Optimized Face Mask Recognition With Edge Deployment Framework In Guard Bots Using Deep Learning

Aarthy R[1] , Navaneethakrishnan S[2],Vishnupriyan V[3],Prasanth S[3] , Vijay M[3]

[1]Assistant Professor, Department of Computer Science and Engineering, Dhanalakshmi Srinivasan Engineering College, Perambalur
[2]Student, Department of Computer Science and Engineering, Dhanalakshmi Srinivasan Engineering College, Perambalur
[3]Student, Department of Computer Science and Engineering, Dhanalakshmi Srinivasan Engineering College, Perambalur

***Abstract*—** The most recent outbreak to prompt an international health emergency is coronavirus illness. It primarily spreads via airborne transmission from person to person. The number of cases worldwide has increased as a result of community transmission. Several machine learning-based techniques have been used in the health field. Lack of data is one problem preventing machine learning techniques from discovering COVID-19 cases. A Mask-RCNN that can accurately identify face masks can alert users to wear one. To do this, Mask-RCCN employs two cutting-edge techniques. First, we present a unique residual context attention module (RCAM) to extract rich context data, focus on key face mask related regions, and detect mask region from the face using RPN. To develop more distinguishing traits for faces with and without masks, second. This method can distinguish between masked and uncovered faces, making it easier to utilize face masks and keeping an eye out for safety violations.

I. Introduction

The most reliable algorithms only successfully verify a person 0.3% of the time when using unmasked photos. Even these top algorithms had a failure rate of roughly 5% when faced with masked images, whereas the failure rate of many other, more capable algorithms ranged from 20% to 50%.The technical term for this is "failure to enroll or template" (FTE), which occurs more commonly when algorithms are unable to process a face in masked images. In order to identify faces, face recognition algorithms often evaluate a face's attributes, such as their size and distance from one another, and then compare these measurements to measurements from a different photo. An FTE indicates that the algorithm was unable to accurately extract a face's features in order to perform a comparison in the first place.

II. FRAMEWORK

1.1.1. Neural networks

There are different kinds of neural networks,

CNN

Which are mainly utilized for computer vision or image processing tasks. CNNs are very good at modelling spatial data, including two-dimensional or three-dimensional images and videos, and they can extract features and patterns from an image to support tasks like classification of images or object detection.

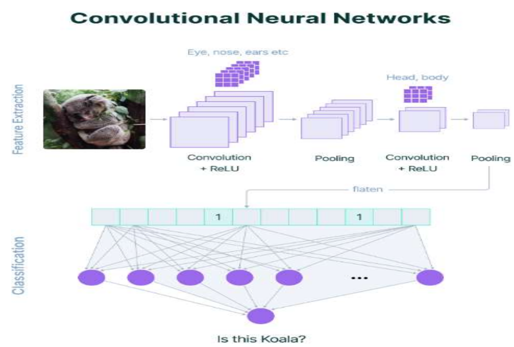

Figure 1 . CNN

III. SYSTEM ANALYSIS

3.1. Existing System

A large number of intelligent models for masked face recognition (MFR) has been recently presented and applied in various fields, such as masked face tracking for people safety or secure authentication. Exceptional hazards such as pandemics and frauds have noticeably accelerated the abundance of relevant algorithm creation and sharing, which has introduced new challenges.

3.2. Disadvantages

- It is only capable of detecting faces up close.
- Handmade element
- Highly Complicated Computing

3.3. Proposed System

Face masks, which are available in a variety of designs, sizes, textures, and hues, have recently been one of the most widely used products to conceal the facial features. This makes it much more important to train deep learning algorithms to recognize masks accurately. The process of masks detection involves tuning and research into the majority of the currently used detection techniques, which are often introduced for object detection.

Region Proposal Network

The backbone layer-generated convolution feature map is the input for this area proposal network, which outputs the anchors produced by the window sliding convolution on the feature map.

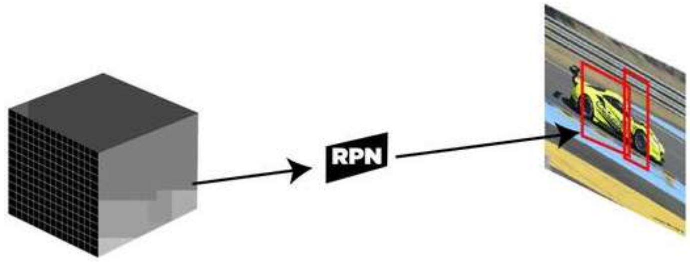

Figure 2. RPN

Mask R-CNN

Mask R-CNN, sometimes known as Mask RCNN, is the most advanced Convolutional Neural Networks, for instance and picture segmentation. Faster R-CNN, a region-based convolutional neural network, served as the foundation for Mask R-CNN.

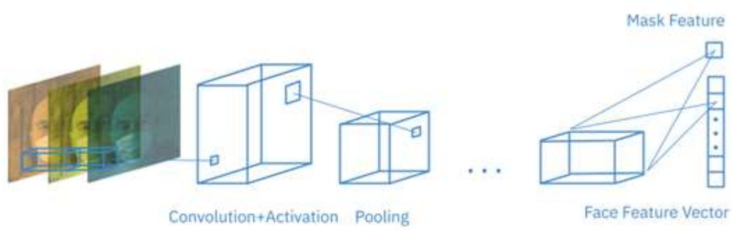

Figure 3. Mask R-CNN

The mask R-CNN is adjusted based on the quicker R-CNN, as illustrated in Fig. 6. In particular, the region of interest (RoI) aligning layer is substituted for the region of interest pooling layer by the mask RCNN. This region that is important alignment layer, which is better suited for pixel-level prediction, uses bi linear interpolation to maintain the location data on the feature maps.

Futures of Mask R-CNN

Mask RCNN is a type of deep neural network designed to address instance segmentation issues in computer vision or machine learning. A simple, adaptable, and all-encompassing framework for object instance segmentation is Mask R-CNN. It is capable of accurately identifying items in a picture while also producing a top-notch segment mask for every instance.

3.4. Advantage

- Both the detection's speed and precision have significantly increased.
- Lowered the computational complexity significantly.
- Estimate the Face and masks accurately.
- Warning Generation

3.5. System Architecture

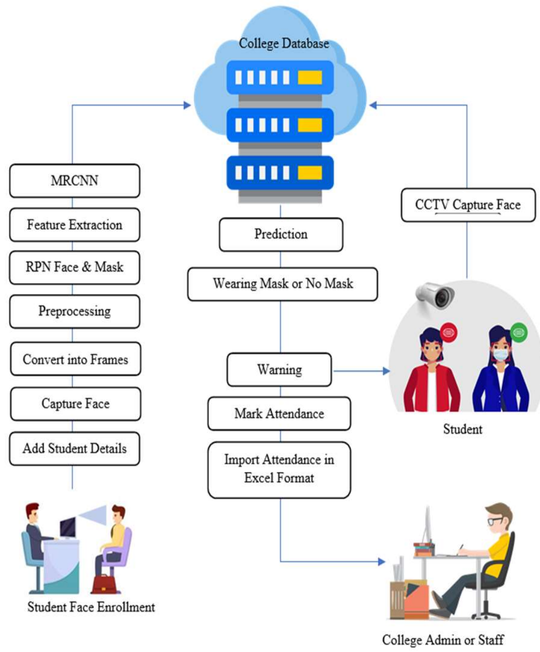

Figure 4. System Architecture

3.6. Flow Diagram

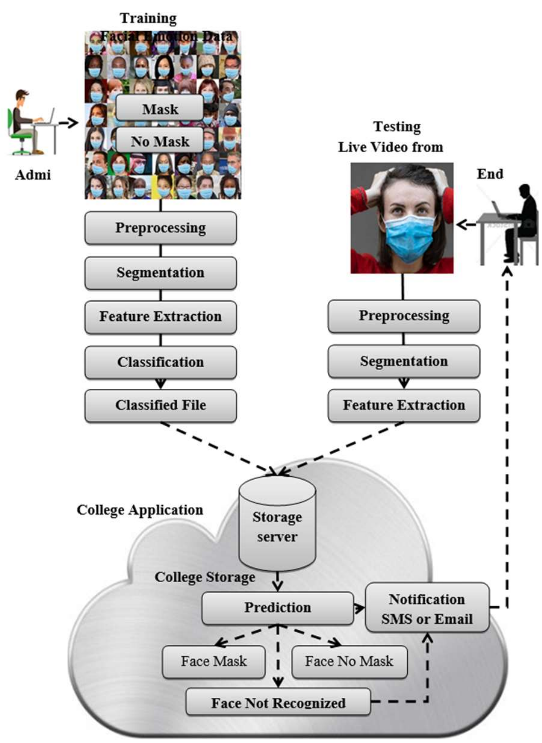

Figure 5. Flow Diagram

IV. SYSTEM IMPLEMENTATION

4.1. Problem Description

The widely utilized Mask - Region Convolutional Neural Network (M-RCNN) is the specific Deep Learning model employed in this research. they are then processed to meet the following functional requirements:

Mask Detection : To determine if a person is wearing a mask or not, partition the captured face into regions of interest. From there, you can detect the feature (mask) present in those regions of interest.

Facial recognition with mask on: Mark attendance, automatically identify staff members without removing their face masks, and deliver matching identification results from a chosen database.

4.2. Module Description

1. College Management Dashboard

In this module, we created a web-based interface for the attendance system that is intended to be used by college administrators, teaching staff, and students who are not skilled in data science.

2.Student Management Module

The student module enables you to keep track of a student's parents' personal, academic, professional, and historical information.

1.2. Student Face Enrolment Phase

The student information database, the face database, and the attendance database are all kept on the student database server in this system. Database of student information includes each student's roll number, name, and class. Student attendance status is recorded in the attendance database daily.

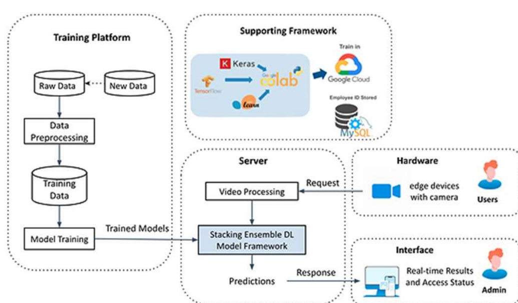

Figure 6 .Student Face Phase Model

4.2.1. Face Image Acquisition-

This module is initial part of the system. Logitech C270 (3MP) is used for image acquisition.

4.2.1.1. Frame Extraction

From the video input, frames are extracted. The video has to be cut up into sequences of images for later processing. From this, we may infer that 20–30 frames are typically captured every second and forwarded to the following stages.

4.1.2. Pre-processing

Face Image pre-processing are the steps taken to format images before they are used by model training and inference.

The steps to be taken are:

- Read image
- RGB to Grey Scale conversion
- Resize image
- Remove noise (Denoise).
- Binarization

Image binarization reduces a grayscale image's 256 shades of grey to just two: black and white, or a binary image. It does this by taking the image and turning it into a black-and-white image.

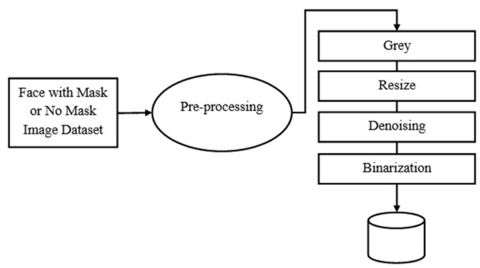

Figure 7 . Pre Processing

4.1.3. Face and Mask Detection

As a result, the Region Proposal Network (RPN) creates RoIs in this module by swiping windows on the feature map over anchors with various scales and aspect ratios. Based on improved RPN, this approach segments faces and masks.

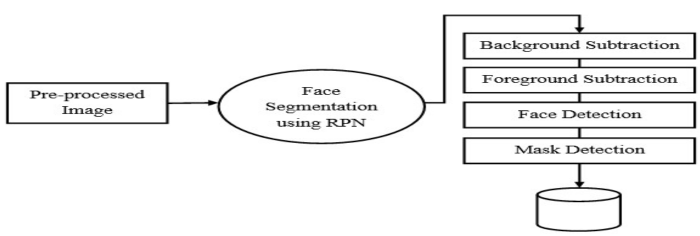

Figure 8 . Face Segmentation

RPN

A fully convolutional network known as an RPN predicts object limits and objectless values at each place at the same time. The RPN has received comprehensive training to produce superior region suggestions. Each feature (point) on the CNN feature map, on which it operates, is referred to as an Anchor Point. We overlay the image with nine anchor boxes (combinations of various sizes and ratios) for each anchor point.

These anchor boxes are cantered at the location in the image that corresponds to the feature map's anchor point.

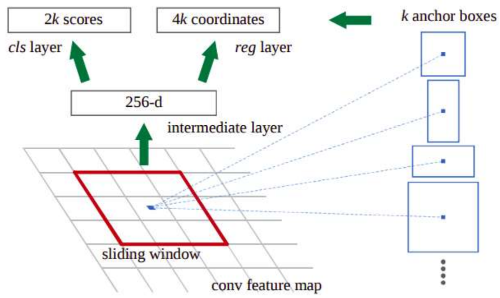

Figure 9 . RPN

Training of RPN.

Therefore, depending on each anchor box's Intersection over Union (IoU) with the provided ground truth, assign a label to each one for training purposes. We essentially give each anchor box one of the three labels (1, -1, 0).

Label 1: (In the foreground) The following circumstances allow an anchor to have label 1:

If the anchor's IoU with the ground truth is the highest.

If the ground truth IoU is greater than 0.7. (IoU>0.7).

Label = -1 (Context): If IoU 0.3, an anchor is given a -1 assignment.

Label = 0: These kinds of anchors don't help with training and are disregarded if they don't fit within either of the two parameters mentioned above.

The processing steps are

- Select the initial seed point
- Append the neighboring pixels—intensity threshold
- Check threshold of the neighboring pixel
- Thresholds satisfy-selected for growing the region.
- Process is iterated to end of all regions.

5.1.4. RoIAlign Layer

The segmentation procedure in G-Mask, in contrast to standard face detection algorithms, necessitates more accurate spatial quantization for feature extraction. The common method for extracting small feature maps from RoIs, which include two quantization operations and lead to misalignments between the RoI and the retrieved features, in the classic region-based approaches is RoIPool. Classification and localization may not be impacted by conventional detection methods, but they are for our approach, where it has a significant effect on pixel-accurate mask prediction and small object identification.

Mask Branch

The G-Mask model, which predicts the segmentation mask in a pixel-to-pixel way by applying Full Convolutional Network (FCN) to each RoI, uses the mask branch to realize the segmentation of the face object and background picture. The FCN scheme, which comes from CNN but differs from standard CNN, is one of the answers for things like segmentation. The convolutional layer is typically connected with several full connection layers in the traditional CNN network architecture in order to obtain the feature vector of fixed dimensions. Finally, the output is a numerical description of the input, which is typically applicable to tasks like image recognition and classification, object detection, and positioning.

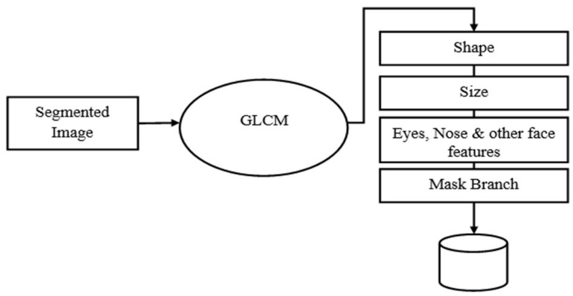

Figure 10 . GlCM

5.1.5. Masked Face Classification

MRCNN

The basic region-based model has two stages, and the proposed G-Mask model also has two stages. RPN suggests potential bounding boxes for the object face in the initial stage. The second stage, using the Fast R-CNN architecture, performs classification and bounding box localization after extracting features from each candidate box. In addition, we paralleled the classification branch and the bounding box location branch with a mask branch similar to the Mask R-CNN.

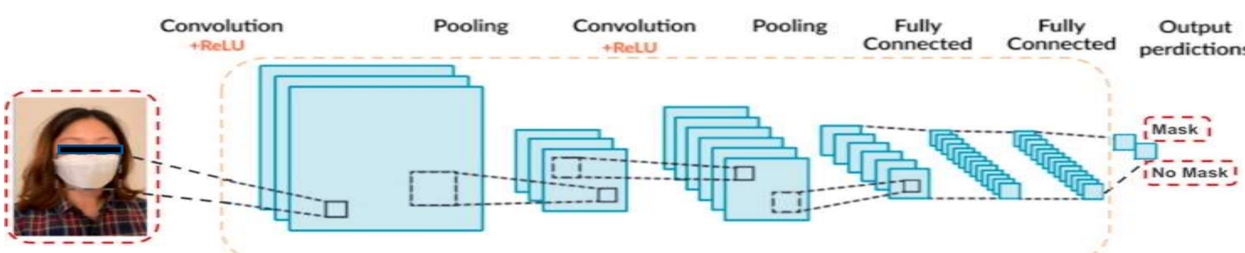

Figure 11 . MRCNN

2. Attendance System

After successful face recognition and face verification, the student's attendance is noted next to his or her roll number. An error page is shown if the face cannot be recognised.

3. End User Module

3.1. Admin

On the admin end, a user interface (UI) displaying real-time results with timestamp and access status is used to showcase the detection results. Python, CSS, and HTML were used to create the user interface. The system will be run by admin users, who can view the interface's real-time results.

3.2. Staff

Management of the staff through registration, attendance tracking, and class reports. The staff is able to update information on student attendance, internal grades, and any information pertaining to the courses they teach.

3.3. Student

The student is the main focus because they play such a vital part in any college. The college's information, subject information, training and placement cell information, and exam section information are all accessible to students.

4. Customized Reports

The college management software will provide hundreds of reports to aid in making various decisions.

V. RESULTS AND DISCUSSION

5.1 Evaluation Metrics

According to the project's context, the key topics associated with performance indicators are discussed: True Positive (TP): The algorithms identify a face and both a mask-covered and -uncovered face. False Positive (FP): The algorithms identify a face with a mask or one without even when there isn't any. False Negative (FN): Although a face exists, the algorithms do not face it with a mask or without one. True Negative (TN).Neither a Face nor anything else is being detected.

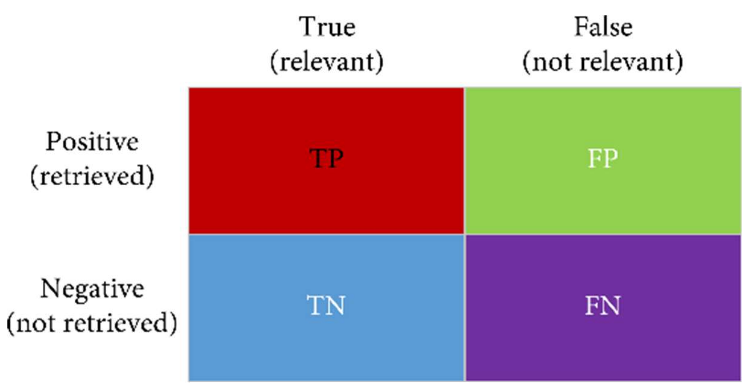

Figure. 12 . Parameter Definition

F1 Score

It is sometimes referred to as a balanced F-measure or F-score. A model's accuracy is evaluated using its F1 score, which combines precision and recall.

If a model's or algorithm's F1 score is 1, it is deemed perfect. This formula is used to calculate it.

F1 = 2 × (Precision × Recall /Precision + Recall)

F1_score: 0.9977122020583142

VI. Conclusion

Under the spread of COVID-19 pandemic, wearing a protective face mask has become a normal and requirement for many public services, colleges and essential business providers. Both WHO and CDC also stress on the importance and effectiveness of wearing correct masks for personal and public health. Therefore, face mask detection and facial recognition with a mask are essential for our society. To serve the above purpose, this paper proposed a Face Mask Detection and Facial Recognition with Mask System for access, attendance and health check under the pandemic. In this project, a Mask Region Convolutional Neural Network method was proposed for face detection and recognition with mask or no masked face. The approach can extract features by GLCM, generate RoIs by RPN, preserve the precise spatial position by RoI Align, and generate binary masks through the full convolutional network (FCN). In doing so, the proposed framework is able to detect faces correctly while also precisely segmenting and predict each face and mask or no mask in an image. Based on the testing results, each model performed relatively well and were able to accurately detect their classes for each feature. An average of 97% accuracy is achieved with proposed models and it indicates that the models can be integrated together and tested on real-time videos. The results show the balance of limited computing resources and high performance. Along with this project, an end-to-end solution is provided with video acquisition, database design and high-level data analytics. The system and the solution can be easily used by small businesses, organizations and universities with minimum cost under the COVID-19 and help to practice social distancing.

Future Enhancement

In future work, we would like to build face mask detection datasets with no, correct and incorrect mask wearing states, or use a zero-shot learning approach to make the model able to detect incorrect mask wearing states. New deep learning detectors may be used to further improve the performance.

REFERENCES

1. F. Boutros, N. Damer, J. N. Kolf, K. Raja, F. Kirchbuchner, R. Ramachandra, A. Kuijper, P. Fang, C. Zhang, F.Wang, and D. Montero, ``MFR 2021: Masked face recognition competition," in Proc. IEEE Int. Joint Conf. Biometrics (IJCB), Aug. 2021, pp. 1-10.
2. J. Deng, J. Guo, X. An, Z. Zhu, and S. Zafeiriou, ``Masked face recognition challenge: The InsightFace track report," 2021, arXiv:2108.08191.
3. Y. Martinez-Diaz, M. Nicolas-Diaz, H. Mendez-Vazquez, L. S. Luevano, L. Chang, M. Gonzalez-Mendoza, and L. E. Sucar, ``Benchmarking lightweight face architectures on specific face recognition scenarios,"Artif. Intell. Rev., vol. 54, pp. 6201-6244, Feb. 2021.
4. Y. Li, K. Guo, Y. Lu, and L. Liu, ``Cropping and attention based approach for masked face recognition," Int. J. Speech Technol., vol. 51, no. 5, pp. 3012-3025, May 2021.

5. P. Kumari and K. Seeja, ``A novel periocular biometrics solution for authentication during COVID-19 pandemic situation," J. Ambient Intell. Humanized Comput., vol. 12, pp. 10321-10337, 2021.
6. Anwar.A and A. Raychowdhury, ``Masked face recognition for secure authentication," 2020, arXiv:2008.11104.
7. N. Damer, J. H. Grebe, C. Chen, F. Boutros, F. Kirchbuchner, and A. Kuijper, ``The effect of wearing a mask on face recognition performance: An exploratory study," in Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG), Sep. 2020, pp. 1-6.
8. J. Deng, J. Guo, E. Ververas, I. Kotsia, and S. Zafeiriou, ``RetinaFace: Single-shot multi-level face localisation in the wild," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 5203-5212.
9. I. Q. Mundial, M. S. U. Hassan, M. I. Tiwana, W. S. Qureshi, and E. Alanazi, ``Towards facial recognition problem in COVID-19 pandemic," in Proc. 4rd Int. Conf. Electr., Telecommun. Comput. Eng. (ELTICOM), Sep. 2020, pp. 210-214.
10. W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, ``SphereFace: Deep hypersphere embedding for face recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 212-220.
11. C. Li, S. Ge, D. Zhang, and J. Li, ``Look through masks: Towards masked face recognition with de-occlusion distillation," in Proc. 28th ACM Int. Conf. Multimedia, Oct. 2020, pp. 3016-3024.
12. Y. G. Jung, J. Park, C. Y. Low, L. C. O. Tiong, and A. B. J. Teoh, ``Periocular in the wild embedding learning with cross-modal consistent knowledge distillation," 2020, arXiv:2012.06746.
13 . Martindez-Diaz, L. S. Luevano, H. Mendez-Vazquez, M. Nicolas-Diaz, L. Chang, and M. Gonzalez-Mendoza, ``ShuffleFaceNet: A lightweight face architecture for efficient and highly-accurate face recognition," in Proc. IEEE/CVF Int. Conf. Comput. Vis. Workshop (ICCVW), Oct. 2019, pp. 1-8.