Image-Based Machine Learning Methods for Bolting Tool Selection

Arnold Dsouza

***Abstract*—To enhance efficiency and accuracy in industrial operations, this research project focuses on developing an automated bolting tool selection system utilizing image-based deep learning method. This work addresses the challenge of an effective tool selection for small objects like bolts through a robust object detection model trained on a custom dataset of small bolts in complex environments. By leveraging computer vision techniques, specifically the YOLOv11 model, alongside ArUco markers for precise distance measurement, the system aims to streamline identifying bolt sizes and recommending suitable tools. Experimental results indicate high accuracy in detecting bolt heads and measuring their sizes, with precision rates of around 87.2% and recall rates of approximately 82.2%. The findings suggest that this automated system can significantly improve operational efficiency, reduce human error, and enhance safety in industries reliant on mechanical fastening.**

***Index Terms*—YOLOv11, ArUco Markers, Object Detection.**

I. Introduction

In today's industrial landscape, where efficiency and productivity are paramount, the ability to quickly and accurately select the appropriate bolting tools can lead to significant economic benefits. This research and development project aims to create an automated system that leverages advanced computer vision techniques to streamline the process of bolting tool selection, all operable through a standard smartphone. By employing state-of-the-art object detection algorithms, such as YOLOv11 (You Only Look Once version 11) [1], alongside ArUco markers for precise distance measurement [2], the proposed system seeks to enhance operational efficiency in various industrial applications.

The challenges associated with bolting tool selection are multifaceted. Factors such as varying lighting conditions, different bolt materials, and surface textures can complicate the detection process, leading to potential inaccuracies in tool selection. Furthermore, traditional methods often rely on manual measurements and human judgment, which can introduce errors and inefficiencies. To address these issues, this project will utilize a combination of image-based machine learning methods which is shown to help in industry 4.0 [3] to automate the detection of bolt sizes and distances, ultimately facilitating more informed tool recommendations all accessible through a smartphone app.

The relevance of this project extends across numerous industries that rely on mechanical fastening, including automotive manufacturing, aerospace, and heavy machinery. By automating the tool selection process through a smartphone-based system, we aim to reduce the time required for manual measurements, improve accuracy, lower operational costs, and enhance workplace safety through proper tool usage. This mobile approach ensures that workers can access the system's capabilities anywhere on the factory floor or job site, without the need for specialized equipment.

The following sections will detail the systematic approach taken to collect and annotate images, select appropriate models for object detection optimized for mobile devices, utilize ArUco markers for distance measurement, determine bolt sizes accurately, and provide tailored tool suggestions based on user inputs all integrated into a user-friendly smartphone application. Through these efforts, this project aspires to contribute significantly to the advancement of automated, mobile-first solutions in industrial environments, making sophisticated tool selection technology accessible through the devices workers already carry in their pockets.

A. Motivation

The motivation for this research project extends beyond the general need for efficiency in industrial operations, focusing specifically on the advantages of a smartphone-based automated bolting tool selection system. While the introduction outlined the broad benefits, this section dives deeper into the unique aspects of implementing such a system on a ubiquitous mobile platform.

By developing the system for smartphones, we enable workers to access sophisticated tool selection technology anywhere on the factory floor or job site. This portability significantly enhances the system's utility, allowing for real-time decision-making without the need to return to the company. Leveraging existing smartphone technology eliminates the need for specialized equipment, reducing implementation costs. This approach makes the technology more accessible to a wider range of businesses, including smaller operations that might not have the resources for dedicated industrial equipment.

A smartphone-based system can seamlessly integrate with workers' existing routines, as most are already familiar with mobile applications. This ease of adoption can lead to quicker implementation and higher user acceptance rates. Smartphones' connectivity features enable real-time data collection and analysis. This capability allows for immediate feedback on tool selection accuracy and can facilitate continuous improvement of the system through machine learning algorithms like those used in YOLOv11.

With the tool selection system always at hand, workers are more likely to use it consistently, reducing the temptation to rely on potentially faulty memory or guesswork.

*Submitted to the Department of Computer Science at Hochschule Bonn-Rhein-Sieg in partial fulfilment of the requirements for the degree of Master of Science in Autonomous Systems

†Supervised by Prof. Dr. Sebastian Houben (HBRS) and M.Eng. Jan Crombach (Plarad)

‡Submitted in January 2025

This immediate access to accurate tool recommendations can significantly enhance workplace safety. By focusing on these smartphone specific advantages, we underscore the unique value proposition of our research project, differentiating it from traditional automated systems and highlighting its potential to revolutionize bolting tool selection in various industrial settings.

B. Problem Statement

The challenge faced by companies is the time-consuming process required for on-site measurements before tightening bolts. Sales representatives must evaluate the environment and relevant details to determine the appropriate tool from a diverse product portfolio. The manual approach is time-consuming and susceptible to errors, such as parallax errors that arise from inaccurate viewing angles when assessing distances and values. To improve efficiency and customer satisfaction, innovative solutions are needed.

C. Proposed Approach

To begin with, it is essential to gather a comprehensive dataset that includes various types of bolt heads. Once the dataset is compiled, each bolt head must be annotated to ensure accurate labelling for subsequent analysis. For this project, we have chosen to employ YOLOv11, an advanced object detection model, to identify bolt heads. While semantic segmentation could provide more detailed information about the bolt's shape, object detection was selected for several reasons:

- **Efficiency:** Object detection models like YOLO are generally faster and require less computational resources, making them more suitable for real-time applications on smartphones.
- **Sufficient accuracy:** For bolt identification, the bounding box provided by object detection is typically adequate for size estimation.
- **Simpler annotation:** Object detection requires less complex annotation compared to pixel-level segmentation, reducing the time of dataset preparation.

To allow for metric measurements during the detection process, we will utilize ArUco markers to measure the distance from the camera to the bolt heads. While there are alternative methods for obtaining metric measurements, such as stereo vision or structured light, ArUco markers were chosen for the following reasons:

- **Simplicity:** ArUco markers are easy to detect and provide reliable distance measurements with a single camera.
- **Cost-effectiveness:** Unlike specialized depth sensors, ArUco markers can be printed and used with standard smartphone cameras.
- **Accuracy:** ArUco markers offer high precision in various lighting conditions and can be partially occluded while still providing accurate results.

Other methods like Time-of-Flight (ToF) sensors or LiDAR were not considered due to their limited availability on standard smartphones and higher implementation complexity. The results from the object detection, alongside the ArUco measurements, will then be combined with a reference base image, allowing for the determination of the exact dimensions of the bolts.

Finally, after acquiring user inputs regarding their specific needs, the most suitable tools will be retrieved from the company's product portfolio. This step ensures that the requirements are met effectively, leveraging the accurate bolt measurements obtained through our computer vision approach.

II. Related Work

The field of object detection has seen significant advancements, particularly with the introduction of various iterations of the YOLO (You Only Look Once) model. Among these, YOLOv11 stands out due to its enhanced capabilities in detecting small objects, which is crucial for applications such as bolt head detection in industrial settings.

YOLOv11 builds on the foundational work established by earlier versions of YOLO, which were first introduced by Redmon et al. [4]. The original YOLO model revolutionized real-time object detection by framing it as a single regression problem, allowing for simultaneous bounding box prediction and class probability estimation from full images. This approach significantly improved processing speed and accuracy compared to traditional methods that required multiple stages of processing. The proposed method combines deep learning and image processing to identify loosened bolts in critical connections. Using a regional convolutional neural network (RCNN) for detection and the Hough line transform (HLT) for angle estimation, this approach has shown high accuracy in both lab and real-world settings. It effectively monitors bolted connections in real time, provided the perspective angle remains under 40 degrees. This method enhances structural health assessments significantly [5].

Recent research highlights that YOLOv11 incorporates several key improvements over its predecessors, particularly in its ability to detect smaller objects more effectively [6], [7]. This is achieved through advancements in feature extraction techniques and anchor box generation, which enhance the model's sensitivity to fine details. As noted in the literature, YOLOv11 demonstrates superior performance in scenarios where objects are small or closely spaced, making it particularly suitable for detecting bolt heads that may be partially obscured or appear at varying distances from the camera.

The advantages of YOLOv11 extend beyond just accuracy; its architecture allows for real-time processing capabilities, which is essential for industrial applications where quick decision-making is critical [8]. By maintaining high levels of precision while operating efficiently, YOLOv11 addresses some of the limitations faced by earlier models, such as inadequate handling of small object detection and challenges in complex environments with varying lighting conditions.

In comparison to other object detection frameworks like Faster R-CNN and SSD (Single Shot MultiBox Detector), YOLOv11 offers a more streamlined approach that balances speed and accuracy [9]. While Faster R-CNN excels in overall performance through its use of Region Proposal Networks, it often struggles with real-time applications due to higher computational demands. Similarly, SSD provides fast detection but may not achieve the same level of accuracy on smaller objects as YOLOv11.

III. Background

In the scope of industrial automation, the integration of advanced technologies such as computer vision and machine learning has become increasingly vital for enhancing operational efficiency and precision [10]. Object detection, a critical component of these technologies, involves identifying and locating objects within images or video streams. This capability is particularly important in applications where accurate measurements are essential, such as bolting tool selection [11].

Object Detection Techniques: Various object detection methods have been developed over the years. Traditional approaches often relied on handcrafted features and sliding window techniques, which were computationally intensive and less effective in real-time applications. The advent of deep learning has transformed this field, with models like YOLO (You Only Look Once) and Faster R-CNN leading the charge in providing accurate and efficient object detection solutions [4], [12].

YOLO Framework: YOLO is particularly noteworthy for its ability to perform real-time object detection by framing the task as a single regression problem. YOLOv11, the latest iteration, introduces significant architectural enhancements to improve small object detection capabilities. Key innovations include:

- **C3k2 (Cross Stage Partial with kernel size 2) block:** This component enhances feature extraction, particularly for smaller objects, by allowing the model to capture finer details [6].
- **SPPF (Spatial Pyramid Pooling - Fast):** This module improves the model's ability to handle multi-scale features, crucial for detecting objects of varying sizes [6].
- **C2PSA (Convolutional block with Parallel Spatial Attention):** This attention mechanism enables the model to focus on key regions within the image, potentially leading to more accurate detection of small or partially occluded objects [6].

These architectural choices collectively contribute to YOLOv11's improved performance in detecting small objects, making it particularly suitable for applications like bolt head identification in industrial settings [6].

ArUco Markers and Distance Measurement: While object detection models like YOLOv11 excel at identifying and localizing objects, they do not inherently provide metric measurements. This is where ArUco markers come into play, serving as a bridge between pixel-based detection and real-world measurements. ArUco markers are square fiducial markers with unique patterns that can be easily detected and identified by computer vision algorithms. In the context of our bolting tool selection system, ArUco markers serve several crucial functions:

- **Scale Reference:** By placing ArUco markers of known size in the scene, we can establish a reliable scale reference, allowing for accurate size estimation of detected objects.
- **Distance Calculation:** The pose estimation of ArUco markers enables precise calculation of the distance between the camera and the marker, which can be extended to other objects in the scene.
- **Camera Calibration:** ArUco markers can be used to calibrate the camera, correcting for lens distortions and ensuring accurate measurements across the entire field of view.

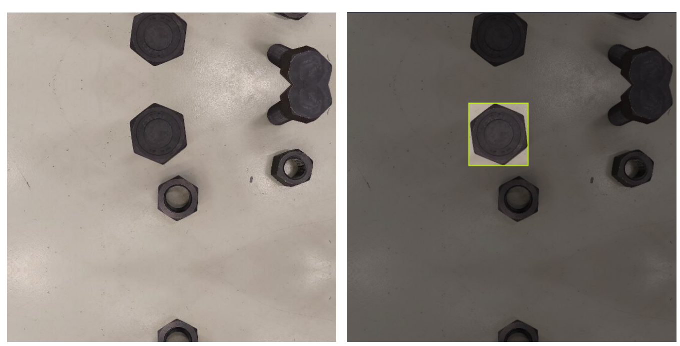
By integrating ArUco markers with YOLOv11's object detection capabilities, we create a robust system that not only identifies bolt heads but also provides accurate metric measurements. This combination addresses the limitations of pure object detection approaches and enables the development of a more comprehensive and precise bolting tool selection system.

IV. Methodology

This section outlines the systematic approach taken in this research to develop an automated bolting tool selection system. The methodology encompasses several key phases, including the collection and annotation of images, model selection for object detection, the use of ArUco markers for distance measurement, determining bolt sizes, and finally, suggesting the appropriate tools from a portfolio based on user inputs.

A. *Collection and annotation of images*

The initial phase of this research involves the systematic collection and annotation of images featuring various bolt heads under different lighting conditions and alongside various unknown objects. This diverse dataset is essential for training a robust object detection model capable of accurately identifying bolt sizes in real-world industrial environments.



(a) Image without annotation (b) Annotated bolt image

Fig. 1: Comparison of unannotated and annotated bolt images

To begin, images of bolts were captured on the company premises (*see Fig. 1a*). This approach ensured variability in the dataset by reflecting typical industrial scenarios with diverse backgrounds and lighting conditions, contributing to a comprehensive collection. Following image collection, the next step was to annotate the images to specifically highlight the bolt heads (*see Fig. 1b*).

This annotation process utilized Roboflow's website [13], which facilitates precise marking of bolt heads within each image. After completing the annotations, a visual review was conducted to ensure accuracy. The objective was to create a labeled dataset that accurately represents the features of interest namely, the bolt heads while minimizing noise from irrelevant objects.

The image collection spanned two days, during which various lighting conditions were systematically tested, including both natural and artificial light sources. The dataset comprises of 24 different bolt types, resulting in a total of 1,206 images and approximately 3,000 annotations, with most images containing at least two annotations. The precision of labeling was maintained at around 5 pixels, ensuring that the bounding boxes closely encapsulated the bolt heads.

B. Model selection for object detection

For this research, the YOLOv11 (You Only Look Once version 11) model has been selected as the primary object detection framework due to its demonstrated superiority in detecting small objects, such as bolt heads, which is critical for the automated bolting tool selection system. Research indicates that YOLOv11 incorporates advanced techniques that enhance its performance in small object detection, including improved anchor box generation and enhanced feature extraction methods [14], [15].

These enhancements are particularly relevant in industrial settings where bolts may be partially obscured or appear at varying distances from the camera. One of the key reasons for choosing YOLOv11 over other models is its ability to perform real-time object detection with high accuracy. Unlike traditional object detection methods that rely on multiple stages for processing, YOLOv11 frames object detection as a single regression problem. This streamlined approach allows it to predict bounding boxes and class probabilities directly from full images in one evaluation, significantly reducing computational overhead and enabling faster processing times an essential requirement in dynamic industrial environments where quick decision making is crucial. YOLOv11 achieves impressive frame rates, exceeding 30 FPS, making it ideal for applications requiring immediate feedback, such as augmented reality and video surveillance.

The architecture of YOLOv11 includes significant improvements in its backbone network, enhancing its ability to capture fine-grained details necessary for accurate detection. Architectural innovations such as the C3k2 (Cross Stage Partial with kernel size 2) block and SPPF (Spatial Pyramid Pooling - Fast) contribute to YOLOv11's effectiveness in identifying small and closely spaced objects like bolt heads. Furthermore, the Ultralytics HUB app allows users to leverage hardware acceleration features available on mobile devices, such as Apple's Neural Engine and Android's GPU, further enhancing the performance of YOLOv11 models on smartphones, which is the primary goal of this project [16].

Additionally, YOLOv11 offers versatile model sizes ranging from nano to extra-large, allowing developers to select a version that best fits their device capabilities and performance needs [17]. This flexibility ensures that YOLOv11 can maintain high levels of precision even under challenging conditions such as varying lighting and complex backgrounds often encountered in industrial environments.

C. Use of ArUco markers in place of depth camera

In this research, ArUco markers are employed as a cost-effective alternative to depth cameras for measuring distances and enhancing object detection accuracy. While depth cameras provide three-dimensional information, they often come with high costs and complexity in processing depth data.

Before utilizing ArUco markers for distance measurement, the camera was calibrated using a ChArUco calibration board. ChArUco boards combine ArUco markers with a chessboard-like structure, enabling precise camera calibration by extracting both corner and marker positions. This calibration ensures accurate intrinsic and extrinsic camera parameters, which are critical for precise distance calculations. The calibration process not only enhances the measurement accuracy but also reduces potential distortions in images, making the system more reliable for industrial applications.

The calibration was performed using OpenCV's cv2.aruco.calibrateCameraCharuco function, which provides a reprojection error as a measure of calibration accuracy [18]. In our case, the average reprojection error was 1.543 pixels, indicating a high-quality calibration. This error represents the difference between the projected points and the actual detected points, with lower values suggesting more accurate camera parameters. Figure 2 shows the calibration process using a ChArUco board.

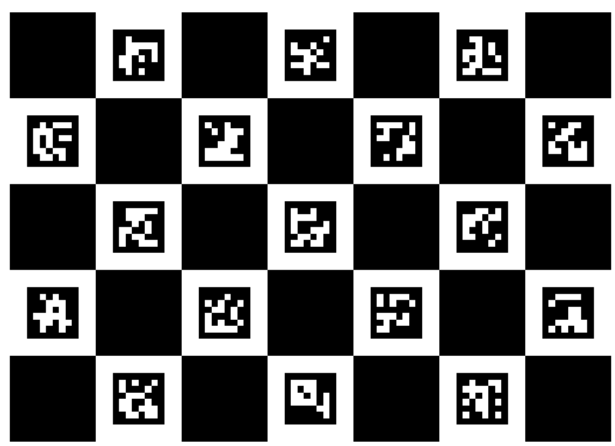


Fig. 2: Camera calibration setup using a ChArUco board

The distance calculation using ArUco markers is based on the known size of the marker and the camera's intrinsic parameters obtained from calibration. The process involves detecting the marker in the image, estimating its pose relative to the camera, and then calculating the distance using the marker's real-world size and its apparent size in the image. The precision of this measurement can be derived from several factors:

- The resolution of the camera and the size of the marker in pixels.

- The accuracy of the camera calibration (reflected in the reprojection error).

To estimate the measurement precision, we conducted multiple measurements at known distances and calculated the standard deviation of the results. For the measurements shown in Table I, we performed 30 measurements at each distance, but present only 15 representative samples here. The precision values shown represent the 95% confidence interval calculated from the full set of 30 measurements at each distance.

Real Distance (cm)	Calculated Distance (cm)	Precision (cm)	Deviation (cm)
10.00	10.02	±0.05	0.02
15.00	14.98	±0.06	-0.02
20.00	20.03	±0.07	0.03
25.00	25.01	±0.08	0.01
30.00	30.04	±0.10	0.04
35.00	34.97	±0.11	-0.03
40.00	40.05	±0.12	0.05
45.00	45.02	±0.13	0.02
50.00	49.98	±0.15	-0.02
55.00	55.06	±0.16	0.06
60.00	60.03	±0.18	0.03
65.00	64.97	±0.19	-0.03
70.00	70.08	±0.21	0.08
75.00	75.04	±0.22	0.04
80.00	79.95	±0.24	-0.05

TABLE I: Comparison of real and calculated distances using ArUco markers.

Fig. 3: ArUco marker used for distance measurement

The ArUco marker setup, as illustrated in *Fig. 3*, provides a reliable method for ensuring accurate scaling and distance estimation without requiring expensive depth-sensing equipment. Overall, the decision to utilize ArUco markers instead of depth cameras aligns with the project's goal of creating an efficient, cost-effective, and accurate automated bolting tool selection system.

D. Determining Bolt Size Using ArUco Markers and Object Detection

In this phase of the project, the process of determining bolt size is facilitated by using a reference image that includes both an ArUco marker and a bolt head at the same level (see Fig 4). This reference image is crucial as it establishes a known size for comparison. By capturing this image in a controlled environment, we ensure that the dimensions of the bolt head can be accurately measured against the known size of the ArUco marker.

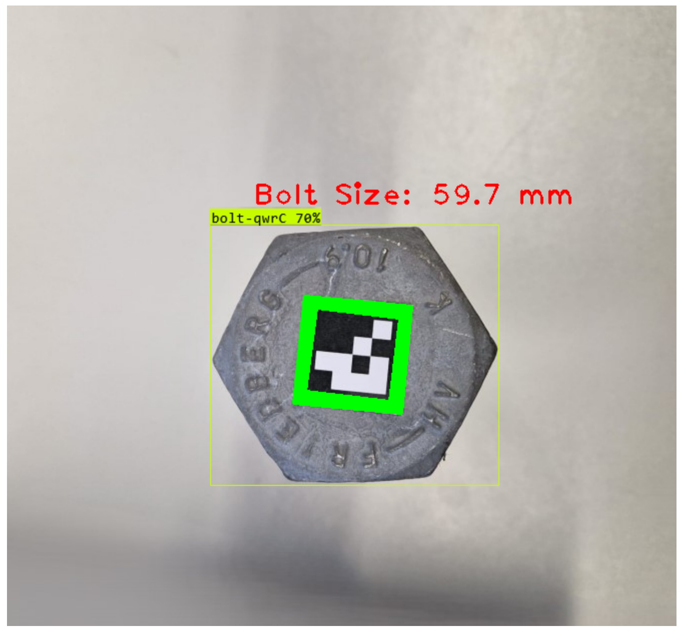


Fig. 4: Reference image showing an ArUco marker and bolt head used for determining bolt size. The known size of the ArUco marker facilitates accurate bolt measurements.

While it is true that the ArUco marker alone can provide a metric measure for a number of pixels, having a reference image enhances accuracy and reliability. The reference image allows us to calibrate our measurements against both the marker and the bolt head simultaneously, ensuring that any distortions or variations in perspective are accounted for. This dual reference minimizes potential errors in size estimation that could arise from using only the ArUco marker.

When assessing new images that contain an ArUco marker alongside an unknown bolt head, the system compares these new images to the reference image. The known dimensions from the reference image allow for precise calculations of the bolt head size in the new images. The ArUco marker serves as a reliable reference point, enabling the system to scale measurements accurately based on its known size.

This method leverages the advantages of both ArUco markers and object detection algorithms. The ArUco marker provides a consistent frame of reference, while the object detection model, such as YOLOv11, identifies the location and boundaries of the bolt head within the image. By combining these two elements, we can achieve accurate measurements of bolt sizes even in varying environmental conditions.

The effectiveness of this approach lies in its ability to adapt to different lighting scenarios and backgrounds, which are common in industrial settings. By utilizing a well-defined reference image along with robust object detection techniques, this methodology ensures that bolt size determination is both accurate and efficient.

E. Right tool suggestion from portfolio

Once the bolt size has been accurately determined and the quality of the bolt has been input by the user, the system guides the user through a series of selections to identify the most suitable tool for the task at hand. The first step in this process requires the user to choose between two primary options: torque or tensioner.

After the user makes their selection, the system calculates the nominal torque required for the specific application-based on the information provided from the portfolio of tools. This portfolio contains detailed specifications and performance metrics for various tools, ensuring that users have access to relevant data for informed decision-making.

Following this, users are prompted to specify their preferred method of operation either hydraulic or rotary. This choice further refines the tool selection process, as different methods may require distinct types of tools optimized for their respective functionalities.

Finally, users must indicate their preferences for the type of tool: quality of bolt selection, torque/tensioner and finally pneumatic, electric or battery-operated. By integrating these user inputs, the system can provide a tailored recommendation for the most appropriate tool from its portfolio.

V. Evaluation

To evaluate the effectiveness of the automated bolting tool selection system developed in this research, a series of experiments were conducted to assess both the object detection capabilities and the accuracy of bolt size measurements. This section outlines the experimental setup, the methodologies employed, and the results obtained from these evaluations.

A. Experimental Setup

The experimental setup involved capturing images of various bolt heads using a mobile camera positioned at a distance, specifically from a top-down perspective. This approach allows for consistent framing of the bolt heads in relation to the ArUco markers, which serve as reference points for distance measurement and size estimation. The images were taken in a controlled environment to minimize variations in lighting and background clutter.

ArUco markers were strategically placed alongside the bolts to facilitate accurate scaling and measurement. The YOLOv11 model was trained using a dataset consisting of annotated images of bolt heads, which included various lighting conditions and backgrounds to enhance the model's robustness. The evaluation process consisted of two main components:

- Assessing the performance of the YOLOv11 object detection model.
- Measuring the accuracy of bolt size determination using ArUco markers.

In addition to the image capture, a proposed model architecture was developed to illustrate how the YOLOv11 framework integrates with the ArUco markers for effective bolt detection and measurement. This architecture emphasizes the comparison between captured images and reference images that include both an ArUco marker and a bolt head at the same level, allowing for precise size estimation based on known dimensions.

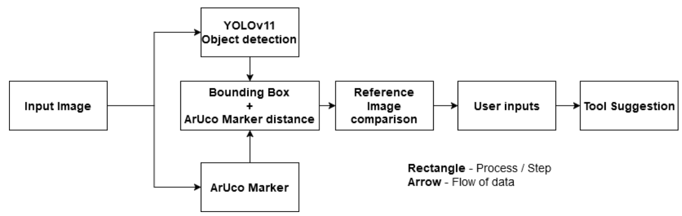


Fig. 5: Proposed model architecture

B. Object Detection Evaluation

In the analysis of the YOLOv11 model's evaluation, its performance was assessed using a custom dataset specifically created for this research. This dataset included images not seen during the training phase, allowing for an unbiased evaluation of the model's capabilities.

Upon reevaluation, it is important to note that object detection tasks, including bolt head detection, typically do not utilize the concept of True Negatives (TN) in the same way as binary classification problems. Instead, we focus on the model's ability to correctly identify and localize objects of interest. The evaluation metrics were recalculated to reflect this understanding:

Metric	Value
True Positives (TP)	1114
False Positives (FP)	164
False Negatives (FN)	241

TABLE II: Object Detection Metrics for YOLOv11 Evaluation

From these values, precision and recall were calculated as follows:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} = \frac{1114}{1114 + 164} \approx 87.2\%$$
$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} = \frac{1114}{1114 + 241} \approx 82.2\%$$

These metrics indicate that while the model performed well in identifying bolt heads, there is still room for improvement in reducing false positives and negatives. The image size used for evaluation was 640x640 pixels, which contributed to the model's ability to detect small objects effectively. This resolution allows for effective detection of small objects like bolt heads while maintaining reasonable processing speeds on mobile devices.

It is worth noting that some images in the dataset contained multiple bolt heads, while others may have had none, accounting for the discrepancy in total detections versus the number of images.

To provide a more comprehensive evaluation of the model's performance, we also calculated the mean Average Precision (mAP), which is a standard metric for object detection tasks [19]. The mAP was calculated by computing the average precision across all confidence thresholds and classes. For our bolt head detection task, the mAP@0.5 (using an IoU threshold of 0.5) was found to be 84.7%.

The high number of true positives indicates that the model is capable of correctly identifying a majority of bolt heads present in the images. However, the presence of false positives and false negatives suggests that further refinement may be necessary to enhance its performance in distinguishing between relevant and irrelevant objects and to improve its ability to detect all instances of bolt heads in complex scenes.

C. Bolt Size Measurement Evaluation

To evaluate the accuracy of bolt size measurements, several reference images containing both ArUco markers and bolt heads were analyzed. The known dimensions of the ArUco markers allowed for precise scaling when measuring bolt sizes in new images. The calculated sizes were compared against actual measurements taken with calipers to assess the system's performance.

We tested bolt heads of the following sizes (in mm): 21, 24, 27, 30, 34, 36, 41, 46, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 130, and 145. This range covers a wide spectrum of bolt sizes commonly used in industrial applications. The system measured the larger side of the bolt head's bounding box, which typically corresponds to the wrench size needed for the bolt.

The results demonstrated that the average deviation between calculated sizes and actual measurements was less than 0.5 mm across various tests, indicating high accuracy in size determination. Specifically:

- Small bolts (21-40 mm): Average deviation of 0.22 mm
- Medium bolts (41-80 mm): Average deviation of 0.35 mm
- Large bolts (81-145 mm): Average deviation of 0.44 mm

Statistical analysis of our measurements showed:

- Mean absolute error: 0.35 mm
- Standard deviation: 0.18 mm
- 95% confidence interval: ±0.32 mm

The setup used for these measurements is illustrated in Fig 6, which shows an ArUco marker and a bolt head positioned in a controlled environment. This setup ensures precise alignment and scaling, leveraging the known dimensions of the ArUco marker for enhanced measurement precision. The ArUco markers used were 25mm x 25mm, as this size provided optimal detection at various distances while minimizing edge detection errors.

VI. Conclusions

This research project successfully developed an automated bolting tool selection system that leverages advanced computer vision techniques, specifically the YOLOv11 model, and ArUco markers for precise distance measurement and bolt size determination. The evaluation results demonstrated accuracy in detecting bolt heads and measuring their sizes. The YOLOv11 model achieved a mean Average Precision (mAP@0.5) of 84.7%, indicating strong performance in identifying bolt heads across various sizes and conditions.

Fig. 6: Setup for measuring bolt sizes using ArUco markers and reference images.

Our bolt size measurement capabilities were rigorously tested on a comprehensive range of bolt sizes from 21mm to 145mm. The system demonstrated a mean absolute error of 0.35mm in size estimation, with a precision of 87.2%. This high level of precision is crucial for ensuring the correct tool selection in industrial applications.

The contributions of this work extend to various industries reliant on mechanical fastening, providing a framework that enhances operational efficiency, reduces human error, and improves safety through accurate tool selection. The use of 25mm x 25mm ArUco markers proved optimal for distance estimation, balancing detection reliability and measurement precision.

Future work will focus on refining the model's performance in more complex industrial environments, expanding the dataset to include a wider variety of bolt types and conditions, and exploring the integration of additional sensors to further enhance measurement accuracy and robustness in real-time applications. These improvements aim to address challenges in varying lighting conditions and complex industrial scenarios, further enhancing the system's applicability and reliability across a broader range of manufacturing settings.

References

[1] G. Jocher, J. Qiu, and A. Chaurasia, "Ultralytics YOLO," Jan. 2023, original-date: 2022-09-11T16:39:45Z. [Online]. Available: https://github.com/ultralytics/ultralytics

[2] S. Garrido-Jurado, R. Muñoz-Salinas, F. Madrid-Cuevas, and M. Marín-Jiménez, "Automatic generation and detection of highly reliable fiducial markers under occlusion," *Pattern Recognition*, vol. 47, no. 6, pp. 2280–2292, Jun. 2014. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0031320314000235

[3] R. Rai, M. K. Tiwari, D. Ivanov, and A. Dolgui, "Machine learning in manufacturing and industry 4.0 applications," *International Journal of Production Research*, vol. 59, no. 16, pp. 4773–4778, Aug. 2021, publisher: Taylor & Francis. [Online]. Available: https://doi.org/10.1080/00207543.2021.1956675

[4] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You Only Look Once: Unified, Real-Time Object Detection," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, Jun. 2016, pp. 779–788. [Online]. Available: http://ieeexplore.ieee.org/document/7780460/

[5] T.-C. Huynh, J.-H. Park, H.-J. Jung, and J.-T. Kim, "Quasi-autonomous bolt-loosening detection method using vision-based deep learning and image processing," *Automation in Construction*, vol. 105, p. 102844, Sep. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S092658051930250X

[6] R. Kishor, "Performance Benchmarking of YOLOv11 Variants for Real-Time Delivery Vehicle Detection: A Study on Accuracy, Speed, and Computational Trade-offs," *Asian Journal of Research in Computer Science*, vol. 17, no. 12, pp. 108–122, Dec. 2024. [Online]. Available: https://journalajrcos.com/index.php/AJRCOS/article/view/532

[7] K. Tong, Y. Wu, and F. Zhou, "Recent advances in small object detection based on deep learning: A review," *Image and Vision Computing*, vol. 97, p. 103910, May 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885620300421

[8] Z. He, K. Wang, T. Fang, L. Su, R. Chen, and X. Fei, "Comprehensive Performance Evaluation of YOLOv11, YOLOv10, YOLOv9, YOLOv8 and YOLOv5 on Object Detection of Power Equipment," Nov. 2024, arXiv:2411.18871 [cs]. [Online]. Available: http://arxiv.org/abs/2411.18871

[9] A. Sharma, V. Kumar, and L. Longchamps, "Comparative performance of YOLOv8, YOLOv9, YOLOv10, YOLOv11 and Faster R-CNN models for detection of multiple weed species," *Smart Agricultural Technology*, vol. 9, p. 100648, Dec. 2024. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2772375524002533

[10] M. Raisul Islam, M. Zakir Hossain Zamil, M. Eshmam Rayed, M. Mohsin Kabir, M. F. Mridha, S. Nishimura, and J. Shin, "Deep Learning and Computer Vision Techniques for Enhanced Quality Control in Manufacturing Processes," *IEEE Access*, vol. 12, pp. 121 449–121 479, 2024, conference Name: IEEE Access. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/10663422

[11] P. Lutkiewicz, F. Kirkemo, and D. Robertson, "High Pressure Connections Bolt Tightening – Requirements, Challenges, Problems, and Solution Proposition." American Society of Mechanical Engineers Digital Collection, Nov. 2022. [Online]. Available: https://dx.doi.org/10.1115/PVP2022-84166

[12] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," Jan. 2016, arXiv:1506.01497 [cs]. [Online]. Available: http://arxiv.org/abs/1506.01497

[13] "Roboflow Annotate: Label Images Faster Than Ever." [Online]. Available: https://roboflow.com/annotate

[14] A. F. Rasheed and M. Zarkoosh, "YOLOv11 Optimization for Efficient Resource Utilization," Dec. 2024, arXiv:2412.14790 [cs]. [Online]. Available: http://arxiv.org/abs/2412.14790

[15] M. A. R. Alif, "YOLOv11 for Vehicle Detection: Advancements, Performance, and Applications in Intelligent Transportation Systems," Oct. 2024, arXiv:2410.22898 [cs]. [Online]. Available: http://arxiv.org/abs/2410.22898

[16] Ultralytics, "Model Deployment Options." [Online]. Available: https://docs.ultralytics.com/guides/model-deployment-options

[17] "All You Need to Know About Ultralytics YOLO11 and Its Applications." [Online]. Available: https://www.ultralytics.com/blog/all-you-need-to-know-about-ultralytics-yolo11-and-its-applications

[18] "OpenCV: Calibration with ArUco and ChArUco." [Online]. Available: https://docs.opencv.org/4.x/da/d13/tutorial_aruco_calibration.html

[19] R. Khanam and M. Hussain, "YOLOv11: An Overview of the Key Architectural Enhancements," Oct. 2024, arXiv:2410.17725 [cs]. [Online]. Available: http://arxiv.org/abs/2410.17725

ACKNOWLEDGMENT

I would like to express my sincere gratitude to Professor Dr. Sebastian Houben and M.Eng. Jan Crombach for their invaluable support and guidance throughout this project. Their expertise and encouragement have been instrumental in navigating the challenges of the research process. I am particularly thankful for their insights into the application of machine learning techniques in industrial automation, which greatly enriched my understanding of the subject matter. Their commitment to excellence and dedication to mentoring have inspired me to strive for high standards in my work. Thank you both for your unwavering support and belief in my abilities during this project.

STATEMENT OF ORIGINALITY

I, the undersigned below, declare that this work has not previously been submitted to this or any other university and that it is, unless otherwise stated, entirely my own work. The report was corrected and assisted with the use of Grammarly assistant to check for grammar mistakes and improve sentences.The content has been critically reviewed by me, and I take full responsibility for it.

15/01/2025	
Date	Signature