Research on Autonomous Recognition and Gripping Method for Robotic Fabrication of Heterogeneous Masonry Based on Computer Vision

A full-stage validation experiment

Sijia Gu[1], Philip F. Yuan[2]*
[1,2]Tongji University
[1] 2230081@tongji.edu.cn [2] philip007@tongji.edu.cn

The emphasis on material diversity in robotic fabrication processes enhances the freedom of design in form and function, enabling the possibility of masonry working as functionally graded materials. However, in the robotic fabrication process based on offline programming, the lack of autonomous judgment of brick materials restricts the fabrication of multi-material masonry, resulting in additional labor and equipment costs. In this context, improving the autonomous judgment ability of construction robots on materials becomes an important breakthrough point, for which computer vision is a possible solution. However, current research on brick materials based on object detection mainly focuses on crack inspection and cannot distinguish multiple types of bricks in the same fabrication process. Therefore, the research aims to establish a methodology for an automatic multi-material brick grasping process based on the plane. The method consists of three parts: target detection, data conversion, and robot grasping. In this process, the research aims to innovate in four aspects: targets of object detection, derivation of dataset structure, introduction of design models, and real-world physical validation. Based on the proposal, a full-stage validation experiment was conducted. The experimental results validate the feasibility of the proposed method, hoping to bring new insights to robotic fabrication and parametric masonry design.

Keywords: *Robotic Fabrication, Heterogeneous Masonry, Computer Vision, Deep Learning.*

INTRODUCTION

In the research field of robotic fabrication, the exploration of materials has attracted considerable attention. Materials with diverse physical properties could be combined within the same fabrication process, ultimately yielding varied appearances and structures. Attention to the freedom of material properties can enhance the adaptability of constructions to structural, environmental, and design standards (Pajonk et al., 2022), thereby fostering richer architectural forms and functionalities. Meanwhile, it opens up possibilities for functionally graded materials (FGM) (Naebe and Shirvanimoghaddam, 2016).

The intervention of robotic fabrication processes endows the discrete assembly of materials with unprecedented potential. Materials can be freely combined according to machine encoding. By introducing variations of bounding forms and material properties, structures can undergo smooth transitions in uniform physical properties. In the field of robotic masonry (Bonwetsch et al., 2016), the research mainly focuses on forms of bricklaying with the same material. Bricks are rotated and translated, assembled into parametric brick masonry with unique appearances. On the contrary, rare practical robotic masonry projects are conducted in terms of diverse material properties.

The lack of such practice is partially due to the inflexibility of fabrication techniques, especially for robotic fabrication methods based on offline programming. Data of material distribution can be proposed in advance during the design stage. Designers can directly convert 3D design models into robotic tool paths using plugins like FURobot in Grasshopper. The relevant techniques have become increasingly mature and have demonstrated their potential in multiple real-world projects.

For multi-material heterogeneous brick masonries, the way of offline programming is theoretically feasible. However, in practice, the increase in the cost of labor and equipment creates obstacles to preset fixed points for the diverse bricks one by one. There are two solutions under the offline programming approach. One is to manually sort the bricks based on complex and unpredictable drawings and place them in a specified location. The second is to add the number of brick-feeding devices based on the amount of material. Either strategy results in additional labor, equipment, and space costs. Therefore, despite of its potential benefits, implementing multi-material brick masonry in practical engineering projects is not straightforward.

In this context, enhancing the autonomous judgment ability of construction robots regarding brick materials becomes a crucial breakthrough. In the field of computer vision, existing research on visual recognition provides possibilities for addressing this issue (Pathak et al., 2018). Visual recognition problems can be categorized into three types: image classification, object detection, and instance segmentation (Wu et al., 2020). Image classification aims to give the semantic category of the captured image. Object detection gives the location of the object in the form of a bounding box while predicting the image category. Semantic segmentation on the other hand aims at assigning specific category labels to each pixel. For the task of grasping multi-material bricks, the recognition system should be capable of both classification and location detection. That is to say, object detection becomes a good solution.

Regarding the implementation methods of object detection, previous studies tended to use traditional detectors such as Support Vector Machine (SVM). However, in recent years, deep learning (DL) methods have gained increasing attention. Unlike traditional machine learning methods, deep learning does not require manually designed features. It provides end-to-end classifiers that can learn features internally and automatically detect objects. And convolutional neural network (CNN) is one of the representative algorithms. Related studies have demonstrated its effectiveness in object detection (Zhao et al., 2015). Thanks to the powerful learning capabilities of CNN, some classic computer vision challenges can be reshaped into high-dimensional data transformation problems and addressed from different perspectives (Zhao et al., 2019). Furthermore, compared to traditionally designed visual descriptors, CNNs are more flexible. They can obtain better feature representations by transforming datasets and even develop detection results for different application scenarios.

In summary, deep learning-based computer vision methods provide possibilities for the robotic fabrication of multi-material heterogeneous brick masonry. More importantly, it will introduce a new paradigm under the context of future Industry 4.0 (Ramsgaard Thomsen et al., 2020), creating a production method with adaptability and flexibility in different situations.

LITERATURE REVIEW

To further explore the application of computer vision in the field of robotic bricklaying, a literature review was conducted. When using robots for bricklaying operations, the following four aspects are strongly emphasized as follow (Pritschow et al., 1996).

- Handling different types and specifications of bricks.
- Detecting and compensating for material tolerances.
- Calibrating brick positions relative to the Tool Center Point (TCP).
- Applying adhesive material spraying.
- Cost-effective solutions for specific production scenarios.

Vision systems based on deep learning are often used for material handling. Mika Litti's study introduced a classification framework for brick quality identification. This study classified bricks into three quality categories: intact, slightly defective, and unusable. A CNN-based solution was proposed, employing the pre-trained VGG-16 deep learning architecture (litti et al., 2021). The dataset was collected in the laboratory, consisting of images of the top, bottom, and side of 1105 bricks. The final accuracy ranged between 88% and 93%.

Similarly, a system capable of providing accurate segmentation masks for close-range images of brick walls was developed in another research. The system aims to classify materials in images of brick walls, removing mortar parts to represent bricks accurately. The study innovatively combined various machine learning algorithms and fused the separate segmentation outputs of eight classifiers using a weighted voting scheme (Kajatin and Nalpantidis, 2021).

Furthermore, a method for augmented reality (AR) bricklaying was proposed. In this method, the outlines of bricks are identified using point cloud scanning sensors. The centroid points of bricks under this operation method are input into the robot's operating system as end-effector grasp points (Song et al., 2023).

Semantic analysis of masonry wall images was also investigated as a topic. For obscured and damaged wall areas, researchers developed a fully automated method based on machine learning and a four-stage algorithm. The algorithm included three interactive deep neural networks and a contour extraction step based on watershed transformation (Ibrahim et al., 2020).

Moreover, the Faster R-CNN object detection architecture was used to detect cracks in brick images. In this architecture, researchers tested three networks—Mobilenetv2, Resnet50, and ZF512—under limited system resources and proposed a progressive detection method to improve robustness. The results showed that Mobilenetv2 achieved high accuracy and fast response speed. Mobilenetv2 reduced response speed by approximately 50% compared to Resnet50 while maintaining similar classification performance and outperforming ZF512 (Marin et al., 2021).

In addition to the considerable attention paid to the issue of brick quality, a few studies have focused on categorizing bricks. A study was conducted focused on distinguishing different types of bricks in masonry debris for reuse and recycling. The study used an image capture system to capture and process images of different bricks. Various artificial intelligence methods were applied to classify images in the study, including SVM, MLP, and k-NN classifiers. Classification results achieved recognition rates of 98% or higher. Additionally, the study aimed to test the acquired dataset using deep learning classification methods (Linß et al., 2023).

PROPOSAL

The literature research indicates that in the classification of brick materials, most studies focus on the quality of bricks, such as cracks. There is also research on classifying material types, but the materials are used in different scenarios rather than the same design and fabrication task. Additionally, in terms of classification methods, machine learning methods remain mainstream. There is still a question on how to leverage the adaptability advantages of deep learning methods in multiple scenarios. Furthermore, the limited resources of computer systems are also considered in the selection of algorithms, providing possibilities for the implementation of embedded systems. Based on the result of the literature review, the research aims to innovate in the following aspects.

- In terms of research content, focus on brick material types and propose a process method suitable for the design and fabrication of multi-material heterogeneous parametric brick masonries. The data exchange of the platforms involved will also be considered.
- In terms of research methods, the R3Det algorithm based on deep learning was adopted. The research also tries to propose a method to build a more effective and accurate dataset structure with a way of asymptotic derivation through experiments. Considering the limited system resources, the dataset will be designed from scratch and streamlined as much as possible. The results of this process will be summarized as dataset building experience.
- The data of 3D models, such as material order, is introduced into the recognition system, leading to an architectural masonry prototype-based process of identification and fabrication.
- Compared to previous software simulation-based experiments, a physical experiment based on a real UR10 robot is conducted to judge the accuracy and feasibility of the method.

The research aims to propose a computer vision-based automatic brick-grasping method for the robotic fabrication process of multi-material heterogeneous brick masonry. This method can be divided into three steps (Figure 1).

1. Establishment of a visual identification system. The system will have the ability to classify diverse materials and generate rotated check boxes for the bricks to pick.

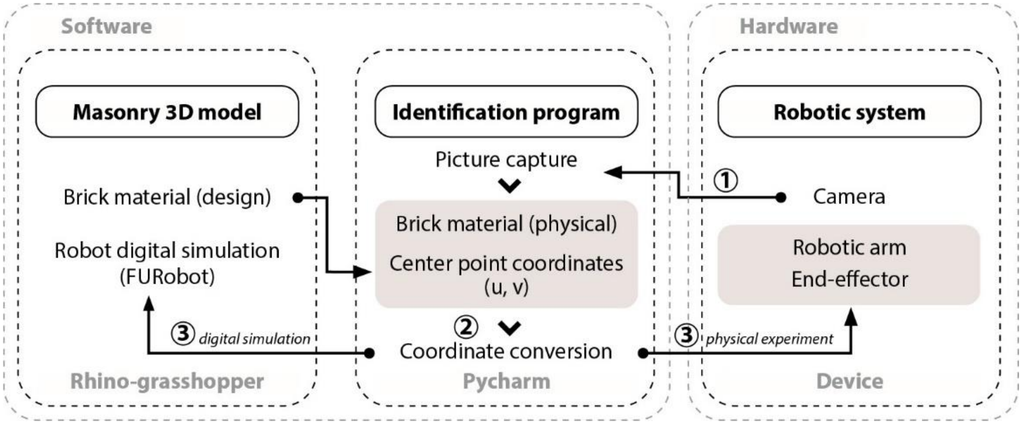

Figure 1
An illustration of the proposal

2. Convert target detection results. The most important thing is to convert the pixel coordinates of the recognized object to the 6-dof robot base coordinates through camera calibration and matrix calculation. How to minimize the error in this process also needs to be considered.

3. Transmission of the generated positional data to the robot system. According to the obtained coordinates, the robot conducts the grasping action. The classification results from the previous stage will be used in this process to compare with the material order code from 3D design model to determine whether the robot grasps or not. The final grasping action can be done in a digital simulation environment or a physical environment.

EXPERIMENT

The experiment is divided into three stages: recognition system building, data transformation, and robot grasping test.

Recognition system

The recognition experiment consists of dataset setup, recognition framework configuration, model training, and error testing.

Dataset setup. For deep learning-based object detection models, the quality of the dataset determines the number of recognized categories and the accuracy of the results. To maximize the number of recognized categories and quantities within limited computational resources, a multi-stage dataset experiment was conducted. In the experiment, bricks were grabbed in a laboratory scenario. Based on image features, data samples were categorized. Different batches of data samples were used for training and validation, and the trained weight files were used to test the same batch of photos to display recognition results.

To collect data samples, the experiment collected 10 red bricks and 10 gray bricks. The bricks were placed in different contexts, and their quantity and relative positions varied. The number of bricks, shooting height, and lighting conditions also varied. In the end, 92 original images were obtained and processed into 276 standard images using a Python algorithm. These images were divided into 4 groups, with each group of images continuously accumulating based on the previous group, including monochrome single-brick simple scene (data 1, 7 images), bicolor single-brick simple scene (data 2, 11 images), bicolor multi-brick simple scene (data 3, 21 images), bicolor single-brick complex scene (data 4, 30 images), bicolor multi-brick complex scene (data 5, 92 images), and lighting variation scene based on data 4 (data 6, 276 images).

The RoLabelImg annotation software was used in the experiment. This software records the planar position and category of objects in each photo with bounding boxes and labels (Figure 2). The planar position is recorded as five values, including the x-coordinate of the center point (cx), the y-coordinate of the center point (cy), the width of the image box (w), the height of the image box (h), and the rotation angle (angle). Finally, XML annotation files were obtained. In these files, bricks that were not fully photographed, not on the horizontal plane, or difficult to extract were not labeled.

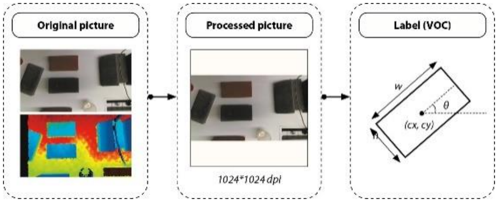

Figure 2
An illustration of the data sample calibration process

Recognition algorithm. The focus of the research is on bridging the gap between the design model and the robotic arm grasping through computer vision recognition methods, rather than improving the visual recognition algorithms. Therefore, products from MMLab are adopted as a proven algorithmic framework. MMDetection and MMRotate are successively used for the recognition of brick types and plane positions. Since MMDetection does not have the function of detecting rotation angles, the MMRotate framework is used for the final dataset construction and physical grasping experiments.

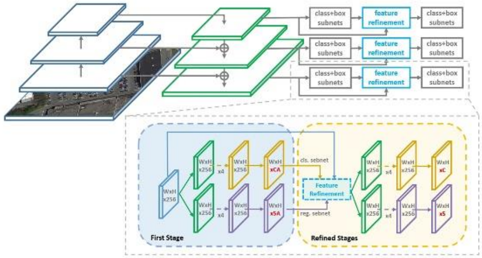

Figure 3
The architecture of the proposed Refined Rotation Single-Stage Detector (Yang et al., 2021)

MMRotate (Zhou et al., n.d.) is an open-source toolbox for rotation box detection based on PyTorch, providing various model libraries (Figure 3). In this study, the r3det_tiny_r50_fpn_1x_dota_oc.py configuration file was used. This file employs the

R3Det algorithm (Yang et al., 2021), which is a fine single-stage detector for rotated object detection. The main components of the model include ResNet50 (serving as the backbone network for feature extraction), FPN (feature pyramid network used for feature extraction at different scales), and rotated retina head (used for generating rotated anchors and performing classification and regression).

Training and testing. The existing datasets from data 0 to data 5 were partially utilized as training data, with another portion designated as validation data. Additionally, for each dataset category, 5 images with distinct features were captured to serve as the testing data after training completion. These images were processed into three different brightness levels. Consequently, the final testing data comprised 15 images, labeled sequentially as 1-1, 1-2, 1-3, 2-1, 2-2, 2-3, 3-1, 3-2, 3-3, 4-1, 4-2, 4-3, 5-1, 5-2, and 5-3. The Intel RealSense D415 camera integrated into the end-effector of the UR 10 robot was employed to capture the original testing images. Simultaneously, the robot poses corresponding to the validation images were recorded for repeatability in experiments.

The training process employed CUDA 11.6 and PyTorch 1.13.1. With six datasets, a total of 6 .pth files were obtained for validating the algorithm's accuracy. The final validation results are illustrated (Figure 4), with their accuracies documented (Table 1, Figure 5). Categories that were inaccurately recognized were assigned a precision of 0. This approach might result in an overall lower accuracy when compared computationally. However, the differences among datasets are notably discernible. The conclusions are as follow.

- As the data samples incorporating variations in brick color, quantity, environmental complexity, and lighting conditions were added, the recognition rates and stability increased for the same testing data. The improvement in dataset structure was positively correlated with recognition rates and stability, serving as a reference for future dataset structuring in research.

- Increasing the number of brick samples not only enhanced recognition rates in images with multiple bricks but also improved recognition rates in single-brick images. Conversely, using only single-brick images as samples may not yield satisfactory recognition rates in both single-brick and multiple-brick images, albeit depending on the sample size of the dataset.
- While the recognition program demonstrated near-perfect recognition in images with moderate to high brightness, achieving accurate recognition, both in terms of category and detection box, for images with lower brightness proved challenging. However, analyzing the recognition rates of images 1-3 to 5-3 reveals that data samples with complex lighting conditions can ameliorate this issue. Nevertheless, achieving similar recognition rates for darker images as for brighter ones would require a larger number of data samples. Simply adjusting the brightness of the same data is insufficient.

Table 1
The results of the recognition rate obtained by the models training with 6 datasets respectively

Dataset	Min-max	Average	Variance
1	0.00-0.00	0.00	0.00
2	0.16-0.34	0.11	0.02
3	0.11-0.96	0.61	0.08
4	0.10-0.89	0.41	0.06
5	0.48-1.00	0.70	0.07
6	0.66-1.00	0.87	0.02

Data transmission
Due to the eventual transmission of data for executing pick-and-place actions by the robotic arm, the image coordinates in the recognition results need to be transformed into the robotic arm base coordinates. In this process, the camera coordinates need to be calibrated to determine the relative position between the robotic arm's TCP (Tool Center Point) and the camera (Figure 6).

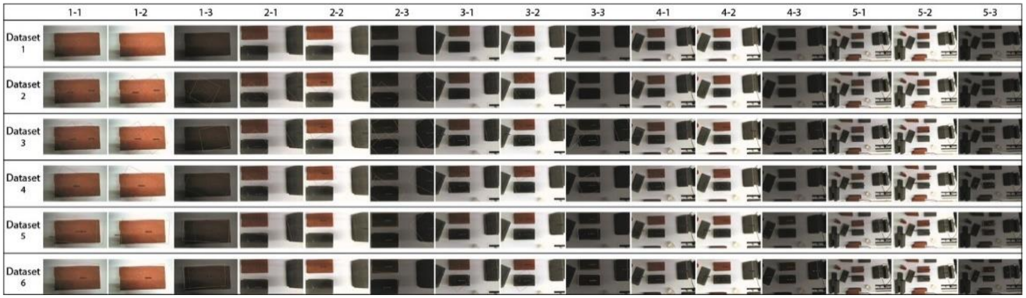

Figure 4
The results of the detection frame obtained by the models training with 6 datasets respectively

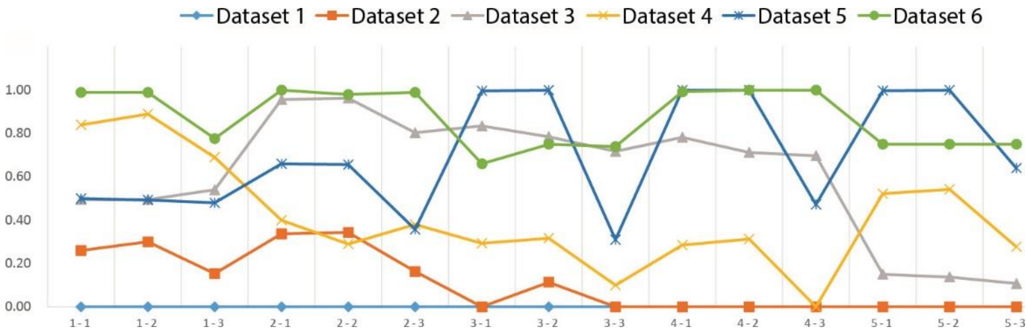

Figure 5
The results of the recognition rate obtained by the models training with 6 datasets respectively

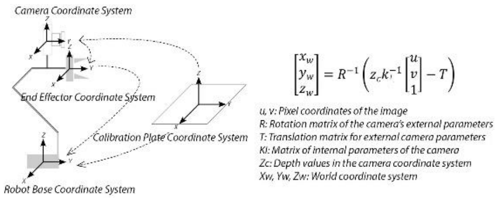

Figure 6
An illustration of the conversion process and formulas for coordinates

Figure 7
Pictures of the camera calibration process

The chessboard calibration method was employed (Figure 7). The study captured 22 images from different angles and heights using an Intel RealSense D415 camera. The chessboard in the images had dimensions of 10*7 squares. Four coordinate systems were involved in the calibration process: the robotic arm base coordinate system, the end effector coordinate system, the camera coordinate system, and the calibration board coordinate system. Since the relative position between the robotic arm base and the calibration board remains unchanged, the transformation matrix of the camera relative to the TCP can be obtained by varying the end effector coordinate system and the camera coordinate system. This constitutes the extrinsic matrix of the camera, which consists of a 3*3 rotation matrix with a 3*1 translation matrix.

Once the extrinsic matrix of the camera is obtained, it can be combined with the camera's intrinsic parameters to perform the transformation from pixel coordinates in the recognition image to the coordinates of the robotic arm's end effector. This transformation can be understood sequentially as two processes: from the pixel coordinate system to the camera coordinate system and from the camera coordinate system to the world coordinate system. In the process, the OpenCV python package is called in pycharm.

Phygital grasping experiment

The robotic arm grasping experiment is divided into two categories: digital simulation and physical grasping. Digital simulation experiments are primarily conducted to compare the three-dimensional model with the recognition classification results. Recognition is performed in Pycharm, and upon completion, the material information (label ID) and position information (coordinates) of the bricks are transmitted to Rhino's Grasshopper platform via TCP/IP protocol using local socket communication (Figure 8).

Parametric masonry models were designed with Grasshopper (Figure 9, Figure 10). Designers need to pre-write corresponding Grasshopper entries according to the labels in the dataset, and utilize the advantages of Grasshopper's parametric modeling to classify the brick materials based on the label content. After each recognition, the label information of the physical brick is compared with the label information of the 3D model. If they match, the coordinates obtained from the recognition are used for grasping. The FURobot plugin is used as the robotic simulation platform in Grasshopper.

Physical experiments are mainly conducted to prepare for actual production test the errors in the recognition results and coordinate the transformation process. The RTDE interface is used again. The robot arm is connected to the personal computer with the recognition program via Ethernet cable. The coordinates of the object relative to the robot base are transmitted to PolyScope and executed by the robot. Through measurements, the error between the center point of the robotic arm end effector and the actual center point of the brick is approximately 5mm, which is considered an acceptable result. Comparing the images of the recognition results, the accuracy of the recognition results becomes the main source of error, while the calibration process has a smaller error (Figure 11).

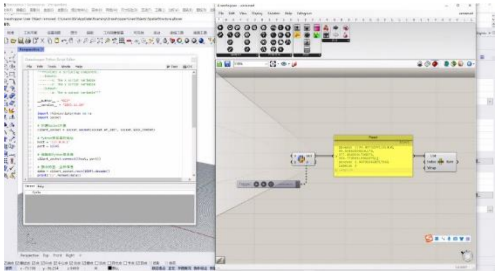

Figure 8
Establishment of socket communication between grasshopper and pycharm with the detection result data

Figure 9
A rendering image and pictures of multi-material parametric masonry

Figure 10
Multi-material parametric masonry prototypes

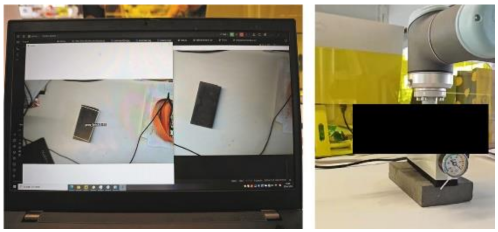

Figure 11
Results of the UR10 robot physical grasping experiment

CONCLUSION AND DISCUSSION

In the same robotic fabrication process, the focus on different materials provides the possibility for the implementation of functionally graded materials and increases the diversity of fabrication forms and adaptability to the environment and functions. However, in the bricklaying mode based on offline programming, the increase in brick types will bring additional costs in manpower and equipment, thereby challenging the fabrication process.

A robot bricklaying method based on computer vision active recognition is proposed in the study, aiming to reduce the cost and fabrication difficulty of multi-material heterogeneous bricks through the method of active recognition by the machine. To this end, a workflow method consisting of three stages—model training, data transformation, and robot grasping—is proposed and validated through experiments. Firstly, a homemade dataset is used to train the recognition model, which obtains information about brick materials and their pixel-represented positions. Secondly, the pixel coordinates of the brick center points are transformed into world coordinates through marker calibration and matrix calculations. Finally, robot simulation and actual grasping experiments are conducted.

Through experimental verification, the computer vision-based deep learning brick grasping method shows feasibility in the design and fabrication of heterogeneous parametric masonries, enabling the recognition and actual grasping of brick materials within an acceptable range of error. The study obtains dataset architecture for brick recognition and general data regularities through progressive dataset construction. In addition, the study obtains the dataset architecture and general data laws for brick recognition by documenting the progressive dataset building process, and successfully introduces the model data into the digital simulation stage.

However, the experiments also revealed some issues that need improvement in subsequent stages. Due to the insufficient quantity of datasets and limitations of equipment, although progress has been made in recognition rate under the current dataset structure, errors still exist in classification and rotation angle detection. On the other hand, the performance of the current recognition results in building brick wall prototypes indicates the need for a stable grasping feedback mechanism. In addition, obstacle avoidance issues also need to be addressed.

ACKNOWLEDGMENT

This work is supported by the Shanghai Science and Technology Committee (Grant No. 21DZ1204500).

REFERENCES

Bonwetsch, T., Willmann, J., Gramazio, F., Kohler, M., 2016. Robotic Brickwork: Towards a New Paradigm of the Automatic. Bricks/Systems 51.

Ibrahim, Y., Nagy, B., Benedek, C., 2020. Deep Learning-Based Masonry Wall Image Analysis. Remote Sens. 12, 3918. https://doi.org/10.3390/rs12233918

Iitti, M., n.d. BRICK CLASSIFICATION USING A CONVOLUTIONAL NEURAL NETWORK.

Iitti, M., Gronman, J., Turunen, J., Lipping, T., 2021. Classification of Masonry Bricks Using Convolutional Neural Networks – a Case Study in a University-Industry Collaboration Project, in: 2021 IEEE International Conference on Progress in Informatics and Computing (PIC). Presented at the 2021 IEEE International Conference on Progress in Informatics and Computing (PIC), IEEE, Shanghai, China, pp. 125–129. https://doi.org/10.1109/PIC53636.2021.9687077

Kajatin, R., Nalpantidis, L., 2021. Image Segmentation of Bricks in Masonry Wall Using a Fusion of Machine Learning Algorithms, in: Del Bimbo, A., Cucchiara, R., Sclaroff, S., Farinella, G.M., Mei, T., Bertini, M., Escalante, H.J., Vezzani, R. (Eds.), Pattern Recognition. ICPR International Workshops and Challenges, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 446–461. https://doi.org/10.1007/978-3-030-68787-8_33

Linß, E., Walz, J., Könke, C., 2023. Image analysis for the sorting of brick and masonry waste using machine learning methods. Acta IMEKO 12, 1–5. https://doi.org/10.21014/actaimeko.v12i2.1325

Marin, B., Brown, K., Erden, M.S., 2021. Automated Masonry crack detection with Faster R-CNN, in: 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE). Presented at the 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE), IEEE, Lyon, France, pp. 333–340. https://doi.org/10.1109/CASE49439.2021.9551683

Naebe, M., Shirvanimoghaddam, K., 2016. Functionally graded materials: A review of fabrication and properties. Appl. Mater. Today 5, 223–245. https://doi.org/10.1016/j.apmt.2016.10.001

Pajonk, A., Prieto, A., Blum, U., Knaack, U., 2022. Multi-material additive manufacturing in architecture and construction: A review. J. Build. Eng. 45, 103603. https://doi.org/10.1016/j.jobe.2021.103603

Pathak, A.R., Pandey, M., Rautaray, S., 2018. Application of Deep Learning for Object Detection. Procedia Comput. Sci. 132, 1706–1717. https://doi.org/10.1016/j.procs.2018.05.144

Pritschow, G., Dalacker, M., Kurz, J., Gaenssle, M., 1996. Technological aspects in the development of a mobile bricklaying robot. Autom. Constr. 5, 3–13. https://doi.org/10.1016/0926-5805(95)00015-1

Ramsgaard Thomsen, M., Nicholas, P., Tamke, M., Gatz, S., Sinke, Y., Rossi, G., 2020. Towards machine learning for architectural fabrication in the age of industry 4.0. Int. J. Archit. Comput. 18, 335–352. https://doi.org/10.1177/1478077120948000

Song, Y., Koeck, R., Agkathidis, A., 2023. Augmented Bricklayer: an augmented human-robot collaboration method for the robotic assembly of masonry structures, in: Blucher Design Proceedings. Presented at the XXVI International Conference of the Iberoamerican Society of Digital Graphics, Editora Blucher, Lima, Peru, pp. 713–724. https://doi.org/10.5151/sigradi2022-sigradi2022_30

Wu, X., Sahoo, D., Hoi, S.C.H., 2020. Recent advances in deep learning for object detection. Neurocomputing 396, 39–64. https://doi.org/10.1016/j.neucom.2020.01.085

Yang, X., Yan, J., Feng, Z., He, T., 2021. R3Det: Refined Single-Stage Detector with Feature Refinement for Rotating Object, in: Proceedings of the AAAI Conference on Artificial Intelligence. pp. 3163–3171.

Zhao, Z.-Q., Xie, B.-J., Cheung, Y., Wu, X., 2015. Plant Leaf Identification via a Growing Convolution Neural Network with Progressive Sample Learning, in: Cremers, D., Reid, I., Saito, H., Yang, M.-H. (Eds.), Computer Vision -- ACCV 2014, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 348–361. https://doi.org/10.1007/978-3-319-16808-1_24

Zhao, Z.-Q., Zheng, P., Xu, S.-T., Wu, X., 2019. Object Detection With Deep Learning: A Review. IEEE Trans. Neural Netw. Learn. Syst. 30, 3212–3232. https://doi.org/10.1109/TNNLS.2018.2876865

Zhou, Y., Yang, X., Zhang, G., Wang, J., Liu, Y., Hou, L., Jiang, X., Liu, X., Yan, J., Lyu, C., Zhang, W., Chen, K., n.d. MMRotate: A Rotated Object Detection Benchmark using PyTorch. ArXiv Prepr. ArXiv220413317.