iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** V **Month of publication:** May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82592

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Intelligent Candidate Monitoring and Interview Surveillance System for Virtual Interview Using Machine Learning

Dr. T. Manoj Prasath[1], V. Sakthi Varshini[2], D. Dhana Lakshmi[3], N. Sumuthuradevi[4]
Department of Computer Science and Engineering, SSM Institute of Engineering and Technology, Dindigul, Tamil Nadu

Abstract: ***The Secure Interview Monitoring System is a web-based interviewing tool that helps keep track of the test environment with the assistance of a web camera as a real test proctor. It works in the background and identifies suspicious actions like head movements, use of mobile phones, and even the presence of other persons and issues alerts on the spot to ensure integrity. Every activity is documented to be reviewed. Automated analysis recognizes possible red flags and transforms an ordinary video call into a safe and monitored space, which is defined as an intelligent system. It is created in a way that guarantees equity and dependability of remote employment without the need to have a human monitoring it all the time.***
Keywords: remote proctoring, cheating detection, online interview security, facial pose estimation, object recognition.

I. INTRODUCTION

It seems like a blind shot to retain fairness in a video interview. Does the other end receive assistance? Do they read the notes off camera? It's a real worry. It is not another human being staring at the screen but intelligent technology that is able to watch without blinking. This system can be compared to a silent digital proctor. It is standing in the corner of the interview, with the help of the web camera only, and listens to those little details that a person would not have noticed. A glimpse of a second screen, the rustling of a paper, even an additional individual in the reflection on one monitor, it notices. And rather than wait to enter a report, it does a gentle, instantaneous beep and a screen pop-up, a slight reminder that one needs to concentrate on the task. The entire situation is documented, of course, but the actual worth of it is that live, automated nudge.

II. LITERATURE REVIEW AND BACKGROUND

Remote proctoring is not a new concept, although most of the solutions still lack the mark [3]. At the beginning of things, people were only able to watch live feeds, and this was costly and unreliable. Then automated record-and-review systems were introduced, that only detect cheating when it was already complete, not so useful in the timeframe. Since then, tools have begun to utilize simple motion detection or infrequent screenshots [6]. Well, they ran up easier but flagged everything. Stretch your neck? Flagged. Look down at your keyboard? Flagged. False alarms accumulated, and confidence in the technology began to wear out. Then there was more developed AI, face recognition, gaze tracking, etc. [5]. Precise, yes, but heavy. They require intensive computing power, tend to gather a lot of data more than is comfortable, and have difficulties with such aspects as bad lighting or outdated webcams. Moreover, many of them are cloud-based and broadcast videos feeds wherever you go. The genuine privacy nightmare, particularly with the tightening of the laws on data transfer everywhere [8]. Here is the dilemma; then we have to either have heavy, invasive systems or simple, shakey systems. What is really required is a balanced one which is accurate enough to detect genuine problems, yet intelligent enough to overlook harmless behavior. Local enough to prevent the need to move data offsite, yet open enough that the candidates are not made to feel intimidated. That's what we built SIMS to do. It is locally based, does not intrude on privacy, reduces false alerts, and maintains clarity in communication hence integrity remains high, without panic.

III. RELATED WORKS/APPROACHES

In terms of maintaining remote test honesty, there has been technology development that has been quite interesting and, in some cases, invasive [3]. It became easy, hit the record button and the webcam to make a recording [6]. However, it required someone to spend hours of footage later, with the hope of trying to capture something. Next, there was face-tracking and motion detection [5]. These devices would know whether you left your seat or not or whether you were not seen on your face, that is all. They would mark your routine stuff, scratch your chair, and stare at the keyboard, and create a lot of false alarms.

And there is nothing like being pinged nothing to kill trust. The actual change occurred when systems became more proficient in cognizing context. They began to measure not only whether you had your face around, but where you were gazing, head tilt, gaze direction, more than a glance off the screen, etc. with tools such as MediaPipe [2], and OpenFace. Combine that with object detection (imagine YOLO spotting cars or papers in the background) and then you had something even resembling actual supervision.

After that, it became even more personal. Certain systems started to scan the way you type, your mouse movements, even the rhythm of your typing, and came up with a kind of behavioral fingerprint. In case your patterns were different on the test, it may indicate that there is another person who may be assisting. Clever? Naturally, all this is irrelevant if you cannot measure the success of this. That is where benchmarks are able to assist, simulations of cheating scenarios, labeled data, helping to distinguish what is actually effective, and what is simply being too concerned [20]. It is not to trap all, but to trap the correct things, and make the test-taker does not feel as though he is in the surveillance state.

IV. EXISTING SYSTEM

It is still a dilemma as to how to make remote interviews fair. At this point, you have two options; the cumbersome automation that raises the alarm about every single head turn, or the post hoc recordings that are only noticing any cheating when it is too late. Neither one really works. The crude instruments lack what is important, a phone in the background, a look at a second monitor. The smarter ones tend to look like overreach, making no commotion in the cloud and copying sensitive video without explicit permission. It causes people to feel guarded and not safeguarded. False alerts start to overwhelm admins, and they learn to ignore them [7]. Applicants become afraid and do not know what is being followed and when they will be dinged because of just stretching or glancing at the keyboard.

V. PROBLEMS IDENTIFIED

It goes without saying: many of the remote proctoring tools [1] are lagging behind. They are forced to tick boxes rather than knowing what is really going on in the room. When a candidate wears a small earpiece, holds a phone right below the camera, or has something written on their virtual background, it is not even captured with simple cameras. They were only not developed to identify smart cheaters but awkward ones. That's the core issue. Most tools monitor either one aspect at a given time, either your face or your movement, but not both aspects together. Thus, a valid stretch is flagged, and a person reading notes off a second screen slides past. Cheaters are aware of how to pull off these eye closures. They have progressed, with AI voices, screen sharing, hidden tabs, proctoring technology [1] is lagging behind, seeking the tricks of last year. And then there is a human review problem. Of course, it is required that somebody look at the recording after the test. Trusting human surveillance as an ex-post facto is not only inefficient; but also, a loophole in the entire system. So, what's the real fix? Not more cameras. Not tighter rules. Smarter ones. Weapons that do not merely observe but comprehend. There should be systems to distinguish between looking away to thinking and looking away to cheat. Real-time adaptable tools that do not suspect honest individuals.

VI. PROPOSED SYSTEM

The system will act as a silent listener when conducting remote interviews. It monitors both the screen and the activities of the candidate using the webcam of the candidate. Based on live computer vision, it identifies objects that cause concern, in case an individual looks off-screen, phone or notes are displayed, or another individual enters the room. When it does, it will provide an immediate, obvious prompt at the time and place as well as save an entire recording. Then there is no waiting and guessing post factum. That is assisted by this system that wants to highlight things that seem out of order, and you are still speaking to the other person. When they continue to look off- camera, or when you notice that someone has a phone in their hand, you are immediately informed. You are able to deal with it immediately rather than in hindsight to feel as though something was amiss.

And, frankly, the more you use it the better it gets. The software gets to know what a real red flag is and what a nervous tic is. It discourages the tendency to cry wolves over any slight movement. To the people who are conducting the interviews, it creates a solid piece of what occurred, no gray areas. In the end, it's about trust. Candidates would like to feel that they had a good opportunity, and you should be aware that the process was good. It is only a device to ensure this and not to make the interview seem like an interrogation.

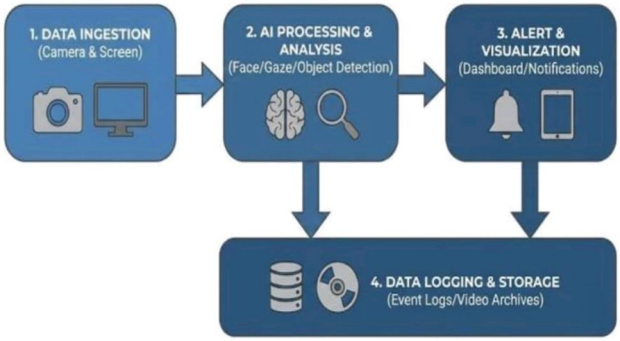


Fig. 1. System Architecture

It monitors live information of the screen and web camera of the applicant. The inputs are processed to identify face movement, gaze direction, and suspicious objects on AI models. Visual prompts and alerts are created immediately on an interactive dashboard. All things are logged down safely to be viewed, analyzed and have transparency.

VII. WORKING DATASETS

This is the way it really goes, straight down the line. During the interview, the system is monitoring two aspects; webcam and what the candidate is doing on the screen. The camera is not merely a photographer of events, but it seeks to detect any signs that are telltale, such as whether their eyes are constantly going off-screen or whether a phone has appeared in their palm. It also looks into the presence of other people in the room. When something appears suspicious, the system indicates it immediately. There's no waiting around.

Meanwhile, it is capturing the screen, as well, the windows open, what they are clicking on, etc. Hence at the end you can have a full picture of the live behavior in addition to a full history of their screen activity. It does not refer to surveillance, it refers to generating an open, truthful record of the session. In that respect, you can tell just what has happened, and there is no conjecture as to equity. Internet interviews are difficult to maintain impartial.

That is assisted with this system as it detects objects that are not fitting as you continue talking to the person. When they continue to peep somewhere out of the camera, or you realize they have a phone in their hand, it makes you well aware immediately. You get to address it in the present and do not regret afterwards whether something has gone awry.

And the more you use it the better it gets. The software gets to know what is actually a red flag and what is a simple nervous tic. It prevents crying wolf over any slightest movement. To the individuals in charge of the interviews, it creates a clear, solid, no gray areas record of what occurred.

In the end, it's about trust. You would like the candidates to believe that they were given a fair chance and you would want to know that the process was sound. This is simply an aid to ensure it, and does not turn the interview into an interrogation.

It operates quietly in the background with no distraction of the candidate or interruptions to the interview process. All data recorded is also stored in a secure place that is not readily available to other persons. The system enhances transparency since verifiable evidence can be made whenever there is need to review.

VIII. SIMULATION SETTINGS

You know that uneasy moment in a remote interview where you wonder if things are really fair? This system is built to fix that. It works in the background, picking up on subtle red flags, like a candidate repeatedly looking off-screen, or a phone suddenly in their hand, and alerts you immediately on your dashboard. If something serious occurs, you can even pause the interview. No waiting, no guessing after the fact.

That real-time alert gives you a chance to address things while they're happening, which keeps the process honest. But it's bigger than just flagging cheaters. It's about creating a level playing field that everyone understands. For hiring teams, it takes the stress.

IX. INTERVIEW MALPRACTICE RISK LEVELS

Risk Level	Probability Range	Description
Low Risk	$P < 0.30$	Candidate looking normally, no suspicious objects detected
Medium Risk	$0.30 \leq p < 0.60$	Frequent eye movement, looking sideway, possible assistance
High Risk	$P \geq 0.60$	Mobile phone, book, another person detected

Table 1. Malpractice Risk Levels

This table classifies the candidate's behavior at various levels of risk during the interview process. Low risk shows a normal behavior that is free of suspicion. Medium risk implies potential malpractice related to abnormal movements or behavior [13]. High risk verifies the good indicators such as the use of devices or external assistance.

X. PERFORMANCE EVALUATION

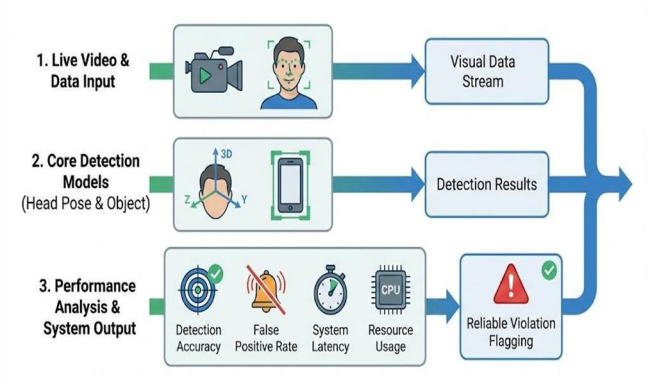


Fig. 2. The Monitoring Pipeline Analysis

Figure 2 illustrates a real-world example of the way the monitoring of the system works. It is monitoring and examining three things simultaneously, which are: the live image on the webcam (where the candidate has his or her head), and whether any prohibited items come into view. The rate at which it is false raises an alarm, responsiveness, and the amount of computer processing required. The system works well when the videos are good, and sensitivity is set accordingly; it captures what it is supposed to with the system not making anyone fret about nothing.

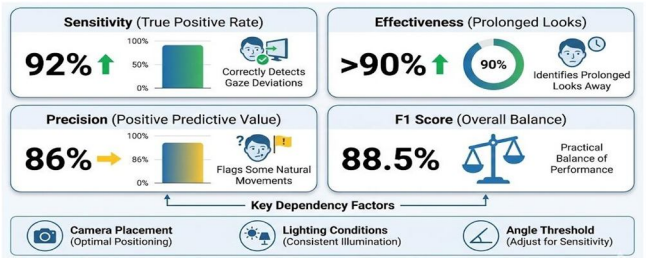


Fig. 3. The Head Pose Detection Analysis

Figure 3, the system monitors a candidate paying attention to the screen. In case a person does not look at it for a longer period than a glance, it identifies that approximately 92 percent. It particularly makes a good catch of those longer, conscious gazes elsewhere. But now it does have a minor tendency to be trigger-happy. Almost every few seconds, a completely normal action, such as adjusting in the chair or looking at the keyboard, is considered suspicious. That is why its accuracy, or the number of times when a flag was a real issue, is more likely to be 86%. Against the occasional false alarms, you consider their active detection, and the total is 88.5%. They rely heavily on the working arrangements. It works well because of good lighting, a centered webcam, and reasonable rules regarding the meaning of looking away.

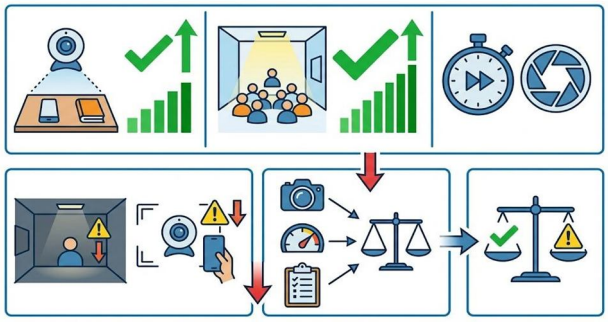


Fig. 4. The Object Detection & Person Counting Analysis

This is demonstrated in figure 4 below which illustrates how the system identifies the presence of things that are not supposed to be in the room, such as a phone or a notebook and counts the number of people in the room. The reason why the short story works is because it works. It recognizes phones and books nearly 89 percent with a standard webcam, and it is virtually never mistaken on the number of people on screen when the light is good, with an accuracy of approximately 96 percent [20]. It is fast as well, at less than a fifth of a second to scan through each frame, hence the feedback is virtually instant. But it's not flawless. The picture is worse when the lighting is poor, or when it is someone holding something to the extreme right or left, hardly in the frame. In the end of the day its performance is reduced to a couple of pragmatic issues: the quality of the camera it uses, the severity of the detection guidelines, and the clarity of what we are looking to find out.

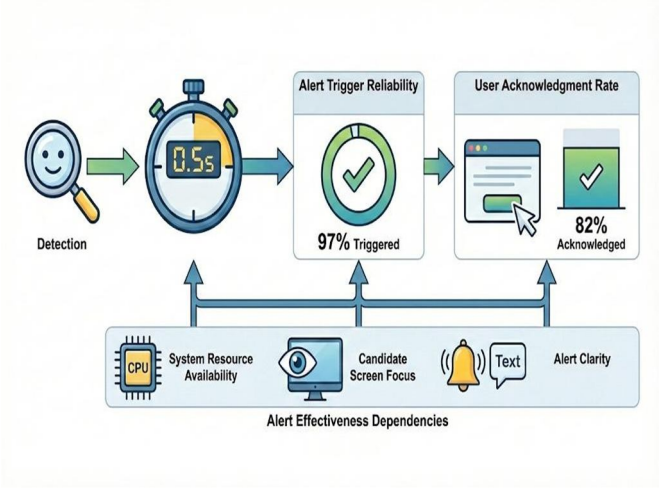


Fig. 5. The Alert System Latency Analysis

In Figure 5, everything concerns the speed and dependability with which the system will alert you when something goes wrong. Its test performance has been good. When an actual violation occurs, it will detect and alert in 97 percent of the tests, and it is fast.

The alert can only be effective now when the candidate actually sees the alert [18]. They recognized the pop-up during our experiments approximately 82 times. But that number isn't fixed. They may miss it when they are away staring at the computer which is running slowly, or when they are looking away at that very time. It is a matter of letting the warning be conspicuous, a loud call, the warning board, and not so noticeable as to cause them to become disconcerted. It's about a nudge, not a shout.

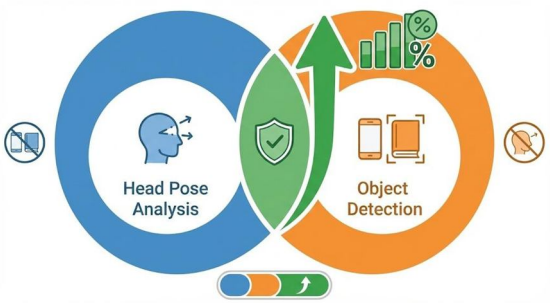


Fig. 6. Comparison of Various Detection Methods

Figure 6: you cannot count upon a single method of determination that there are problems. When you are just following where one is gazing, you will fail to spot that they are holding a phone in their laps. When you search only, you may fail to see them reading out of a second monitor.

This is the reason why the system makes multiple checks simultaneously. These are overlapping and thus when one method fails to capture what another method may pick it up. This integrated method detected 40 percent more potential problems during testing in reality than any one method alone. Of course, nothing is perfect. There is also performance based on issues such as quality of the webcam, room lighting, and even the natural movement of the candidate. However, the combination of several approaches makes monitoring less frail and much more reliable.

XI. DISCUSSION

This interview surveillance system is essentially a webcam spying system which monitors your location and the surrounding environment [12]. It does not sneak in; it is right in your face on your screen so that you are aware of its existence. When there is something wrong, you receive a warning on the dashboard. Frankly speaking, it is quite efficient in what it does. Most blatant rule breakages are caught without obstructing, and since it is a local program, you do not require a high bandwidth connection or some complex cloud infrastructure. At this point, according to prior research, such systems, including one which combines head tracking with object recognition, can be expected to work better than less complex instruments that merely examine whether you are in front of the camera [15]. But we know that: intelligent cheaters work around all things. Keep your phone to a side but not in sight, change the virtual background to scribble notes, peep into a second monitor; these tricks are bound to give it a glitch. This is why more recent strategies bring a sense of context, such as timing rules and confirmation checks, and the system can be made wiser, as to what is a genuine violation and what is a nervous habit.

And this is where it becomes tricky. First, the system can only search for a specified list of objects, phones, books, and additional laptops [15]. The new cheating tools appear regularly, and the system cannot keep up with them unless a person does it manually. Second, it is hard to refer to this as AI proctoring. It does not get to know you, as you use it over the years, or scan to see if you are colluding with someone via messaging apps or even whether you are on a fake webcam feed. Third, there is no real information on whether it was tested effectively. Simulated demonstrations will not find intelligent, flexible cheats; you would require an actual simulation with real-life individuals attempting to overcome the laws. Fourth, the majority of the measures you listen to, such as detection speed and accuracy, are not the whole story. We must be aware of the frequency of where it raises a red flag over innocent moves, does it put candidates under stress and how it responds to various lighting, cameras or even cultural variations in body language.

XII. RESULT

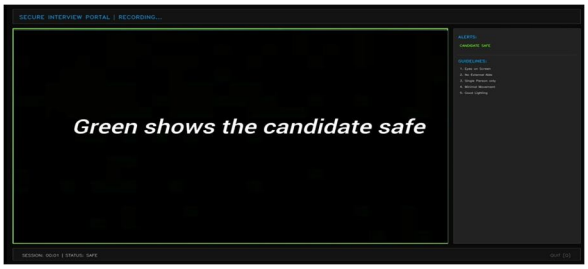


Fig. 7. Candidate Safe Detection

Figure 6: In case there is no suspicious activity, the system indicates the screen in green and shows the status of candidate safe.

Fig. 8. Candidate Safe Detection

Figure 7: When more than one person is detected, the system draws the attention of the screen with red color and shows an alert of other person detected.

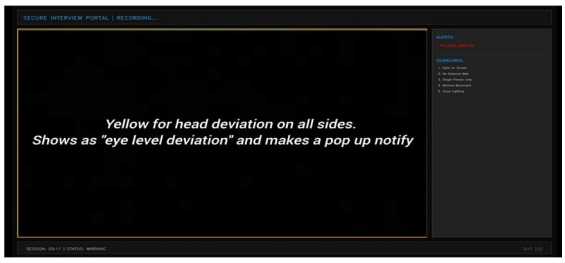


Fig. 9. Eye Level Deviation

Figure 8: In case head deviation has been detected, the user is indicated by the system by highlighting the screen with yellow color & informing them that they have deviated the eye level.

XIII. GRAPHS

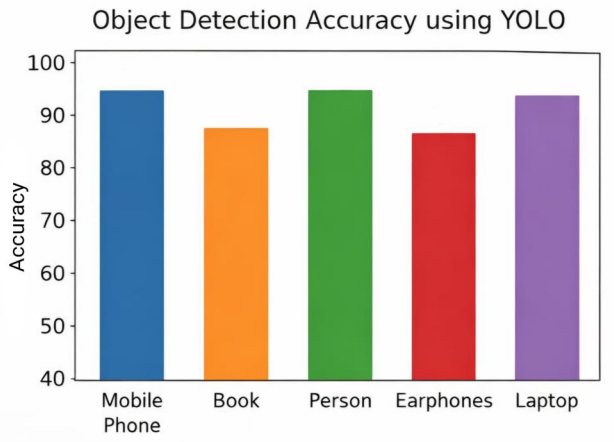


Fig. 10. Object Detection Accuracy using YOLO Model

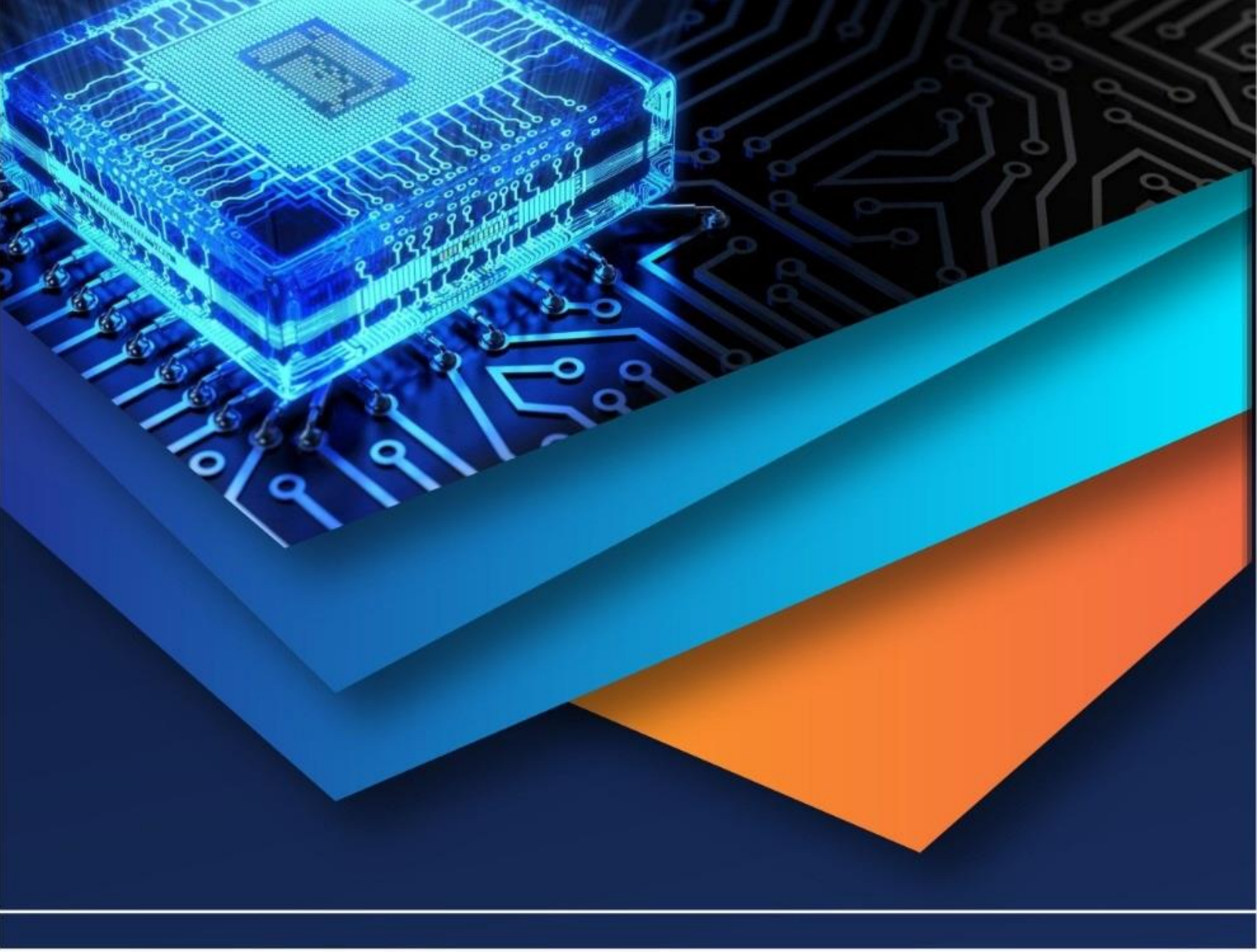
This graph illustrates the precision of detection of various objects with the help of the YOLO algorithm. The model achieves the best results in recognizing individual persons at 96 percent. The mobile phones and laptops also demonstrate high detection accuracy with more than 90. The maximum accuracy of books and earphones is slightly lower than in other objects.

XIV. CONCLUSION

That nagging doubt during a remote interview? It's there for a reason. Is this actually fair? Could someone slip something by? For the longest time, the answer was pretty unsatisfying. But keeping things honest doesn't have to mean watching every blink or listening to every breath. Sometimes, it's just about paying smart attention. So, that's what this system does. It simply notices where someone's looking and what's in the room, then speaks up right away if something's not right. No complicated setups, no delays. It just works.

REFERENCES

[1] A. S. Rathore and T. N. Goh, "AI-based Proctoring Systems: A Survey of Methods, Challenges and Ethics," IEEE Access, vol. 10, pp. 34521–34542, 2022.
[2] L. Martinez, J. Rodriguez, and R. Chen, "Real-Time Head Pose Estimation Using MediaPipe for Online Exam Monitoring", Proc. Int. Conf. on Computer Vision and Pattern Recognition Workshops, 2021, pp. 410–418.
[3] K. Sharma and P. Verma, "YOLO-Based Object Detection for Unauthorized Item Recognition in Remote Assessments," Journal of Artificial Intelligence in Education, vol. 32, no. 3, pp. 521–540, 2021.
[4] M. Lee and S. Park, "Fairness and Privacy in Automated Exam Proctoring: A Student- Centric Analysis," Computers & Education, vol. 178, 104301, 2022.
[5] B. R. Smith, T. O'Neil, and J. A. Kim, "Behavioral Integrity Monitoring in Remote Testing Using Computer Vision and Gaze Tracking," Int. Journal of Distance Education Technologies, vol. 20, no. 2, pp. 33–52, 2022.
[6] R. J. Wilson et al., "Screen Recording and Video Audit Trails in Online Proctoring: Legal and Ethical Implications," Journal of Cybersecurity and Privacy, vol. 3, no. 1, pp. 78–95, 2023.
[7] N. Gupta and A. Desai, "False Positives in AI Proctoring: Causes, Impact, and Mitigation," Proc. ACM Conf. on Learning @ Scale, 2022, pp. 215–224.
[8] C. Torres and E. Fernandez, "Local-Only Proctoring: A Privacy-Preserving Approach to Remote Exam Integrity," IEEE Security & Privacy Workshops, 2023, pp. 122–130.
[9] S. Zhao and H. Li, "Multimodal Cheating Detection in Online Exams Using Facial Landmarks and Object Recognition," Proc. Int. Conf. on Educational Data Mining, 2021, pp. 320–329.
[10] P. Zhang and L. Wang, "Real-Time Alert Systems for Academic Integrity: Design and Usability Study," Behaviour & Information Technology, vol. 41, no. 9, pp. 1889–1905, 2022.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY