



Integrating AI and UAV Technologies for Enhanced and Workplace Safety

Cosmin Badea^(✉), Antonia Costea, Alexandra Halmaghi, Alexandra Hancu,
Ioana Corina Bogdan, and Horia Alexandru Modran

Department of Electronics and Computers, Transilvania University of Brasov, Braşov, Romania
{cosmin.badea, antonia.costea, alexandra.halmaghi,
alexandra.hancu}@student.unitbv.ro, {corina.bogdan,
horia.modran}@unitbv.ro

Abstract. The integration of artificial intelligence (AI) and unmanned aerial vehicle (UAV) technologies presents a significant advancement in enhancing safety in traffic, workplace, and healthcare environments. This study explores the application of AI-driven computer vision algorithms in UAVs to detect and mitigate risks associated with substance abuse, fatigue, and health impairments. Utilizing sophisticated image processing techniques, such as edge detection and support vector machine (SVM) algorithms, drones are equipped to autonomously monitor and analyze ocular characteristics and facial expressions of individuals. The research employs a mobile phone camera and Python-based libraries to conduct real-time assessments, providing critical data to medical and industrial professionals. The study demonstrates the potential of drones to enhance safety by checking sobriety and monitoring worker health. The experimental setup includes a detailed workflow for real-time video detection and facial analysis, leveraging pre-trained models and convolutional neural networks. The results confirm the effectiveness of this approach, highlighting significant progress in AI and UAV technology. Future work aims to transition these innovations from laboratory conditions to practical, real-world applications, continuously enhancing the algorithms and expanding their applicability across various safety-critical scenarios.

Keywords: Drones · Surveillance · Computer Vision · Artificial Intelligence · Image Processing · Healthcare

1 Introduction

In the contemporary era, artificial intelligence (AI) is progressively establishing its prominence across a diverse range of sectors, including but not limited to engineering, education, and healthcare. This sophisticated technology significantly influences the quality of human life by augmenting numerous facets. From medical and social applications to industrial and aerial robotics, these innovative advancements contribute substantially to the enhancement of human life quality. Most robotic systems use intelligent technologies. They streamline everyday tasks, facilitate efficient management of voluminous data, and accelerate processes, culminating in an improved quality of life for individuals [1, 2].

Artificial Intelligence (AI) and intelligent devices represent critical subjects with immense potential implications for our society. Ensuring a secure environment, both in terms of traffic and workplaces (for instance, industries, production lines, etc.), is of paramount importance. Impairments resulting from substance abuse or fatigue not only jeopardize individual health but also pose a threat to the safety and productivity of all stakeholders. Consequently, the integration of AI methodologies for the detection and prevention of these conditions guarantees reliable and precise surveillance, thereby augmenting safety and mitigating risks in traffic and occupational settings. Employing sophisticated ocular analysis, with a particular emphasis on 2D or 3D eye imagery, the engineered system is capable of identifying and categorizing various indicators associated with fatigue, inattention, or the individual's health status.

The majority of robotic systems are engaged in image analysis and processing. A survey of medical robots and other medical equipment (such as noninvasive surgery da Vinci robots, ultrasound scanners, MRI machines), along with social, industrial, and aerial robots, reveals that these entities operate based on image processing and computer vision technology [3–6]. Drones, which have found applications in a wide array of fields ranging from automation to delivery, surveillance, topography, and espionage [7], have undergone development in the realm of image processing and data recognition [8].

For vehicle operators, the integration of a validation system that permits the initiation of the car engine and subsequent driving only after executing a preliminary check of the individual could prove advantageous [9, 10]. This preliminary examination could encompass evaluations of several critical factors such as ocular characteristics (color, pupillary dilation, indications of fatigue), facial expressions, or cardiac rhythm [11, 12]. The objective is to alleviate health or traffic hazards and promote safe driving practices by circumventing impaired driving or maintaining awareness of road conditions and traffic [13, 14]. The same strategies employed in the design and implementation of such preliminary check validation systems could be beneficial for workers in industrial environments (such as production lines), not only to ensure optimal safety conditions for individuals but also for equipment and the environment [15, 16]. Conversely, this issue can also surface in healthcare, for instance, in the emergency room. With the availability of video equipment or a drone that oversees the area, the images of the patients can be evaluated, analyzed in real time, and relayed to the medical teams responsible for emergency care [17].

In this study, a mobile phone camera is employed for image detection. By utilizing Python and its associated libraries, a comparative analysis can be conducted between a healthy individual or eye and an individual afflicted with a disease.

The camera can be strategically positioned in the Emergency Room or ambulance, facilitating improved and expedited usage. This arrangement enables doctors to receive real-time information, thereby allowing nursing staff to promptly identify individuals requiring urgent care.

Various datasets are utilized for expedited detection of patient injuries. The algorithm was trained on subjects for the eye-tracking component of the project.

The paper is structured into five distinct sections. The initial section is dedicated to the articulation of the research objective and a review of related studies. The subsequent section provides an in-depth description of the principal technologies employed in this

project. The third section elucidates the detection algorithm utilized. The fourth section details the experimental setup, while the final section is devoted to the presentation of the results obtained, including conclusions drawn and potential future enhancements to this study.

The objectives of this research are multi-faceted, encompassing (1) traffic monitoring, (2) workplace environment, and (3) healthcare. The study aims to use drones equipped with intelligent technologies to identify traffic participants under the influence of alcohol or psychoactive substances. This methodology could also be beneficial for workplace environments or healthcare. This same testing methodology could prove advantageous for workplace environments or healthcare. For vehicle operators, a validation system that checks drivers before starting the car engine could improve safety. This preliminary examination could encompass evaluations of several critical factors such as ocular characteristics (color, pupillary dilation, indications of fatigue), facial expressions, or cardiac rhythm. The objective is to alleviate health or traffic hazards and promote safe driving practices by circumventing impaired driving or maintaining awareness of road conditions and traffic. Consequently, this project presents potential for the implementation of Unmanned Aerial Vehicle (UAV) technologies in global security measures.

2 Smart Technologies

2.1 Computer Vision and UAVs

Computer Vision in the context of Unmanned Aerial Vehicles (UAVs) represents an advanced intersection of artificial intelligence and drone technology. The primary objective is to enable drones with the capability to perceive and comprehend their environment through the use of smart algorithms and sensors. Equipped with these technological advancements, UAVs gain the ability to navigate autonomously, identify surrounding objects, and make real-time decisions.

Computer Vision applied to Unmanned Aerial Vehicles (UAVs) enables drones to perceive and comprehend their environment. Algorithms empower drones with capabilities like object detection, recognition, and motion tracking. Leveraging techniques such as deep learning and image processing, drones equipped with this technology find extensive applicability across a multitude of industries. Applications of Computer Vision-enabled Unmanned Aerial Vehicles (UAVs) span a wide range of sectors, including agriculture, environmental surveillance, infrastructure inspection, and emergency response. These technologies enhance safety, expedite processes, and augment efficiency, thereby transforming the perception and interaction with the environment from an aerial perspective. As research in this domain progresses, it is anticipated that further innovative advancements in this field will be witnessed.

UAVs can be categorized into fixed-wing and rotary-wing types. Fixed-wing UAVs maintain high-speed flight for extended durations but require a launcher for takeoff and lack the ability to hover. On the other hand, rotary-wing UAVs have the capacity to hover but are characterized by lower speeds and shorter flight ranges. Rotary-wing UAVs often come equipped with integrated cameras. The incorporation of image processing algorithms into the UAV allows for the potential of conducting image analysis during

flight, thereby introducing a higher degree of intelligence to the system and automating the interpretation of these images.

Certain factors influencing the utilization of Rotary Unmanned Aerial Vehicles (RUAVs) include their ability to fly autonomously, the maximum duration of flight they can sustain, and the maximum load they can carry.

2.2 Artificial Intelligence

Three major elements contribute to the evolution of Artificial Intelligence (AI): computational power, data availability, and algorithms. In the context of the drone industry, it is evident that there is effortless access to a vast quantity of data. Consequently, discussions around AI predominantly revolve around the applications of models or algorithms. Computer Vision, an interdisciplinary field, deals with enabling computers to derive high-level comprehension from digital images and videos. A fundamental challenge is to ascertain whether the images encompass the target of interest. This can be further classified into object recognition (also known as classification), or object detection. Recognition aims to answer a binary question, such as confirming the presence of eyes in an image. Detection, conversely, amalgamates location information. It not only identifies the presence of an eye in the image but also pinpoints its location within the image.

In the context of this study, a Support Vector Machine (SVM) algorithm has been utilized to detect and locate specific targets in the images captured by drones. The algorithm maps the input vectors into a high-dimensional feature space where a hyperplane is constructed to separate the different classes based on a defined criterion. The SVM model is trained using a set of labeled training data and then used to detect these targets in new, unseen images. Datasets utilized in this study are sourced from various public domain repositories [18–20] and include synthetic data as well. An example image employed for training the Machine Learning model is presented in Fig. 1.



Fig. 1. Sample image used for training.

The unique aspect of this research is the application of the SVM algorithm for the detection of substance abuse, fatigue, or other diseases of the human eye. This combination of recognition and precise location detection is crucial for enhancing workplace safety with the integration of AI and UAV technologies. The use of SVM for detection

tasks in this research has proven to be highly effective, demonstrating the potential of AI algorithms in the drone industry. This approach underscores the significant role of AI in promoting safety and well-being in the workplace.

3 Detection Algorithms

Segmentation via edge detection is used to delineate boundaries or edges within an image. The objective is to identify sudden shifts in intensity, color, or texture that signify the limits of objects. Edge detection is frequently utilized as an initial step in image processing and computer vision tasks such as object recognition, image segmentation, and feature extraction. The following is a succinct explanation of how segmentation by edge detection, based on algorithms, operates:

- **Image Preprocessing:** The procedure typically commences with image preprocessing to augment the quality of the image and diminish noise. The process involves image preprocessing (grayscale conversion and noise reduction) followed by edge detection.
- **Edge Detection:** The principal phase of the algorithm involves detecting edges in the preprocessed image. There exists a variety of edge detection algorithms, with some of the well-known ones including:
 - **Sobel Operator:** The Sobel operator employs convolution masks to compute the gradient of the image. The gradient signifies the rate of alteration of pixel intensity in the x and y directions, facilitating the detection of edges.

The approach adopted for the project development, carried out by the research team, is executed in several stages. The initial aspect of the methodology was to establish a function for facial detection and analysis with high precision. To achieve this, a pre-trained model with complex datasets was utilized, which employs advanced algorithms to identify facial landmarks. Specifically, the Cascade of Regression Trees algorithm was used, designed for accuracy and precision in detecting the shape and position of facial features. Additionally, they trained an AI model to distinguish individuals under the influence of psychoactive substances based on eye analysis. The differentiation between these two categories of individuals is achieved by analyzing the person's eyes. These two aspects will be integrated to result in real-time detection and analysis of individuals. The strategy of designing and implementing these pre-check validation systems could prove beneficial for workers in industrial environments (such as production lines), ensuring optimal safety conditions not only for individuals but also for equipment and the environment. Conversely, this issue can also emerge in healthcare, for instance, in the emergency room. With the availability of video equipment or a drone to monitor the area, patient images can be evaluated, analyzed in real time, and transmitted to medical teams in emergency situations.

4 Experimental Setup

For implementing the application and testing its functionality, as well as the logic, we used hardware-equipped equipment, with an Intel® Core™ i5-12600H processor, 16 GB of RAM memory, NVIDIA GeForce MX550 and 512 GB SSD.

To test the accuracy of real-time video detection, the camera from the device was used solely to evaluate functionality. Subsequently, to replicate a scenario as close to real-world application as possible, a network connection was established to connect the system to a phone's camera and hover over a room to validate its effectiveness. From a software perspective, the application was developed and run on the Windows operating system without encountering any incompatibilities, being in a functional and stable state.

This section describes the workflow for developing the application, as well as its functionality. The whole process is segmented into a few steps, that are put together as an application.

The whole workflow of the application is described in Fig. 2.

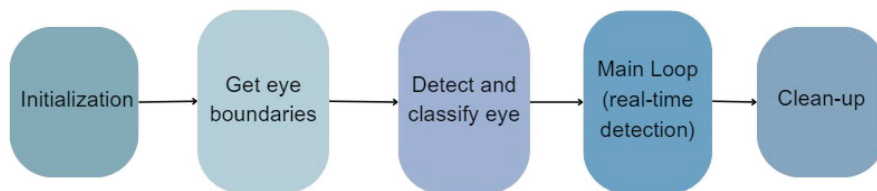


Fig. 2. Application Workflow.

Initially, the drones hover at low altitude, slightly above the cars, so it can detect the eyes of drivers. This is where it comes in the actual logic for eye detection. Based on a pre-trained model, the drone detects faces, implicitly eyes, and that leads to the next process, precisely eye analysis. With authorized datasets, the team trained a convolutional neural network to analyze whether, according to the eye, the driver in question is under the influence of psychoactive substances or alcohol. The model was designed with a classification for each type of substance to provide a clear response.

The process begins with the initialization phase, where several key components are loaded. The face detector is loaded using the `dlib.get_frontal_face_detector()` function, and the landmarks facial predictor is loaded using the `dlib.shape_predictor` function based on the `shape_predictor_68_face_landmarks.dat` file. The video stream from the drone is opened using the `cv2.VideoCapture` function, and the Support Vector Machine (SVM) trained model is loaded using the `joblib.load` function for classification of each type of substance.

The `get_landmarks(frame, bbox)` function is then defined. This function extracts the bounding box coordinates of the face from the `bbox` variable, creates a `dlib.rectangle` object using these bounding coordinates, and returns the predicted landmarks for the face using the predictor. As Fig. 3 illustrates, the eye is detected by the application by using a 6 points boundary.

Next, the `classify_eye(landmarks, classifier, frame, image_size)` function is defined. This function begins by defining the indices for the right and left eyes and initiating an empty list for classification storage. For each set of indices (left and right), the function performs several steps: it extracts the landmark points for the eyes, creates a bounding box for the eyes using the `cv2.boundingRect` function, extracts the eye image using the bounding box coordinates, converts the eye image to greyscale using the `cv2.cvtColor`



Fig. 3. Detecting the eye boundaries in 6 points.

function, resizes the eye image to the specified size using the `cv2.resize` function, flattens the eye image into a feature vector, verifies if the processed image size is correct (4096 features), utilizes the SVM model to predict the eye status, and adds the classification result to the list of classifications. The function then returns this list of classifications.

The main loop of the process involves reading a frame from the video stream. If reading fails, the loop ends. Otherwise, the frame is converted to greyscale using the `cv2.cvtColor` function, and faces are detected from the frame using the face detector. For each detected face, the bounding coordinates of the face are extracted, the face landmarks are obtained using the `get_landmarks` function, and the eyes are classified using the `classify_eye` function. The classification results are then displayed on the frame using the `cv2.putText` function. The processed frame is displayed, and if the user presses the 'q' key, the loop ends.

Finally, the cleanup phase involves releasing the video resources using the `cap.release()` function and closing all windows opened by OpenCV using the `cv2.destroyAllWindows()` function. This concludes the whole process.

The developed application, is designed to detect substance abuse, fatigue, and various health issues through advanced image processing and machine learning techniques. Utilizing a mobile phone camera and Python-based Machine Learning algorithms, the system performs real-time analysis of ocular characteristics and facial expressions. By employing a pre-trained convolutional neural network and Support Vector Machine (SVM) classifier, the application can accurately identify signs of impairment such as pupil dilation, scleral shading, and vascular changes, as well as facial indicators of fatigue. This technology allows for immediate and precise detection of potential health risks, enabling timely intervention in traffic, workplace, and healthcare environments. The integration of these capabilities into UAVs enhances their utility in safety surveillance, providing a robust tool for ensuring individual and public safety.

A few volunteers with IRB agreements participated in the testing phase of this application (Fig. 4). To comply with the General Data Protection Regulation (GDPR), the data collected was anonymized and will be used only for this purpose.

5 Conclusions and Future Works

This study utilized image processing and training based on artificial intelligence to analyze datasets, differentiating individuals with normal ocular conditions from those impacted by ailments such as fatigue, conjunctivitis, or blepharitis. It also identified

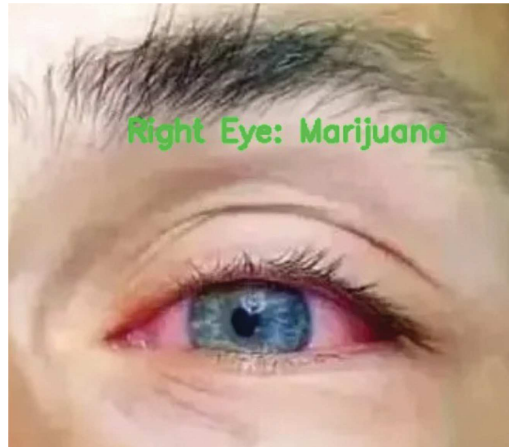


Fig. 4. Detection result.

features like pupil dilation, scleral shading, and vascularity. Additional images were procured using a smartphone to emulate conditions of drone flight. The efficacy of the identification method was corroborated through testing on a representative sample group. This investigation signifies a considerable progression in drone technology, simultaneously fostering the evolution and implementation of intelligent solutions aimed at significantly mitigating workplace mishaps and traffic accidents, thereby enhancing safety in pertinent environments.

The research highlights the considerable potential of Unmanned Aerial Vehicle (UAV) technology in identifying fatigue and impairment induced by psychoactive substances in road users and workers. It also affirms that progress in research related to automated traffic safety surveillance can be realized without significant financial or risk implications. The objective is to eventually transition this project, in conjunction with other related initiatives, from laboratory environments to practical applications. Moreover, the study accentuates the significance of continuous efforts to enhance data processing algorithms and evaluate their effectiveness across a variety of traffic situations, underscoring the constant scope for innovation in this field.

References

1. Shen, Y., et al.: Robots under COVID-19 pandemic: a comprehensive survey. *IEEE Access* **9**, 1590–1615 (2020). PMID: 34976569; PMCID: PMC8675561. <https://doi.org/10.1109/ACCESS.2020.3045792>
2. Aditya, S., Singh, R., Singh, P.K., Kalla, A.: A survey on blockchain in robotics: issues, opportunities, challenges and future directions. *J. Netw. Comput. Appl.* **196** (2021). ISSN: 1084-8045. <https://doi.org/10.1016/j.jnca.2021.103245>
3. D'Ettoire, C., et al.: Accelerating surgical robotics research: a review of 10 years with the da Vinci research kit. *IEEE Robot. Autom. Mag.* **28**(4), 56–78 (2021). <https://doi.org/10.1109/MRA.2021.3101646>

4. Azizian, M., Liu, M., Khalaji, I., DiMaio, S.: The Da Vinci surgical system. In: The Encyclopedia of Medical Robotics, Chapter 1 (2019). https://doi.org/10.1142/9789813232266_0001
5. Jecu, S.A., Demeter, A.G., Bogdan, I.C.: Face-tracking mount display for medical and engineering applications. In: 2023 17th International Conference on Engineering of Modern Electric Systems (EMES), Oradea, Romania (2023). <https://doi.org/10.1109/EMES58375.2023.10171630>
6. Habib, A., Das, S.K., Bogdan, I.C., Hanson, D., Popa, D.O.: Learning human-like facial expressions for Android Phillip K. Dick. In: 2014 IEEE International Conference on Automation Science and Engineering (CASE) (2014). <https://doi.org/10.1109/CoASE.2014.6899473>
7. Perera, G.S.N., Nalani, H.A.: UAVS for a complete topographic survey. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.* **XLIII-B2-2022**, 441–447 (2022). <https://doi.org/10.5194/isprs-archives-XLIII-B2-2022-441-2022>
8. Samadzadegan, F., Dadrass Javan, F., Ashtari Mahini, F., Gholamshahi, M.: Detection and recognition of drones based on a deep convolutional neural network using visible imagery. *Aerospace* **9**, 31 (2022). <https://doi.org/10.3390/aerospace9010031>
9. Premalatha, M., Heymath Kumar, A., Manoj Kumar, M., Pavithran, P., Shatyadeep, K.: Drugged eye detection using image processing. *Ijrasat J. Res. Appl. Sci. Eng. Technol.* (2023). <https://doi.org/10.22214/ijrasat.2023.50427>
10. Pyingkodi, M., Thenmozhi, K., Wilfred Blessing, N.R., Archana, K., Arun Hiruthik, X., Dinesh Kumar, V.: Drug consumption detection by eye using image processing techniques. In: 2023 International Conference on Computer Communication and Informatics (ICCCI), pp. 1–6 (2023). <https://doi.org/10.1109/ICCCI56745.2023.10128445>
11. Chandrasekaran, S.N., Ceulemans, H., Boyd, J., Carpenter, A.: Image-based profiling for drug discovery: due for a machine-learning upgrade? *Nat. Rev. Drug Discovery* **20**, 145–159 (2021). <https://doi.org/10.1038/s41573-020-00117-w>
12. Ting, H., Chung, S., Chen, C., Chiu, H., Hsieh, Y.: A drug identification model developed using deep learning technologies: experience of a medical center in Taiwan. *BMC Health Serv. Res.* **20**, 312 (2020). <https://doi.org/10.1186/s12913-020-05166-w>
13. Kayadibi, I., Güraksın, G.E., Ergün, U., et al.: An eye state recognition system using transfer learning: AlexNet-based deep convolutional neural network. *Int. J. Comput. Intell. Syst.* **15**, 49 (2022). <https://doi.org/10.1007/s44196-022-00108-2>
14. Ktistakis, E., et al.: COLET: a dataset for COgnitive workLoad estimation based on eye-tracking. *Comput. Methods Programs Biomed.* **224**, 106989 (2020). ISSN: 0169-2607. <https://doi.org/10.1016/j.cmpb.2022.106989>
15. Fuentes-Hurtado, F., Diego-Mas, J.A., Naranjo, V., Alcañiz, M.: Automatic classification of human facial features based on their appearance. *PLoS ONE* **14**(1), e0211314 (2019). <https://doi.org/10.1371/journal.pone.0211314>
16. Chandrika, K.R., Amudha, J.: Eye movements dataset for objective-based assessment of object-oriented programming knowledge. *Data Brief* **50**, 10955 (2023). <https://doi.org/10.1016/j.dib.2023.109558>
17. Porta, S., Bossavit, B., Cabeza, R., Larumbe-Bergera, A., Garde, G., Villanueva, A.: U2Eyes: a binocular dataset for eye tracking and gaze estimation. In: 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW), Seoul, Korea (South), pp. 3660–3664 (2019). <https://doi.org/10.1109/ICCVW.2019.00451>
18. NIR Iris images under alcohol effect. <https://ieee-dataport.org/documents/nir-iris-images-under-alcohol-effect>. Accessed 27 May 2024

19. Spot alcohol face – Image Dataset. <https://www.primroselodge.com/help-guides/how-to-spot-an-alcoholic-face/>. Accessed 27 May 2024
20. How to Tell if Someone is on Drugs by Their Eyes – Image Dataset. <https://sobercollege.com/addiction-blog/tell-if-someone-is-on-drugs-by-eyes/>. Accessed 27 May 2024