AUTOMATIC INVIGILATION SYSTEM USING YOLOv7 WITH DEEP LEARNING TECHNIQUES

M. Anita Indu[1], P. Girisha[2]

[1]Assistant Professor, PG & Research Department of Computer Science Shri Shankarlal Sundarbai Shasun Jain College for Women T.Nagar, Chennai, TamilNadu, Email : m.anitaindu@shasuncollege.edu.in
[2]UG Student, Department of Computer Applications Shri Shankarlal Sundarbai Shasun Jain College for Women T.Nagar, Chennai, TamilNadu, Email : girishapradeep@gmail.com

Article History:Received: 18.04.2023 **Revised:**07.05.2023 **Accepted:** 16.06.2023

Abstract: *Examination is one of the traditional methods of assessment in India. During invigilation face recognition system was used to detect the students cheating activities. The traditional method takes more time, requires human resource and there is more chances of fault. The continuous monitoring is done by the exam supervisor for avoiding the malpractice in traditional method. The existing system can be replaced by deep learning method to improve the efficiency of automatic invigilation system. Deep learning is used to automate predictive analysis through which we can gain a lot in education sector. The few applications in education sector are language translation, chat bots and service bots. This paper discusses about web cam based invigilation system which is followed through the face detection and hand recognition methods. The main objective is to develop a model to detect mischievous activities among the students during real time examination in the classroom. This can be detected through object detection, hand movement recognition along with the face recognition. This model will lead to automatic invigilation system using deep learning approach.*

Keywords : *Deep Learning Algorithms, Face detection, Hand Movement Recognition, YOLOv7, Mediapipe.*

DOI: 10.48047/ecb/2023.12.10.685

INTRODUCTION

Evaluation of the learning ability of a student is done through examination. In traditional method, the invigilator is allotted to examine the students in the classroom. The invigilator continuously monitors the student for avoiding malpractice like whispering, head movement and impersonation. In spite of this, the students are involved in several malpractice during the exam.

Today, Deep Learning techniques are used to monitor the students in the classroom efficiently. Face detection, Biometrics, hand recognition and object detection can be done easily through Deep Learning Algorithms. In the earlier research work, the automatic invigilation system was done. The existing system used some of the deep learning modules like Faster RCNN, MTCNN, YOLOV4, Haar Cascades, HoG for face detection and object recognition. The model was trained on the invigilation dataset using CNN algorithm. MTCNN algorithm was used for student identification and face recognition.

EXISTING SYSTEM

The existing system using YOLOv4 does not include the hand gestures and hand contact detection while passing the sheets or any device. Thus in the proposed system, the researcher used the deep learning modules like object detection, hand movement recognition and the face recognition to improve the automatic invigilation system. This can be done with the help of YOLOv7.

A. YOLO V7

YOLO (You Only Look Once) is one of the deep learning algorithms, used in real time application, makes the detection in a single forward propagation through a neural network. The speed and accuracy of this algorithm is very high when compared to others. Among deep learning algorithms, it requires cheaper hardware which made YOLOv7 to become more popular and hence it is used in current scenario.
There are many version of YOLO like YOLOv1 to YOLOv6. YOLOv7 was released in 2022. YOLO was first introduced by Joseph Redmon et al. in 2016 and after several versions YOLOv7 was released. The implementation is done in Pytorch instead of darknet. According to Upesh Nepal and Hossein Eslamiat[2], CSPDarknet53 backbone solves the repetitive gradient information in large backbones. It integrates gradient change into feature map. This integration reduces the inference speed, increases accuracy, and reduces the model size by decreasing the parameters.
Path aggregation network (PANet) is used to boost the information flow. It adopts a new feature pyramid network (FPN) to improve the propagation of low level features in the model. The localization was improved to enhance the accuracy of the object. This algorithm is popular because of its speed and accuracy with minimum background errors. The image background is eliminated by image pre-processing. Then Fast YOLO model was trained for object detection to obtain the object information in less time. The four models, namely s, m, l, and x are used in YOLOv7 for offering different detection accuracy and performance as shown below:

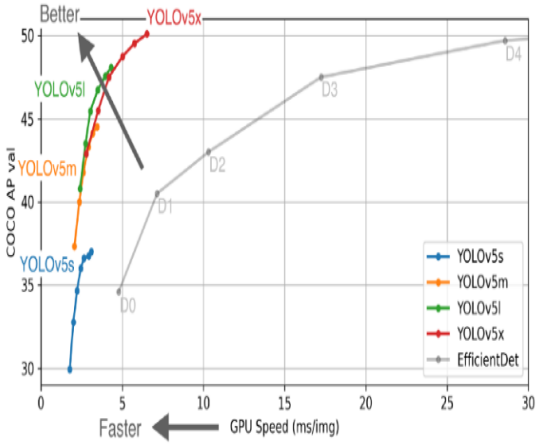

Fig 1 : Models of YOLOv7 (Source : https://github.com/ultralytics/yolov5)

YOLOv7 can effectively reduce about 40% parameters and 50% computation of state-of-the-art real-time object detections, and achieve faster inference speed and higher detection accuracy[1].

REVIEW OF LITERATURE

To determine the objects present in the image or video and specify their location in image through classification and localization, object detection method is used. The most challenging and important problem in computer vision is to detect small and dense objects in complex environment with high accuracy and precision. This problem was rectified by V. Arulalan and Dhananjay Kumar(2022). They suggested Hyperbolic Tangent based YOLO v4(HTYOLOV4) with a Modified Manta-Ray Foraging Optimization-based Convolution Neural Network(M2 RFO-CNN)to detect and classify objects.
In pre-processing stage, the data in the form of video were converted into sequence of images. With the help of Polynomial Adaptive Edge preserving Algorithm (PAEPA), image resizing and noise removal were done. In prediction process, HTYOLOV4 was utilized to bring accurate bounding box prediction as

well as reduced Loss Function. Also, Grasp configuration was employed for detecting smaller objects with higher accuracy. The analysis shows the improved efficiency for detection and classification.

The paper published by Lifeng Zhang , Hongyan Cui, Anming Hu, Jiadong Li, Yidi Tang and Roy Elmer Welsch proposed an improved detection algorithm for Ischemic Stroke Non Contrast Computer Tomography (NCCT) system based on YOLOv5[13]. An Ischemic Stroke(IS) occurs due to the interruption of blood supply to part of the brain. This prevents brain tissue from getting oxygen and nutrients which make the brain cells to die in minutes. Due to lack of effective treatment, preventive measures such as timely and clinical diagnosis are conducted. NCCT is a common detection method for stroke. The characteristics of NCCT for acute IS are relatively simple which leads to a lower feature detection.

Ziwen Chen ,Lijie Cao and Qihua Wang in their paper proposed an application of the adaptive clipping algorithm in the vehicle detection. The improved YOLO algorithm shows advancement in the precision of the algorithm[14].

Hyun-Ki Jung and Gi-Sang Choi conducted a research to improve the performance of the model in their paper[15].

The YOLO v7 algorithm achieves the highest accuracy among all other real-time object detection models – while achieving 30 FPS or higher using a GPU V100[1]

METHODOLOGY

Data preparation

A dataset of physical exams was created by recording the classroom activities through a camera of high resolution. The videos contained both normal and suspicious activities of the students. For frame extraction from the recorded videos, the python script with OpenCV library was used by the researcher. 5000 plus frames were extracted from the videos. The extracted frames were pre-processed by removing blurred frames from the folder manually.

Proposed Model

YOLOv7 is used for Object detection in the classroom using Pytorch. The model is trained and deployed to detect the object. Initially the video is captured in the camera.

The captured date is converted into sequences of images. Polynomial Adaptive Edge preserving Algorithm (PAEPA) is used to resize the image and remove the noise from the data. This pre processing is followed by Contrast Limited Adaptive Edge Preserving Algorithm (CLAHE) to enhance the contrast of the image. Cross Stage Partial Network(CSP) is used to reduce the computation cost.

YOLO divides an image into a grid system which detects the object within itself for real-time object detection based on the data streams. Few computational resources were used. All the objects were detected at the end of this step with their bounding box and the loss was calculated. Finally, the M2 RFO-CNN algorithm executed the object detection and classification step.

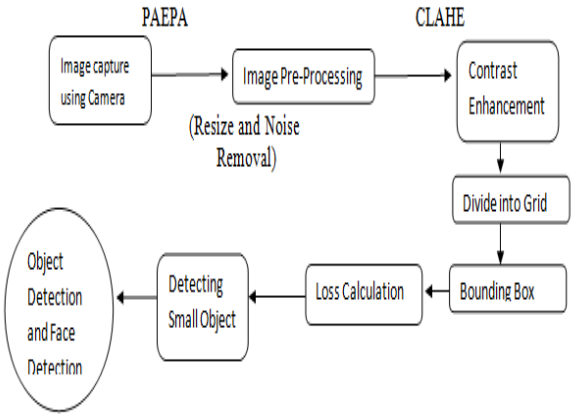

Fig 2 : Steps for object detection using YOLOv7

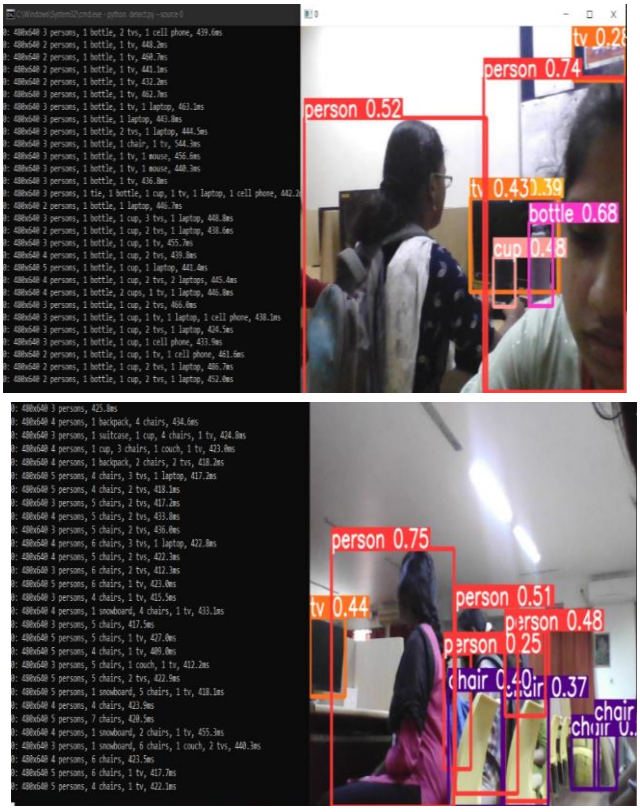

Fig 3 : Object Detection using YOLO v7

C. Mediapipe

During examination, passing of papers and showing the symbols takes place as a result of malpractice. This can be avoided using MediaPipe[3]. The MediaPipe uses a single-shot palm detection model in which twenty one 3D palm coordinates are used to perform precise key point localization in the detected hand region. An oriented hand bounding box from the full image was drawn in the palm. The cropped

image region is fed to a hand landmark model defined by the palm detector and returns high-fidelity 3D hand key points.

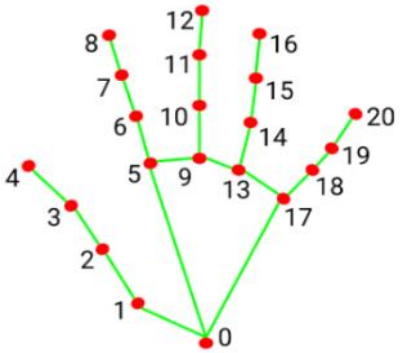

Fig 4 : Hand Land Mark Model [Source: https://google.github.io/mediapipe/solutions/hands.html]

Instant motion tracking is done through mediapipe. The image or video input data is fetched as separate streams and analyzed. The face, hands and pose detection are done through holistic tracking along with face landmarks through the components available in Mediapipe.

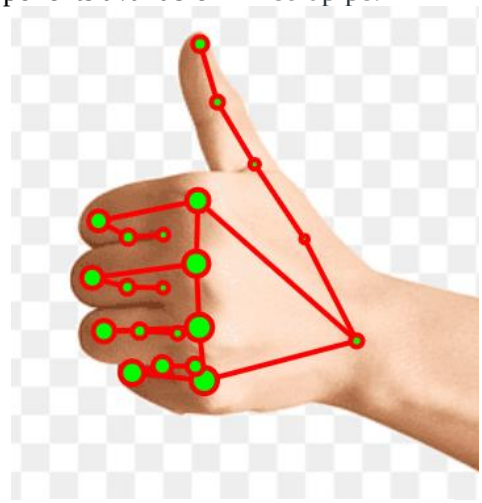

Fig 5 : Motion Tracking in Hand

With the help of YOLOv7 and the mediapipe, the students are being tracked if any mischief behaviour was found during the examination inside the exam hall.

CONCLUSION

The model was developed using YOLOv7 and mediapipe to control unethical activities held during real time examination through like object detection, hand movement recognition along with the face recognition. This model leads us to automatic invigilation system using deep learning approach.

REFERENCES

1. Gaudenz Boesch,"YOLOv7 : the most powerful object detection algorithm". Read more at: https://viso.ai/deep-learning/yolov7-guide/
2. Upesh Nepal and Hossein Eslamiat[2022]," Comparing YOLOv3, YOLOv4 and YOLOv5 for Autonomous Landing Spot Detection in Faulty UAVs". Retrieved from https:// doi.org/10.3390/s22020464
3. Real-time Hand Gesture Recognition using TensorFlow & OpenCV . Retrieved from https://techvidvan.com/tutorials/hand-gesture-recognition-tensorflow-opencv/
4. Ahmed Fawzy Gad(2020). Faster R-CNN Explained for Object Detection Tasks. Retrieved from https://blog.paperspace.com/faster-r-cnn-explained-object-detection/
5. Q. Yu, B. Wang and Y. Su, "Object detection-tracking algorithm for unmanned surface vehicles based on a radar□photoelectric system," IEEE Access, vol. 9, pp. 57529–57541, 2021.
6. Ahmed, S. Din, G. Jeon, F. Piccialli and G. Fortino, "Towards collaborative robotics in top view surveillance: A framework for multiple object tracking by detection using deep learning," IEEE Chinese Association of Automation Journal of Automatica Sinica, vol. 8, no. 7, pp. 1253–1270, 2021.
7. S. Yi, H. Ma, X. Li and Y. Wang, "WSODPB weakly supervised object detection with PCS net and box regression module," Neurocomputing, vol. 418, no. 12, pp. 232–240, 2020
8. Kumar and S. Srivastava, "Object detection system based on convolution neural networks using single shot multi box detector," Procedia Computer Science, vol. 171, no. 1, pp. 2610–2617, 2020.
9. J. Kim, J. Koh and J. W. Choi, "Video object detection using motion context and feature aggregation," in Int. Conf. on Information and Communication Technology Convergence (ICTC), Jeju, Korea (South), pp. 269–272, 2020.
10. Y. Zhu, J. S. Wu, X. Liu, G. Zeng, J. Sun et al., "Photon-limited non-imaging object detection and classification based on single-pixel imaging system," Applied Physics B, vol. 126, no. 1, pp. 1–8, 2020.
11. N. V. Kousik, Y. Natarajan, A. R. Raja, S. Kallam, R. Patan et al., "Improved salient object detection using hybrid convolution recurrent neural network," Expert Systems with Applications, vol. 166, no. 3, pp. 114064, 2020.
12. V. Arulalan and Dhananjay Kumar(2022), "Efficient Object Detection and Classification Approach Using HTYOLOV4 and M2 RFO-CNN", Computer Systems Science & Engineering, DOI: 10.32604/csse.2023.026744.
13. Zhang, L., Cui, H., Hu, A., Li, J., Tang, Y., & Welsch, R. E. (2022). An Improved Detection Algorithm for Ischemic Stroke NCCT Based on YOLOv5. Diagnostics, 12(11), 2591. MDPI AG. Retrieved from http://dx.doi.org/10.3390/diagnostics12112591.
14. Ziwen Chen ,Lijie Cao and Qihua Wang(2022). YOLOv5-Based Vehicle Detection Method for High-Resolution UAV Images. Mobile Information Systems. Volume 2022, Article ID 1828848, 11 pages.
15. Jung, H.-K.; Choi, G.-S(2022). Improved YOLOv5: Efficient Object Detection Using Drone Images under Various Conditions. 12, 7255. https://doi.org/10.3390/ app1214725.
16. David Cochard(2021), "YOLOv5 : The Latest Model for Object Detection". Resource taken from YOLOv5 : The Latest Model for Object Detection | by David Cochard | axinc-ai | Medium