A Computer Vision-Based System for Monitoring Liveness in Remote Work

Filipe Oliveira de Saldanha da Gama, Rafael Senra Donner Jorge, Caio Rangel Fernandes, Marcio Moreira do Nascimento Filho, Rigel P. Fernandes, and Thiago Silva de Souza

***Abstract*— This paper proposes a home office time clock system that uses images captured from the user's webcam to detect the liveness of employees. Growing concerns about productivity and accountability in remote work environments have created a strong demand for effective monitoring solutions. This work evaluates computer vision techniques for tracking employee activity in such settings, offering a reliable system that operates with minimal interruption to the user's workflow, suitable for real-world deployment. Experimental results from tests performed under challenging conditions, including low-light environments, demonstrate that the system achieves a 100% detection rate in captured frames at ambient illumination levels near 7 lux.**

***Keywords*— Liveness Detection, Computer Vision, Remote Monitoring.**

I. INTRODUCTION

Monitoring employee work hours has long been a standard practice in traditional in-person workplaces. As an example, a Swiss court ruling allows companies to monitor employees' bathroom visits during working hours as a contributing factor in measuring productivity [1]. Although such practices can be controversial, they underscore the importance placed on accurately tracking employee time and activity. However, with the growing shift toward home office arrangements, this task has become significantly more complex. Unlike physical workspaces, remote environments require new strategies and technologies to ensure that employees are present and actively engaged during work hours.

Remote work has grown exponentially in recent years [2]. With this growth, there is a demand for technologies to monitor employees. During the COVID-19 pandemic, many companies were impacted and had to transition their employees to remote work [3]. As a result, there was a demand for new ways for companies to monitor employee engagement remotely.

Webcams, which are present on most laptops, capture the user's image primarily for video calls, recordings, photos, etc. However, with the advent and evolution of computer vision, it is possible to use the camera to assess employee engagement during work at a low cost.

Filipe Oliveira de Saldanha da Gama, Rafael Senra Donner Jorge, Caio Rangel Fernandes, Marcio Moreira do Nascimento Filho, Rigel P. Fernandes, and Thiago Silva de Souza are part of the Computer Engineering and Software Engineering (Tech) programs at the Brazilian Institute of Capital Markets (Ibmec), Rio de Janeiro, RJ, Emails: filipesgama@gmail.com, rafaeldonner14@gmail.com, cbnr1644@gmail.com, warmarcio617@gmail.com, rigelfernandes@gmail.com, t.souza@ibmec.edu.br.

This study evaluates the use of computer vision to monitor remote workers. In addition, the performance of the system is assessed under low-light conditions, an important scenario to consider for employees who prefer or need to work during nighttime hours.

This work is organized as follows. Section II examines the computer-vision-based liveness monitoring system, while Section III details the system's implementation. We then analyze the system's performance in liveness detection and discuss encountered issues in Section IV, and Section V concludes the work.

II. COMPUTER VISION-BASED MONITORING SYSTEM

Computer vision techniques that can be used for monitoring users are: face detection and recognition, blink detection, and hand gesture identification.

Face detection and recognition [4] can be used in various applications. Blink detection [5] depends on face segmentation and the detection of specific eyelid pixels. These techniques have applications in detecting driver drowsiness.

Hand detection [6] can be used for robot interaction. This area has been growing to bridge human-robot barriers. The goal is to develop techniques that make communication with robots more natural.

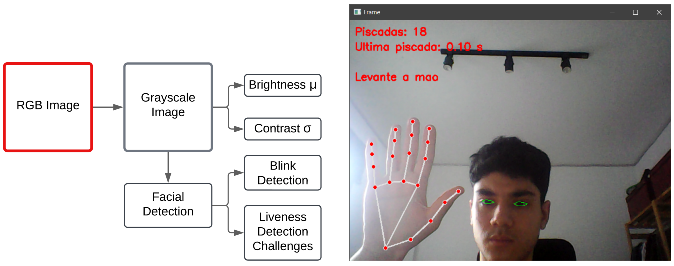
III. THE PROPOSED METHOD

The proposed scheme was implemented in Python 3, using the dlib [4], [5], OpenCV [4], and mediapipe libraries. Figure 1(a) shows the system block diagram. Initially, the system converts the RGB color space to grayscale to reduce processing time. The first task is to detect and recognize the face using the dlib library by comparing a reference image (user's selfie) with the detected face. The second task is to detect the facial landmarks. To count eye blinks, the system uses facial landmarks and calculates the distance between the eyelids to determine whether the eyes are open or closed.

Figure 1(b) shows the blink counting process and the moment when the system prompts the user with a liveness detection challenge. The system supports multiple configurations: face detection only, a combination of face and blink detection, or a full integration of face, blink, and liveness detection.

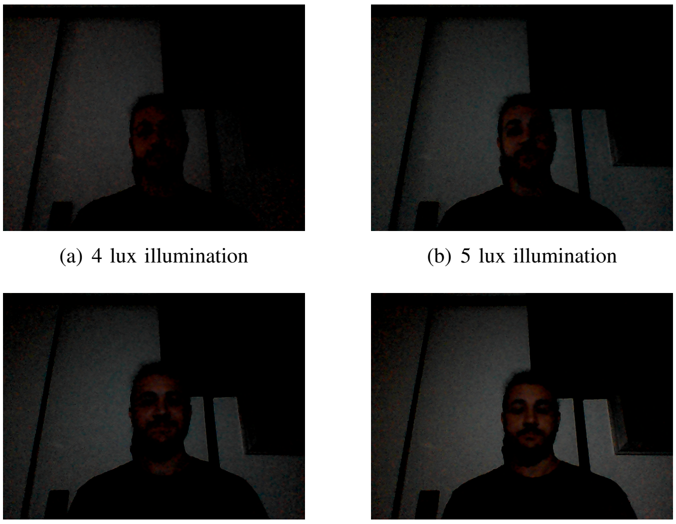
IV. EXPERIMENTAL RESULTS

Figure 2 shows the setup to measure the performance of the system. We used a low-intensity light source to illuminate



(a) (b)

Fig. 1. Diagram and system interface. (a) Block diagram of the system and (b) system with real-time visual annotations. Bottom left: hand detection using MediaPipe. Top left: blink counter, elapsed time since the last blink, and hand-raising challenge prompt. Center: eye detection implemented with dlib.



(a) 4 lux illumination
(b) 5 lux illumination
(c) 6 lux illumination
(d) 7 lux illumination

Fig. 2. Test images under varying illumination levels.

the subject and placed a smartphone with a built-in lux meter behind the user to measure light levels. We captured and stored 20 frames, along with respective average contrast and brightness, for each of the four lighting scenarios, with illumination varying from 0 to 8 lux.

We measure brightness through the mean pixel intensity [7]:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} I_i \tag{1}$$

where μ is the average brightness, N is the total number of pixels, and I_i is the value of each pixel.

The contrast is the standard deviation of pixel intensities [7]:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(I_i - \mu)^2} \tag{2}$$

where σ represents the contrast of the image.

For each scenario, we evaluated the accuracy of the proposed face detection method. As shown in Figure 3, the system

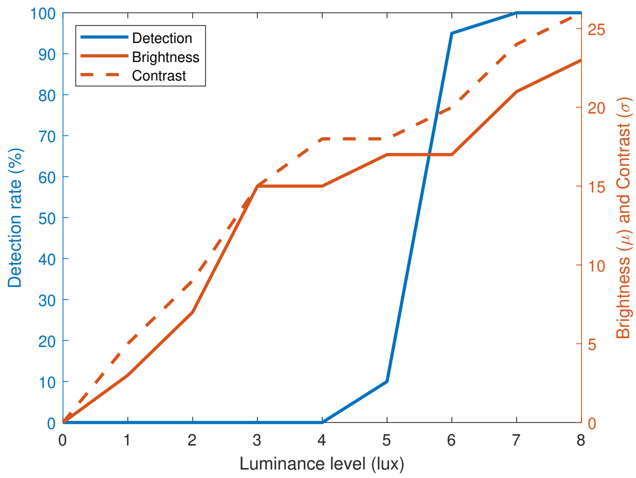


Fig. 3. Graph of detection efficacy based on brightness and contrast.

begins detecting faces at 5 lux and achieves a 100% detection rate at 7 lux or higher.

V. Conclusion

The system has a solid performance in detecting employee engagement and liveness in remote work scenarios, including low-light conditions. It also features mechanisms that notify employers of suspicious behavior. As an automated, low-cost solution, it offers companies a low-cost solution with enhanced automated anti-evasion, oversight, and monitoring of remote employees. Future work includes integrating image pre-processing techniques to enable operation below the current 7 lux threshold. Further plans involve developing a user interface for the human resources team and a back-end system to support advanced data processing and analysis.

References

[1] C. Morris, "Swiss workers must clock out for bathroom breaks, court rules," *Fortune*, 2024. [Online]. Available: https://fortune.com/2024/10/10/swiss-workers-bathroom-breaks/

[2] L. Martin, L. Hauret, and C. Fuhrer, "Digitally transformed home office impacts on job satisfaction, job stress and job productivity. COVID-19 findings," *Plos one*, vol. 17, no. 3, p. e0265131, 2022. [Online]. Available: https://doi.org/10.1371/journal.pone.0265131

[3] R. Săvescu, Ş. Kifor, R. Dănuţ, and R. Rusu, "Transition from office to home office: Lessons from romania during covid-19 pandemic," *Sustainability*, vol. 14, no. 10, p. 5758, 2022. [Online]. Available: https://doi.org/10.3390/su14105758

[4] N. Boyko, O. Basystiuk, and N. Shakhovska, "Performance evaluation and comparison of software for face recognition, based on dlib and opencv library," in *2018 IEEE Second International Conference on Data Stream Mining & Processing (DSMP)*. IEEE, 2018, pp. 478–482. [Online]. Available: https://doi.org/10.1109/DSMP.2018.8478556

[5] S. Mohanty, S. V. Hegde, S. Prasad, and J. Manikandan, "Design of real-time drowsiness detection system using dlib," in *2019 IEEE International WIE Conference on Electrical and Computer Engineering (WIECON-ECE)*. IEEE, 2019, pp. 1–4. [Online]. Available: https://doi.org/10.1109/WIECON-ECE48653.2019.9019910

[6] J. Qi, L. Ma, Z. Cui, and Y. Yu, "Computer vision-based hand gesture recognition for human-robot interaction: a review," *Complex & Intelligent Systems*, vol. 10, no. 1, pp. 1581–1606, 2024. [Online]. Available: https://doi.org/10.1007/s40747-023-01173-6

[7] S.-C. Huang and C.-H. Yeh, "Image contrast enhancement for preserving mean brightness without losing image features," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 5-6, pp. 1487–1492, 2013. [Online]. Available: https://doi.org/10.1016/j.engappai.2012.11.011