Single-Camera Mirror-Assisted Detection of Quality Marks for Automotive Windshield Control

Bruno Ferreira, João Ferreira, Mohammad Zarei, Paulo Menezes

University of Coimbra, Dept. of Electrical and Computer Engineering
Institute of Systems and Robotics
Coimbra, Portugal
{bruno.ferreira, joao.ferreira, mohammad.zarei, paulo}@isr.uc.pt

***Abstract*—In today's industry, efficient and precise vision systems that support robotic operations are essential to ensure continuous and smooth production, and the automotive industry is no exception. In this paper, we present a creative approach for detecting and inspecting quality features on windshields by coupling a single-camera with a programmable moving mirror system. Our method addresses the challenges posed by limited space and constraints for multi-camera setups that are often overlooked in production environments. The experimental apparatus of this system uses a camera with a 50mm focal length lens that captures different areas of the windshields at multiple positions via a precisely controlled mirror. We introduce a calibration technique that uses homography-based distortion correction to account for perspective changes caused by different mirror positions. In addition, the algorithm includes an OCR stage with an image filtering bank optimized for glass surfaces that enables accurate code recognition under different windshield types and curvatures. The experimental results show that the system is capable of reliably recognizing and verifying quality features. Compared to other scanning systems with multiple cameras or robots, our approach offers significant advantages in terms of cost efficiency, space savings and possible adaptation to different configurations of production lines.**

***Index Terms*—Automotive quality control, windshield inspection, computer vision, artificial intelligence, single-camera system**

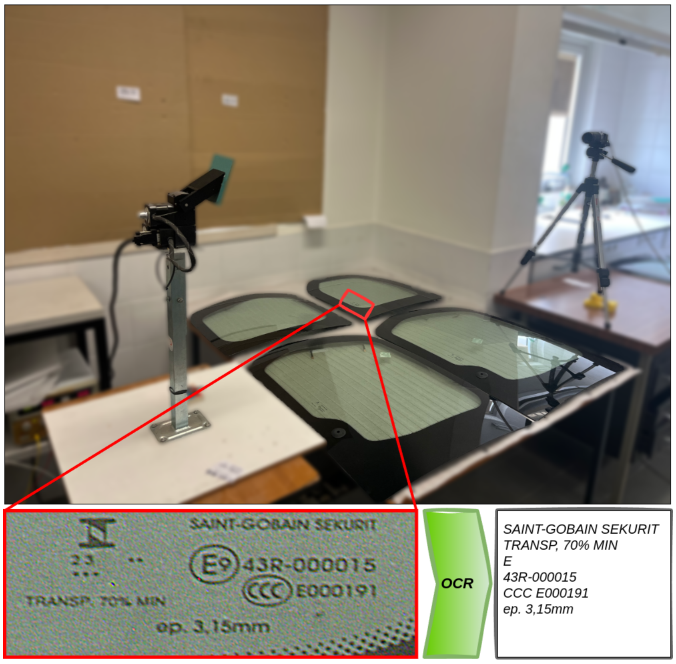

Fig. 1. Proof-of-concept system for the automatic recognition of engraved codes on windshields.

I. Introduction

The industry is constantly evolving and increasingly focused on automation and quality control, with efficiency and reliability at the forefront. In fact, the automotive industry is one of the main drivers of technological innovation, particularly in the fields of automation and robotics [1].

As we move into the age of Industry 5.0, there is an increasing focus on digital technologies that address the challenges of precision, flexibility and robustness in automotive production lines. More specifically, advanced vision systems that automate complex tasks such as positioning, identification, verification, measurement or defect detection throughout the production line. Fortunately, these systems are able to operate continuously while ensuring consistent and precise inspection, resulting in improved product quality and adherence to strict quality standards that are unfortunately different from a human reality [2], [3]. In the automotive industry, some of the existing processes are still performed by humans, which are a variant of the factory, but to some of them are transverse [4]. For example, the visual inspection of quality standards carried out by humans before the windshields are installed in the vehicles to ensure that it meets all the required parameters.

To facilitate this, advanced vision systems can be used to accurately read and interpret these engraved codes. Integrating such systems into the robotic workflow ensures that only windshields that meet all required criteria are selected for installation, increasing the efficiency and reliability of the production process. Therefore, we propose a creative approach for detecting and reading engraved quality marks on windshields using a single camera system with a moving mirror controlled

This work is funded and part of the PRR - Recovery and Resilience Plan and by the Next Generation EU European Funds, following NOTICE No. 02/C05-i01/2022, Component 5 - Capitalization and Business Innovation - Mobilizing Agendas for Business Innovation, under the project Greenauto - PPS10/PPS12/PPS13 with the reference 7255, C629367795-00464440. It is also partially funded by the project FCT/MCTES UIDB/00048/2020 DOI:10.54499/UIDB/00048/2020.

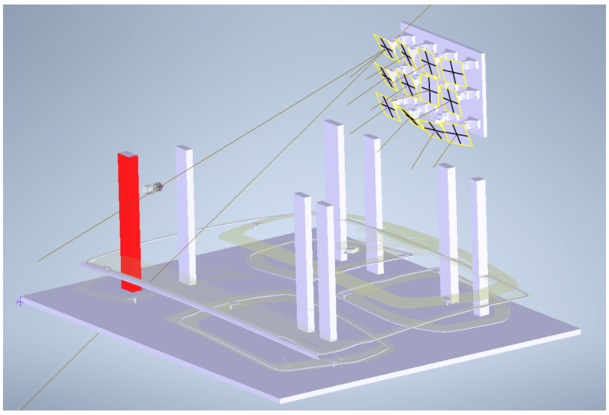

Fig. 2. Simulation of the ideal configuration for a single camera system that observes N target areas with adjustable mirrors.

by a Pan-Tilt Unit (PTU) as shown in Figure 1, which supports one of the main goals of the GreenAuto project. This enables the acquisition of multiple Region of Interests (ROIs) that include the relevant codes that are recognised and further processed by the Optical Character Recognition (OCR) algorithm. In addition, although this work is a proof of concept, it is primarily intended to serve as the basis for a system consisting of a single camera aimed at a grid of mirrors specifically tuned and oriented towards the desired target areas (Figure 2). This grid of mirrors fills the entire camera sensor and ultimately offers several compelling advantages over conventional multi-camera setups or robotic scanning methods.

First, by using a single camera and mirrors, this system significantly reduces hardware costs compared to multi-camera setups. Secondly, the introduction of new types of systems in existing production lines or inspection areas is usually associated with a major space problem. The compact design of a single-camera system with a grid of mirrors allows for easy integration into these areas. Thirdly, this PTU-based mirror positioning enables processing multiple positions in just one pass, which significantly reduces scanning time. Furthermore, this system is flexible as it can be customised by adjusting the calibration phase for scanning additional positions without the need for extensive reconfiguration.

The rest of the paper is organized as follows: Section II presents other work developed in an industrial context, Section III describes the proposed system, highlighting each module considered, and Section IV presents the defined test protocol supported by the quantitative results obtained. Section V concludes the paper and presents follow-up work.

II. Related Work

With the rise of automation and increasing production demands, machine vision systems have become essential for industrial inspection, helping detect defects and ensure product quality in recent years. Semeniuta et al [5] began working with machine vision systems in the automotive industry and proposed solutions to improve the systems, working with data storage and data analysis based on the training of machine learning models. Liu et al [6] have proposed an algorithm using Extreme Learning Machine (ELM) to create a quality database for glue application. The goal of obtaining an online inspection in real time was achieved, as the algorithm learns very quickly with a training time of less than 0.1s.

Würschinger et al [7] worked with machine vision systems to monitor the quality control of a vertical turning process of a piston rod, and use transfer learning to detect whether there are chips in it or not. Shan et al [8] proposed an algorithm for measuring involute gear tooth profile errors using images. It extracted pixel information in transition zones using Gaussian filtering and threshold segmentation to recognise the spur gear with small or medium modulus, reaching 1 tooth shape per second. In addition, Akundi et al [9] presented a machine vision system for dimensional inspection of different shaped samples, which determines various aspects (perimeter, area, rectangularity and circularity) and inspection of surface defects (scratches, dents and marks) and is used in a quality control system in the automotive industry.

More recently, Leberruyer et al [10] further explored the application of Artificial Intelligence (AI) in industrial settings to determine the essential requirements for achieving Zero Defect Manufacturing (ZDM). A semi-supervised learning method was used to identify the vibration characteristics that distinguish confirmed defects from approved products. Schoch et al [11] used deep neural networks to predict the results of quality assurance testing based on vehicle configuration in the automotive industry. Similarly, Cavaliere et al [12] worked on a real-world integration of an inspection system into an industrial production scene to perform surface defect inspections on aluminium components used in hybrid vehicles.

Yet, to the best of our knowledge, there is little to no work that could serve as a direct comparison for the approach presented. In fact, our proposal for detecting quality marks for automotive windshield inspection with a single camera is rather creative and a practical application of computer vision principles to artificial perception systems, with a touch of AI-based OCR methods, aligned with the ZDM policy.

III. Proposal

The proposed method for detecting quality marks on automotive windshields employs a single-camera system coupled with a mirror manipulated by a PTU. This uses a high-resolution RGB camera (5320×3032) with a 50mm lens that is 1.5m away from a square mirror with a size of 10 cm. The mirror is located 1.2m from the center of the 2.02m^2 inspection area and serves as a dynamic reflective surface to navigate over 4 predefined positions within the working area. The implementation of the system is divided into a calibration phase and a deployment phase, as shown in Figure 3, which ideally take place in this order in the industrial environment.

A. Preliminaries: Distance vs Readability

The challenge of reliably reading engraved codes on windshields illustrates an important trade-off between camera distance and OCR performance. A camera positioned close to

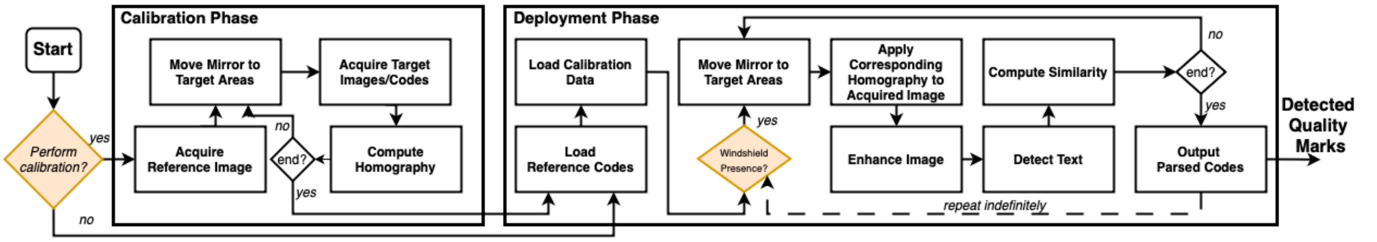

Fig. 3. System flowchart illustrating the Calibration and Deployment Phases for quality mark detection using a movable mirror setup. Yellow diamond shapes represent external input signals: calibration trigger and windshield presence check. The calibration phase computes homographies for multiple target areas, while the deployment phase applies them to enhance, detect, and compare with reference codes, repeating indefinitely.

Fig. 4. Distance vs readability of the engraved codes on the windshields.

the windshield provides high-resolution ROIs of the codes, allowing the OCR system to perform with near-perfect accuracy. However, creating an arrangement of multiple cameras or robotic scanning approaches to observe different target areas is often impractical in real factory environments due to space constraints or cost considerations. But, if the camera is positioned away from the working area, the resolution of the ROI decreases significantly, which has a negative impact on the text recognition of the engraved codes as seen in Figure 4.

This reduced resolution makes it difficult to accurately recognisethe fine details of the markings on the windshields, which depend heavily on the position in the working area. Thus, this trade-off between camera distance and the ability to reliably detect the quality marks is a critical design challenge. However, inspired by fundamental optical principles, such as telescopes or periscopes for indirect vision, the proposal of a mirror-based approach to overcome this can be described as creative. The mirror acts as an optical mediator that overcomes the physical limitations resulting from the proximity to the area of interest for our imaging system. Combined with the observation of different mirrors to reach any target area, the idea of a cost-effective, compact and flexible solution for modern production environments is supported.

B. Calibration Phase

During the calibration phase, as shown in Algorithm 1, the PTU systematically orients the mirror to acquire the N possible targets according to the imposed industrial constraints, while dynamically creating the baseline positions for the future deployment phase. These mirror positions depend on the possible positions of the quality marks in a given area, and therefore, it is necessary to calibrate the system according to the existing processes of the production line and their respective requirements.

First, the process of focusing a camera while adjusting the aperture takes place. By setting the camera's aperture to the widest setting, the maximum amount of light enters the camera, resulting in a shallower depth of field, and one can focus on the target object. The aperture is then closed at the smallest setting, creating a pinhole camera where the light enters at a central point and practically everything is focused on the image plane. In addition, the exposure time adjusts the visibility of the objects in the image. This step is important not only to capture the N possible positions, but also to obtain different calibration patterns seen by the focused camera.

These images are subjected to rigorous processing to compute the homography matrices that describe the geometric transformation between the camera view and the reflected region of interest on the windshield. It allows to link spatial perspectives between planar surfaces captured from different viewpoints, which is essential for the following pipeline tasks that require distortion correction. Consequently, this Algorithm 1 performs an iterative acquisition process for N mirror positions of the workspace. Between iterations, the system takes ≈ 2 seconds to move the mirror, detect features with Scale-Invariant Feature Transform (SIFT), match the features of the reference image with RANdom SAmple Consensus (RANSAC) and calculate the homography.

C. Deployment Phase

In the deployment phase, the system uses the calculated homography matrices to account for both the lens distortion and the different perspectives of the engraved observable codes on the mirror. The Algorithm 2 describes a real-time process that performs the desired inspections by using the previously defined PTU sequential positions that guide the mirror reflect each target area onto the camera sensor.

As soon as the system receives an input control from the external production pipeline, signalling the presence of a windshield, the mirror scans every possible position to detect the engraved codes while each acquired image is subjected

Algorithm 1 Calibration Phase for N Mirror Positions

```
procedure CalibrateSystem(N)
    P ← {}                              ▷ Initialize set of homography matrices
    F_r ← DetectFeatures(ReferenceImage)
    for i ← 1 to N do
        M_i ← SetMirrorPosition(i)
        I_i ← CaptureCalibrationImage()
        F_i ← DetectFeatures(I_i)
        M_i ← MatchFeatures(F_i, F_r)
        H_i ← EstimateHomography(M_i)
        P ← P ∪ {H_i}
    end for
    return P
end procedure
```

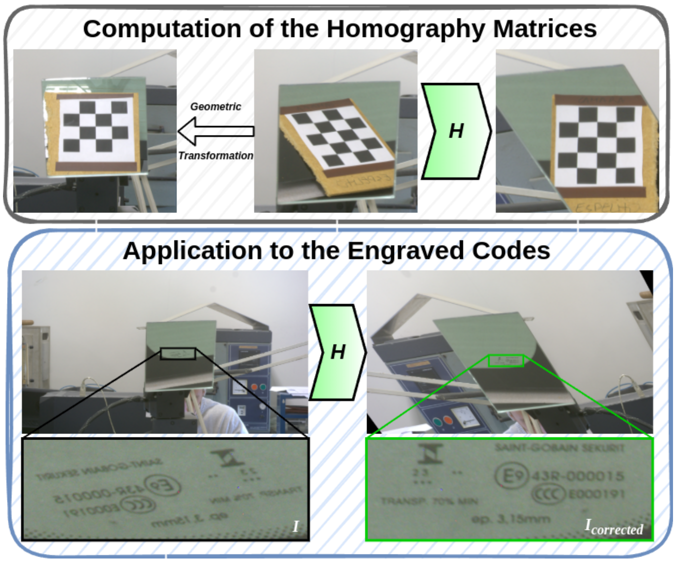

Fig. 5. Depiction of the homography-based process to rectify target text areas.

Algorithm 2 Deployment of Single-Camera Mirror-Assisted Quality Mark Detection System for Automotive Windshields

```
procedure DeployCodeDetectionSystem
    C ← InitializeCamera(w, h)
    M ← InitializeMirror()
    H ← LoadCalibrationData()           ▷ Load homographies
    R ← LoadReferenceCodes()            ▷ Load target codes
    while ProductionLineActive() do
        if InputSignalWindshieldPresence() then
            for i ← 1 to N do           ▷ N mirror positions
                SetMirrorPosition(M, i)
                I ← CaptureImage(C)
                I_corrected ← ApplyHomography(I, H[i])
                I_denoised ← EnhanceImage(I_corrected)
                codes ← DetectCodes(I_denoised)
                scores ← ComputeSimilarity(codes, R)
            end for
            OutputParsedResults(codes, scores)
        end if
    end while
end procedure
procedure DetectCodes(image)
    regions ← LocalizeCharacterRegions(image)
    characters ← DetectCharacters(regions)
    codes ← PostProcessing(characters, regions)
    return codes
end procedure
procedure ComputeSimilarity(code, references)
    S ← {}
    for each ref in references do
        S ← ∪ CompareCode(code, ref) > α
    end for
    return S
end procedure
```

to the corresponding precomputed homography. This effectively reduces distortions and paves the way towards the preprocessing stage for optimizing image quality.

Namely, these ROIs pass through a denoising filter bank that applies different filter(s) and possible combinations to influence the outcome of the OCR algorithm. The most optimal image corresponds to the best average confidence score. Furthermore, for each detected code, a similarity score is computed against each expected/known quality mark on the systems' priors. The Levenshtein distance can be used to measure the minimum of single-character edits required to change the detected code into the ground-truth. The next step of the system is providing the detected codes to the external flow and output a signal that denotes whether the windshield has passed or failed the inspection.

D. Understanding the AI-based OCR Algorithm

The application of advanced OCR algorithms allow to decipher potentially low-contrast or etched characters typically of windshields engravings. Specifically, we leverage the use of state-of-the-art modules, such as the EasyOCR [13] that easily automates text-related computer vision applications.

For example, its architecture is based on Deep Learning (DL) models, where the first step encompasses a preprocessing of the image and then the Character-Region Awareness For Text detection (CRAFT) [14] method is used to detect the existing text despite its different size, orientation or fonts, by creating two binary maps. One indicates the likelihood of a pixel belonging to a character region and it is used to localize individual characters in the image. The other indicates likelihood of two adjacent characters belonging to the same text instance, and it is used to group characters into words. Naturally, this creates a spatial map containing the text, and therefore, reduces the dimensionality of the original image.

Each crop containing text is fed through a Convolutional Recurrent Neural Network (CRNN) model presented in [15], which is composed of three components, including convolutional and recurrent layers, and decoding. A VGG [16] with the fully-connected layers removed, constitutes the convolutional block that extract a sequential feature representation from the input image, while preserving the spatial relationships among the pixels. These feature representations are fed as input to the recurrent layers that capture long-term dependencies in this sequential information, by selectively remembering and forgetting parts that are not important for the process.

The final layer outputs a sequence of probabilities for each character in the input text, which is then processed by the decoding step. In fact, the Connectionnist Temporal Classification (CTC) algorithm transforms it into a written text sequence. However, there is an additional post-processing step that applies certain techniques to refine this text sequence as it aims to correct any errors or inconsistencies (e.g., spell checking or language-specific rules).

Given the target application area for the automotive industry, the detected text encodes the engraved codes on the windshields, which are typically sequences of characters that do not follow a language model. Yet, these follow specific standards and rules, primarily governed by regulatory entities in different regions. The Economic Commission for Europe (ECE) standard is used, indicated by an ”E” followed by a number representing the certifying country (E9-Spain). The ”43R” marking indicates compliance with ECE safety guard number 43 for motor vehicle safety glass. Additionally, other codes typically include manufacturer’s name, part number, production details such as glass thickness and tint, etc.

IV. Experimental Results

In this section, we present the experimental results focusing on two key aspects: i) the accuracy of geometric calibration using homographies when simulated changes in final mirror position are introduced, and ii) the performance of the OCR pipeline in extracting text from the captured images.

A. Calibration Validation

To evaluate the effectiveness of the geometric correction based on homography, we performed experiments on the correctness of the distorted reflections of the observed patterns. In addition, previously calculated homographies that align the ideally positioned reflected checkerboard with the reference are used to correct new checkerboard positions.

Figure 6 shows a graphical representation of the deviations caused by an under- or over-rotation of the PTU compared to the ideal, which ultimately affect the final position of the mirror. In particular, we measured the reprojection error, which is defined as the Root Mean Square Error (RMSE) between the expected and observed position of the checkerboard corners after applying homography. As can be seen in Figure 7, the trend line for 3 of the 4 positions shows little vertical change, which is a positive effect of the rectification process by these homography matrices. These results suggest that these transformations are generally effective, with mean errors remaining below a tolerable threshold for geometric alignment and boosting further text recognition on the ROIs.

B. Considerations on OCR Performance

After geometric calibration, we evaluated the performance of the OCR system in reading engraved quality features. The captured images presented significant challenges, including low contrast between the engraved text and the background, varying depths of engraving, and uneven surfaces. To enhance text recognition accuracy, we applied various image processing techniques aimed at improving the visibility of the characters. These methods, demonstrated in Figure 8, include contrast enhancement, adaptive thresholding, and histogram equalization.

Table I presents the quantitative evaluation of the system by means of the Character Error Rate (CER) metric[1]. Initially, we tested the EasyOCR algorithm [13], which was expected to perform well, but the results were lower than anticipated across all four code types. As a result, we explored alternative methods, including Tesseract [17]. However, this also struggled to recognize the engraved text on the windshields, despite the text being clearly visible to the naked eye.

The final approach involved using a Large Vision Model

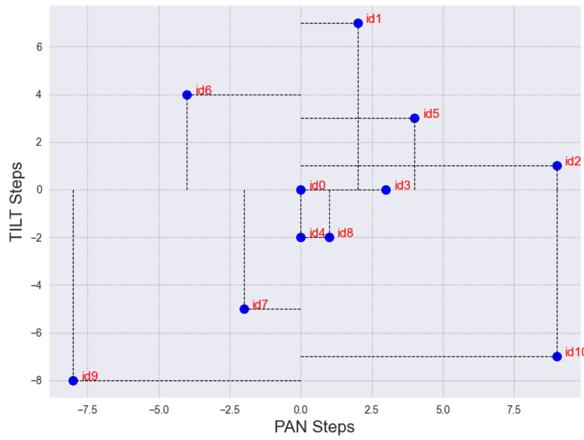

Fig. 6. Simulated non-ideal PTU positions by varying each axis.

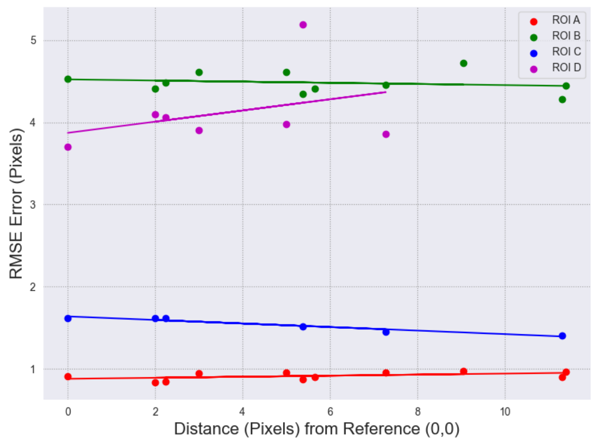

Fig. 7. Distance to the ideal position vs RMSE computed between the expected and observed checkerboard corner positions after homographies.

[1] CER is a relevant metric used in the evaluation of natural language processing systems that quantifies the similarity between reference and candidate texts. It is formulated as the sum of substitutions, insertions and deletions operations, divided by the number of characters existing in the reference.

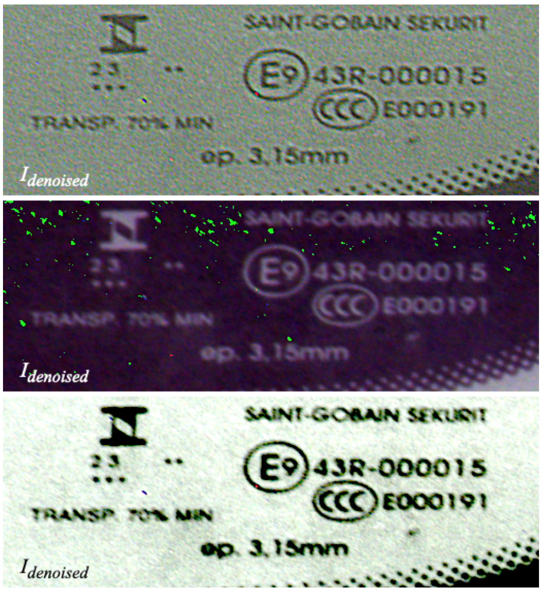

Fig. 8. Examples of the applied image operations to improve text readability.

TABLE I
OCR RESULTS FOR CODES ON 4 CONSIDERED POSITIONS.

	CER			
OCR Algorithm	**A**	**B**	**C**	**D**
EasyOCR [13]	18.18%	78.79%	72.73%	93.94%
Tesseract [17]	18.18%	63.64%	88.64%	87.88%
LLaVA-NeXT [18]	0.00%	3.03%	6.82%	54.55%

(LVM)-based solution, specifically the LLaVA-NeXT [18], configured for 4-bit precision to suit consumer GPUs. It was used as-is, without the need for further fine-tuning or to be prompted with specific instructions. Yet, this model demonstrates significantly better performance when compared to the widely used EasyOCR. Its advantages stem from its design natureas these models can understand and process images more comprehensively, recognizing not just text but the broader visual context. In addition, they are often trained on vast datasets that help it generalize across different fonts, writing styles, and visual imperfections.

V. CONCLUSIONS AND FUTURE WORK

In this paper, we propose a creative approach for quality marks detection in automotive windshield inspection using a system with a single-camera and a mirror mounted on a PTU. Our approach covers different constraints, namely the distance camera to mirror and mirror to the engraved codes, but also low-contrast image areas. On the other hand, the homography matrices computed in the calibration phase mitigate the distortion caused by the positioning of the mirror and support the following OCR-based phase. Different detectors are used to recognize the text on the engraved codes, showing different performances, with the LVM-based solution being the best that can be considered. However, the camera-mirror pair can also be positioned in a different location in order to reduce the reflection angles on the mirror and improve the overall results achieved. For example, placing the setup vertically, with the camera facing upwards towards the mirrors will not only save space, but also minimize the degrees of the reflection angles, which results in a more parallel view of the working area. This detail would potentially boost the OCR performance over the ROI D, which exhibits the biggest reflection angle and lowest contrast. Future work encompasses additional testing in more optimal configurations and deployment on the production line.

REFERENCES

[1] M. Cinal, A. Sioma, and B. Lenty, "The quality control system of planks using machine vision," *Applied Sciences*, vol. 13, no. 16, p. 9187, 2023.

[2] S. Wang and R. J. Jiao, "Smart in-process inspection in human–cyber–physical manufacturing systems: A research proposal on human–automation symbiosis and its prospects," *Machines*, vol. 12, no. 12, p. 873, 2024.

[3] R. Nascimento, I. Martins, T. A. Dutra, and L. Moreira, "Computer vision based quality control for additive manufacturing parts," *The International Journal of Advanced Manufacturing Technology*, vol. 124, no. 10, pp. 3241–3256, 2023.

[4] F. K. Konstantinidis, S. G. Mouroutsos, and A. Gasteratos, "The role of machine vision in industry 4.0: an automotive manufacturing perspective," in *2021 IEEE international conference on imaging systems and techniques (IST)*. IEEE, 2021, pp. 1–6.

[5] O. Semeniuta, S. Dransfeld, K. Martinsen, and P. Falkman, "Towards increased intelligence and automatic improvement in industrial vision systems," *Procedia cirp*, vol. 67, pp. 256–261, 2018.

[6] K. Liu, Y. Qiao, H. Yang, Y. Li, and W. Yu, "Research on construction method of glue spreading quality database based on extreme learning machine," in *2019 Chinese Control And Decision Conference (CCDC)*. IEEE, 2019, pp. 5675–5680.

[7] H. Würschinger, M. Mühlbauer, M. Winter, M. Engelbrecht, and N. Hanenkamp, "Implementation and potentials of a machine vision system in a series production using deep learning and low-cost hardware," *Procedia CIRP*, vol. 90, pp. 611–616, 2020.

[8] Z. Shan, M. Xin, and W. Di, "Measuring method of involute profile error based on machine vision," in *2021 IEEE International Conference on Real-time Computing and Robotics (RCAR)*. IEEE, 2021, pp. 1272–1276.

[9] A. Akundi, M. Reyna, S. Luna, and E. Chumacero, "Automated quality control system for product dimensional and surface analysis–an industry case study," in *2022 17th Annual System of Systems Engineering Conference (SOSE)*. IEEE, 2022, pp. 60–65.

[10] N. Leberruyer, J. Bruch, M. Ahlskog, and S. Afshar, "Toward zero defect manufacturing with the support of artificial intelligence—insights from an industrial application," *Computers in Industry*, vol. 147, p. 103877, 2023.

[11] A. Schoch, R. Refflinghaus, and P. Zivkovic, "Deep learning based predictive testing strategy in the automotive industry," *Procedia CIRP*, vol. 118, pp. 1108–1113, 2023.

[12] G. Cavaliere, Y. Borgianni, and E. Savio, "Performances of an in-line deep learning-based inspection system for surface defects of die-cast components for hybrid vehicles," *Procedia CIRP*, vol. 126, pp. 999–1004, 2024.

[13] J. AI, "Easyocr: Ready-to-use ocr with 80+ supported languages," https://github.com/JaidedAI/EasyOCR, 2024, accessed: March 10, 2025.

[14] Y. Baek, B. Lee, D. Han, S. Yun, and H. Lee, "Character region awareness for text detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 9365–9374.

[15] B. Shi, X. Bai, and C. Yao, "An end-to-end trainable neural network for image-based sequence recognition and its application to scene text recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 11, pp. 2298–2304, 2016.

[16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[17] R. Smith *et al.*, "Tesseract ocr engine," *Lecture. Google Code. Google Inc*, vol. 40, 2007.

[18] F. Li, R. Zhang, H. Zhang, Y. Zhang, B. Li, W. Li, Z. Ma, and C. Li, "Llava-next-interleave: Tackling multi-image, video, and 3d in large multimodal models," *arXiv preprint arXiv:2407.07895*, 2024.