Design and Evaluation of a Real-Time Barcode Detection System Using Video Analysis

Alysson P. Nascimento, Otacílio de A. Ramos Neto and Ruan D. Gomes

***Abstract*— This paper presents the BarcodeVision system, designed for real-time barcode detection and reading in uncontrolled environments. The first version of the system implements four reading approaches, ranging from methods that use only the ZBar library to strategies with additional processing to improve detection and reading, such as multithreaded frame rotation, the use of You Only Look Once (YOLO) detection, and region-of-interest (ROI) reading. The experiments were performed using a video recorded on an automated treadmill, resized to 11 different resolutions, ranging from VGA to DCI 4K. In total, 218 frames containing readable barcodes were counted, which were used as the basis to calculate the percentage of reading and the average processing time of each method. The results showed that the method that combines ZBar with multithreaded rotation presented the highest reading percentage, evidencing the effectiveness of rotation compensation. YOLO-based methods demonstrated competitive performance, especially in scenarios with codes present, with shorter response times.**

***Keywords*— Barcode detection, computer vision, Industry 4.0 applications.**

I. Introduction

The use of barcodes has become a standard for traceability and logistics management, as they provide a high degree of reliability in item identification throughout the production chain [4]. Barcodes can be read using video as input through computer vision algorithms, which has become a strategic technology for the automation of industrial processes, enabling more accurate and agile monitoring and control of inputs [9].

Real-time detection and reading of barcodes represent a key solution to the challenges of Industry 4.0, where efficiency, automation, and connectivity must converge to support dynamic and integrated operations [9]. Deep learning-based approaches, such as YOLO detector adaptations to recognize and localize one-dimensional (1D) and two-dimensional (2D) barcodes, have already demonstrated detection rates exceeding 99% on standard benchmarks [1]. Hybrid methods that combine classical processing techniques, orientation estimation, and specialized filtering have achieved high robustness in low-resolution or blurred images, while maintaining real-time performance [3]. Meanwhile, approaches that integrate drones and ground-based systems leverage barcodes as landmarks for autonomous navigation, optimizing inspection trajectories in complex warehouse environments [2], [4].

Despite these advances, there is still a lack of comparative studies that evaluate multiple reading strategies within the same capture pipeline under different resolution and orientation conditions. In this context, this work presents the BarcodeVision system and analyzes four distinct reading methods: (i) direct reading using the ZBar library[1]; (ii) ZBar with *multithreaded* frame rotation; (iii) YOLO-based detection followed by ROI extraction and reading using method (ii); and (iv) YOLO-based detection followed by full-frame reading using method (ii). These approaches were evaluated to identify the most effective solution for simulated industrial scenarios on an automated conveyor belt.

In this paper, experiments were conducted using a video recorded on an automated conveyor belt, resized to 11 different resolutions, with the goal of evaluating four different reading approaches implemented in the BarcodeVision system. The analysis was based on two main metrics: the percentage of reading, which indicates the effectiveness of barcode detection and reading, and the average processing time per frame, which reflects the feasibility of the system for real-time applications. The results show that the method combining the ZBar library with multithreaded rotation achieved the best performance in terms of reading accuracy, while YOLO-based approaches stood out for their speed.

II. Related Work

The use of computer vision for barcode detection and reading has been widely explored in various industrial applications, especially in the context of logistics automation and Industry 4.0. Several approaches have been proposed in the literature to address challenges such as orientation variation, low resolution, the presence of multiple codes, and real-time execution.

In [1] a barcode detector is described, which uses YOLO to perform the simultaneous detection and classification of 1D and 2D barcodes in real time. The system also included a method to predict the orientation of the barcode, which enables automatic image rotation prior to reading. The results showed a detection rate above 99% in the Muenster dataset [8], demonstrating its viability in the detection aspect.

In the context of automated warehouse inventory, in [2] the WareVision system is described, which employs Convolutional Neural Networks (CNN) for barcode detection combined with drone trajectory optimization. The approach aims to reduce

Alysson P. Nascimento, PPGTI, Instituto Federal da Paraíba (IFPB), João Pessoa-PB, e-mail: alysson.nascimento@polodeinovacao.ifpb.edu.br; Ruan D. Gomes, Instituto Federal da Paraíba (IFPB), João Pessoa-PB e-mail: ruan.gomes@ifpb.edu.br; Otacílio de A. Ramos Neto, Instituto Federal da Paraíba (IFPB), João Pessoa-PB e-mail: otacilio.ramos@ifpb.edu.br. This work is supported by EMBRAPII (BFA 2301.0001) and the companies Cisco, Prysmian, and MPT Cable. The authors also thank CPQD, Inatel, Taggen, Data Machina, CNPq (305536/2021-4), IFPB, and the IFPB Innovation Hub.

[1] `https://pypi.org/project/zbar/`

inspection time in large storage facilities by using barcodes as visual markers for aerial navigation.

In [3], the Bars Detection method is described, based on filtering, binarization, and geometric analysis to identify barcode bars. The barcode orientation is estimated using least squares and refined through the Hough transform. The method achieved 100% accuracy in real time on images captured by a webcam, even under adverse conditions such as blur or low resolution.

Approaches that use videos as a data source are also being explored. In [4] a solution composed of three integrated techniques is proposed, such as key frame selection, recognition of multiple barcode regions by geometric connectivity, and orientation adjustment to facilitate reading. This approach aims to enable the use of drones for inventory in warehouses, eliminating the need for point-by-point scanning.

In [5] a hybrid approach for the detection of 1D barcodes is proposed, integrating deep learning techniques with geometric methods. The proposal combines a CNN-based detector with geometric refinement to accurately locate barcodes in challenging images with low resolution, noise, or distortions. The model was trained with a realistic dataset and demonstrated robust performance in terms of detection and speed, outperforming traditional methods based solely on computer vision. This intelligent combination of deep learning and geometric analysis stands out as an effective strategy for applications in industrial and commercial environments, where reliability and efficiency in barcode reading are crucial.

In [6] an efficient approach is proposed for real-time detection of 1D barcodes using spontaneously captured unframed images, such as in videos from handheld cameras. Unlike traditional methods that rely on manual framing and good lighting, their technique uses extremal regions combined with filtering and geometric clustering to identify barcode patterns. Compared with the state of the art at the time of the investigation, the solution showed significant improvements in both accuracy (+14%) and speed (-20%). This work is relevant for passive and autonomous barcode reading systems, especially in scenarios with minimal user interaction.

Finally, in [9] a comprehensive review of barcode localization techniques is described, and the authors introduce the BarBeR, a benchmark with more than 8000 annotated images and a standardized set of metrics. The study highlights current limitations, such as the scarcity of robust public datasets and the low reproducibility of many works, in addition to demonstrating the recent impact of deep learning-based methods, especially YOLO and Faster R-CNN.

Different from previous studies that focus on specific scenarios, such as drones or environments with fixed resolution, this work proposes an experimental evaluation of barcode reading in a simulated industrial environment, exploring eleven different resolutions, from VGA to DCI 4K. The BarcodeVision system adopts a hybrid approach, combining classical methods using a reading library with multithreaded rotation and detection via YOLO, allowing for a comparison of the individual and combined performance of these techniques. Thus, the proposal contributes to the literature by offering a more comprehensive analysis adaptable to industrial contexts with high visual variability.

III. Description of the BarcodeVision System

The BarcodeVision system was developed with the goal of performing real-time barcode detection and reading, addressing traceability, management, and logistics needs in the context of Industry 4.0. The current version of the system was designed to allow for a comparative evaluation of the performance of different reading strategies.

The BarcodeVision is composed of four distinct barcode reading methods, all operating on frames extracted from a video source, whether from a live camera feed or a recorded video file. Each method has its own specific characteristics:

- Method 1 (M1) – Direct reading with the ZBar library: each captured frame is processed directly by the ZBar library. When a barcode is identified, the method returns the decoded information;
- Method 2 (M2) – Multithreaded rotation with ZBar: this method distributes the processing of each frame across different threads, applying rotations at various angles up to 180°, in 15° increments, to increase detection robustness in cases of misalignment. The version implemented employs 4 threads;
- Method 3 (M3) – Detection with YOLO + ROI with rotated reading (M2): the customized YOLO model in version 4 detects ROIs in the frames, indicating possible barcode locations. Each ROI is then processed using the M2 logic. The padding value applied to the ROIs is configurable and is set by a parameter during execution;
- Method 4 (M4) – Detection with YOLO and rotated reading (M2): similar to M3, this method uses YOLO to identify the presence of barcodes in the frame. When detection occurs, the entire frame is processed according to the M2 approach.

Each method was developed with the aim of exploring different trade-offs between reading accuracy, processing time, and robustness under adverse conditions, such as variations in code orientation, image quality, and lighting. The metrics used for this evaluation are detailed in the next section.

IV. Experimental Methodology

For the experiments, a previously recorded video was used in a conveyor belt scenario. A camera was positioned above the belt, simulating an automated visual inspection industrial environment. The camera's perspective can be seen in Figure 1.

The original video was resized to 11 different resolutions, ranging from simpler formats to high definition resolutions. The resolutions considered in the experiment were VGA, WVGA, SVGA, XGA, SXGA, HD, UXGA, Full HD (FHD), Quad HD, Ultra HD, and DCI 4K. The main objective of this variation is to evaluate the impact of resolution on the performance of the four reading methods implemented in the system.

As part of the analysis, a total of 218 frames in the original video were identified as containing barcodes. This value was adopted as a reference to calculate the reading rate obtained by each method. It should be noted that the frames analyzed in

Fig. 1. Evaluation scenario of the BarcodeVision system.

each test are dynamically extracted during execution, that is, the methods process the frames in real time from the playback of the original video, considering the resolution and scenario evaluated in each experiment.

In addition, a complementary experiment was conducted for M3, in which the influence of different padding margins applied to ROI was evaluated. Three variations were tested: 15%, 25%, and 35%, with the aim of observing how different ROI values affect the reading rate and processing time.

The experiments were designed to analyze two main aspects:

- Reading Accuracy: the ability of the methods to correctly identify and decode the barcodes present in the frames;
- Processing time: the total time required to process all frames.

V. Results and Discussion

The results obtained from experiments conducted with different video resolutions are presented in this section. The information obtained from these results form a fundamental basis for analyzing the current performance of BarcodeVision and enable important insights into the strengths and limitations of each tested approach. In addition, the results will serve as a reference for guiding future improvements, with the aim of optimizing both the reading rate and processing time of the system, in order to apply the BarcodeVision in Industry 4.0 scenarios.

Figure 2 shows the average processing time when using Method 1 at different resolutions. The time remains low at lower resolutions but gradually increases from Full HD onward, due to the fact that ZBar processes the entire frame, making the required time a function of the number of pixels in the image. In Figure 3, it can be seen that Method 1 presents a reading rate below 45% at most resolutions. A significant drop in the reading rate is observed for resolutions above UXGA, despite the increase in processing time. This may be due to challenges in locating barcodes within larger image areas, where the abundance of pixels can act as noise and hinder pattern recognition, especially in the absence of ROI detection or optimized scaling. This reveals its limitation when facing variations in barcode orientation and the absence of pre-processing steps, making it less effective.

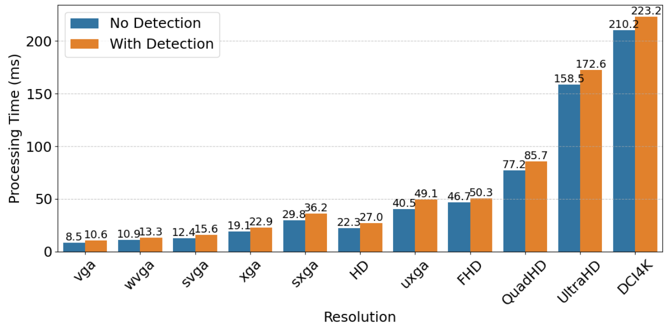

Fig. 2. Average processing time per resolution with Method 1.

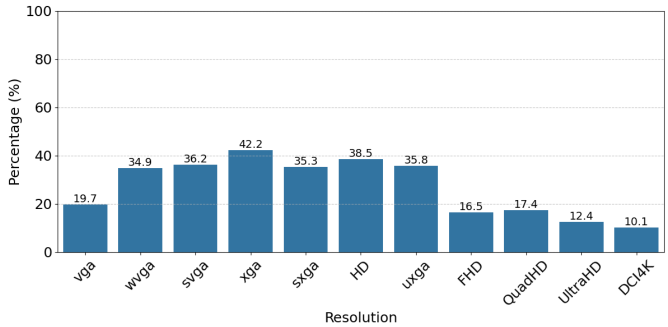

Fig. 3. Reading rate per resolution with with Method 1.

Figure 4 indicates that the average time when using Method 2 increases significantly compared to M1 due to the multiple parallel rotations. However, it can be observed that the increase in processing time follows a pattern similar to that of Method 1, with a more pronounced rise starting from FHD resolution. As shown in Figure 5, the Method 2 achieves up to 71% reading rate at XGA resolution, demonstrating that the multithreaded rotation strategy substantially improves robustness, compensating for the higher computational cost with relevant accuracy gains.

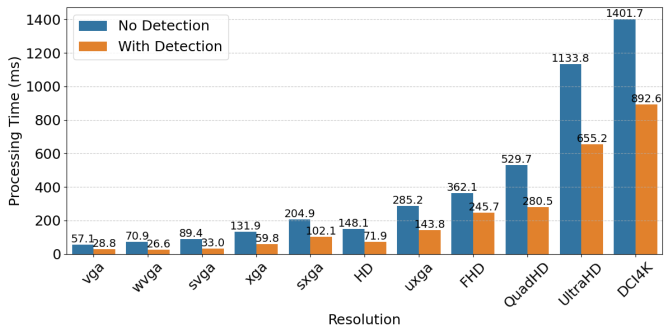

Fig. 4. Average processing time per resolution with Method 2.

The processing time graphs in Figures 6, 7, 8, and 10 present the following categories, used to more accurately assess the performance of each method: *(i)* false/false - no detection and no reading; *(ii)* true/true - detection by YOLO and successful reading; *(iii)* true/false - detection by YOLO but no valid reading.

Figures 6, 7, and 8 present the average processing time

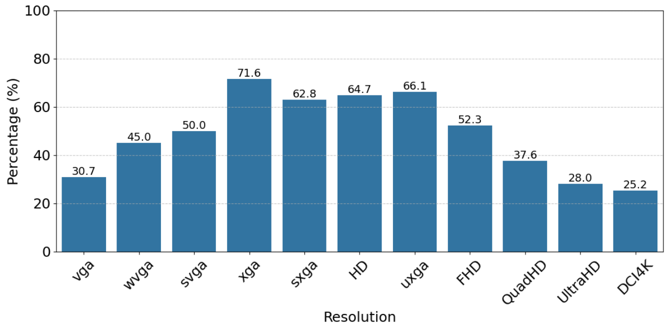

Fig. 5. Reading rate per resolution with Method 2.

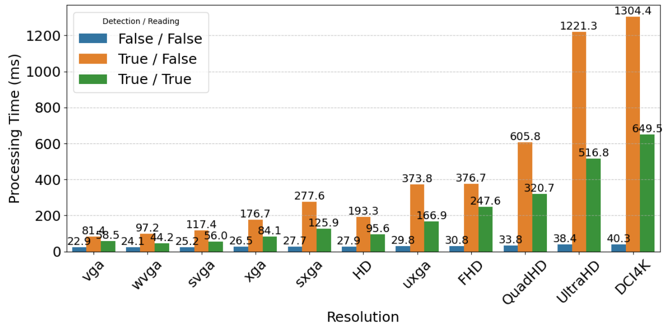

Fig. 8. Average processing time per resolution with Method 3 and 35% padding on ROI.

when using Method 3 with 15%, 25%, and 35% padding, respectively. Time increases as ROI expands, requiring more processing. However, M3 is faster than M2 at several resolutions, especially when no code is present in the frame, since the barcode reading algorithm is not executed when the YOLO model does not identify the presence of a barcode in the image.

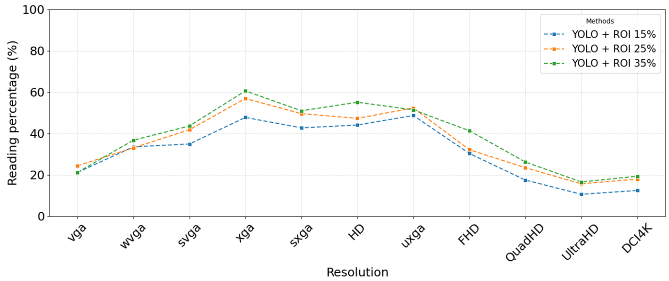

Fig. 9. Reading rate per resolution with Method 3, and 15%, 25%, 35% padding on ROI

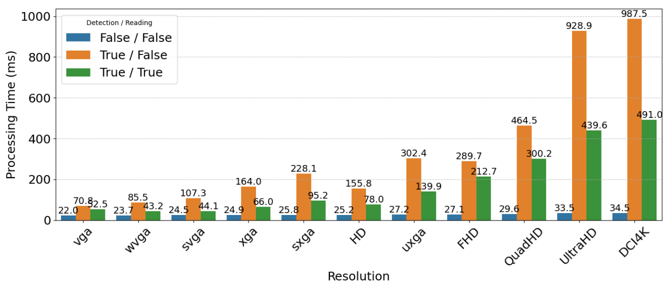

Fig. 6. Average processing time per resolution with Method 3 and 15% padding on ROI.

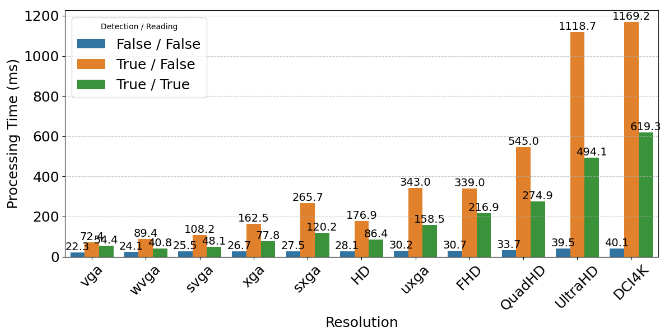

Fig. 7. Average processing time per resolution with Method 3 and 25% padding on ROI.

Figure 9 reveals that increasing the padding improves the performance of Method 3, indicating that expanding the ROI favors reading.

Figure 10 shows that the processing time when using Method 4 is higher when compared to Method 3, especially when no barcode is present, as the system executes the full rotation logic after detection but does not obtain successful readings.

According to Figure 11, Method 4 does not surpass Method 2 in reading rate, but approaches the performance of Method 3 with 35% padding, making it a viable alternative when seeking a balance between processing time and accuracy.

Figure 12 compares all methods in terms of reading rate. Method 2 is the most effective, followed by Method 3 with padding 35% and Method 4, which show similar performance. Method 1, although fast, is the least reliable. In frames without barcodes, Method 3 demonstrates a better response time by focusing only on ROIs.

The frame-by-frame analysis revealed limitations in the YOLO model. In some cases, as shown in Figure 13, YOLO was unable to detect codes that Method 2 read successfully. In another case, shown in Figure 14, there were false positives, with detection in frames without codes. These cases show that the detector is not entirely accurate, being sensitive to noise, resolutions, or conditions outside the training domain. Such failures reinforce the need for complementary strategies or

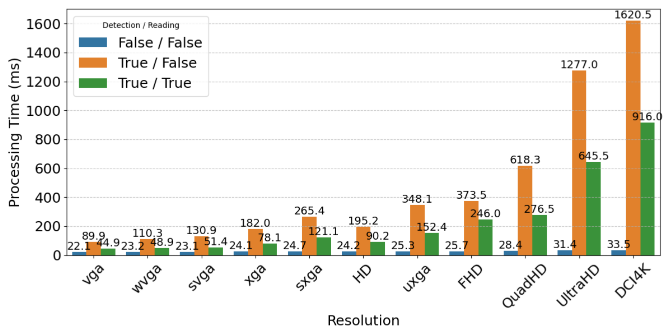

Fig. 10. Average processing time per resolution with Method 4.

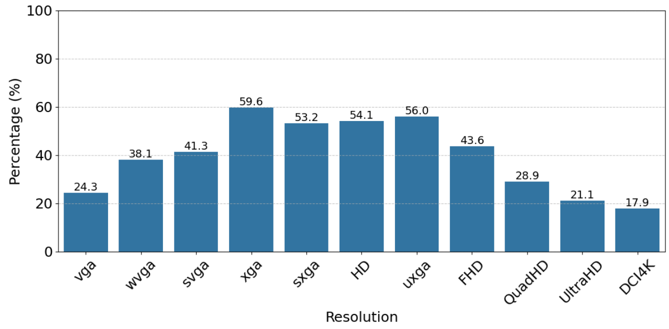

Fig. 11. Reading rate per resolution with Method 4.

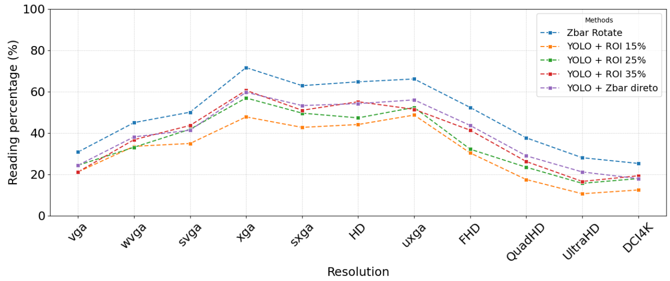

Fig. 12. Reading rate comparison among all methods.

model adjustments to ensure greater robustness in uncontrolled environments, such as industrial ones.

Fig. 13. Frame not detected by YOLO but successfully read by Method 2.

VI. Conclusion and Future Work

The experiments carried out with the first version of BarcodeVision enabled analysis of the performance of different real-time barcode reading methods at multiple video resolutions. The results indicate that Method 1 (direct use of ZBar) presents the lowest processing time, but with a low reading rate, especially at higher resolutions. Conversely, Method 2 (ZBar with multithreaded rotation) achieved the best reading performance, compensating for the higher computational cost.

Methods using YOLO detection (M3 and M4) demonstrated a greater balance between reading rate and processing time, particularly in the absence of barcodes in the frames. The use of extended ROIs proved to be effective in accelerating

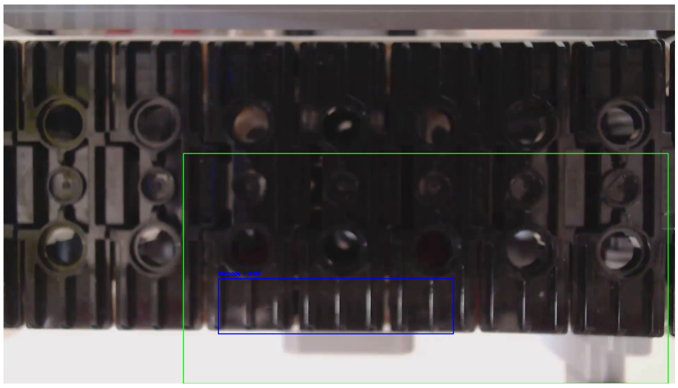

Fig. 14. Frame in which YOLO indicated barcode presence even though it was not present.

execution without significantly compromising the reading rate. The general comparison shows that approaches with localized detection and optimized reading areas are promising for dynamic industrial scenarios.

In future work, it is proposed to investigate optimization techniques aimed at processing high-resolution images, such as 4K, maximizing the cameras' field of view to enable reading multiple barcodes per frame. In this direction, the use of YOLO as a tool for prior ROI selection stands out as a relevant strategy, which requires further studies on its integration with new methods that can provide improvements in accuracy and performance. In addition, experiments will be carried out in real industrial environments.

Referências

[1] Hansen, D., Zhao, X., & Schindler, K. (2017). *Real-Time Barcode Detection and Classification using Deep Learning*. Proceedings of the International Conference on Computer Vision Theory and Applications (VISAPP), 2017. https://doi.org/10.5220/0006508203210327

[2] Kalinov, M., Burian, T., Smutný, P., & Zolotová, I. (2023). *WareVision: CNN Barcode Detection-Based UAV Trajectory Optimization for Autonomous Warehouse Stocktaking*. Sensors, 23(1), 274. https://doi.org/10.1109/LRA.2020.3010733

[3] Namane, A., & Arezki, M. (2017). *Fast Real-Time 1D Barcode Detection From Webcam Images Using the Bars Detection Method*. Proceedings of the World Congress on Engineering (WCE), Vol I, London, UK. ISBN: 978-988-14047-4-9.

[4] Xu, L., Kamat, V. R., & Menassa, C. C. (2018). *Automatic extraction of 1D barcodes from video scans for drone-assisted inventory management in warehousing applications*. International Journal of Logistics Research and Applications, 21(3), 243–258. https://doi.org/10.1080/13675567.2017.1393505

[5] Yunzhe Xiao and Zhong Ming, *1D Barcode Detection via Integrated Deep-Learning and Geometric Approach Applied Sciences*, vol. 9, no. 16, article 3268, 2019. https://doi.org/10.3390/app9163268

[6] C. Creusot and A. Munawar,*Real-Time Barcode Detection in the Wild*, in *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*, Waikoloa, HI, USA, Jan. 2015, pp. 239–245. https://doi.org/10.1109/WACV.2015.39.

[7] A. Zamberletti, I. Gallo, and S. Albertini, *Robust Angle Invariant 1D Barcode Detection*, in *Proceedings of the 2nd IAPR Asian Conference on Pattern Recognition (ACPR)*, Naha, Japan, Nov. 2013, pp. 160–164. https://doi.org/10.1109/ACPR.2013.17.

[8] Wachenfeld, Steffen and Terlunen, Sebastian and Jiang, Xiaoyi, *Robust recognition of 1-D barcodes using camera phones*, in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR)*, pages 1–4, 2008. https://doi.org/10.1109/ACPR.2013.17.

[9] Vezzali, E., Bolelli, F., Santi, S., & Grana, C. (2025). *State-of-the-art review and benchmarking of barcode localization methods*. Engineering Applications of Artificial Intelligence, 147, 110259. https://doi.org/10.1016/j.engappai.2025.110259