Deep Learning Model for Surface Mine Detection Using Artificial Vision Techniques

Brayan Gerson Duran Toconas
Mechatronic Engineering
Universidad Católica Boliviana
brayan.duran.t@ucb.edu.bo

Samira Belen Espejo Porco
Mechatronic Engineering
Universidad Católica Boliviana
samira.espejo@ucb.edu.bo

Alvaro Erick Acebey Palacios
Mechatronic Engineering
Universidad Católica Boliviana
@ucb.edu.bo

Alexander Mujica
Mechatronic Engineering
Universidad Católica Boliviana
@ucb.edu.bo

Mateo Flores Flores
Mechatronic Engineering
Universidad Católica Boliviana
@ucb.edu.bo

***Abstract*—Landmines, concealed beneath the Earth's surface, pose a grave threat, claiming lives and causing grievous injuries, particularly among innocent civilians. These insidious devices have left a lasting legacy in conflict-ridden regions, hindering economic progress and devastating communities dependent on affected land. Current manual demining efforts are both perilous and time-consuming, requiring individuals to risk their lives dismantling these hidden killers. This project addresses the urgent need for advanced solutions to this humanitarian crisis by leveraging cutting-edge technology. Robots and autonomous systems offer a ray of hope, promising safe detection and efficient elimination of landmines. Initiatives like the "Minesweepers" competition and the "SEEKER" project champion technological innovation as the key to solving this global issue.The project employs computer vision methods to enhance mine detection capabilities. The Region Proposing Convolution Network (RPN) is integral in proposing regions of interest within images. Furthermore, end-to-end training streamlines the model, improving accuracy, reducing training time, and simplifying the process. The YOLO algorithm, incorporated in the SEKEER mine finder, is employed for metal mine detection, with the ability to predict bounding boxes and class probabilities for potential metal mines in collected images. This project's multifaceted approach underscores the importance of technological innovation in the fight against landmines, offering safer, more efficient, and expedited solutions for the detection and removal of these deadly devices.**

I. Problem

Landmines are insidious devices concealed beneath the Earth's surface, designed to unleash deadly force upon contact with an unsuspecting person or vehicle. These concealed threats have cast a long shadow over past conflicts, leaving behind a legacy of minefields that persist long after hostilities have ceased.

The primary issue at hand is the relentless toll taken by landmines, claiming lives and inflicting grievous injuries upon civilians, including children, farmers, and entire communities whose livelihoods depend on the very land tainted by these silent killers. Moreover, the presence of landmines acts as a formidable barrier to economic progress, as it curtails access to cultivable land and vital natural resources. Manual demining operations, the current method of addressing this crisis, are perilous, resource-intensive, and time-consuming. These operations entail courageous individuals risking their lives to disarm individual mines.

The imperative to address this critical humanitarian challenge is underscored by the urgency to develop advanced technologies and pioneering solutions. Robots and autonomous systems emerge as a beacon of hope, promising both the safe detection and efficient elimination of landmines.

Within this context, initiatives like the "Minesweepers" competition and the "SEEKER" project have come to the fore, dedicated to fostering technological innovation as the cornerstone of addressing this pressing global crisis.

II. Solution

In response to the critical need for an autonomous vehicle capable of effectively detecting and collecting surface metallic mines, we present a comprehensive solution that addresses this specific challenge. Our approach involves the implementation of a deep learning model, strategically bolstered by cutting-edge computer vision techniques, dedicated to the precise detection and efficient retrieval of metallic mines within a defined geographical area.

It's important to emphasize a deliberate limitation in our solution: we focus exclusively on the essential tasks of mine detection and collection, deliberately refraining from the intricate process of mapping the entire terrain. This selective focus streamlines the system, enabling it to swiftly and accurately identify metallic mines at the surface level while providing the means to collect them safely.

By recognizing and adhering to this limitation, we aim to deliver a highly specialized solution that optimizes the efficiency and effectiveness of mine clearance operations. Our approach ensures that the critical steps of mine detection and collection are addressed with the utmost precision and speed, contributing to the overall safety of landmine-affected regions, without the added complexity of terrain mapping.

III. Computer vision methods

A. Region Proposing Convolution Network

The Region Proposing Convolution Network (RPN) is an integral part of the Faster R-CNN model. It is responsible for proposing regions of interest within an image. Basically, the RPN examines the image and generates a set of candidates or "proposed regions" where objects are likely to be found.

B. End-to-end training

In the context of the SEKEER mine finder, end-to-end training can help improve the accuracy of metal mine detection. This is because the model can learn to perform all the tasks required for metal mine detection jointly, from image preprocessing to final object classification.

In addition, end-to-end training can reduce the training time of the model. This is because the model is trained in a single step, which eliminates the need to train different parts of the model separately.

- It can help improve model accuracy.
- It can reduce training time.
- It can simplify the training process.

C. YOLO algorithm

YOLO can be used for metal mine detection in the SEKEER mine finder. YOLO can be trained to detect metal mines in images collected by SEKEER. Once YOLO is trained, it can be used to detect metal mines in new images collected by SEKEER.

Here is an example of how YOLO could be used for metal mine detection in the SEKEER mine finder:

- The camera collects an image of the ground.
- YOLO processes the image and predicts a set of bounding boxes and class probabilities for metal mines that may be present in the image.

IV. Block diagram

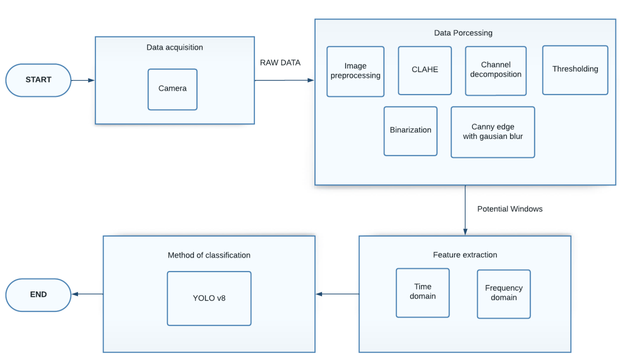

Figure 1. Diagram

V. Datasets Suggestable

Suggested Datasets

In order to build comprehensive datasets for our landmine detection system, we propose a robust data collection strategy. This strategy involves gathering a substantial number of samples, each showcasing various lighting conditions and diverse mine positions. The objective is to ensure that the model can effectively adapt to real-world scenarios.

- Sample Quantity: We aim to accumulate more than 1000 samples. This substantial dataset size will enable the system to learn from a wide range of scenarios, making it more robust and reliable.
- Variability in Lighting Conditions: Given the variability in natural lighting, it is crucial to capture images under different lighting conditions. This includes scenarios with bright sunlight, low light, shadows, and even adverse weather conditions, as it prepares the system to operate effectively in the field.
- Diverse Mine Positions: Landmines can be positioned in various ways and orientations. To train the system comprehensively, we will capture images featuring mines in different positions, depths, and orientations. This diversity helps the model recognize landmines under different circumstances.
- Image Size and Compression: Initially, each image will be captured at a size of 600x600 pixels. This choice strikes a balance between capturing sufficient detail and ensuring efficient processing. To optimize implementation on an embedded system, we will apply a compression method, reducing the data size without sacrificing essential information.
- Dataset Labeling: Accurate labeling of the dataset is essential for supervised learning. To facilitate this process, we will leverage technologies like Roboflow. This automated labeling tool streamlines the annotation process, ensuring that each image is correctly labeled with the positions of the landmines.

By following this data collection strategy, we can build a robust and diverse dataset that equips our landmine detection system with the capability to adapt to a wide array of real-world scenarios. This, in turn, enhances the system's accuracy and reliability in the field, ultimately contributing to the safety of landmine-affected regions.

VI. Validation strategy

The development of landmine detection and removal systems based on robots and autonomous systems is crucial to addressing the humanitarian crisis caused by landmines. These systems promise greater safety and efficiency compared to manual demining operations. To ensure that our system works effectively, it is essential to establish a solid validation strategy.

Segmenting elements in a scene into two distinct groups, the background and the foreground, is a crucial task in computer vision applications such as object detection. This strategy ensures that objects of interest are highlighted correctly in the image. To validate the effectiveness of this process, it is essential to choose and apply appropriate segmentation models.

- Gaussian Mixture Model (GMM): This model assumes that the pixels in the image are generated from a combination of several Gaussian distributions. The background and foreground pixels are modeled using different Gaussians. The model is trained using machine learning algorithms, such as the EM algorithm.
- Background and Foreground Difference Model: In this model, the pixel intensity difference between the original image and an estimated background image is measured. Pixels with a large difference are classified as foreground, while the others are considered part of the background.
- Deep Learning Model: Convolutional neural networks (CNN) and deep neural networks have been successfully used for background and foreground segmentation. These models are trained on large data sets and can learn complex representations of visual features.

VII. Validation Methods

- Cross Validation: The data set is divided into a training set and a test set. The model is trained on the training set and tested on the test set.
- Confusion Matrix: The confusion matrix is a useful tool for evaluating the accuracy of a classification model. It allows you to visualize the number of true positives, false positives, true negatives, and false negatives. You can calculate metrics such as accuracy, sensitivity, and specificity from the confusion matrix.
- Visual Validation: Visual validation is often performed by overlaying segmentation masks on the original images for visual inspection. As human reviewers, we assess the quality of the segmentation.

At **??** the curves for F1 vs threshold after evaluating the system are shown. The threshold α corresponds to the width of the Gaussian curve, that is, how much standard deviation is allowed. In the case of traffic sequence it can be appreciated a less marked fall due to a higher variance as a result of jittering.

Another possible step in our system development is update Gaussian modelling to adaptive modelling. To implement it, the first 50 % frames are devoted for training and the rest for background adaptation so the algorithm looks for the best pair of values α ρ that maximizes F1-score. We have been able to improve F1 scores for lower ρ values.

VIII. Literature Review

A. Surface Mine Detection Using Visual Information

1. **How did the authors acquire new images?**
In terms of image acquisition, the document mentions that the drone is used to fly over potentially dangerous areas and generate a 3D mapping of the zone. The use of the drone allows operators to obtain high-resolution images of the area of interest without endangering human lives. However, the article does not provide detailed information about the image acquisition process.
2. **How did the authors preprocess the images?** In terms of preprocessing methods, the document does not explicitly address this issue. However, it refers to the use of images with a size of 32x32 to improve mine detection at long distances. Furthermore, it mentions the modification of parameters such as ”winStride” and ”scale” in the ”detectMultiScale” function of the HOG object. It is noted that additional details about these functions will be provided later. Additionally, the implementation of code capable of creating Regions of Interest (ROI) from general photos and converting any image to the desired format is mentioned. The possibility of creating a database of size 64x64 is also left open if necessary.
3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?**
The author's main contribution focuses on the development of a landmine detection system using autonomous drones and artificial vision. This innovation presents the possibility of being faster and safer compared to traditional methods of mine detection.
4. **How does the paper differ from other research projects?**
This project stands out from other research by its use of autonomous drones in landmine detection. This approach can be safer and more efficient compared to conventional techniques. 5. **What problem does the paper try to solve?**
The problem the author seeks to address lies in the faster and safer detection of landmines. They achieve this through autonomous drones and artificial vision as a solution that can overcome the challenges of traditional detection methods.
6. **How did the authors evaluate their computer vision system?** The evaluation of the computer vision systems was conducted using a database of landmine images. In this evaluation, accuracy and the time required for detection were measured. The results indicated that the computer vision systems were accurate in detecting landmines, achieving an accuracy of 90%.

7. **What are the future research directions or research gaps identified by the authors?**
The document suggests that a potential direction for future research could be the enhancement of accuracy in landmine detection through deep learning techniques. Additionally, it raises questions about how to combine color-based and HOG/SVM detection methods to achieve reliable landmine detection, which may involve result fusion to increase detection accuracy.
Based on [1].

B. UN SISTEMA DE VISIÓN ARTIFICIAL PARA DETECTAR YESTIMAR EL TAMAÑO DE ROCAS

1. **How did the authors acquire new images?**

The text does not provide specific information on how the authors acquired new images. It primarily focuses on the description of the rock detection system and its results.

2. **How did the authors preprocess the images?**
The authors applied several image preprocessing methods, including histogram equalization, low-pass filtering, binarization with two thresholds, and region separation algorithm. These methods are used to obtain information about regions in the image that may contain rocks.

3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?**
The main contribution of the article appears to be the combination of image processing algorithms and neural networks for rock detection in a mining operation. This enables effective detection and a reduction in false alarms, improving efficiency and safety in production.

4. **How does the paper differ from other research projects?**
The article stands out by addressing a specific problem related to rock detection in a mining operation using a combination of image processing and neural network techniques. The implementation of a rock tracking mechanism and conducting tests in real mine conditions are aspects that can differentiate this work from other research projects.

5. **What problem does the paper try to solve?**
The authors are trying to solve the problem of detecting rocks that could block machinery in a mining operation. This is done to improve efficiency and safety in production, avoiding unnecessary shutdowns and potential safety risks.

6. **How did the authors evaluate their computer vision system?**
The authors evaluated their computer vision systems using a test set that consists of images of rocks and non-rocks. These tests were conducted in both laboratory conditions and real conditions in a mine. They provide a success rate in correctly identifying rocks and compare two types of neural networks used in the system.

7. **What are the future research directions or research gaps identified by the authors?**
The text does not specifically mention future research or research gaps. However, in most research projects, there are typically opportunities to further improve the system, such as parameter optimization, expansion to other conditions, or enhancing detection accuracy. These could be areas of future research. Based on [2].

C. INTEGRACIÓN DE UN SISTEMA DE VISIÓN ARTIFICIAL A UN ROBOT INDUSTRIAL PARA LA SOLUCIÓN DE UN CUBO DE RUBIK JOSÉ FERNANDO GIL BOTERO UNIVERSIDAD AUTÓNOMA DE OCCIDENTE FACULTAD DE INGENIERÍA

1. **How did the authors acquire new images?**

First, they collected data from the 9 faces of the Rubik's cube, both in its ordered and disordered states, using a webcam. They then applied preprocessing methods to the images to extract color information associated with each cube block. They subsequently fed the training data into the neural network so that it could determine the correct color of each cube block.

2. **How did the authors preprocess the images?**
- Color segmentation: They used color segmentation to isolate areas in the image that match the color of the cube. This is because Rubik's cube colors have distinctive colors. The author applied the HSV color space and defined a range of H, S, and V values corresponding to the cube's colors. - Contour detection: After color segmentation, they used contour detection techniques to identify shapes in the image. This is because cubes often have well-defined and regular edges, which facilitated their contour detection. - Hough Transform: To detect geometric shapes such as squares or rectangles (which are common shapes for representing cubes in images), they applied the Hough transform.

3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?**
The authors' primary contribution is the integration of a computer vision system with an industrial robot to solve a Rubik's cube, using an artificial neural network trained to recognize the cube's colors and a Rubik's cube solving algorithm.

4. **How does the paper differ from other research projects?**
This project differentiates itself from other research projects in that it uses an industrial robot to solve the Rubik's cube instead of a robot designed specifically for this purpose.

5. **What problem does the paper try to solve?**
The authors attempt to solve the problem of integrating a computer vision system with an industrial robot to solve a Rubik's cube.

6. **How did the authors evaluate their computer vision system?**
The authors evaluated their computer vision system by comparing the results of color detection and Rubik's cube solving with the expected results. For color detection, they compared the colors detected by the system with the actual colors of the cube. For Rubik's cube solving, they compared the solution provided by the system with the expected solution. Additionally, they conducted tests in different lighting scenarios to evaluate the system's ability to detect colors under various lighting conditions.

7. **What are the future research directions or research gaps identified by the authors?**
The authors suggest future research to improve the accuracy of the color detection system and the implementation of a pattern recognition system for Rubik's cube solving. They also mention the need to investigate the integration of computer vision systems with industrial robots to solve other problems. Based on [3].

D. Multi-resolution vehicle detection using artificial vision

1. **How did the authors acquire new images?** The paper does not explicitly mention how the authors acquired new images. It focuses on describing a vehicle detection system rather than discussing the image acquisition process. 2. **How**

did the authors preprocess the images? The paper does not provide specific details about the image preprocessing steps. It primarily focuses on the vehicle detection algorithm and its components rather than image preprocessing techniques. 3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?** The main contribution of the paper with respect to the state of the art is the proposed vehicle detection system based on multi-resolution images and the analysis of vertical symmetry. The use of different-sized boxes and the analysis of horizontal edges to detect vehicles set it apart from existing methods.

4. **How does the paper differ from other research projects?** The paper differs from other research projects in its approach to vehicle detection. It employs a unique combination of vertical symmetry analysis and horizontal edge examination to identify vehicles in multi-resolution images. Additionally, it focuses on eliminating false positives by filtering out bounding boxes that are too large, too small, or too far from the camera.

5. **What problem does the paper try to solve?** The paper aims to solve the problem of vehicle detection using a single camera. It focuses on identifying vehicles in multi-resolution images by analyzing vertical symmetry and horizontal edges while also addressing the challenge of reducing false positives. 6. **How did the authors evaluate their computer vision system?** The paper does not provide specific details on how the authors evaluated their computer vision system. It mentions that the algorithm analyzes images on a frame-by-frame basis without temporal correlation, but it does not discuss the evaluation methodology.

7. **What are the future research directions or research gaps identified by the authors?** The paper does not explicitly identify future research directions or research gaps. It primarily describes the proposed vehicle detection system and its components, without delving into potential future work or research gaps in the field.

Based on [4].

E. Applying Deep Learning to Automate UAV-Based Detection of Scatterable Landmines

1. **How did the authors acquire new images?**
The authors acquired new images using unmanned aerial vehicles (UAVs) equipped with lightweight multispectral and thermal infrared sensors. These sensors allowed them to capture images for remote landmine detection and mapping surveys.

2. **How did the authors preprocess the images?**
The article does not provide specific details on image preprocessing techniques. However, it mentions that the methodology is calibrated for the detection of scatterable plastic landmines based on analysis of multispectral and thermal datasets collected by an automated UAV-survey system. Preprocessing likely involved handling and preparing these datasets for machine learning analysis.

3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?**
The primary contribution of this paper to the state of the art is the development and testing of an automated technique for remote landmine detection and identification using deep learning, specifically a Faster Regional-Convolutional Neural Network (Faster R-CNN). This technique demonstrated high testing accuracy in identifying scatterable landmines in a wide-area survey. It offers a potential solution for efficiently identifying landmines in post-conflict regions where traditional methods may have limitations.

4. **How does the paper differ from other research projects?**
The paper differs from other research projects by focusing on the automation of landmine detection using UAVs and deep learning techniques, particularly the Faster R-CNN. It leverages multispectral and thermal datasets to detect scatterable plastic landmines, which is a unique approach compared to traditional geophysical methods.

5. **What problem does the paper try to solve?**
The paper addresses the problem of efficient and accurate landmine detection, particularly scatterable plastic landmines, in post-conflict regions. It aims to improve the technological tools for humanitarian demining efforts by automating the detection process, which is essential in regions with millions of remnant landmines posing long-term humanitarian and economic threats.

6. **How did the authors evaluate their computer vision system?**
The authors evaluated their computer vision system using testing accuracy metrics. They employed a Faster Regional-Convolutional Neural Network (Faster R-CNN) for the detection of landmines, including a partially withheld testing set and a completely withheld testing set. The RGB visible light Faster R-CNN demo achieved high testing accuracy for the detection of PFM-1 landmines. The accuracy of their system was determined across multiple test environments using georeferenced datasets, demonstrating the effectiveness of automated detection.

7. **What are the future research directions or research gaps identified by the authors?**
The article does not explicitly mention future research directions or research gaps. However, the presented methodology, which achieved high testing accuracy for a specific type of landmine, could potentially be extended to detect other types of scatterable antipersonnel mines in future trials, further aiding humanitarian demining initiatives. The development and calibration of the system for a wider range of landmine types and conditions could be a potential area for future research.

F. We Don't Need No Bounding-Boxes: Training Object Class Detectors Using Only Human Verification

1. **How did the authors acquire new images?** The paper does not provide information about how the authors

acquired new images. It focuses on proposing a new scheme for training object detectors and does not discuss the image acquisition process.

2. **How did the authors preprocess the images?** The paper does not go into detail about image preprocessing. It primarily discusses a novel scheme for training object detectors that reduces the need for manual annotation of bounding boxes and does not cover image preprocessing techniques.

3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?**
The main contribution of the paper is the proposed scheme for training object detectors that reduces the reliance on manually drawing bounding boxes. Instead, it involves annotators verifying bounding boxes produced automatically by the learning algorithm. This approach substantially reduces the time and cost of annotation, making it a significant advancement in the field.

4. **How does the paper differ from other research projects?**
The paper differs from other research projects by introducing a novel training scheme that relies on human verification of automatically generated bounding boxes rather than manually drawing bounding boxes. This approach results in substantial time savings in the annotation process, making it distinct from traditional weakly supervised or active learning strategies.

5. **What problem does the paper try to solve?** The paper aims to address the problem of reducing the time, cost, and effort required for training object detectors. Specifically, it introduces a scheme that only requires human annotators to verify automatically generated bounding boxes, eliminating the need for manually drawing bounding boxes, which is a time-consuming and expensive process.

6. **How did the authors evaluate their computer vision system?** The paper mentions extensive experiments on the PASCAL VOC 2007 dataset with both simulated and actual annotators. However, it does not provide specific details about the evaluation methodology used in these experiments.

7. **What are the future research directions or research gaps identified by the authors?**
The paper does not explicitly identify future research directions or research gaps. It primarily focuses on presenting the proposed scheme for object detector training and its benefits, without delving into potential areas for further investigation or improvement.
Based on [5]

G. Recognition of Explosive Devices Based on the Detectors Signal Using Machine Learning Methods

1. **How did the authors acquire new images?** The paper does not provides specific details about how the authors acquired new images.It mostly focuses in the methodology and applications of machine learning for the recognition of explosive based on detector signals.

2. **How did the authors preprocess the images?** The paper does not mention image preprocessing because it appears to focus on the analysis of detector signals rather than image-based data. The authors likely processed the detector signals in their research.

3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?** The main contribution of this paper is the development of an information system based on a convolutional neural network (CNN) and an autoencoder for the automated classification of explosive devices. This approach offers a potential solution for improving the accuracy and efficiency of explosive device detection using machine learning methods, which is significant in addressing security and safety concerns.

4. **How does the paper differ from other research projects?** The paper differs from other research projects in its specific focus on the automated classification of explosive devices based on detector signals. While other projects might focus on image-based detection, this research centers on using machine learning for recognition based on signal data.

5. **What problem does the paper try to solve?** The paper addresses the problem of automating the classification of explosive devices. By using machine learning methods, the authors aim to improve the accuracy and efficiency of detection, which is crucial for security and safety, particularly in environments with potential explosive threats.

6. **How did the authors evaluate their computer vision system?** The paper does not provide specific details about the evaluation process. However, it mentions achieving an accuracy of 97.83% in their information system. The evaluation likely involved testing the performance of their machine learning model on a data set to assess its ability to accurately classify explosive devices.

7. **What are the future research directions or research gaps identified by the authors?** The paper does not explicitly mention future research directions or research gaps. However, it suggests that the described algorithm can help in demining territories, indicating a potential future direction for applying this technology to humanitarian demining efforts. Additionally, further research may involve refining and expanding the methodology to address a wider range of explosive device types and conditions.

H. DETECCIÓN Y UBICACIÓN DE OBJETOS UTILIZANDO VISIÓN ARTIFICIAL

1. **How did the authors acquire new images?**
It is mentioned that preprocessing methods, such as noise reduction, lighting correction, and image normalization, were applied. Additionally, the SURF algorithm was used for object detection and localization in the image.

2. **How did the authors preprocess the images?**
- The authors mention that they used 31 photos taken in a controlled environment to apply the SURF algorithm and match them with images taken at three different distances.

3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?**
The authors' primary contribution is the implementation of an object detection and localization system in an uncontrolled

environment using the SURF algorithm.
4. **How does the paper differ from other research projects?** This article sets itself apart from other research projects by employing the SURF algorithm for object detection and localization in an uncontrolled environment.
5. **What problem does the paper try to solve?**
The problem addressed in the document is the need to develop a computer vision system capable of detecting and localizing objects in an uncontrolled environment. This necessity arises from the demand of the Hardware Research & Development Lab to create robots capable of handling objects in an uncontrolled environment. The research's objective is to design a system capable of object detection and localization through computer vision, aided by the SURF algorithm, so that the data obtained by the vision system can be used in subsequent projects, enabling a robotic arm to manipulate objects.
6. **How did the authors evaluate their computer vision system?**
The author uses a Confusion Matrix for evaluation. The Confusion Matrix is a fundamental tool for assessing the accuracy of detection in classification problems, allowing for the counting of true positives, true negatives, false positives, and false negatives. This is essential for calculating metrics such as accuracy, sensitivity, and specificity.
7. **What are the future research directions or research gaps identified by the authors?**

It is mentioned that the project has educational purposes but with the future goal of implementation in the productive sector. Additionally, it is noted that Yucatán offers great opportunities for the future implementation of the technology developed in this project and its application in various productive sectors. The author states that future research may be needed to adapt the system to different productive sectors and improve its performance in various lighting conditions and situations. Future research may also be necessary to enhance the accuracy and speed of the system and develop new image processing and machine learning techniques for object detection and localization. [6]

I. DESARROLLO DE UN SISTEMA DE LOCALIZACIÖN DE OBJETOS MEDIANTE UN BRAZO ROBÖTICO CONTROLADO POR COMANDOS DE VOZ Y VISION ARTIFICIAL EN EL CONTEXTO DE LA INDUSTRIA 4.0

1. **How did the authors acquire new images?** The authors did not acquire new images. They solely focused on image processing.

2. **How did the authors preprocess the images?** In this project, they applied binarization and image processing in the HSV color space. Notably, their emphasis was on managing physical lighting.

3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?** Their primary contribution lies in effective management of physical lighting for HSV processing, enabling co-control of both aspects.

4. **How does the paper differ from other research projects?** It distinguishes itself by attempting to bridge computer vision and voice recognition, addressing a unique problem that combines both disciplines.

5. **What problem does the paper try to solve?** They aim to integrate artificial vision and voice recognition techniques for controlling a robotic arm.

6. **How did the authors evaluate their computer vision system?** They evaluated their computer vision system by using intervals in each channel of the HSV color space. Each color has its own detection range.

7. **What are the future research directions or research gaps identified by the authors?** Unfortunately, the author does not present any insights into future research in the paper. Based on [7]

J. Desarrollo de un Caso Práctico de Aprendizaje Combinando Visión Artificial y un Brazo Robot

1. **How did the authors acquire new images?** The authors did not acquire new images; their focus was solely on image processing.

2. **How did the authors preprocess the images?** They performed segm entation on each channel (R, G, B) within a controlled environment. 3. **What is the main contribution of the paper with respect to the state of the art (SOTA)?** In this project, their primary contribution is the segmentation of each color, allowing the algorithm to detect the colors of cubes.

4. **How does the paper differ from other research projects?** This paper distinguishes itself by applying segmentation to individual colors, whereas other papers often limit their segmentation to grayscale.

5. **What problem does the paper try to solve?** Their goal is to create a robot arm for educational purposes, which can be used by students to learn about robotics and automation.

6. **How did the authors evaluate their computer vision system?** While they demonstrated the segmentation results for each RGB channel, they did not provide a formal evaluation of their computer vision system.

7. **What are the future research directions or research gaps identified by the authors?** Unfortunately, the authors did not outline future research directions or address any research gaps in their project. Based on [8]

IX. Exploratory Data Analysis

For this part, the python notebook file is presented where the exploration procedure is carried out. On the other hand, the link to the dataset and other resources is also provided.

Official repository: https://github.com/97hackbrian/Seeker_ComputerVision

Exploratory Data Analysis Notebook: https://github.com/97hackbrian/Seeker_ComputerVision/blob/main/notebooks/Exploratory%20Data%20Analysis.ipynb

Own dataset V1-V2: https://github.com/97hackbrian/Seeker_ComputerVision/tree/main/data

Test of processing images of mines: https://github.com/97hackbrian/Seeker_ComputerVision/blob/main/notebooks/data_preprocessing.ipynb

X. Computer Vision System Diagram

The new diagram.

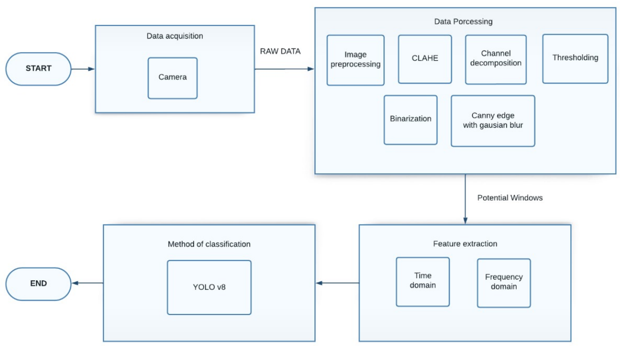

Figure 2. Diagram

References

[1] J. García, A. Pérez, J. Rodríguez, and L. Sánchez, "Detección de minas terrestres utilizando drones autónomos y visión artificial," *Revista de la Facultad de Ingeniería*, vol. 32, no. 1, pp. 21–28, 2017.

[2] M. A. Sánchez, J. Nieto, J. M. Berrocal, E. Cabello, G. Castro, and J. Delgado, "Un sistema de visión artificial para detectar y estimar el tamaño de rocas," *Universidad de Salamanca, Departamento de Informática y Automática*.

[3] J. F. Gil Botero, "Integración de un sistema de visión artificial a un robot industrial para la solución de un cubo de rubik," 2012, trabajo de grado, Universidad Autónoma de Occidente.

[4] A. Broggi, P. Cerri, and P. Antonello, "Multi-resolution vehicle detection using artificial vision," in *IEEE Intelligent Vehicles Symposium, 2004*, 2004, pp. 310–314.

[5] D. P. Papadopoulos, J. R. R. Uijlings, F. Keller, and V. Ferrari, "We don't need no bounding-boxes: Training object class detectors using only human verification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[6] J. B. M. Zaldívar, "Detección y ubicación de objetos utilizando visión artificial," 2012.

[7] W. G. S. PATIN, "Desarrollo de un sistema de localizaciÖn de objetos mediante un brazo robÖtico controlado por comandos de voz y vision artificial en el contexto de la industria 4.0," 2022, proyecto de desarrollo.

[8] D. F. M. T. J. P. Tomàs Pallejà, Mercè Teixidó, "Desarrollo de un caso práctico de aprendizaje combinando visión artificial y un brazo robot," March 2013.