# Integration of Computer Vision and CNC Machining for Automated Object Measurement and Customized Packaging

Pablo Guacho, Angélica Quito*, Andrea Pilco, Samuel Peña*, and Marcelo Moya

School of Mechatronics, Faculty of Digital Engineering and Emerging Technologies,
International University of Ecuador UIDE, Quito, Ecuador
Email: davidguacho16@hotmail.com (P.G.); anquitoca@uide.edu.ec (A.Q.); anpilcoat@uide.edu.ec (A.P.); spena@uide.edu.ec (S.P.); mmoya@uide.edu.ec (M.M.)
*Corresponding author

***Abstract*—This paper presents an integrated computer vision—Computer Numerical Control (CNC) workflow that measures an object's planar footprint and automatically generates G-code to cut customized cardboard packaging, thereby reducing manual sizing and setup in low-cost manufacturing environments. Specifically, a monocular camera within a controlled enclosure captures a single image, after which contour extraction, centroid estimation, and pixel-to-millimeter calibration are applied to convert pixel distances into metric length and height estimates; these measurements are then translated into CNC laser toolpaths and validated in Laser G-code Real-Time Blink Loop (LaserGRBL) prior to cutting. The approach was evaluated on 59 heterogeneous objects against calliper/ruler measurements, yielding sub-millimetric performance (Mean Absolute Error (MAE): 0.793 mm in length and 0.861 mm in height; RMSE: 1.057 mm and 1.394 mm, respectively). Moreover, most items exhibited errors below 2%, confirming suitability for practical packaging tolerances; nevertheless, occasional outliers (e.g., a 7.49% height error for a small snap-fastener case) indicate that further gains may be achieved through improved illumination control and enhanced contour robustness. Overall, these findings evidence a scalable route to on-demand packaging automation that effectively bridges perception and fabrication for small enterprises.**

***Keywords*—computer vision, Computer Numerical Control (CNC) machining, object measurement, G-code automation, packaging system**

## I. INTRODUCTION

Computer vision encompasses the acquisition, processing, analysis, and interpretation of visual data to extract meaningful information from physical environments. This discipline has become a fundamental enabler for the measurement of three-Dimensional (3D) objects through image capture under controlled or natural illumination conditions [1]. Following image acquisition, camera calibration establishes a precise mapping between pixel coordinates and real-world dimensions, allowing measurements to be expressed in metric units with high reliability.

Among active vision techniques, Fringe Projection Profilometry (FPP) remains one of the most widely adopted methods for high-precision 3D measurement. Several variants of FPP have been developed, including Fourier Transform Profilometry (FTP), Phase-Shifting Profilometry (PSP), and Computer-Generated Moiré Profilometry (CGMP), each addressing challenges related to noise robustness, acquisition speed, and surface reflectivity [2, 3]. Within these approaches, phase unwrapping plays a critical role in reconstructing continuous phase distributions from wrapped phase data, enabling accurate depth estimation [4]. The resulting depth maps facilitate the reconstruction of geometric representations such as point clouds or meshes, from which dimensional features including length, height, and volume can be extracted.

Accurate geometric measurement is a key requirement for the integration of perception systems with physical manufacturing processes, particularly within Computer Numerical Control (CNC) environments. In this context, computer vision extends beyond passive observation, functioning as a sensing layer that supports object detection, dimensional estimation, and process control [5]. The convergence between perception and fabrication has enabled the development of increasingly autonomous manufacturing systems capable of reducing human intervention while improving repeatability and precision.

Several studies have demonstrated the feasibility of integrating vision-based measurement systems with CNC machining workflows. Mileski *et al.* [6] proposed a vision-based 3D scanning system for workpiece positioning using photogrammetric techniques, while Bustos *et al.* [7] developed a methodology for automatic G-code generation from fringe projection-based geometric reconstruction. Similarly, Nguyen *et al.* [8] highlighted the integration of Computer-Aided Process Planning (CAPP)

Manuscript received November 5, 2025; revised December 24, 2025; accepted March 9, 2026; published September 11, 2026.

doi: 10.18178/joig.14.5.787-795

with automated G-code generation, emphasizing the role of visual inputs in machining decision-making. These contributions collectively confirm that vision-based measurement can serve as a direct input for CNC programming and execution.

In parallel, the role of computer vision in CNC setup and calibration has expanded significantly. Hou *et al*. [4] introduced an automatic tool-setting system based on edge detection with submicrometric precision, whereas Zhou *et al*. [9] and He *et al*. [5] provided comprehensive analyses of segmentation, recognition, and positioning algorithms in manufacturing environments. Furthermore, vision-based automation has been successfully applied to cutting operations, including glass positioning in water-jet systems, reducing the need for manual alignment and handling [10].

Beyond geometric measurement, recent research has incorporated artificial intelligence and hybrid modelling approaches to enhance machining performance. Pan *et al.* [11] reviewed intelligent architectures for ultra-precision manufacturing systems, while Fu *et al*. [12] proposed data-driven frameworks for predicting surface quality in grinding processes. Hybrid approaches combining physics-based models with machine learning techniques have demonstrated improved prediction accuracy and robustness [13]. Additionally, Alseedi *et al.* [14] explored the interpretation and translation of G-code for open CNC controllers, supporting the development of flexible and interoperable machining systems.

Despite these advances, the integration of computer vision and CNC machining in low-cost and scalable environments remains a challenge. In particular, achieving reliable dimensional measurement without relying on complex 3D reconstruction or specialized hardware is still an open research problem. This work addresses this gap by proposing a planar measurement framework based on monocular vision, enabling accurate dimension estimation and direct G-code generation for customized packaging applications.

## II. Literature Review

The evolution of computer vision for object measurement and recognition has progressed from classical image processing techniques toward advanced learning-based methodologies. Early approaches relied on fundamental operations such as edge detection, contour extraction, and size estimation [15], typically implemented using OpenCV-based frameworks [16, 17]. These methods emphasized computational efficiency through preprocessing strategies, including grayscale conversion and noise reduction [18], which improved segmentation performance and reduced processing time.

Subsequent developments introduced machine learning and deep learning techniques for object detection and classification. Neural network-based approaches enabled more robust feature extraction and segmentation, particularly in complex or unstructured environments [19–21]. The integration of these methods with geometric measurement techniques facilitated more reliable estimation of object dimensions, effectively bridging the gap between classical image processing and data-driven models [20].

In industrial contexts, computer vision has been extensively applied to dimensional inspection, workpiece positioning, and process monitoring. Calibrated vision systems have demonstrated the ability to reduce setup time while improving measurement accuracy and repeatability [4, 6]. These systems typically rely on controlled illumination conditions and precise camera calibration to ensure consistent performance across varying operating conditions.

Beyond industrial manufacturing, computer vision techniques have been applied in diverse domains, highlighting their adaptability. In healthcare, vision-based approaches have been used to design support systems for fall prevention based on balance recovery characteristics [22]. In agriculture, deep learning-based vision systems deployed on mobile platforms have enabled the estimation of biological properties and growth parameters [23]. Similarly, robotic systems have incorporated vision-based object detection and classification for manipulation and automation tasks [7].

More recent research has focused on improving robustness and accuracy through enhanced preprocessing, calibration, and hybrid modelling strategies. Advanced segmentation techniques and calibration methods have been proposed to address limitations in traditional recognition approaches, particularly in real-world scenarios with varying lighting and environmental conditions [18]. Additionally, the integration of vision-based measurement with process-aware modelling has been shown to improve machining quality by considering both geometric and physical process variables [8].

Vision-based automation has also been successfully implemented in cutting processes, such as water-jet machining systems, where accurate object positioning is critical for process efficiency [24]. These applications demonstrate the potential of computer vision to enable fully automated workflows that integrate perception, decision-making, and actuation within a unified framework.

Overall, the literature reveals a clear transition from isolated measurement techniques toward integrated perception–manufacturing systems. However, there remains a need for simplified and cost-effective solutions capable of delivering reliable measurements without the complexity of full 3D reconstruction. Addressing this challenge is essential for extending the adoption of vision-based automation to small-scale and resource-constrained manufacturing environments.

### *Contributions*

The main contributions of this work are summarized as follows:

(1) An integrated computer vision-CNC framework for automated object measurement and customized packaging is proposed, enabling a closed-loop workflow from visual perception to physical fabrication without manual intervention. The

proposed method relies on a lightweight monocular vision setup suitable for low-cost and scalable manufacturing environments.

(2) A monocular vision-based measurement approach using contour extraction, centroid computation, and pixel-to-millimeter calibration is proposed, achieving accurate non-contact dimensional measurements without requiring depth sensors or structured-light techniques.

(3) An automatic G-code generation method is developed to directly translate measured object dimensions into CNC laser cutting instructions for customized cardboard packaging.

(4) A statistically validated planar measurement framework is presented, including quantitative evaluation using mean absolute error, root mean square error, and dispersion analysis across 59 heterogeneous objects.

## III. System Overview

The system consists of a CNC machine designed for cutting cardboard, an additional module referred to as the enclosure box, and a computer, as shown in Fig. 1.

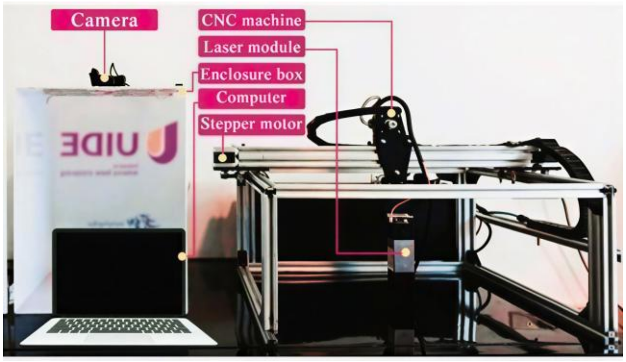



Fig. 1. CNC machine used for cutting cardboard objects.

The CNC machine integrates electronic and mechanical components to achieve precise cutting within a working area of 800 × 600 × 400 mm. Structural elements are fabricated from aluminum and carbon steel to enhance durability and mechanical stability. The device features two axes of motion (X and Z), each actuated by a NEMA 17 stepper motor for accurate positioning. Control is provided by an MKS DLC32 board, which is specifically designed for desktop laser and CNC applications. This board includes a 32-bit ESP32 module and is compatible with both MKS DLC firmware and GRB Laser software. The system incorporates a laser module with a nominal machine power of 80 W and a true optical output of 10 W. The laser employs a dual-compressed focus point, allowing for high-precision engraving at a resolution of 0.003 × 0.003 inches.

The enclosure is constructed from 3 mm-thick white acrylic and houses a camera mounted at 40 cm. System operation is supported by a computer equipped with a 2.2 GHz Apple i7 processor, 16 GB of DDR4 RAM, and an NVIDIA GeForce GTX 1050 Ti graphics card with 4 GB of dedicated memory, which ensures efficient data processing and reliable performance.

## IV. Methodology

The object measurement process begins when the item is placed inside the enclosure box, and the camera captures an image as shown in Fig. 2.

Fig. 2. Item placement inside the enclosure box.

The enclosure box is designed to accommodate objects with maximum dimensions of 400 × 400 mm. Image processing techniques, including preprocessing, contour detection, centroid calculation, and calibration, are employed to measure the object. Following these procedures, the object's dimensions are determined and converted into G-code. The generated G-code is then transmitted to the GRB Laser software to initiate the cardboard-cutting process, which enables packaging of the detected object. The overall workflow is illustrated in Fig. 3.

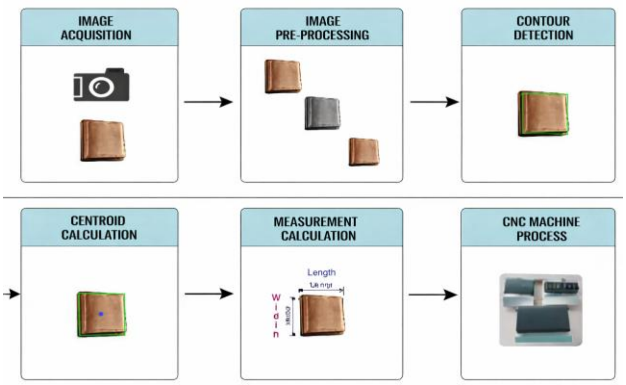



Fig. 3. Diagram of the CNC machine process.

To facilitate the precise extraction of object dimensions prior to CNC cutting, this study implements a computer vision-based measurement strategy. In contrast to methods that utilize three-dimensional reconstruction or depth sensing, the proposed system emphasizes planar geometry acquisition within controlled environments. The subsequent subsection outlines the computer vision methodology employed, including its theoretical foundation, operational assumptions, and relevance for producing dimensional data suitable for automated CNC machining.

## A. Computer Vision Method

The computer vision method employed in this research is a monocular two-Dimensional (2D) measurement approach based on classical image processing and geometric calibration. The system does not reconstruct three-dimensional geometry. Instead, it estimates planar object dimensions (X–Y) by converting pixel distances into real-world units through pixel-to-millimeter calibration.

The implementation of a static imaging setup within a controlled planar environment eliminates the need for volumetric analysis or RGB-D hardware. By mitigating geometric distortion, the system achieves reliable metric calibration using standard 2D optics. This design aligns with the requirements of economical CNC packaging solutions, as planar footprint extraction effectively satisfies the input parameters for the cutting operations.

The algorithm, developed using Python and the OpenCV library, performs the measurement directly without the need for a pre-existing dataset. The measurement is achieved through the following steps.

## B. Image Acquisition

The image of the object is captured when it is placed inside the enclosure box, as shown in Fig. 2. The camera is positioned 400 mm above the object and is connected to the computer via USB. A single image is obtained. The camera specifications are presented in Table I.

TABLE I. SPECIFICATION OF THE MACHINE VISION SYSTEM

| Description Camera | |
|---|---|
| Resolution | 1920 × 1080 color image |
| Optical zoom | 2 MP |
| Frame rate | Up to 30 fps |

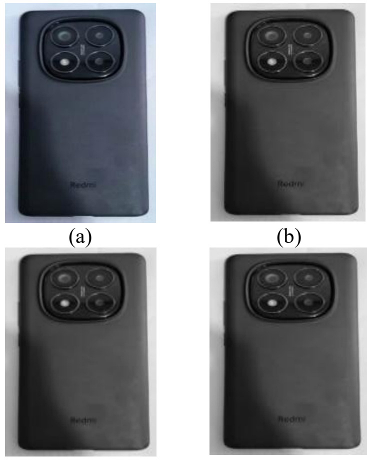
## C. Image Preprocessing

The input image had a resolution of 1920 × 1080 pixels and was resized to 800 × 800 pixels to reduce computational load. The image was then converted to grayscale, followed by the application of a Gaussian Blur smoothing filter. This filter was selected for its effectiveness in enhancing image segmentation [20], especially in contexts that require edge-sharpness analysis. The smoothing process utilized the cv2.GaussianBlur function, which implements the Gaussian function as defined in Eq. (1) [25].

$$g(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (1)$$

where $\sigma$ is the standard deviation of the Gaussian distribution, and $(x, y)$ are the Cartesian coordinates of the image. This function generates a kernel of weights based on the spatial distance from the center and the specified $\sigma$.

In this study, a kernel size of 7 × 7 was selected as it offers a balanced compromise between noise reduction and detail preservation. The choice of an odd-sized kernel ensures that a central pixel is well-defined, and the 7 × 7 dimension is sufficiently large to capture the Gaussian distribution's characteristics. The entire process carried out during the image preprocessing phase is shown in Fig. 4.



(a) (b) (c) (d)

Fig. 4. Image preprocessing of item placement inside the enclosure box: (a) Resized image; (b) Colour to gray; (c) Smoothing filter; (d) Gaussian Filter.

To reduce computational load, the acquired images were downsampled to a fixed resolution prior to further processing. While a systematic parametric study of downsampling ratios was not performed, the chosen resolution was determined empirically to maintain contour definition and enhance processing efficiency. Since pixel-to-millimeter calibration occurs after downsampling, any scale reduction is accounted for during calibration. Experimental results, with average measurement errors below 2%, demonstrate that the implemented downsampling does not substantially impact final measurement accuracy.

## D. Contour Detection

After an initial smoothing step, image noise was substantially reduced, enabling more effective edge analysis using cv2.Canny function. The cv2.The Canny function identifies intensity transitions corresponding to object boundaries by applying two threshold values: a lower threshold of 50 and an upper threshold of 100. Pixels with gradient magnitudes exceeding the upper threshold are classified as strong edges, while those with values between the thresholds are considered weak edges and are retained only if they are connected to strong edges [15].

After edge detection, contour extraction was conducted using cv2.findContours, which returns only the external contours of the image. The detected contours were subsequently sorted and filtered using contours.sort, which arranges them from left to right and removes small, irrelevant contours. This refinement ensures that subsequent analysis focuses exclusively on significant object boundaries, as illustrated in Fig. 5.


Fig. 5. Contour detection of item placement inside the enclosure box.

### E. Centroid Calculation

In the task of determining the dimensions of an item, a critical step is identifying its center point, or centroid [17]. To achieve this, after detecting the item's edges, each object was enclosed using a rotated bounding box, calculated with the cv2.minAreaRect function. This method estimates the smallest-area rectangle that can encompass the contour, while also allowing for rotation.

Once the bounding rectangle was determined, the function cv2.boxPoints was used to retrieve the four vertices of the rotated box. As these points may be returned in an arbitrary order, the perspective.order_points function was employed to organize them in the order: top-left, top-right, bottom-right, and bottom-left.


Fig. 6. External contour with marked centroid.

Additionally, the centroid of the object, representing its geometric center, was computed to support object localization as shown in Fig. 6. The centroid coordinates $C_x$, $C_y$ was derived from image moments using the Eqs. (2) and (3).

$$C_x = \frac{M_{10}}{M_{00}} \tag{2}$$

$$C_y = \frac{M_{01}}{M_{00}} \tag{3}$$

where $M_{00}$, $M_{10}$, and $M_{01}$ are spatial moments of the detected contour.

The bounding box is determined by employing a minimum-area rotated rectangle, which permits the bounding region to rotate and thereby minimize its enclosed area. This method removes the requirement for object alignment with the principal image axes and facilitates precise dimension estimation irrespective of object orientation.

### F. Measurement Calculation

Before measuring any unknown object, a one-time pixel-to-millimeter calibration is performed using a planar reference with known physical dimensions placed inside the enclosure. This calibration is conducted under fixed imaging conditions, including a constant camera position, fixed working distance, and a planar measurement surface. The resulting calibration factor remains constant and is reused for all subsequent measurements of unknown objects, provided that the imaging geometry does not change. After detecting the object and calculating its centroid using image moments, the next step was to measure its real-world dimensions. The first part of the process was the Pixel-to-Millimeter Calibration, which involves determining how many pixels correspond to one unit of measurement (millimeters). This was done by calculating the Euclidean distance between two reference points with a known real-world distance [23] as shown in Eq. (4).

$$d = \sqrt{(x_2 - x_1)^2 + (y_1 - y_2)^2} \tag{4}$$

The number of pixels per unit, called pixelsPerMetric, was computed using the Eq. (5).

$$pixelsPerMetric = \frac{d_{pixels\_know}}{L_{know}} \tag{5}$$

where $d_{pixels\_know}$ is the Euclidean distance in pixels between two reference points of a calibration object with known physical dimensions, and $L_{know}$ is the corresponding real-world distance between those points.

Once the calibration factor is known, the real-world dimensions of the object can be calculated. The *width* of the object was obtained by dividing the Euclidean distance between the top-left and top-right points by pixelsPerMetric, following Eq. (6).

$$width = \frac{d_{pixels\_width}}{pixelsPerMetric} \tag{6}$$

The *height* of the item was calculated by measuring the Euclidean distance between the top-left and bottom-left points, as shown in Eq. (7).

$$height = \frac{d_{pixels\_height}}{pixels\ Per\ Metric} \tag{7}$$

The sequence from centroid extraction to pixel calibration and final dimension estimation ensures accurate measurement of objects within images, effectively transforming pixel data into meaningful physical metrics, as can be seen in Fig. 7. Image annotation was employed to visually display the measurements obtained. This was achieved using the cv2.putText function, which allows text to be overlaid directly onto the image. Specifically, the measured width and height of the object.


Fig. 7. Object contour with length and width in millimeters.

### G. G-Code Generation and Validation

Once the object dimensions were obtained through the computer vision module, the corresponding G-code instructions were automatically generated based on the measured planar dimensions. Before executing the machining process on the CNC hardware, the generated G-code was validated using the LaserGRBL software.

LaserGRBL was employed as a verification and visualization environment that allows pre-execution inspection of the cutting paths. The software provides a graphical preview of the toolpath, enabling verification of geometry, motion sequence, cutting limits, and potential trajectory errors.

The generated programs incorporate standard CNC operations, including homing and initialization commands ($H, $X, G90), laser power control instructions (S1000 for activation and S0 for deactivation), and linear interpolation commands (G1) to define the cutting trajectory. For example, a rectangular cutting path is produced, with X and Y coordinates automatically determined based on dimensions provided by the computer vision algorithm.

Programmatic pauses were implemented during execution to ensure adequate thermal dissipation, mechanical synchronization, and precise laser-material interaction. CNC system performance was validated by comparing theoretical dimensions from the vision system with physical measurements of the cut boxes, assessing dimensional repeatability across multiple executions, and verifying correct homing, trajectory execution, and laser activation timing. These validation procedures demonstrate that the CNC system consistently executes G-code generated from computer vision measurements, thereby enabling accurate and customized box cutting.

### H. CNC Machining Process

After successful validation, the G-code was transmitted to an ESP32 control board compatible with MKS DLC firmware and LaserGRBL software [26]. The ESP32 interprets the validated instructions and controls the motion of the CNC system accordingly. Using this information, the CNC machine performs precise cardboard cutting operations, ensuring that the resulting cut accurately matches the object dimensions obtained from the vision-based measurement process.

## V. RESULT

During the testing phase, a comprehensive evaluation was conducted on objects with both circular and square geometries. This evaluation involved precise measurements using a dual methodology that combined advanced computer vision techniques with traditional physical measuring instruments, such as rulers and callipers. Computer vision facilitated the acquisition of theoretical dimensions, which were subsequently validated against direct physical measurements. The integration of these two methods enabled an accurate assessment of the objects' geometric properties and ensured that the findings were based on empirical data. This comprehensive assessment strengthened the study's integrity and reinforced a commitment to precision. The systematic evaluation process established a solid foundation for future project phases. Thorough validation of the measurements is expected to enhance the credibility of the results and contribute to the project's overall success.

Table II presents measurement accuracy tests. In these tests, algorithms estimated the length and height of each item in millimetres and compared these values with actual measurements obtained using precise manual tools. The purpose of these tests was to evaluate the efficacy of the computer vision-based measuring systems and to compare their effectiveness with direct physical measurements.

The vision system calculations are presented in the columns labelled "Algorithm Length" and "Algorithm Height", while "Actual Length" and "Actual Height" represent the ground truth measurements. The percentage error was calculated from these dimensional values to assess the accuracy of the automated predictions.

The algorithm-estimated length and height were compared to the corresponding actual values. The absolute difference between the estimated and actual dimensions was divided by the actual value, then multiplied by 100, to determine the error. This calculation was performed separately for length and height. The resulting percentage errors were plotted sequentially for all objects, following the order presented in the results table, to illustrate the variation in measurement accuracy across the dataset.

The measurement errors for the objects do not exhibit consistent trends. In some cases, length measurements are more accurate than height, while in others, the opposite is true. Factors such as object shape, reflective surfaces, curvature, and calibration conditions may influence these measurements.

This visualization identifies the most challenging areas of the measurement system, thereby facilitating targeted algorithm refinement and laying a foundation for a comprehensive investigation.

The results presented for the proposed measurement system in Table II indicate sufficient accuracy, as algorithmic estimates closely match the actual particle diameters. For most tested objects, length and height errors fall below 2%, demonstrating the robustness of the computer vision techniques. The lowest error values were observed for objects with regular geometries and well-defined edges, such as notebooks, folders, and passports. This finding confirms that the algorithm performs with high accuracy in controlled environments where object edges are clearly defined and surface illumination is uniform. In contrast, higher error rates were observed for objects with irregular geometries, reflective surfaces, or small sizes, such as company seals, mobile ring cases, and decorative items. These discrepancies are likely due to challenges in contour detection, camera alignment, and external lighting conditions. Overall, the findings suggest that the system provides consistent measurements suitable for real-world applications, particularly where non-contact size determination is required. The method is effective within the relevant measurement range and achieves an average error margin of approximately 1–3%. It is applicable in industrial, educational, and research contexts. However, further improvements in preprocessing and calibration are necessary to enhance performance, especially for irregular or complex objects.

TABLE II. RESULTS OF THE PROJECT

| Test Number | Object | Algorithm Length (mm) | Algorithm Height (mm) | Actual Length (mm) | Actual Height (mm) | Error Length (%) | Error Height (%) |
|---|---|---|---|---|---|---|---|
| 1 | Film Case | 135.00 | 128.00 | 134.00 | 127.00 | 0.74 | 0.53 |
| 2 | Photo Folder | 137.55 | 183.77 | 137.00 | 183.00 | 0.40 | 0.42 |
| 3 | Snap Fastener Case | 54.92 | 84.92 | 54.00 | 79.00 | 1.70 | 7.49 |
| 4 | Passport | 124.76 | 89.96 | 124.00 | 89.00 | 0.61 | 1.07 |
| 5 | Acer Aspire Notebook Holder | 276.26 | 216.77 | 275.00 | 216.00 | 0.45 | 0.35 |
| 6 | KFC Magic Toy | 77.08 | 81.45 | 77.00 | 81.00 | 0.10 | 0.55 |
| 7 | Tablet Case | 184.01 | 115.57 | 182.00 | 114.00 | 1.10 | 1.37 |
| 8 | Tecno Pova 4 Case | 176.91 | 85.97 | 176.00 | 85.00 | 0.51 | 1.14 |
| 9 | Tecno Pova 4 | 174.03 | 84.45 | 174.00 | 84.00 | 0.01 | 0.53 |
| | …… | …….. | …… | ….. | ….. | ….. | …… |
| 54 | Energy Drink Lid JL | 59.59 | 59.42 | 59.00 | 59.00 | 1.00 | 0.71 |
| 55 | Toy Spring | 30.50 | 30.85 | 30.00 | 30.00 | 1,67 | 2.83 |
| 56 | Glue Stick | 18.05 | 18.02 | 18.00 | 18.00 | 0.27 | 0.11 |
| 57 | Dickie Ruskie | 82.85 | 82.83 | 82.00 | 82.00 | 1.04 | 1.01 |
| 58 | Drink Lid | 63.05 | 63.08 | 62.00 | 62.00 | 1.69 | 1.74 |
| 59 | Perfume Packaging | 123.00 | 123.00 | 122.00 | 122.00 | 0.82 | 0.82 |

To provide quantitative validation beyond individual percentage errors, a statistical analysis was performed using the complete dataset of 59 measured objects. The objective was to evaluate global accuracy, systematic deviation, dispersion, and linear agreement between the proposed vision-based measurements and reference manual measurements.

Let $L_a$ denote the length estimated by the algorithm and $L_r$ the corresponding reference length obtained using manual instruments. Similarly, $H_a$ represents the algorithm estimated height and $H_r$ the reference height. The measurement error for each object $i$ was defined as the difference between the estimated and reference values, expressed as $e_{L,i} = L_{a,i} - L_{r,i}$ for length and $e_{H,i} = H_{a,i} - H_{r,i}$ for height.

Based on these error definitions, three primary statistical indicators were computed to characterize the accuracy and stability of the proposed system:

Mean Absolute Error (MAE): quantifies the average magnitude of the measurement deviation, regardless of sign. It is defined as the average of the absolute values of the individual errors [27]:

$$MAE_L = \frac{1}{n}\sum_{i=1}^{n}\left|L_{a,i} - L_{r,i}\right| \tag{8}$$

$$MAE_H = \frac{1}{n}\sum_{i=1}^{n}\left|H_{a,i} - H_{r,i}\right| \tag{9}$$

where $n$ represents the total number of evaluated objects ($n = 59$). The computed MAE for length was 0.793 mm, whilst the MAE for height was 0.861 mm. These results indicate that, on average, the proposed system produces sub-millimetric deviations relative to manual measurements. This level of precision is suitable for cardboard packaging applications, where tolerances typically exceed 1 mm.

Root Mean Square Error (RMSE): provides a global measure of accuracy that penalises larger deviations more strongly than MAE. It is calculated as [27].

$$RMSE_L = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(L_{a,i} - L_{r,i})^2} \tag{10}$$

$$RMSE_H = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(H_{a,i} - H_{r,i})^2} \tag{11}$$

The resulting RMSE values were 1.057 mm for length and 1.394 mm for height. The slightly higher RMSE observed in height measurements reflects the influence of a limited number of irregular or reflective objects that generated localized contour deviations. Nevertheless, the overall magnitude remains below 1.5 mm, confirming stable performance across varied geometries.

Standard Deviation of Measurement Error: Measurement dispersion was quantified using the sample standard deviation, defined as [27]:

$$\sigma_L = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(e_{L,i} - \bar{e}_L)^2} \tag{12}$$

$$\sigma_H = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(e_{H,i} - \bar{e}_H)^2} \tag{13}$$

where $\bar{e}_L$ and $\bar{e}_H$ represent the mean errors for length and height, respectively. The standard deviation was 0.868 mm for length and 1.117 mm for height. Considering that the evaluated objects ranged from 18 mm to 276 mm, this dispersion is relatively small, indicating consistent system behavior across different object scales.

These results provide quantitative evidence that planar dimension estimation using contour-based monocular vision, when combined with calibrated pixel-to-millimeter conversion, can achieve reliable metrological performance under controlled imaging conditions. The obtained accuracy levels are suitable for automated CNC-based cardboard packaging applications, where dimensional tolerances typically exceed the observed error margins.

## VI. Conclusions

The computer-vision measurement system shows strong accuracy across 59 objects[1], with mean errors of ~1.02% in length and ~1.32% in height, and median errors below 1% (≈0.69% and ≈0.77%). Most items remain within practical limits (≈88% ⩽ 2% length error; ≈86% ⩽ 2% height error), supporting use in automated packaging/inspection. However, a few small or geometrically ambiguous objects generate outliers, notably the Snap Fastener Case, where the height error reaches 7.49%. This elevated error is likely due to the object's small scale and less well-defined vertical edges, which increase sensitivity to pixel quantization, thresholding/contour segmentation uncertainty, and slight perspective or placement deviations during image capture. Similar behavior is observed for the Square Base Toy Cone, whose height error exceeds 10%. Despite these exceptions, overall performance is reliable; further gains should come from tighter camera calibration (including lens distortion), improved illumination control, and more robust contour refinement for small parts. The integration of computer vision and CNC machining in this project marks a significant advancement in the convergence of technologies for industrial applications. Beyond effective dimensional measurements, the collaborative operation between the computer vision module and the CNC system represents a key achievement. The precision attained through the CNC platform, managed by a Python-based human–machine interface and supported by robust object identification and measurement techniques, positions the proposed system as a viable solution for scalable, efficient, and customized packaging.

In conclusion, the challenges associated with G-code macroprogramming were successfully addressed despite the lack of native support in the GRBL firmware. Generating adaptable G-code instructions from formatted text strings and via serial communication proved effective, enabling flexible CNC control based on measured object dimensions. As future work, improvements to the experimental setup will include the incorporation of fixed and controlled lighting conditions to further minimize the influence of external illumination. Additionally, a systematic study of measurement error fluctuations under varying lighting scenarios will be conducted to enhance the robustness and repeatability of the proposed system.

Subsequent research will aim to extend the proposed system to accommodate objects with more complex geometries, such as irregular shapes, concave boundaries, and internal contours or holes. Addressing these challenges will require enhancements to the existing computer-vision pipeline, including advanced contour segmentation, shape decomposition, and robust methods for handling partial occlusions and non-uniform edges.

From a computational perspective, processing complex geometries will increase algorithmic complexity and demand additional resources. Higher-resolution image acquisition may be necessary to preserve fine geometric details, resulting in greater memory usage and longer processing times. Furthermore, the adoption of advanced image-processing techniques—such as adaptive thresholding, multi-scale contour analysis, or graph-based contour representation—may further increase CPU utilization. In addition, real-time or near-real-time processing of complex shapes may benefit from hardware acceleration, such as multi-core CPUs or GPU-based parallelization. Future implementations are therefore likely to require more capable embedded platforms or dedicated graphics processing units to maintain processing efficiency while ensuring accurate dimension estimation and reliable G-code generation for CNC machining.

A highly practical application of this research is an on-demand packaging station for Small and Medium-sized Enterprises (SMEs) and e-commerce fulfilment. An operator places a product inside the imaging enclosure, the system measures its planar footprint automatically, and a CNC cutter generates and executes G-code to produce a correctly sized cardboard insert or packaging layout. This reduces manual measuring, speeds up packing, and improves consistency across batches. It also helps minimize material waste by avoiding oversized

[1]https://docs.google.com/spreadsheets/d/1UZaqey6gKdMvh7IaTw7dSWaeIvjWpf0VjeU6nLFsIvs/edit?gid=64701813#gid=64701813

boxes and unnecessary void fill. The same workflow is valuable for returns (reverse logistics), where items vary in size and must be repacked quickly and reliably.

## CONFLICT OF INTEREST

The authors declare no conflict of interest.

## AUTHOR CONTRIBUTIONS

A.Q. conducted the research; P.G. analyzed the data; S.P. wrote the paper; A.P. developed the methodology and M.M. analyze the results. All authors have read and agreed to the published version of the manuscript.

## REFERENCES


[1] H. Shang, C. Liu, and R. Wang, "Measurement methods of 3D shape of large-scale complex surfaces based on computer vision: A review," *Measurement*, vol. 197, 111302, 2022. doi: 10.1016/j.measurement.2022.111302

[2] J. Xu and S. Zhang, "Status, challenges, and future perspectives of fringe projection profilometry," *Opt. Lasers Eng.*, vol. 135, 106193, 2020. doi: 10.1016/j.optlaseng.2020.106193

[3] S. Feng, L. Zhang, C. Zuo *et al.,* "High dynamic range 3-D measurements with fringe projection profilometry: A review," *Meas. Sci. Technol.*, vol. 29, 122001, 2018. doi: 10.1088/1361-6501/aae4fb

[4] B. Hou, C. Zhang, and S. Yang, "Computer vision tool-setting system of numerical control machine tool," *Sensors*, vol. 20, no. 18, 5302, 2020. doi: 10.3390/s20185302

[5] W. He, Z. Jiang, M. Wuyi *et al.,* "A critical review for machining positioning based on computer vision," *Measurement*, vol. 184, 109973, 2021. doi: 10.1016/j.measurement.2021.109973

[6] Y. R. Mileski, A. J. Souza, and H. J. Amorim, "Development of a computer vision-based system for part referencing in CNC machining centers," *J. Braz. Soc. Mech. Sci. Eng.*, vol. 44, 243, 2022. doi: 10.1007/s40430-022-03531-w

[7] S. Bustos, C. Aguilera, and A. Martínez, "Automatic generation of codes for routine CNC machining based on three-dimensional information obtained by fringe projection," in *Proc. Adv. Intell. Syst. Comput.*, 2017, vol. 475, pp. 195–201. doi: 10.1007/978-3-319-28513-9_27

[8] T. K. Nguyen, L. X. Phung, and N.-T. Bui, "Novel integration of CAPP in a G-code generation module using macro programming for CNC application," *Machines*, vol. 8, no. 4, 61, 2020. doi: 10.3390/machines8040061

[9] L. Zhou, L. Zhang, and N. Konz, "Computer vision techniques in manufacturing," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 53, no. 1, pp. 105–117, 2022.

[10] A. M. Viora-Vilar, C. Molina, F. J. Martín, L. Ramírez-Zamora, and J. C. Romero, "Detection of glass position using computer vision for automation of water-jet cutting machines," in *Proc. IEEE Int. Conf. Automation / XXV Congr. Chilean Assoc. Automatic Control (ICA-ACCA)*, 2022, pp. 1–6. doi: 10.1109/ICA-ACCA56767.2022.10006064

[11] M. Pan, G. Zhang, W. Zhang, J. Zhang, Z. Xu, and J. Du, "A review of intelligentization system and architecture for ultra-precision machining process," *Processes*, vol. 12, no. 12, 2754, 2024. doi: 10.3390/pr12122754

[12] X. Fu, L. Lv, B. Chen *et al.,* "Pre-control of grinding surface quality by data-driven: A review," *Int. J. Adv. Manuf. Technol.*, vol. 133, no. 7, pp. 3081–3104, 2024. doi: 10.1007/s00170-024-13921-0

[13] A. K. Adeleke, T. O. Igunma, and Z. Q. S. Nwokediegwu, "Modeling advanced numerical control systems to enhance precision in next-generation coordinate measuring machine," *Int. J. Multidisciplinary Res. Growth Eval.*, vol. 2, no. 1, pp. 638–649, 2021. doi: 10.54660/.IJMRGE.2021.2.1.638-649

[14] N. H. O. Alseedi, Y. Yusof, A. Z. A. Kadir, K. Latif, and M. A. Mohammed, "Interpreting the G-code of drilling machining to use in open CNC controller machine," *J. Phys.: Conf. Ser.*, vol. 1892, 012014, 2021. doi: 10.1088/1742-6596/1892/1/012014

[15] M. Abdullah-Al-Noman, A. N. Eva, T. B. Yeahyea, and R. Khan, "Computer vision-based robotic arm for object color, shape, and size detection," *J. Robot. Control*, vol. 3, no. 2, pp. 180–186, 2022. doi: 10.18196/jrc.v3i2.13906

[16] W. Ma, X. Qi, Y. Sun *et al.,* "Computer vision-based measurement techniques for livestock body dimension and weight: A review," *Agriculture*, vol. 14, no. 2, 306, 2024. doi: 10.3390/agriculture14020306

[17] A. Pilco, V. Moya, A. Quito, J. P. Vásconez, and M. Limaico, "Image processing-based system for apple sorting," *J. Image Graphics*, vol. 12, no. 4, pp. 362–371, 2024. doi: 10.18178/joig.12.4.362-371

[18] V. Iglesias, A. Pilco, V. Moya, and F. Abedrabbo, "FOMO model based logo detection in embroidery process," *LACCEI*, vol. 1, no. 10, 2024.

[19] M. Basu, "Gaussian-based edge-detection methods—A survey," *IEEE Trans. Syst., Man, Cybern. C*, vol. 32, no. 3, pp. 252–260, 2002. doi: 10.1109/TSMCC.2002.804448

[20] H. Wang, X. Xu, Y. Liu, D. Lu, B. Liang, and Y. Tang, "Real-time defect detection for metal components: A fusion of enhanced Canny–Devernay and YOLOv6 algorithms," *Appl. Sci.*, vol. 13, no. 12, 6898, 2023. doi: 10.3390/app13126898

[21] E. Navarro, V. Moya, A. Quito, and D. Pozo-Espín, "A cost-effective prototype for egg crack detection and labeling with a comprehensive review of existing methods," *Discover Food*, vol. 5, no. 1, pp. 1–30, 2025. doi: 10.1007/s44187-025-00604-2

[22] Z. Chen, Y. Quan, and H. Ji, "Unsupervised deep unrolling networks for phase unwrapping," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2024, pp. 25182–25192. doi: 10.1109/CVPR52733.2024.02379

[23] K. Okarma and M. Grudziński, "The 3D scanning system for machine vision-based positioning of workpieces on CNC machine tools," in *Proc. 17th Int. Conf. Methods Models Autom. Robot. (MMAR)*, 2012, pp. 85–90. doi: 10.1109/MMAR.2012.6347906

[24] S. Matsuda and Y. Takeda, "Strategy for application of support object for fall prevention in the elderly based on balance recovery characteristics," *Machines*, vol. 8, no. 4, 60, 2020. doi: 10.3390/machines8040060

[25] G. Li, H. Lu, H. Wang, Y. Ran, R. Ji, Y. Liu, Y. Zhang, B. Cai, and X. Yin, "A visualization method for cross-scale online monitoring of grinding state based on data-mechanism hybrid-driven digital twin system," *Mech. Syst. Signal Process.*, vol. 225, 112293, 2025. doi: 10.1016/j.ymssp.2024.112293

[26] Z. H. Khan, T. Rahman, S. Arabi, S. Mohammad, and M. M. Khan, "Development of a low-cost CNC machine laser engraver," in *Proc. IEEE 12th Annu. Ubiquitous Comput., Electron. Mobile Commun. Conf. (UEMCON)*, 2021, pp. 376–382. doi: 10.1109/UEMCON53757.2021.9666661

[27] C. Willmott and K. Matsuura, "Advantages of the Mean Absolute Error (MAE) over the Root Mean Square Error (RMSE) in assessing average model performance," *Climate Research*, vol. 30, no. 1, pp. 79–82, 2005. doi: 10.3354/cr030079