

Automatic Quality Inspection of Rebar Spacing Using Vision-Based Deep Learning with RGBD Camera

Songyue Wang¹, Lu Deng^{1,2,*}, Jingjing Guo^{1,2}, Mi Liu¹, and Ran Cao^{1,2}

¹College of Civil Engineering, Hunan University, China

²Hunan Provincial Key Laboratory for Damage Diagnosis for Engineering Structures, Hunan University, China

*Corresponding Author

songyue@hnu.edu.cn, *denglu@hnu.edu.cn, guojingjing@hnu.edu.cn, liumi@hnu.edu.cn, rcao@hnu.edu.cn

Abstract –

Non-compliant rebar placement in reinforced concrete structures directly impacts the strength and durability of the overall structure. Current inspection methods heavily rely on manual techniques, introducing subjectivity and potential errors. This paper proposes a method for automatic quality inspection of rebar spacing using vision-based deep learning combined with RGBD camera. The method consists of three modules: (1) The image preprocessing module applies point cloud plane fitting techniques to eliminate the interference of the background rebar layer and rotate the image of the current rebar layer; (2) The recognition and localization module employs YOLOV8 keypoint detection algorithm to obtain pixel coordinates of rebar crosspoints, which are then transformed into spatial coordinates in the camera coordinate system; (3) The automatic inspection module proposes an automatic spacing measurement method based on the pixel coordinates of rebar crosspoints, enabling the inspection of all rebar spacings in the image and evaluating their compliance. Experimental results demonstrate the robustness of the pixel segmentation method, showcasing its applicability for compliance inspection with an average error of 2.65mm. The study concludes by suggesting potential directions for future research.

Keywords –

Keypoint detection; RGBD camera; Rebar spacing; Distance detection; Quality inspection

1 Introduction

Rebar is a vital structural element in reinforced concrete structures, providing support and ensuring proper load distribution. The correct placement between rebars is vital for the overall quality, strength, and durability of the structure [1]. Building codes in many countries, such as China, specify allowable deviation for

rebar placement, typically $\pm 20\text{mm}$ for bidirectional rebar mesh [2], as shown in Figure 1. Currently, quality inspection for rebar placement is mainly focused on the concealed engineering acceptance stage before concrete casting. However, inadequate quality control during construction process can lead to potential issues such as rework, construction delays, and even catastrophic consequences [3]. Considering the sequential layering construction process involved in Multi-layered bidirectional rebar, monitoring rebar spacing should also adopt a layer-by-layer approach. Therefore, strengthening rebar quality inspection during the construction process is crucial to minimize errors and enhance structural safety.

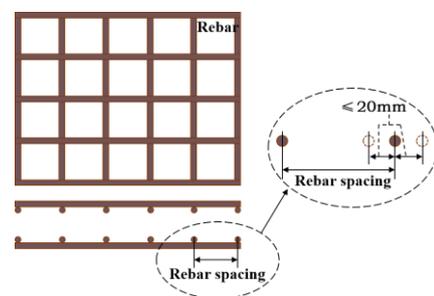


Figure 1. Illustration of the allowable deviation for rebar spacing

Traditionally, rebar placement inspection relies on manual methods involving experienced inspectors, which are cumbersome, time-consuming, and error-prone [4]. Recent advancements in computer vision technology have provided opportunities for automating dimensional quality inspection of reinforced concrete components. Current research has mainly focused on prefabricated components (e.g., precast panels and beams) [5][6][7],[8], with some exploring its application to rebar inspection [9][10],[11]. However, the reliance on 3D computer vision and point cloud models presents challenges such as complexity and limited automation. Therefore, there is an urgent need for a rapid, accurate

method to automatically inspect the dimensional quality of rebar.

The image-based measurement method enables real-time measurements [12], critical for monitoring rebar spacing layer by layer in the construction process. Existing algorithms struggle with interference from background rebar layers when inspecting the current layer. Kardovskyi et al. [13] employed Mask Region-Based Convolutional Neural Network (Mask R-CNN) for single-layer rebar instance segmentation but did not extend it to double-layer bidirectional rebar. An et al. [14] proposed a method using images and a laser rangefinder but faced challenges in layer attribution for bidirectional rebar. Jin et al. [15] introduced a neural network-based method but required consistent camera distance and angle for effective depth filtering, posing limitations in practical applications.

To address the aforementioned problems, this paper proposes a novel method for automatic quality inspection of rebar spacing using vision-based deep learning combined with RGBD camera. The key innovations are as follows: (1) Introducing a method to filter the background rebar layer using point cloud plane fitting, addressing interference in the current layer. (2) Presenting an automatic spacing measurement method based on the pixel coordinates of the rebar crosspoint for inspecting all rebar spacing in the image. (3) Proposing an intelligent end-to-end method for rebar spacing inspection. The feasibility and reliability of the proposed method are validated in physical double-layer bidirectional rebar structures.

2 Theoretical method

This paper presents a method for automatic quality inspection of rebar spacing using keypoint detection algorithm combined with RGBD camera. As shown in Figure 2, the automatic rebar spacing inspection consists of three modules: the image preprocessing, recognition and localization, and automatic inspection module. A detailed explanation is as follows.

2.1 Image preprocessing module

Obtaining the rebar pixels is crucial for spacing inspection. However, because of the feature similarity existing in double-layer bidirectional rebar, the background rebar layer could greatly influence the inspection of rebar spacing in the top layer. To extract the pixels belonging to the top layer, this paper employed a point cloud plane fitting technique based on depth information. Firstly, the RGB and depth images, captured by using the RGBD camera, are transformed into point clouds by incorporating the camera intrinsics, and then the passthrough filter is applied to denoise the point clouds, retaining the region of interest containing the rebar points. Subsequently, the denoised point clouds undergo RANSAC plane fitting algorithm, wherein various planes are segmented to obtain multiple point cloud planes along with their respective plane equations $Ax + By + Cz + D = 0$. Optimal plane parameters are determined through the minimization of aggregate vertical distances from each rebar layer's data points to the plane. To identify the point cloud plane containing the foremost rebar layer, considering that the coordinate reference system of the point cloud has its origin at the camera's position, the distance from the origin (0,0,0) to each point cloud plane is calculated, as shown in Equation (1). The plane with the shortest distance is identified as the point cloud plane of the current rebar layer.

$$dis = \frac{|D|}{\sqrt{A^2 + B^2 + C^2}} \quad (1)$$

To achieve automatic spacing inspection in Section 2.3, it is essential to ensure that the rebar in the image is approximately parallel to the x and y axes. For images that do not meet this requirement, this paper employs a rotation-based processing approach, which involves rotating the point cloud plane of the current rebar layer with the rotation center at the coordinate origin (0,0,0). This paper determines the rotation angle by taking the minimum angle between each rebar and the xz plane in the current layer. Utilizing the RANSAC algorithm, the

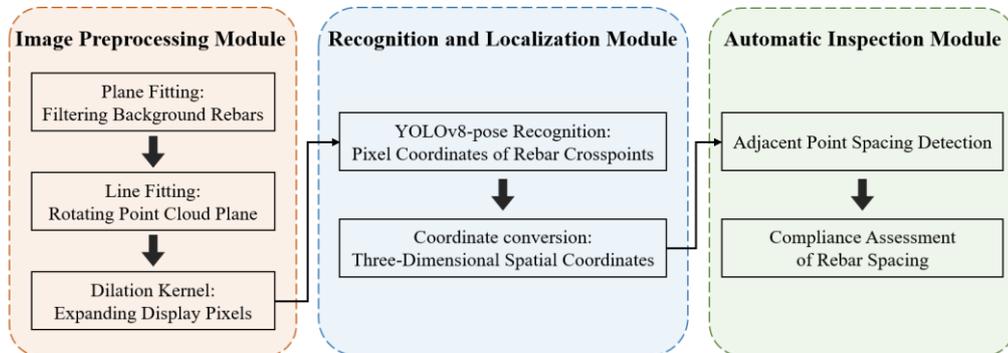


Figure 2. Flowchart of proposed formwork

spatial line fitting is applied to the point cloud plane of the current rebar layer, resulting in multiple rebar lines and their corresponding equations. Subsequently, the angle θ_i between each rebar line and the xz plane is calculated, as shown in Equation (2). By comparing these angles, the minimum value θ_{\min} is identified as the rotation angle. Correspondingly, the direction vector \vec{A}_{\min} of the rebar line associated with θ_{\min} is obtained.

$$\sin \theta_i = |\cos \phi_i| = \left| \frac{\vec{A}_i \cdot \vec{u}}{|\vec{A}_i| \cdot |\vec{u}|} \right| \quad (2)$$

where ϕ_i is the angle between the direction vector of the rebar and the normal vector $\vec{u} = (0,1,0)$ of the xz plane, \vec{A}_i is the directional vector of the rebar line defined by $y = \vec{A}_i x + \vec{B}_i$, \vec{B}_i is the 3D intercept of the rebar line.

$\vec{A}_{\min xz}$ is the projection of \vec{A}_{\min} on the xz plane. The rotation direction of the point cloud is determined by the z-axis component of the cross product $\vec{A}_{\min} \times \vec{A}_{\min xz}$. If it is greater than zero, the rotation direction is counterclockwise; otherwise, the rotation direction is clockwise. Finally, the rotated point cloud plane is mapped into an image of the current rebar layer. After processing, the recognition of rebar crosspoints on this image effectively avoids interference from other rebar layers.

2.2 Recognition and localization module

2.2.1 Introduction to the YOLOv8 framework

YOLOv8 from Ultralytics [16], the latest iteration of the YOLO-based object detection algorithm [17] series, boasts advanced capabilities encompassing object detection, instance segmentation, keypoint detection, tracking, and classification. By replacing the detection head of YOLOv8 with a pose detection head, it can be repurposed for keypoint detection. In contrast to existing keypoint detection algorithms, YOLOv8-pose eliminates the need for post-processing steps inherent in bottom-up algorithms to group detected keypoints into a target object, and also circumvents the top-down approach of first employing object detection algorithms to find bounding boxes and then conducting keypoint detection within each box. YOLOv8-pose introduces an end-to-end training method that associates all keypoints of the detected objects with bounding boxes. This model simultaneously learns the tasks of object detection and keypoint detection, employing a joint loss function to account for the correlation between the two tasks and shared features.

The YOLOv8-pose model comprises the following key components: input, backbone network, neck network,

and detection head, as shown in Figure 3. After undergoing the preprocessing outlined in Section 2.1, the images are scaled and subjected to data augmentation techniques such as mosaic and mixup before being input into the network. In the backbone, the C2f module based on cross stage partial (CSP) [18] and the SPPF module based on spatial pyramid pooling (SPP) [19] are used to extract image features to eliminate redundant operations and expedite feature fusion. The neck network combines the Feature Pyramid Network (FPN) [20] and the Path Aggregation Network (PAN) [21] structure, facilitating bidirectional fusion of low-level and high-level features, thereby improving the model detection performance across different scales. YOLOv8-pose uses a decoupled detection head to calculate the loss of localization and classification for bounding boxes and keypoints through four parallel branches of convolution.

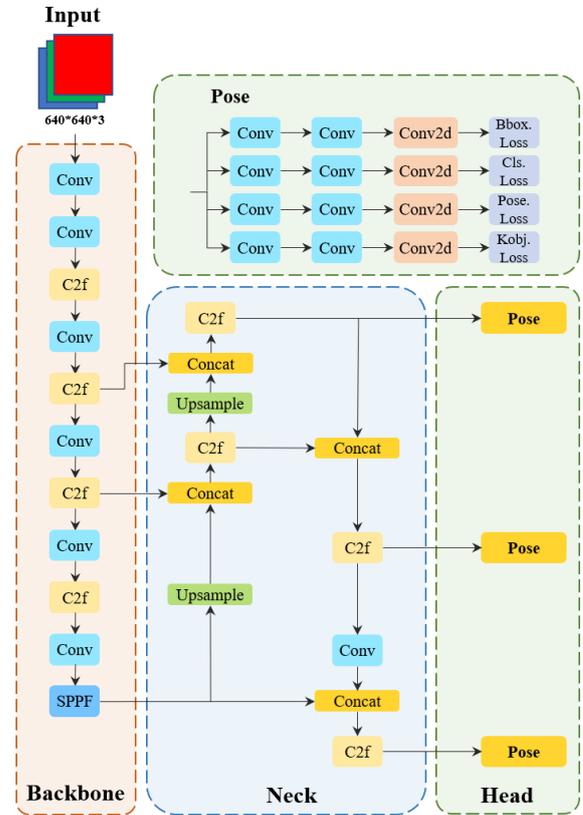


Figure 3. The network architecture of YOLOv8-pose

2.2.2 Keypoint spatial localization

The pixel coordinates of rebar crosspoints in the image can be predicted through the above keypoint detection algorithm. By integrating depth information and camera intrinsic parameters, these pixel coordinates can be transformed into spatial coordinates in the camera coordinate system.

Depth information can be obtained from the depth image. Camera intrinsic parameters are the inherent properties of a camera, including focal lengths (f_x and f_y), principal point coordinates (c_x , c_y), etc. The camera intrinsic matrix K is shown in Equation (3).

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (3)$$

For any given point, the transformation relationship between the camera coordinate system and the pixel coordinate system is described by Equation (4).

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = K \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} \quad (4)$$

where (u, v) is the coordinate in the pixel coordinate system, and (X_c, Y_c, Z_c) is the coordinate in the camera coordinate system.

2.3 Automatic inspection module

Through the aforementioned steps, the 3D spatial coordinates (X_c, Y_c, Z_c) and pixel coordinates (u, v) of all rebar crosspoints in the image have been obtained. This paper proposes a method for automatic spacing measurement based on the pixel coordinates of rebar crosspoints, aiming to obtain the rebar spacing values.

This paper transforms the problem of measuring the rebar spacing into measuring the distance between two rebar crosspoints. Only the distance between adjacent crosspoints accurately reflects the spacing of adjacent rebars, as shown in Figure 4. Therefore, this paper defines the adjacent points for a given point as the four points closest to it in the up, down, left, and right directions. In terms of code implementation, two points are considered adjacent if they satisfy the following conditions: (1) one of the pixel coordinates of the two points is considered extremely close, which is equal to or less than a small threshold; (2) the other pixel coordinate is the minimum among all points satisfying the first condition. It is noted that a point is only connected to adjacent points on its right and above, excluding those on its left and below. This design ensures that there is no re-

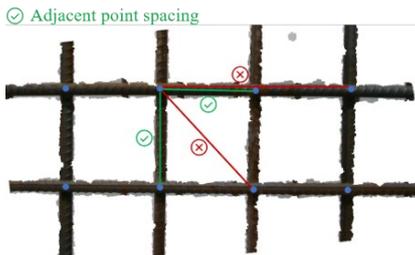


Figure 4. Schematic diagram of adjacent point spacing

petition in distance measurements while traversing all points.

By measuring the distances between all adjacent rebar crosspoints, all rebar spacing in the image can be obtained. Finally, based on the allowable deviation, assess whether these comply with the specified standards.

3 Experiment setup

The proposed method was tested on a double-layer bidirectional rebar cage with 8mm diameter rebar, as shown in Figure 5. The rebar cage, fabricated by spot welding at crosspoints, had dimensions of $2\text{m} \times 0.2\text{m} \times 0.9\text{m}$. The designed rebar spacing was $100\text{mm} \times 10$ and $200\text{mm} \times 5$. In the experiment, data collection was performed using the Intel D435i camera device with the resolution of 848×480 . The camera intrinsic parameters were specified as follows: $f_x = 606.946$, $f_y = 607.077$, and the principal point coordinates $(c_x, c_y) = (418.495, 250.889)$. The shooting distance from the camera lens to the rebar plane was approximately 300mm. Both RGB images and aligned depth images were acquired for analysis. Then the RGB images were depth-filtered and rotated to obtain images containing only the current layer of rebar pixels. However, due to the challenging task of accurately capturing the depth information of small-diameter rebar with the camera, there were discontinuous areas in the segmentation effect, as shown in Figure 6. To expand the display of rebar pixels in the current layer, considering the linear nature of double-layer bidirectional rebar, this paper employed a row-column pixel display method. Specifically, if the number of rebar pixels in a row or column exceeded a certain threshold, all pixels in that row or column were displayed. The improved segmentation result is shown in Figure 6. But it is important to note that this method is primarily designed to address the poor performance of the Intel D435i camera in segmenting small-diameter rebar.

This study utilized a total of 124 images as the dataset, all of which were captured under indoor laboratory conditions. Considering the small size and simplicity of the dataset, it was advisable to increase the proportion of the validation set. The dataset was split into a training set (65%) and a validation set (35%) to ensure accurate and



Figure 5. The double-layer bidirectional rebar cage

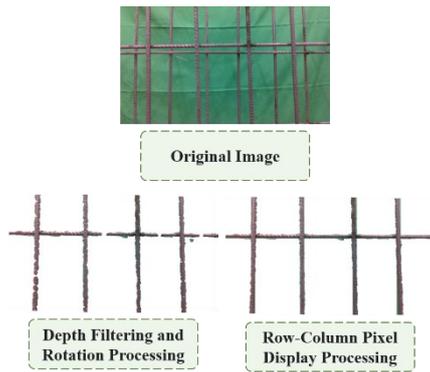


Figure 6. Schematic of the row-column pixel display method

robust prediction of model performance. 81 images were randomly selected for the training set, while the remaining 43 images constituted the validation set. Annotations were conducted on preprocessed RGB images, manually marked using the LabelMe tool to generate JSON files containing information such as image dimensions, object category names, coordinates of the four vertices of the bounding box, and coordinates of keypoints. Since the measurement of spacing is critically dependent on the accurate localization of pixels at rebar crosspoints, manual annotations points should ideally be positioned at the central point of intersection areas between two rebars to the greatest extent possible. Finally, the label files were converted from the JSON format to the TXT format suitable for YOLO training.

The YOLOv8-pose network was trained on Python 3.9 and PyTorch 2.0 environment. The training was conducted on the Windows 10 operating system with hardware specifications including an Intel(R) Core (TM)

i5-10400F CPU @ 2.90GHz and NVIDIA GeForce RTX 3060 GPU. The network training utilized the SGD optimizer with the following settings: learning rate of 0.01, 100 training epochs, and batch size of 8. Furthermore, the images fed into the network were randomly cropped and resized to the uniform size of 640×640 pixels. Data augmentation techniques were employed to enhance the training process.

4 Experimental results and discussion

4.1 Experimental result analysis

First, the effects of image depth filtering and rotation were analyzed. Figure 7 shows the rebar segmentation effects with different camera distances and rebar diameters. To explore the impact of different distances on segmentation effects, the row-column pixel display method was not applied in this case. From Figure 7, it can be observed that at longer distances, both the 8mm and 12mm diameter rebar exhibit pixel fracture or discontinuity. However, for thicker diameter rebar, the distance threshold at which this phenomenon occurs is larger and less frequent. The results indicate that, with increasing distance, the segmentation effect of a individual rebar gradually diminishes; At the same distance, thicker diameter rebar exhibit better segmentation effect compared to their thinner counterparts. This may be attributed to the difficulty of the camera depth sensor in accurately capturing long-distant and smaller-diameter rebar. Additionally, some background pixel points were retained after segmentation, possibly due to the interference of depth information by factors such as ambient noise and lighting variations.

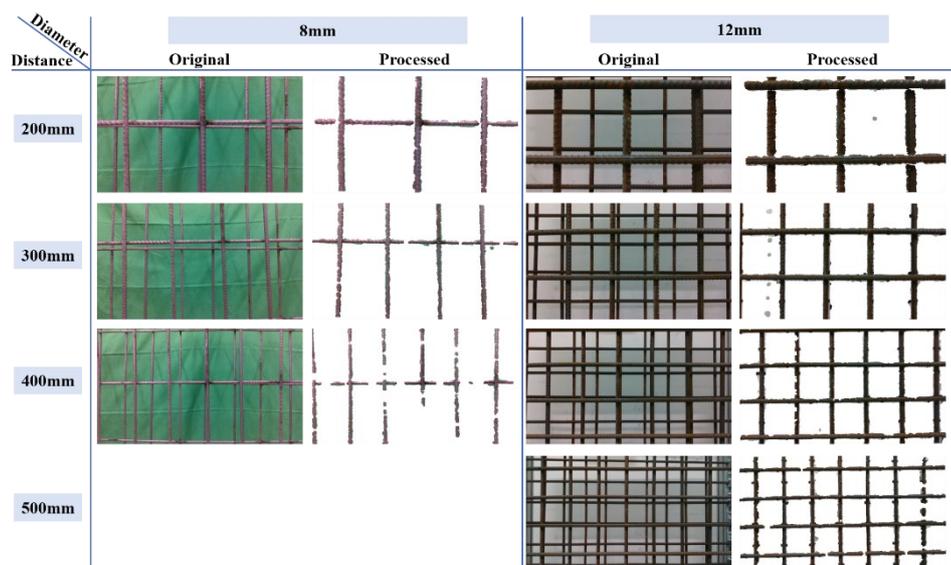


Figure 7. Segmentation effects for different distances and diameters

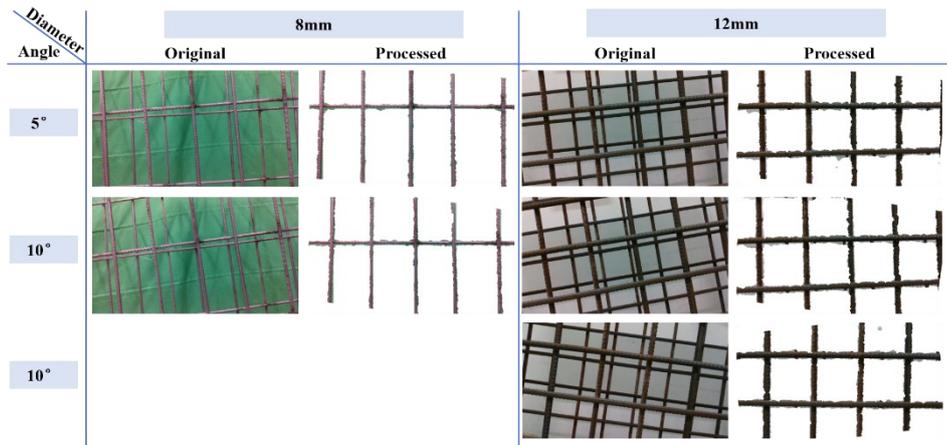


Figure 8. Segmentation effects for different rotation angles

Figure 8 shows the segmentation effects of the rebar at various angles of rotation in the camera plane, with the camera positioned 300mm away from the rebar plane. Here, the row-column pixel display method were employed for the 8mm rebar to better show the effects of plane rotation. It can be observed from Figure 8 that, for different rotation angles and diameters of rebars, the rebars were roughly parallel to the x and y axes after processing, which achieved considerable effects. However, pixel segmentation based on depth information exhibits some background point clouds at the rebar boundary. This may lead to errors in the process of fitting the rebar lines, consequently affecting the accurate calculation of the rotation angle. Therefore, having a clear rebar boundary is crucial for obtaining precise rotation angle.

Subsequently, the rebar crosspoint recognition and localization as well as the automatic spacing inspection methods were validated. After the images were processed by depth filtering and rotation, the trained YOLOv8-pose model was utilized for the prediction of rebar crosspoints, and their pixel coordinates were converted to 3D spatial coordinates relative to the camera coordinate system. Finally, the automatic spacing inspection module was utilized for rebar spacing inspection, with specific results shown in Figure 9. The errors between the rebar spacing inspection results and the manual measurement results are shown in Table 1, with an average error of 2.65 mm. Experimental results demonstrate the proposed method had high accuracy for rebar spacing inspection. By integrating these three modules, an end-to-end inspection process from image to result was achieved.

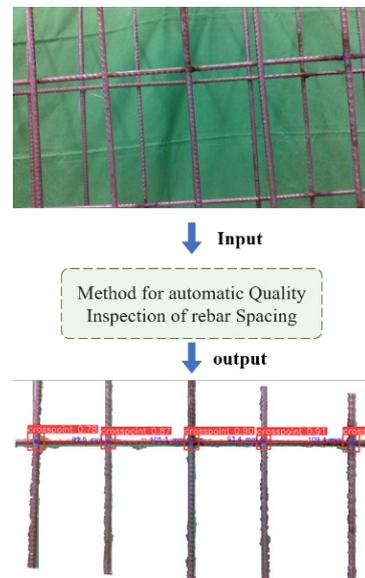


Figure 9. The method for inspecting rebar spacing

Table 1 Predicted rebar spacing and real results, and their errors in the image above

| | 1 | 2 | 3 | 4 |
|------------------|------|-------|------|-------|
| Prediction (mm) | 92.5 | 105.3 | 93.4 | 109.4 |
| Real result (mm) | 95 | 102 | 92 | 106 |
| Error (mm) | 2.5 | 3.3 | 1.4 | 3.4 |

4.2 Limitations

Rebar pixels segmentation based on depth information performs poorly in capturing small-diameter rebar. This is attributed to the challenge of accurately capturing depth information for small-diameter rebar through sensors. Additionally, factors such as ambient noise and lighting variations also introduce interference in the depth image, resulting in a few background pixels

being retained and making it difficult to clearly define the rebar pixel boundary.

In rebar spacing inspection, the limitation in the camera field of view result in each capture covering only a local region of the rebar layer. The challenge lies in effectively integrating and processing these local images.

5 Conclusion

In this paper, a method for automatic detection of rebar spacing quality is proposed, which uses a keypoint detection algorithm combined with RGBD camera. This method comprises three consecutive modules: (1) an image preprocessing module that addresses the challenge of recognizing double-layer bidirectional rebar in the image by filtering the background rebar layer; (2) a module for the recognition and localization of rebar crosspoints, obtaining their spatial coordinates; (3) an automatic inspection module enabling rebar spacing measurement in the image. The integration of these modules allows for an end-to-end inspection from image to result. Experiment testing on rebar cages in the laboratory demonstrates the efficacy of the image preprocessing module in segmenting and rotating the current rebar layer. Importantly, this module exhibits robustness without constraints related to camera distance and rotation angle. Furthermore, the average error of rebar spacing inspection is 2.65mm, which can be used for compliance inspection.

In future research, the proposed method will be validated in more complex scenarios, including construction sites and prefabrication plants. Additionally, advanced camera equipment and onboard platform will be utilized for autonomous image capture and inspection. For acceptance inspection, we will further refine its methodologies.

6 Acknowledgement

This work was supported by the National Natural Science Foundation of China (No. 52278177, No. 52308312, No. 52108136) and the National Key Research and Development Program of China (No. 2023YFC3806804).

References

- [1] Yuan X, Moreu F, Hojati M. Cost-effective inspection of rebar spacing and clearance using RGB-D sensors. *Sustainability*, 13(22): 12509, 2021.
- [2] Ministry of Construction of P.R. China. Code for acceptance of constructional quality of concrete structures: GB50204-2015. 2015.
- [3] Hwang B G, Thomas S R, Haas C T, et al. Measuring the impact of rework on construction cost performance. *Journal of construction engineering and management*, 135(3): 187-198, 2009.
- [4] Ma Z, Liu Y, Li J. Review on automated quality inspection of precast concrete components. *Automation in Construction*, 150: 104828, 2023.
- [5] Kim M K, Cheng J C, Sohn H, et al. A framework for dimensional and surface quality assessment of precast concrete elements using BIM and 3D laser scanning. *Automation in Construction*, 49: 225-238, 2015.
- [6] Kim M K, Wang Q, Park J W, et al. Automated dimensional quality assurance of full-scale precast concrete elements using laser scanning and BIM. *Automation in construction*, 72: 102-114, 2016.
- [7] Kim M K, Sohn H, Chang C C. Automated dimensional quality assessment of precast concrete panels using terrestrial laser scanning. *Automation in Construction*, 45: 163-177, 2014.
- [8] Yoon S, Wang Q, Sohn H. Optimal placement of precast bridge deck slabs with respect to precast girders using 3D laser scanning. *Automation in construction*, 86: 81-98, 2018.
- [9] Liu S L, Ma Z L. Automatic checking algorithm for the number and spacing of reinforcing bars based on point cloud. *Journal of Architecture and Civil Engineering*, 39(4): 90-99, 2022.
- [10] Wang Q, Cheng J C, Sohn H. Automated estimation of reinforced precast concrete rebar positions using colored laser scan data. *Computer - Aided Civil and Infrastructure Engineering*, 32(9): 787-802, 2017.
- [11] Yuan X, Smith A, Sarlo R, et al. Automatic evaluation of rebar spacing using LiDAR data. *Automation in Construction*, 131: 103890, 2021.
- [12] Lin F. Research on measurement technology of structural construction process based on machine vision [D], Southeast University, 2022.
- [13] Kardovskiy Y, Moon S. Artificial intelligence quality inspection of steel bars installation by integrating mask R-CNN and stereo vision. *Automation in Construction*, 130: 103850, 2021.
- [14] An M, Kang D S. The distance measurement based on corner detection for rebar spacing in engineering images. *The Journal of Supercomputing*, 78(10): 12380-12393, 2022.
- [15] Jin J, Zhang W, Li F, et al. Robotic binding of rebar based on active perception and planning. *Automation in Construction*, 132: 103939, 2021.
- [16] Jocher G. Ultralytics YOLOv8. 2023. On-line: <https://github.com/ultralytics/ultralytics>.
- [17] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection. *In proceedings of the IEEE conference on computer*

- vision and pattern recognition*, pages 779-788. 2016.
- [18] Wang C-Y, Liao H-Y M, Wu Y-H, et al. CSPNet: A new backbone that can enhance learning capability of CNN. *In proceedings of the Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, pages 390-391. 2020.
- [19] He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE transactions on pattern analysis and machine intelligence*, 37(9): 1904-1916, 2015.
- [20] Lin T-Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection. *In proceedings of the Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2117-2125. 2017.
- [21] Li H, Xiong P, An J, et al. Pyramid attention network for semantic segmentation. *arXiv preprint arXiv:180510180*, 2018.