AUTOMATED AND SMARTEXAM MONITORING SYSTEM USING MACHINE LEARNING

Mrs . Sanjitha S
Department of Information Science and Engineering
East West Institute of Technology
Bangalore,India
sanjitha94@gmail.com

Harshitha P S
Department of Information Science and Engineering
East West Institute of Technology
Bangalore,India
harshithaps2001@gmail.com

Mahima Manjunath Bhat
Department of Information Science and Engineering
East West Institute of Technology
Bangalore,India
mahimambhat166gmail.com

Niriksha N
Department of Information Science and Engineering
East West Institute of Technology
Bangalore,India
nirikshaneha@gmail.com

Spoorthi M J
Department of Information Science and Engineering
East West Institute of Technology
Bangalore,India
spoorthimj291@gmail.com

***Abstract*— Exams are a significant part of our educational system's assessment and evaluation of student understanding. However, after the release of COVID-19, remote learning has flourished. We are presenting a means of conducting exams online with remote proctoring by utilizing Face detection technology in an effort to make them considerably smoother and more valuable. Eyeball tracking, mouth opening detection utilizing the distance between lips, device detection to discover any instances of the device, instance segmentation to count the number of people, and head posture estimation to determine the direction of the person's gaze are all used in this. These features can be used to forecast the likelihood that someone will cheat or engage in improper behavior in a webcam-proctored situation. To keep track of everything, the entire action of the individual in front of the camera is documented in a file structure**

Keywords*— Face recognition, head pose estimation, eye tracking, mouth opening detection.

I. INTRODUCTION

The development of remote learning has been facilitated by COVID-19. Despite the closure of schools and institutions, students continued their education using programs like Microsoft Teams. Exams have not yet found a solution, though. Some have simply canceled them, while others have altered them to assignment forms that students may malpractice from other sources. To live in the new society we need to find new solutions.We require a significant amount of data to train any face identification machine learning model. In order to identify person and cell phone devices in the webcam feed, we took advantage of the pre-learned weights of YOLOv3 trained on the COCO dataset. An alarm can be set off if the count falls outside of the range. The COCO dataset has 67 records for mobile phones, thus we must first determine if any class indexes match that number before reporting a mobile phone.

There are several applications for facial landmarks identification, also known as facial key points detection, in computer vision, including drowsiness detection, face alignment,, and Snapshot filters, to mention a few. The Dlib's 68 key points landmark predictor is the most well-known model for this task and produces excellent real-time results. However, the issue arises when the face is obscured or angled toward the camera. For jobs like estimating head poses, it's critical to obtain correct answers at angles. Therefore, we are presenting a lesser-known Tensorflow model that can accomplish this goal.

We explored with the issue of stable face landmarks in films, labeled out the current methods, such as Open Face and Dlib's facial landmark recognition, and used the datasets available. and about preparing data to make it usable. The faces must then be extracted, marked with facial landmarks to prepare them for CNN training, and stored as TFRecord files. Then Tensorflow is used to train a model. The learned model can then be used to forecast faces and other requirements.

II. LITRETURE SURVEY

1. "Design of face detection and recognition system to monitor students during online examinations using Machine Learning algorithms"

For a huge number of people, educational institutions administer tests online, making it challenging to manually keep track of the students. The suggested system is focused on creating an appropriate face detection and identification model for keeping an eye on the students during online tests.

2."A Systematic Review on AI-based Proctoring Systems: Past, Present and Future"

We have discussed the many kinds of human-based and AI-based systems that are now available on the market and those that have been proposed in various research publications. databases show a direct route to widespread global smart phone-assisted crop disease detection. Research on online proctoring in education is not brand-new. Many colleges and other institutions used proctoring methods for online courses even before the Pandemic. The GRE, GMAT, and CAT are only proctor-based tests, as are other competitive and adaptive exam.

3. "An intelligent system for online exam monitoring"

E-learning has gained popularity over the past few years due to its adaptability, accessibility, and user-friendliness. The proctoring methods employed are the main issue the research community is grappling with when it comes to online exams. In this study, we describe a technique for avoiding a proctor's physical presence throughout the exam by developing a thorough multi-modal system. We have utilized hardware, such as a webcam, to record audio and video in addition to actively capturing windows. An intelligent rule-based inference system can use this combination as input to determine whether any malpractices have occurred. The face of the examinee is recognized, and feature points are extracted from it to estimate a head pose. On the basis of changes in the yaw angle, the existence of audio, and active window capture, misconduct is discovered. We tested our solution in an e-learning environment and were able to make exam monitoring simple. The results of the experiment demonstrated that our system outperformed the current systems.

4. "E-proctored exams during the COVID-19 pandemic: A close understanding"

During the COVID-19 epidemic, research has concentrated on analyzing and investigating the online examination experience. However, there is still a lack of knowledge regarding how people perceive utilising an e-proctoring tool during an online exam. This study examines the views and concerns of students who used an e-proctoring tool for their final exams during the COVID-19 epidemic for the first time. Additionally, it emphasizes how the e-tools affect students' academic performance in order to direct educational institutions toward suitable practices moving forward, particularly given that the pandemic is anticipated to have a significant impact.

5. "Online exam proctoring system - IJAERD"

Students and institutions all across the world are using online education to access a wide range of information bases. This type of education and learning is spreading quickly, and the evaluation and proctoring of online courses has emerged as a significant barrier to their potential to scale. Manual human supervision is a typical method of exam proctoring and evaluation where the examiner must physically be there or must use a camera to visually and audibly supervise the test taker's testing environment.

6."Dimension Reduction of Hand and Face Feature Level Fusion in Multimodal Biometric Authentication."

The proposed work is a multimodal biometric authentication approach with image texture feature dimension reduction of trained feature vector which leads reduction in memory size and in turn reduces the computational time. In this paper hand and face features are used for person identification.

III. METHODOLOGY

1. Face Detection

Four face detection models were tested. Specifically, a Caffe model using OpenCV's DNN module, MTCNN, dlib frontal face detector, and Haar cascades. Briefly describe each of the below.

Haar Cascades: Paul Viola and Michael Jones first suggested them in their study titled "Rapid Object Detection with a Boosted Cascade of Simple Features" from back in 2001. Like the straightforward CNN, it pulls a lot of characteristics from photographs and is incredibly quick to use. Adaboost is then used to choose the best features. As a result, just 6000 features remain from the original 160000+ features. However, it will still take a long time to implement all of these functionalities in a sliding window. The features are then bundled in a Cascade of Classifiers, which was then introduced.[21]These subsequent characteristics in that cascade are not analyzed if a window fails at the initial level. If it succeeds, the process is repeated for the following feature. A window is categorized as a face region if it can pass all the characteristics.

Frontal Face Detector for Dlib: Dlib is a C++ toolbox that contains machine learning techniques for use in solving practical issues. Despite being built in C++, it can be executed in Python thanks to python bindings. It also features the fantastic facial landmark keypoint detector that I used to create a real-time gaze tracking system in one of my previous posts. The Histogram of Oriented Gradients (HOG)-extracted features used in the frontal face detector offered by dlib are then subjected to an SVM. The distribution of gradient directional information is employed as a feature in the HOG feature descriptor.

MTCNN: It was first discussed in a research titled "Joint Face Detection and Alignment Using Multi-task Cascaded Convolutional Networks" by Kaipeng Zhang et al. in 2016. In addition to the face, it also picks up on five crucial areas. Three CNN stages are used in a cascade arrangement.[20]They first collect candidate windows and their bounding box regression vectors using a fully convolutional network, and then they employ on-maximum suppression to overlap the highly overlapped candidates (NMS). These candidates are then given to a different CNN, which performs bounding box calibration and rejects a substantial proportion of false positives. The face landmark detection is done in the last step.

The Single Shot-Multibox Detector (SSD)-based DNN Face Detector in OpenCV is a Caffe model that employs ResNet-10 architecture as its foundation. It debuted in OpenCV's deep neural network module after version 3.3.Although a quantized Tensorflow version is also available, we'll utilise the Caffe Model instead.

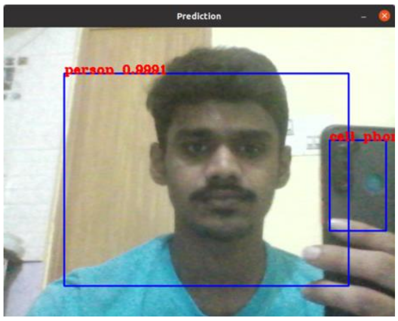

Fig. 1. Face and device detection

The Haar Cascade classifier provided the worst results in the majority of the tests, coupled with a significant number of false positives, according to our analysis of the aforementioned approaches. And while MTCNN somewhat outperformed Dlib in terms of outcomes,[23]Dlib is unable to recognizes very small faces. Additionally, after comparing the images, it became clear that MTCNN might produce the greatest results when the image size is extremely large, the lighting is certain to be good, there is little occlusion, and the majority of the faces are facing forward. Additionally, the Caffe model of the DNN module in OpenCV is the best for solving common computer vision issues. It is effective with occlusion, swift head motions, and can also recognize side faces. Additionally, it provided the highest frame rates of all.

2. Facial Landmark Detection

There are several applications for facial landmarks identification, also known as facial key points detection, in computer vision, including face alignment, drowsiness detection, and Snapshot filters, to mention a few. The Dlib's 68 key points landmark predictor is the most well-known model for this task and produces excellent real-time results. However, the issue arises when the face is obscured or angled toward the camera. For jobs like estimating head poses, it's critical to obtain correct answers at angles. The landmark prediction step in Dlib takes about 0.005 seconds and provides about 11.5 FPS.

Fig. 2. Facial Landmark detection

3. Eye Tracking

Before beginning image processing, we must first locate the eyes, and in order to locate the eyes, we must locate a face. A rectangle object of the dlib module that contains just a face's coordinates is what the facial keypoint detector uses as input. The frontal face detector built into dlib can be used to find faces. Any classifier will work for this purpose. I would advise using a CNN if you want high accuracy and speed is not a concern for you because it will provide much better accuracy, especially for non-frontal facing faces and partially occluded faces.

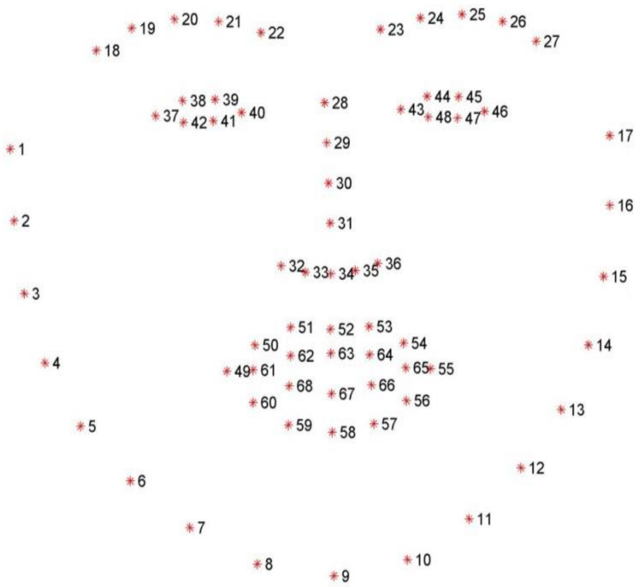

Fig.3. Dlib's Facial Key Points

4. Head Pose Estimation

The many stages needed to perform head posture estimation make it a difficult problem in computer vision. We must first identify the face in the frame before identifying the numerous facial landmarks. Today, identifying a face appears like a simple task, and this is especially true for faces that are facing the camera. When the face is angled, the issue manifests. Additionally, some facial landmarks are obscured by head movement. The many stages needed to perform head posture estimation make it a difficult problem in computer vision. We must first identify the face in the frame before identifying the numerous facial landmarks. Today, identifying a face appears like a simple task, and this is especially true for faces that are facing the camera. When the face is angled, the issue manifests. Additionally, some facial landmarks are obscured by head movement. Then, in order to determine the inclination, we must translate the points into 3D coordinates. Does that sound like a lot of work? Don't worry, we'll take it step by step and recommend two excellent tools that will make our job much simpler.

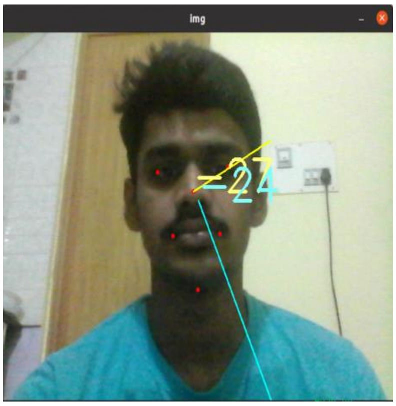

Fig. 4. Head Pose Estimation

5. Count people in webcam

Instance-based segmentation is what we use to count the number of persons in the video. In semantic segmentation, images are divided into different categories based on labels like a person, dog, cat, etc., however it is impossible to tell apart two person objects. This flaw is fixed by instance segmentation, which allows us to distinguish between many objects bearing the same label in addition to distinguishing between distinct labels. 80 objects may be classified using the YOLOv3 pre-trained model, which is extremely quick and almost as accurate as SSD.. It requires RGB images in the input format, so if you're using OpenCV, remember to convert them. The input type is float32, and the image sizes can range from 320x320 to 416x416 to 608x608.

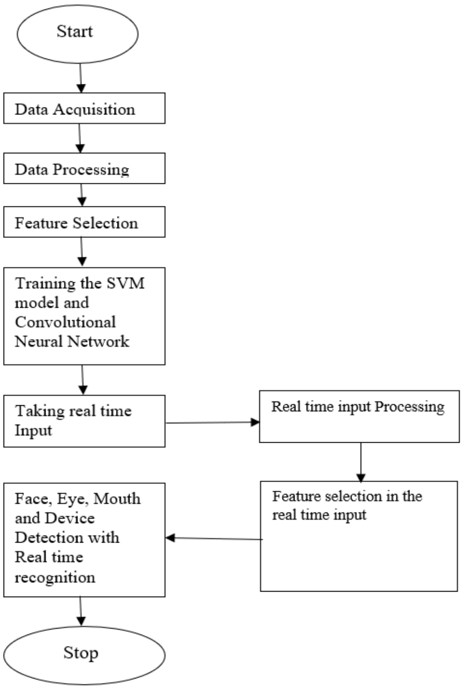

Fig. 5. Data Flow Diagram of training the model

IV. RESULTS

The FPS (frames per second) versus 300/300 image size graph is shown below. for the purpose of speed The MMOD detector can be run on a GPU, however OpenCV still doesn't support NVIDIA GPUs. Therefore, we solely analyze the methods on the CPU and also present the results for MMOD on the CPU and GPU. On the provided image, we execute each procedure a total of ten thousand times. We then average the times for these ten iterations. The outcomes are listed below.

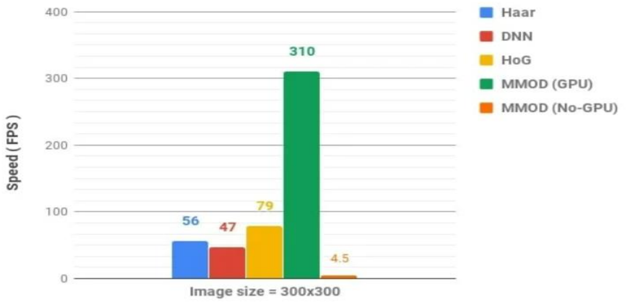

Fig. 6. Speed Comparison Results

As you can see, all methods—aside from MMOD—perform in real-time for images of this size. On a GPU, MMOD detector is quite quick, but on a CPU, it is rather slow.

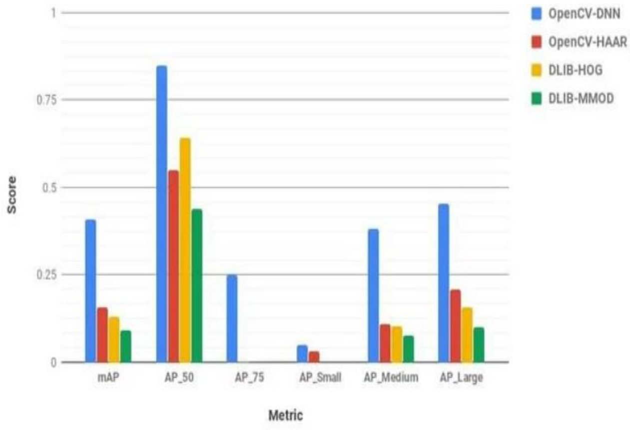

Fig. 7. Accuracy Comparison

It is clear that, in the perspective of accuracy OpenCV-DNN wins the race and gives the better results.

V. CONCLUSION AND FUTURE WORK

In many situations, size of the face in the image is not known in advance, according to the experimentation. It is therefore preferable to utilize the OpenCV-DNN approach because it is quite quick and extremely accurate, especially for small-sized faces. It can also recognize faces from a variety of angles. In most cases, we advise using OpenCV-DNN. The quickest CPU approach for photos with a medium resolution is Dlib HoG. However, it cannot recognize faces that are smaller than 70x70. Additionally, MMOD's face detector is the best choice if you have access to a GPU because it runs quickly and offers detection from a variety of angles. Since providing high resolution images to these algorithms (for processing speed) is not possible, HoG / MMOD detectors may not work properly when you scale the image. However, since the OpenCV-DNN approach can recognize small faces, it can be used to these. The lowest results are typically produced by Haar, which is somewhat outdated. Although the DNN module of OpenCV's face detection model performs well, it can have issues with really large images. Since we rarely use 3000x3000 photos, there shouldn't be a problem. If you're working with little photographs, make sure to upscale them because Dlib cannot recognize faces less than 80x80. However, this can lengthen the processing time. Therefore, several models can be utilized successfully to achieve the goal for various use cases.

REFERENCES

[1] R. Bhatia, "Biometrics and Face Recognition Techniques," Int. J. Adv. Res. Comput. Sci. Softw. Eng., Vol. 3, 2013

[2] S. Syed Navaz, T. D. Sri, P. Mazumder, and A. Professor, "Face Recognition Using Principal Component Analysis and Neural Networks," Int. J. Comput. Networking, Wirel. Mob. Commun., Vol. 3, 2013.

[3] M. K. Dhavalsinh V. Solanki, "A Survey on Face Recognition Techniques," J. Image Process. Pattern Recognit. Prog., Vol. 4, 2013.

[4] AI Maria Joseph and B Asoke Nath. Integration of massive open onlineeducation (mooc) system with in-classroom interaction and assessment and accreditation: An extensive report from a pilot study. In Proceedings of the international conference on e-learning, e-business, enterprise information systems, and e-Government (EEE), page 105. The Steering Committee of The World Congress in Computer Science, Computer ..., 2013

[5] A. Fayyoumi, A. Zarrad, 'Novel solution based on face recognition to address identity theft and cheating in online examination systems', Adv Internet Things, Vol.4, No. 3, 2014, pp. 5–12.

[6] C. Zhu, Y. Zheng, K. Luu, M. Savvides. 'CMS-RCNN: Contextual Multi-Scale Region-Based CNN for Unconstrained Face Detection', In: Bhanu B., Kumar A. (eds) Deep Learning for Biometrics. Advances in Computer Vision and Pattern Recognition. Springer, Cham. 2017.

[7] Helaine M Alessio, Nancy Malay, Karsten Maurer, A John Bailer, and Beth Rubin. Examining the effect of proctoring on online test scores. Online Learning, 21(1):146–161, 2017.

[8] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, M. Nießner, 'FaceForensics: A Large-scale Video Dataset for Forgery Detection in Human Faces', Computer Vision and Pattern Recognition, 2018

[9] Rebecca Kantar, Keith McNulty, Erica L Snow, Richard Wainess, Sonia

[10] D Doshi, Devon B Walker, and Matthew A Emery. The rise ofimmersive cognitive assessments: Towards simulation-based assessment for evaluating applicants. In International Conference on Artificial Intelligence in Education, pages 438–441. Springer, 2018.

[11] A. A. Sukmandhani and I. Sutedja, 'Face Recognition Method for Online Exams', In: International Conference on Information Management and Technology (ICIMTech), Jakarta/Bali, Indonesia, 2019, pp. 175-179.

[12] P. Venkateswar Lal, G. R. Nitta, A. Prasad, 'Ensemble of texture and shape descriptors using support vector machine classification for face recognition' J Ambient Intell Humaniz Comput, 2019.

[13] Design of face detection and recognition system to monitor students during online examinations using Machine Learning algorithms by M.Geetha, R.S.Latha, S.K.Nivetha, S.Hariprasath, S.Gowtham, C.S.Deepak.

[14] S. M Bah, F. Ming, 'An improved face recognition algorithm and its application in attendance management system', Array, Vol. 5, 2020.

[15] B. Kranthikiran, P. Pulicherla, 'Face Detection and Recognition for use in Campus Surveillance', International Journal of Innovative Technology and Exploring Engineering, 2020, Vol. 9, No.3.

[16] Z. Yuan, 'Face Detection and Recognition Based on Visual Attention Mechanism Guidance Model in Unrestricted Posture', Scientific programming towards a smart world, 2020.

[17] Phillip Dawson. Defending assessment security in a digital world: preventing e-cheating and supporting academic integrity in higher education. Routledge, 2020.

[18] Thomas Langenfeld. Internet-based proctored assessment: Securityand fairness issues. Educational Measurement: Issues and Practice, 39(3):24–27, 2020.

[19] Submitted to Uttar Pradesh Technical University.

[20] Submitted to Babes – Bolyai University.

[21] Submitted to CSU, San Jose State University.

[22] M.Geetha, R.S.Latha, S.K.Nivetha, S.Hariprasath, S. Gowtham, C.S. Deepak."Design of face detection and recognition system to monitor students during online examinations using Machine Learning algorithms", 2021 International Conference on Computer Communication and Informatics(ICCCI), 2021.

[23] Pranjali Singh,Amritpal Singh," chapter 43 Unmasking the Masked Face Using Zero-ShotLearning", Springer Science and Business Media LLC, 2022.

[24] Submitted to Cranfield University.

[25] "Hybrid Intelligent System", Springer Science and Business Media LLC,2021.

[26] Aditya Nigma, Rhituik Pasricha, Tarishi Singh Prathamesh Churi. "A Systematic Review onAI-based Proctoring Systems: Past, Present and Future", Education and Information Technologies, 2021.

[27] Faten F.Kharbat,Ajayeb S.Abu Daaber."E.Prathamesh Churi. "A Systematic Review onAI-based Proctoring Systems: Past, Present and Future", Education and Information Technologies, 2021

[28] Shankara Gowda, S. R., and A. N. Nandakumar. "Dimension Reduction of Hand and Face Feature Level Fusion in Multimodal Biometric Authentication." 2019 [IJITEE]

[29] Asha K H., Dr.Manjunathaswamy B.E., "Analysis of biometric modalities using iris face multimodel framework", Journal of Design Engineering (Totonto,Canada), A Scopus indexes Journal,ISSN:0011-9342, Vol,2021, Issue 9, pp. 15153 – 15172.

[30] Asha K H., Dr.Manjunathaswamy B.E., " Challenge responsive multimodel biometric authentication resilient to presentation attacks" International journal of intelligent engineering and systems, A Scopus indexed journal Vol.15 No. 2, 2022.