.1 YOLO v3

Yolo stands for You only look once. It detects objects using CNN. It can detect numerous objects in an image. It predicts not only the class of the object but also its location in the image. For the object detection task, Yolo version 3 employs a neural network of 53 CNN's (Darknet-53) and an additional 53 layers, totaling 106 layers.

This neural network divides images into regions and assigns cell probabilities to each of them.

Yolo estimates the number of bounding boxes needed to cover specific areas of the image and chooses the best ones based on probabilities.

The network receives input in the form of a batch of images with the shape (n,w,h,3), where n is the number of images, w and h are the image's width and height (multiples of 32), and 3 is the RGB channel.

Yolo version 3 is a three-scale detection algorithm, with detection occurring at three different layers of the CNN: layer-82, layer-94, and layer-106. The network down samples the input image by 32,16, and 8 strides of the network at each of these points.

At each of these locations, 1 by 1 convolution is applied to down sampled outputs.(see Fig-1)

Yolo version 3 calculates three bounding boxes for each cell of these feature maps, with various attributes for each bounding box. The center coordinates (abscissa, ordinate), bounding box dimensions (width, height), objectness score, and a list of confidence classes to which the object may belong are among these attributes.

Yolo version 3 uses nine default anchor boxes at each of the three network locations to predict bounding boxes. Yolo v3 calculates offsets to predefined anchors and predicts center coordinates by passing the output through the sigmoid function to predict the real width and height of the bounding box. Finally, the final output is determined by the 'intersection over union' of these three bounding boxes.

1.2 CNN

Neural networks simulate the behavior of the human brain in the application areas of Artificial Intelligence, machine learning, and deep learning, allowing computer programs to recognize patterns and solve common problems.

Object from an image is identified by CNN. First, the image's pixels are fed as input, and then the input layer accepts the image's pixels as input.

Hidden layers extract features by conducting specific calculations and transformations. To detect patterns in the image, this layer employs a matrix filter and convolution operations. Finally, the object in the image is identified by a fully connected layer. Convolutional neural network is a feed-forward neural network that processes data in a grid-like topology to analyses visual images. Convolution layers, ReLu layers, the fully connected layer, and the Pooling layer are the four Hidden Layers. Convolution is the process of coiling or twisting data and altering it to find a new pattern. Convolution operations are carried out by several filters in the convolution layer. The next step is to move the maps and features to the ReLu layer after they have been extracted.

Multiple convolutions and a Relu layer are used to scan the real image and locate features. The rectified feature map now travels via a pooling layer(vast information is reduced to a single response).Pooling is a down sampling process that decreases the feature map's dimensionality. The pooling layer employs many filters to identify various aspects of an image, such as borders ,corners ,body, and so on.

From a parallel pooled featured map, a single long continuous linear vector is generated by flattening the resulting 2D arrays. The fully connected layer receives the flattened matrix from the pooling layer as input to identify the image. (See Fig-2)[15]

2. RELATED WORK

[1] "Automated evaluation of COVID-19 risk factors coupled with real-time, indoor, personal localization data for potential disease identification, prevention and smart quarantining" by J. Barabas, R.Zalman and M. Kochlan. To lower COVID-19 spread, automated systems with sanitary, temperature measurement, individualized person tracking, and examining is necessary. This is important not only for disease surveillance but also for disease intervention. A system is proposed that incorporates all the component's functionality based on easily available components including temperature and mask detection module further strengthened with preliminary RTLS (hardware and software).

This helps in data collection, allowing for post-symptom detection and identification of person interactions to assess the risk of infection vectors and take further precautions.

[2] "Multi-angle Head Pose Classification when Wearing the Mask for Face Recognition under the COVID-19 Coronavirus Epidemic," by Shuang Li, Xin Ning, Lina Yu, Liping Zhang, Xiaoli Dong, Yuan Shi, Wei He .Head pose classification is a preprocessing technique applied prior to the face recognition because it requires a front facing input image. Hence, a new method called HGL is proposed which combines H-channel of HSV (Hue Saturation Vue) with grayscale image of face. CNN is used to train this model for feature extraction. This approach is advantageous to others in terms of accuracy.

[3] "A review of the prevalent ICT techniques used for COVID-19 SOP violation detection" by Talha Ikram, Abdullah Saeed, Noor Ul Ayn, Muhammad Ali Tahir, Rafia Mumtaz. This survey paper addresses the most common methods for detecting violations of government-mandated covid-19 standard operating procedures. It uses computer vision and deep learning algorithms for object detection and distance estimation. The research provides a new method for operating the system on non-GPU-based machines while keeping the frame rate respectable.

[4] "CNN Based COVID-19 Prevention System" by Deepalaxmi R Shenvi, Krishnananda Shet. .The Artificial

Intelligent IoT system proposed in the research paper includes temperature monitoring, auto sanitizing, and mask detection. A Raspberry Pi 4 model with 2GB of RAM has been used in the proposed system. Face recognition has been performed using the RMFRD Dataset. Keras is a Python library that makes artificial neural networks easier to use. This system detects masks with a 98 % accuracy rate. A comparison of the aforementioned processes with previous research is also included in the paper.

[5] "Study of Masked Face Detection Approach in Video Analytics" by Gayatri Deore, Ramakrishna Bodhula, Dr. Vishwas Udpikar, Prof. Vidya More. This paper focuses mainly on security using video analytics which improves video surveillance systems and hence it saves human resources and performance of such systems is also improved. Calculating distance from camera, line detection, facial part detection, and eye detection are the four steps in the masked face detection technique used in this paper. These steps are useful in finding if a person is wearing a mask or not. Here line detection and eye detection can give some false detections as their detection is not that accurate in poor resolution images.

[6] "A Novel AI-enabled Framework to Diagnose Coronavirus COVID-19 using Smartphone Embedded Sensors: Design Study" by Halgurd S. Maghdid, Kayhan Zrar Ghafoor, Ali Safaa Sadiq, Kevin Curran, Danda B. Rawat, Khaled Rabie. The researchers offer a new method for detecting COVID-19 using smartphone sensors in this study.

When compared to other clinical methods of COVID-19 detection, such as CT scans and blood tests, the proposed method is more convenient and cost-effective. These hardware devices like fingerprint sensor, microphone, camera and inertial sensor are employed to analyze the patient's condition.

To assess whether a patient is COVID-19 positive or negative, the proposed approach uses machine learning techniques, RNN and CNN.

[7] "Adversarial Examples – Security Threats to COVID-19 Deep Learning Systems in Medical IoT Devices" by Md. Abdur Rahman, M. Shamim Hossain, Nabil A. Alrajeh, Fawaz Alsolami.

The study looked at nine COVID-19 deep learning applications that can diagnose the virus quickly. They put six open-source applications to the test and watched the models closely to create adversarial cases for each type and pinpoint the flaws in model.

The vulnerability of existing DL applications to AE attacks has been realized, necessitating additional research, attention, and installation of proper security mechanisms, safeguards, and controls before these applications are employed in real-world healthcare settings.

[8] "Detection of Respiratory Infections using RGB-infrared sensors on Portable Device" by Zheng Jiang, Menghan Hu, Zhongpai Gao, Lei Fan, Ranran Dai, Yaling Pan, Wei Tang, Guangtao Zhai,and Yong Lu.

This study proposes a portable non - contact method for examining the health of people wearing protective masks by assessing respiratory characteristics from RGB-infrared sensors. The system uses face recognition to capture those who are wearing masks.

The information passes through a bidirectional GRU neural network with an attention mechanism. The accuracy of the proposed method is 83.69 percent, with a sensitivity of 90.23 percent and a specificity of 76.31 percent.

[9] "Diagnosing COVID-19: The Disease and Tools for Detection" by Buddhisha Udugama, Pranav Kadhiresan, Hannah N. Kozlowski, Ayden Malekjahani, Matthew Osborne, Vanessa Y. C. Li, Hongmin Chen, Samira Mubareka, Jonathan B. Gubbay, and Warren C. W. Chan.

This survey paper proposes the plug and play design of COVD-19 diagnostics. To screen and identify covid-19, an integration of computed tomography imaging, whole genome sequencing, and electron microscopy was initially used. The general intent of this study is to educate the reader on diagnostic and surveillance technologies for covid-19, as well as their performance characteristics.

The proposed study describes upcoming point-of-care diagnostics and encourages academics to take their technologies beyond conception. Creating plug-and-play screening tests to manage the COVID outbreak would be beneficial in preventing future outbreaks.

[10] "Experimental Investigations on Particle Contamination of Masks Without Protective Pellicles During Vibration or Shipping of Mask Carriers" by Se-Jin Yook, Heinz Fissan, Christof Asbach, Jung Hyeun Kim, Dabrina D. Dutcher, Pei-Yang Yan, and David Y. H. Pui.

Since pellicles cannot be used to protect EUVC masks due to high EUV absorption, particle-free handling is a challenge in EUV lithography. Here, to hold a mask, two shipping mask carrier models with 16 supports are used. Online and offline particle detection methods were implemented.

For detection of Particles, Surface particle detection and airborne particle detection methods are used. Some methods for testing particle contamination of masks without protective pellicles include manual shaking, real shipping, and controlled vibration. The particle size in this paper is up to 60 nm PSL equivalent size.

[11] "Explainable AI and Mass Surveillance System-based Healthcare Framework to Combat COVID-19 like Pandemics" by Nadra Guzeni, M. Shamim Hussain and Ghulam Mohammad.

With the help of 5G technology and a network of wireless connectivity COVID 19 can be managed easily. This can be done via hierarchical edge computing which enables scalability, very low latency. Furthermore, the deep learning algorithm has two drawbacks: one is the need for a large amount of dataset, and the other is the need for acknowledgement and clarification by the health care sector in order for the results to be accepted.

This paper proposes B5G tech which can detect COVID symptoms using CT scan and chest X-ray images. Also, deep learning models like ResNet, Deep tree and Inception v3 along with blockchain technology are analyzed to provide security of data.

[12] "MASK R-CNN for Pedestrian Crosswalk Detection and Instance Segmentation" by Mon Arjay Malbog. Mask R-CNN, which is based on deep convolutional networks (CNN) is an object detection model which is used for pedestrian crosswalk.

In this paper, for validation purpose almost 500 pedestrian crosswalk images were used. In this ,80% of the images are used for the training set while 20% is used for the validation set. Mask R-CNN is a two-step process that begins with (RPN) Region Proposal Network and ends with class and box offset prediction. It includes data gathering, Annotation and Construction of data set, Training the Model and Mask R-CNN detection model.

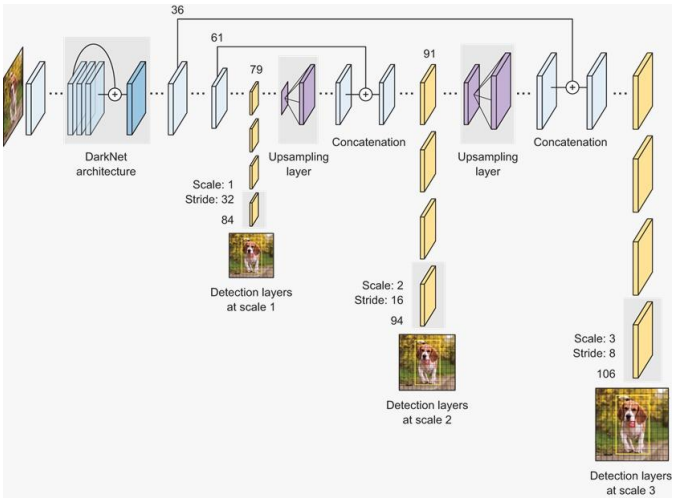

Fig – 1: Yolo V3 architecture

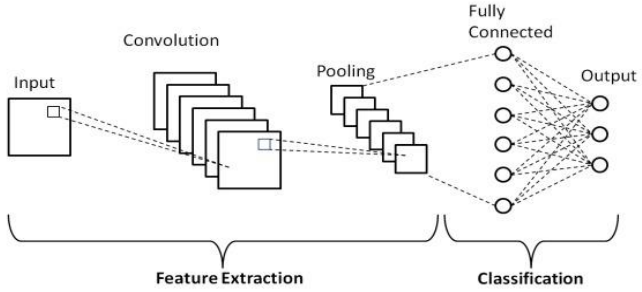

Fig -2: Convolutional Neural Network

3. DATASETS

MAFA-DATA:

The MAFA dataset is utilized to identify masked faces. Benchmark datasets are a grouping of datasets of varying sizes that represent real-world data science scenarios.

Mafa has 35,806 disguised faces and 30,811 photos.

Masked face images also differ in face angles and degrees of occlusion.

The significant identified features of masked faces consist of face orientation, eyes, face location, occlusion degree, masks, and mask types.

Notable offerings by two distinct contributors, the current activity status of MAFA dataset consists of 7150 views, 1239 downloads, and a 0.17 download per view ratio.

4. ARCHITECTURE

The system will use continuous video footage from the IP or CCTV camera as input. Images are created by processing this video footage frame by frame.

The system will identify the persons' faces, masks, and various attributes. The Euclidean distance between the faces will be calculated using the centroid of a rectangular box formed around each of the face. If there are more than three faces in that detection and the calculated distance is less than the prescribed standards, a notification is sent indicating absence of social distance.(See Fig-3)

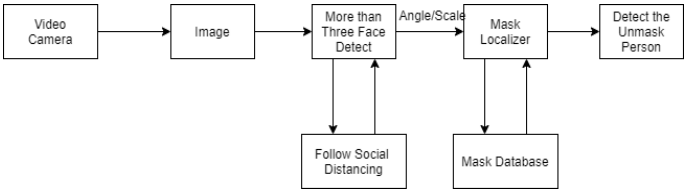

Fig- 3: Proposed System Architecture

Features will be extracted from those images, regardless of their angles or scales. The Yolo v3 algorithm would be used to accomplish this.

The mask localizer, a convolutional neural network-based classification model, receives the processed image. The MAFA dataset will be used to train the model and will help with image analysis. The system will detect if the person is not wearing a mask and notify.

5. EVALUATION METHODS

1] *Confusion Matrix:*

The metrics module in Python's popular Scikit-learn library can be used to determine the performance measures in the confusion matrix. Determination of worn mask is binary-class problem which can be solved using confusion matrix.

We employ two of its accepted parameters:

1. y true: The truth labels.

2. y pred: Predicted labels.

2] *Accuracy*:

Accuracy is a success metric that sums up how well a model performs throughout all the classes. It's helpful when all of the classes are equally important. (see Fig-4)

$$Accuracy = \frac{True_{positive} + True_{negative}}{True_{positive} + True_{negative} + False_{positive} + False_{negative}}$$

Fig -4: Accuracy

3] *Precision:*

Precision of the model is the ratio of True Positives to all positive predictions.

The lower the number of false positives predicted by the model, the higher the precision. (see Fig-5)

$$Precision = \frac{True_{positive}}{True_{positive} + False_{positive}}$$

Fig-5: Precision

4] *Recall*:

The ratio of True Positives to all positives in your dataset is known as recall. It is also known as sensitivity.

The lower the number of False Negatives predicted by the model, the higher the recall. (see Fig-6)

$$Recall = \frac{True_{positive}}{True_{positive} + False_{negative}}$$

Fig-6: Recall

6. CONCLUSIONS

Current research work is helpful in new normal social life and industrial life. The researchers could study prominent

ML algorithms and suggested useability of CNN and Yolo v3 as the simplest and appropriate methods to build an application on. Due to strong library support and being apt language for machine learning, Python can be used for the implementation.

The proposed system architecture identified six important characteristics of custom MAFA dataset to identify the face mask. Model is trained on this MAFA dataset. Research suggests Yolo v3 can be successfully used to check social distancing.

7. FUTURE SCOPE

In further research, an application will be developed that will detect whether someone is wearing a mask and will also identify whether people are maintaining social distance. If a person is recognized to be without mask, alert will be sent. The projected system architecture can be further modified using hardware, sensors to accommodate thermal scanning features.

The proposed framework is not limited to mask and social distancing detection but ,it can be integrated into a high-resolution video surveillance device. The system can also be implemented and expanded in biometric face-recognition systems to detect features of faces with facemasks.

This application can be used in any work context where predictive performance is needed, such as a public place, a station, an industrial environment, streets, commercial buildings, and examination centers.

This method can be applied to smart city innovation, that would help many developing countries ramp up their development. The current research work with actual application development can thus become a part and parcel of new normal and be integrated at all crowdy places.

8. ACKNOWLEDGEMENT

Researchers would like to convey gratefulness to their project mentor, Prof. Anuradha S. Deokar for her competent assistance and advice throughout the survey and implementation process of this project.

9. REFERENCES

[1] J. Barabas, R. Zalman and M. Kochlan, "Automated evaluation of COVID-19 risk factors coupled with real-time, indoor, personal localization data for potential disease identification, prevention and smart quarantining," 2020 43rd International Conference on Telecommunications and Signal Processing (TSP), 2020, pp. 645-648, doi: 10.1109/TSP49548.2020.9163461.M.

Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[2] S. Li et al., "Multi-angle Head Pose Classification when Wearing the Mask for Face Recognition under the COVID-19 Coronavirus Epidemic," 2020 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS), 2020, pp. 1-5, doi: 10.1109/HPBDIS49115.2020.9130585.K. Elissa, "Title of paper if known," unpublished.

[3] T. Ikram, A. Saeed, N. Ayn, M. A. Tahir and R. Mumtaz, "A review of the prevalent ICT techniques used for COVID-19 SOP violation detection," 2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET), 2020, pp. 194-198, doi: 10.1109/HONET50430.2020.9322821.

[4] D. R. Shenvi and K. Shet, "CNN Based COVID-19 Prevention System," 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS),2021,pp.873-878,doi: 10.1109/ICAIS50930.2021.9396004.

[5] G. Deore, R. Bodhula, V. Udpikar and V. More, "Study of masked face detection approach in video analytics," 2016 Conference on Advances in Signal Processing(CASP),2016,pp.196-200,doi: 10.1109/CASP.2016.7746164.

[6] H. S. Maghded, K. Z. Ghafoor, A. S. Sadiq, K. Curran, D. B. Rawat and K. Rabie, "A Novel AI-enabled Framework to Diagnose Coronavirus COVID-19 using Smartphone Embedded Sensors: Design Study," 2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science(IRI),2020,pp.180-187,doi: 10.1109/IRI49571.2020.00033.

[7] A. Rahman, M. S. Hossain, N. A. Alrajeh and F. Alsolami, "Adversarial Examples—Security Threats to COVID-19 Deep Learning Systems in Medical IoT Devices," in IEEE Internet of Things Journal, vol. 8, no. 12, pp. 9603-9610, 15 June15, 2021, doi: 10.1109/JIOT.2020.3013710.

[8] Z. Jiang et al., "Detection of Respiratory Infections Using RGB-Infrared Sensors on Portable Device," in IEEE Sensors Journal, vol. 20, no. 22, pp. 13674-13681,15Nov.15,2020,doi: 10.1109/JSEN.2020.3004568.

[9] Buddhisha Udugama, Pranav Kadhiresan, Hannah N. Kozlowski, Ayden Malekjahani, Matthew Osborne, Vanessa Y. C. Li, Hongmin Chen, Samira Mubareka, Jonathan B. Gubbay, and Warren C. W. Chan* "Diagnosing COVID-19: The Disease and Tools for Detection", American Chemical Society, year 2020, doi: 10.1021/acsnano.0c02624.

[10] S. -J. Yook et al., "Experimental Investigations on Particle Contamination of Masks Without Protective Pellicles During Vibration or Shipping of Mask Carriers," in IEEE Transactions on Semiconductor Manufacturing, vol. 20, no. 4, pp. 578-584,Nov.2007,doi: 10.1109/TSM.2007.907635.

[11] M. S. Hossain, G. Muhammad and N. Guizani, "Explainable AI and Mass Surveillance System-Based Healthcare Framework to Combat COVID-I9 Like Pandemics," in IEEE Network, vol. 34, no. 4, pp. 126-132, July/August 2020, doi: 10.1109/MNET.011.2000458.

[12] M. A. Malbog, "MASK R-CNN for Pedestrian Crosswalk Detection and Instance Segmentation," 2019 IEEE 6th International Conference on Engineering Technologies and Applied Sciences (ICETAS),2019,pp.1-5,doi: 10.1109/ICETAS48360.2019.9117217.

[13] Q. -C. Mao, H. -M. Sun, Y. -B. Liu and R. -S. Jia, "Mini-YOLOv3: Real-Time Object Detector for Embedded Applications," in IEEE Access, vol. 7, pp. 133529-133538,2019, doi: 10.1109/ACCESS.2019.2941547

[14] https://livebook.manning.com/book/grokking-deep-learning-for-computer-vision/chapter-7/v-8/

[15] A High-Accuracy Model Average Ensemble of Convolutional Neural Networks for Classification of Cloud Image Patches on Small Datasets - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/Schematic-diagram-of-a-basic-convolutional-neural-network-CNN-architecture-26_fig1_336805909

10. BIOGRAPHIES

GAURAV GARJE
BE COMPUTER
AISSMSCOE, PUNE

RASIKA LANDE
BE COMPUTER
AISSMSCOE, PUNE

SHARDUL SHINDE
BE COMPUTER
AISSMSCOE, PUNE

NISHANT INAMDAR
BE COMPUTER
AISSMSCOE, PUNE

PROF. ANURADHA DEOKAR
AISSMSCOE, PUNE