

Virtual Voice Assistant for Detecting Drowsiness, Objects and for Entertainment

S. Sowmya¹, A. Anusha², Ooha. P³, Mukta. P⁴, Prof. Sampath Kumar. R⁵

Students, Department of Computer Science^{1,2,3,4}

Assistant Professor, Department of Computer Science⁵

Rao Bahadur Y Mahabaleswarappa Engineering College, Bellary, India

Abstract: *Virtual learning has made students to face many challenges, it has made interactions complicated and difficult to understand the concepts. The proposed system is mainly developed to help students to have an effective learning. The proposed system mainly runs on voice commands given by user and has three modules: drowsiness detection, object recognition and for entertainment. A simple camera or desktop webcam can be used with the application to capture the real-time video and a speaker to accept the voice command. Based on the commands given by user the drowsiness detector will be start monitoring the student, detects the object and plays a music, movie or tells a joke. OpenCV which is an open-source library for image processing and computer vision which processes the captured video and forms frames. YOLOv3 (You Only Look Once) is an image classifier that takes parts of the frames and creates the co-ordinates and rectangular window across the image. In this way the objects are detected and the output is in form of speech. The students are monitored using detector and a warning is generated in form of an alarm to keep the students awoken so that they can concentrate on their studies and understand the concepts easily.*

Keywords: EAR (Eye Aspect Ratio), YOLOv3, OpenCV, face detection, Google API.

I. INTRODUCTION

Virtual learning has made students to face many challenges such as constantly working with a system, reduces concentration and lack of communication. This project is mainly helpful for students which uses voice assistant and includes three modules which are: drowsiness assistant, object detection and for entertainment.

Due to this pandemic all the Educational Institutions have opted virtual learning which has been a challenging task for students to continuously sit, concentrate and be awake for longer duration. This project helps students to be attentive whenever they fall asleep by triggering an alarm. The drowsiness detection system is implemented using OpenCV.

Object detection is a phenomenon in computer vision that involves detection of various objects using OpenCV for image processing and YOLOv3 to detect the object in the frame in an accurate way. This feature helps students to recognize the objects in their surroundings easily.

If students feel bored listening to the classes constantly, they can be entertained just by giving a voice command as an input to the system to play a music, watch a movie, tell a joke etc.

II. LITERATURE SURVEY

The voice assistant which is proposed in this system mainly works on the users' commands either to start a detector to monitor the student continuously so that he/she keeps awoken throughout his studies, to detect the objects in the surroundings and to get entertained. Currently, many researches are going on in the computer vision field for detection of drowsiness and objects.

Kirbana Jai Raman et al. [1] proposed drowsiness detection model based on yawning and head movements. Face extraction is done using Support Vector Machine (SVM) and detects mouth parts using Circular Hough transform (CHT). CHT is been applied on mouth extracted regions to calculate the geometry of mouth. If a person covers the mouth while yawning it becomes difficult to detect drowsiness of a person, this led to the failure of the system.

Mika Sunagawa et al. [2] proposed a system which determines that a person is drowsy based on the behavioral and physiological information. Behavioral information such as facial features, head movements and eye blinking index. Physiological information such as heart rate variability, rate of respiration by direct contact with person. Here the person must be connected to Echocardiography (ECG) and respirometer directly to determine physiological features. If the connectivity is not proper and the person is not monitored continuously then the alarm will not be triggered thus the system fails.

Kun Wang, Mao Zhen Liu et al. [3] proposed an Artificial Neural Network (ANN) model to recognize the object using DCGAN (Deep Convolution Generative Adversarial Network) and faster R-CNN (Region-based Convolution Neural Network). Both night and day images are fed as input and to obtain high-precision detection results they have combined faster R-CNN target detection system through multi-scale ROI (Region of Interest) pooling. As this system uses R-CNN it fails to detect the objects in real-time as R-CNN uses Selective algorithm and also takes more time to find out the region proposals.

Kim et al. [4] compared different Convolution Neural Network (CNN) models for real-time object detection. They compared deep learning-based object detection models such as R-CNN, Fast R-CNN, Faster R-CNN, SSD (Single Shot Detector) and YOLO in their processing speeds and accuracy to determine the best performance. Faster R-CNN has lesser FPS (frames per second) but has better accuracy, while SSD had less accuracy with better FPS. The YOLO model was a middle ground and had high processing speed than all the models.

Wei Fang et al. [5] proposed a real-time object detection model which used Tiny-YOLOv3 improved version of YOLOv3. Object detection performance is enhanced in Tiny-YOLOv3 and has smaller model of size 8.9MB which reduces the system design and improves the runtime speed as Tiny-YOLOv3 is faster. But Tiny-YOLOv3 is about 23.7% accurate whereas YOLOv3 is 51-57% accurate in detecting the objects.

III. ARCHITECTURE DIAGRAM

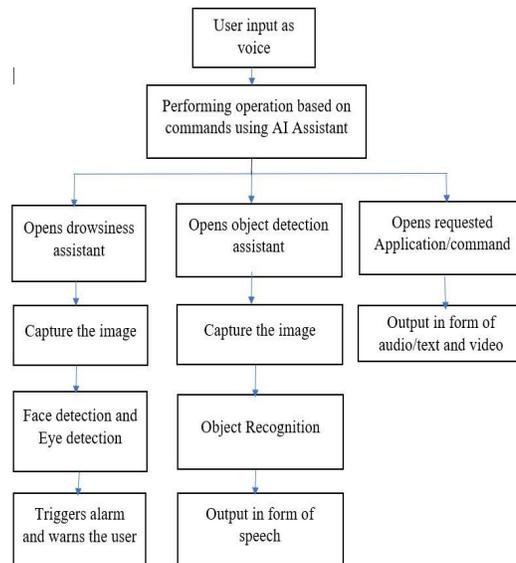


Fig 1. System architecture

IV. PROPOSED SYSTEM

Online classes are gaining importance these days. But they are not as effective as classroom teaching. The proposed project helps students to concentrate on their studies, have good interaction, helps to recognize the objects in their surroundings and to entertain whenever a student feels bored. This system mainly works as a voice assistant and has additional three features: drowsiness detection, object detection and for entertainment purpose.

There are chances of students getting bored and may feel drowsy. The project recognizes student drowsiness and generates an alarm keep them alert so that they can concentrate on their studies.

It also becomes easy for students to detect objects in the surroundings and gain knowledge on them. It also helps blind people in detecting the objects and obstacles. It also helps students to get entertained by listening to music, watch a movie or video and to tell a joke.

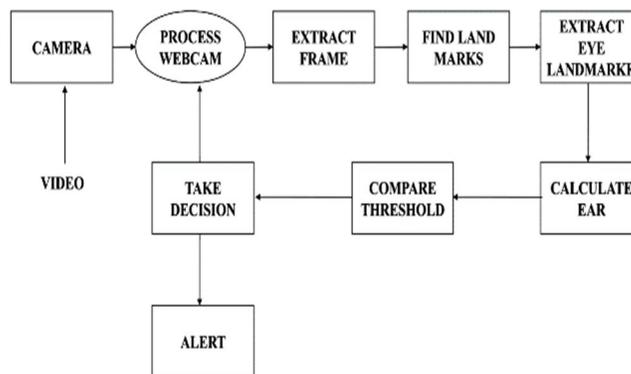
The proposed project is an innovative system which helps students to automate day to day task just by giving voice commands. The main cause of this proposed project is to have a dynamic interactive, effective and comfortable learning environment for students.

IV. METHODOLOGY

As the proposed project uses AI based voice assistant and has three modules: drowsiness detection, object detection and for entertainment. Based on the voice commands given by the user the respective module will be activated and performs the necessary tasks and waits for the further commands from the user. The methodology of each module is detailed below:

Steps involved in drowsiness detection are:

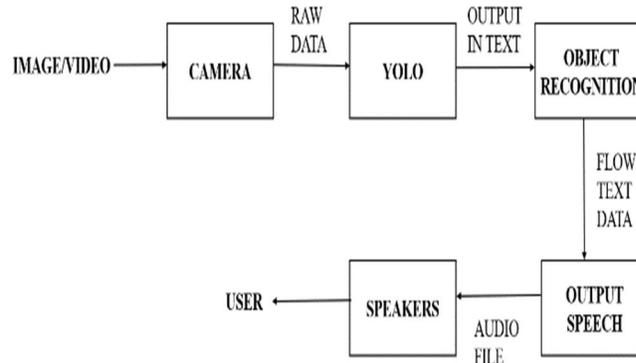
1. In the first step, the video which is captured through the camera is processed and frames are extracted using OpenCV.
2. If a face is found, the facial landmark detection is done.
3. Extract the eye landmarks and create co-ordinates across the eye region.
4. Eye Aspect Ratio (EAR) is computed to determine if eyes are closed.
5. Compare the calculated EAR with the threshold:
 - a. If EAR is below threshold value, then an alarm is triggered to alert the user and repeat from Step 1.
 - b. If EAR is below threshold value, then repeat the process from Step 1.



Steps involved in object detection are:

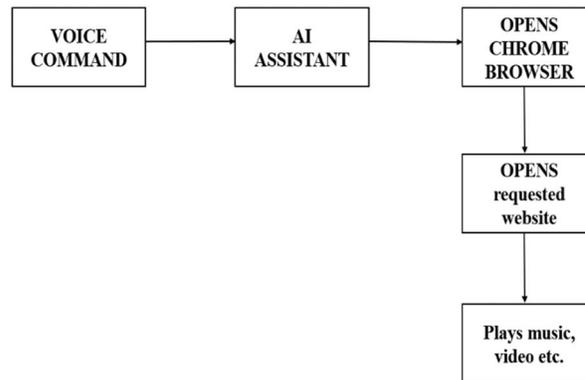
1. The video captured is processed and converted into frames using OpenCV.
2. These frames are passed to YOLOv3 where the objects in the image are identified and its location is given using co-ordinates.

3. The rectangular window is created across the object based on its co-ordinates using OpenCV and the object is recognized.
4. The name of the object in form of text is passed to output speech which is converted to audio file using play-sound (e-speak synthesizer).
5. The audio file is passed to the speakers through which it is outputted to the users.



Steps involved in entertainment module are:

1. The voice commands given by user through the speakers is passed to the AI assistant.
2. The AI Agent analyses the voice and opens the appropriate web browser and in turn opens the respective website.
3. Performs the required task such as playing music or video etc.
4. Waits for further commands from user and repeats the process from Step 1.



VI. CONCLUSION

In conclusion this system mainly tends to focus on detecting the drowsiness of students and keep them alert by triggering an alarm, recognizing the objects and to entertain students. The main purpose of this project is to overcome the problems faced by students during virtual learning and help them to concentrate on their studies. Here the students need not give the commands manually all the processes are carried out using voice assistant which is easy and friendly for students to use. It helps students to have effective and interactive learning.

In future this proposed system can be used by ed-tech (Educational Technology) companies to keep students alert so that they can concentrate on their classes and have effective learning.

REFERENCES

- [1]. Kirbana Jai Raman, Afizan Azman, Venosha Arumugam, Siti Zainab Ibrahim, Sumendra Yogarayan, Kalaiarasi Sonaimuthu, "Fatigue Monitoring Based on Yawning and Head Movement", 2018 6th International Conference on Information and Communication Technology (ICoICT).
- [2]. Shabnam Abtahi, Behnoosh Hariri, Shervin Shirmohammadi, "Driver drowsiness monitoring based on yawning detection", 2011 IEEE International Instrumentation and Measurement Technology Conference, 7th July 2011, Hangzhou, China.
- [3]. Kun Wang, Mao Zhen Liu, "Object Recognition at Night Scene Based on DCGAN and Faster R-CNN", IEEE paper, 22 October, 2020.
- [4]. Wei Fang, Lin Wang, Peiming Ren, "Tinier-YOLO: A Real-Time Object Detection Method for Constrained Environments", January 6th, 2020, IEEE paper.
- [5]. Nawal Alious, Aouatif Amine and Mohammed Rziza, "Driver's Fatigue Based on Yawning Extraction", Hindawi Publishing Corporation, International Journal of Vehicular Technology, Volume August 6, 2014.
- [6]. Wanhua Deng and Ruoxue WU, "Real-Time Drowsiness Detection System Using Facial Features", IEEE, September 5, 2019, China.
- [7]. Smart Eye. (2018). Smarteye. [Online]. Available: <https://smarteye.se/>
- [8]. "Object Recognition System Design in Computer Vision: A Universal Approach", Sampletalk Research, 2016, Andrew Gleibman.
- [9]. Mika Sunagawa, Shin-ichi Shikii, Wataru Nakai, Makoto Mochizuki, Koichi Kusukame, and Hiroki Kitajima, "Comprehensive Drowsiness Level Detection Model Combining Multimodal Information", IEEE Sensors Journal, VOL. 20, NO. 7, April 1, 2020.
- [10]. Feng You, Xiaolong Li, Yunbo Gong, Haiwei Wang, and Hongyi Li, "A Real-Time Driving Drowsiness Detection Algorithm with Individual Differences Consideration", IEEE paper, December 23rd, 2019.
- [11]. Adwitiya Arora, Atul Grover, Raksha Chugh, S. Sofana Reka, "Real-time Multi Object Detection for Blind using Single Shot Multibox Detector, Springer Science, Wireless Personal Communications, 30 March 2019.
- [12]. Tae Rang Choi, Minette E. Drumwright, "OK, Google, why do I use you?", Motivations, post-consumption evaluations, and perceptions of voice AI assistants, Science Direct, Telematics and Informatics 62 (2021), April 24, 2021.
- [13]. S. Subhash, Prajwal. N Srivatsa, S. Siddesh, A. Ullas, B. Santosh, "Artificial Intelligence-based Voice Assistant", IEEE, 2020 fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4), 27-28 July 2020.
- [14]. EV Polyakov, MS Mazhanov, AY Voskov, LS Kachalova, MV and SV Polyakov, "Investigation and development of the intelligent voice assistant for the IOT using machine learning", Moscow workshop on electronic technologies, 2018.
- [15]. Piyush Vashista, Juginder Pal Singh, Pranav Jain and Jitendra Kumar, "Raspberry PI based voice-operated personal assistant", International Conference on Electronics and Communication and Aerospace Technology (ICECA), 2019.
- [16]. Rui Liu, Zhihua Yu, Daili Mo, Yunfei Cai, "An Improved Faster-RCNN Algorithm for Object Detection in Remote Sensing Images", 2020 39th Chinese Control Conference (CCC), IEEE, 27-29 July 2020.
- [17]. Ma Songhui, Shi Mingming, Hu Chufeng, "Objects detection and location based on mask RCNN and stereo vision", 2019 14th IEEE International Conference on Electronic Measurement & Instruments (ICEMI), 1-3 November 2019.
- [18]. Amin Azizi Suhaiman, Zazilah May, Noor A'in A. Rahaman, "Development of an intelligent drowsiness detection system for drivers using image processing technique", 2020 IEEE Student Conference on Research and Development (SCORED), 27-29 Sep. 2020.
- [19]. "SSD: Single Shot Detector for object detection using MultiBox", 30 Nov. 2019, [Online] Available:

<https://towardsdatascience.com/>.

[20]. Ashwani Kumar, Zuopeng Justin Zhang & Hongbo Lyu, “Object detection in real time based on improved single shot multi-box detector algorithm”, EURASIP Journal on Wireless Communications and Networking (2020), SpringerOpen, 17 October 2020.

[21]. Sakshi Gupta, Dr. T. Uma Devi, “YOLOv2 based Real Time Object Detection”, International Journal of Computer Science Trends and Technology (IJCSST)- Volume 8 Issue 3, May-June 2020.