ISSN: 2472-3703 DOI: 10.6911/WSRJ.202307_9(7).0014

Two-stage Transfer learning based on Improved Mask R-CNN for Microscopic Primitive Hole Identification and Segmentation in Aluminum Alloy Materials

Xiaochen Zhang[1], Q. C. Kanhukamwe[2], Zinan Wang[1], Weiying Meng[1] and T. Garikayi[2]

[1]School of mechanical engineering, Shenyang Jianzhu University, Shenyang, Liaoning Province, China

[2]Harare Institute of Technology, Harare, The Republic of Zimbabwe

Abstract

Aluminum alloy is increasingly used in important fields such as aerospace equipment, marine engineering equipment, advanced rail transportation equipment, etc. The current method, using manual experience sampling inspection is low in efficiency and high in cost. In this paper, we propose an improved Mask RCNN convolutional neural network structure to achieve recognition segmentation of microscopic raw holes in aluminum alloys. In addition, we propose a new two-stage migration learning method to address the lack of training data, which utilizes the knowledge learned from sufficient feature source data and auxiliary unlabeled aluminum alloy microscopic raw hole data to migrate to the target domain. The experimental results show that our proposed network architecture achieves the desired performance with an IoU index up to 0.743. The results outperform most advanced convolutional neural network structures by both comparison with other networks and field tests, verifying the effectiveness and accuracy of aluminum alloy microscopic raw hole defect detection and demonstrating the feasibility and advantages of the proposed two-stage migration learning strategy.

Keywords

Aluminum microscopic primitive holes; Mask RCNN recognition segmentation; Two-stage migration learning.

1. INTRODUCTION

Lightweight design is the main development trend of structural design of mechanical equipment and its components today. Aluminum alloys, with their advantages of light weight, high strength and good corrosion resistance, are being used more and more widely in important fields such as aerospace equipment, marine engineering equipment and advanced rail transportation equipment. Due to the interaction of material organization during solidification of aluminum alloy castings, such as hydrogen solubility changes, oxide film involvement, pressure changes in the paste zone, inter-dendrite complementary shrinkage, solidification volume shrinkage, and development of solid phase organization[1]2], resulting in aluminum alloy castings produces the microscopic porosity defects, mainly porosity and shrinkage, seriously reduce the air tightness and mechanical properties of the castings.

A relatively mature system has been formed for the qualitative study of microscopic pores. Due to the complexity of the mechanism of microscopic hole formation in aluminum alloy castings, the understanding is still not very thorough, which is a major obstacle to the development of numerical models for predicting microscopic hole formation[1][2]. Recently,

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202307_9(7).0014

deep learning is rapidly changing many fields, solving many problems and providing new ideas for solutions with its ability to solve complex tasks autonomously. In particular, convolutional neural networks are able to automatically learn distinctive features and exhibit better robustness, which also enables them to achieve better performance in image analysis. Kim. et al.[3] proposed a new method for identifying bridge cracks using a drone combined with a VGG convolutional neural network model. The model was trained on a dataset consisting of about 20,000 cracked images and 20,000 non-cracked images, and the recognition accuracy was up to 95%, which provided good results. However, the dataset is large and the model can only identify whether the images are cracked images and cannot locate crack defects.

2. RELATED WORK

In recent years, deep learning techniques, especially convolutional neural network techniques, have proven to be very successful in the field of computer vision. Compared with traditional hand-described methods, neural networks can efficiently learn more intrinsically informative features from training data. CNNs can be generalized to solve different types of problems using similar designs and have a significant improvement in detection accuracy, which makes it popular in many applications. It is widely used in agriculture, transportation, and architectural image segmentation. In agriculture: Lin et al. [7] used Mask RCNN to classify rice lice and trained data on ResNet50 framework by migration learning to achieve effective and fast recognition of rice lice and non-rice lice with an average recognition accuracy of 0.923; Chen et al. [8] utilized MobileNet-V2 pre-trained on Image Net as the backbone network with an added attention mechanism and, at the same time, optimized the loss function. An average recognition accuracy of 99.67% was achieved on a public dataset compared to other state-of-the-art methods. Even under complex background conditions, the average accuracy of identifying rice plant diseases reached 98.48, proving the better performance of the method.

3. METHOD

3.1. Improved mask RCNN

Based on the faster-trained RCNN framework, mask RCNN adds a parallel semantic segmentation branch for target detection and regression. In this model, a feature pyramid network (FPN) + ResNet101 is used to extract the features of the backbone network. After the features are extracted by this model, the microscopic primitive hole target detection framework is trained end-to-end using a region proposal network. Dilated convolution is also involved in the feature computation and finally RoI (region of interest) alignment is used to address the problem of large actual bias in the original image.

For a given boundary (Px, Py, Pw, Ph), the final regression boundary (Fx, Fy, Fw, Fh) is obtained using the objective boundary regression and made closer to the true boundary (Gx, Gy, Gw, Gh). That is, we need to find a mapping f such that f(Px, Py, Pw, Ph) = (Fx, Fy, Fw, Fh) and (Fx, Fy, Fw, Fh) ≈ (Gx, Gy, Gw, Gh), where the subscripts x, y, w, and h denote the horizontal distance, vertical distance, width, and height of the three types of boundaries at the center point, respectively. The boundary regression learns these transformations and calculates the translation (tx, ty) and scaling (tw, th) based on the parameters of the real and predicted boundaries. The calculations are as follows:

$$t_x = \frac{(G_x - P_x)}{P_w} \tag{1}$$

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202307_9(7).0014

$$t_y = \frac{(G_y - P_y)}{P_h} \tag{2}$$

$$t_w = \log_2^{\frac{G_w}{P_w}} \tag{3}$$

$$t_h = \log_2^{\frac{G_h}{P_h}} \tag{4}$$

3.2. Two-stage transfer learning

A common migration learning approach is single-stage migration, where pre-training is performed on the source domain and then fine-tuning is performed directly on the target dataset. On top of this, with the thought of performing traditional transfer learning methods in addition, we tried to explore the effectiveness of multi-stage migration for the microscopic raw hole classification task. As mentioned earlier intermediate auxiliary data has the potential for migration learning. Therefore, instead of directly transferring the pre-trained model to the target domain, it is first fine-tuned in the intermediate domain. The whole framework of the proposed two-stage transfer learning is shown in Fig. 1. The key element of the two-stage transfer approach is to learn knowledge from the intermediate domain dataset. Since the intermediate auxiliary dataset has no label information, we consider training the network in an unsupervised manner.

4. EXPERIMENTS AND RESULTS

Images of concrete ground holes made from publicly available were used as source domains. Considering the limited number of samples in this dataset, we performed manual data enhancement. Finally, we obtained a total of 2000 image patches for training. All instances were sized to the original size of the concrete ground hole image of 227 × 227. The input size of the latter proposed network was matched.

4.1. Target domain dataset

We used scanning electron microscopy to cut and scan the original aluminum alloy specimens without fatigue test to view the microscopic original hole images of aluminum alloy and to construct a target domain database. The database was collected on the scanning electron microscope equipment at the Analysis and Testing Center of Shenyang University of Architecture. Each image has a pixel size of 640 × 480 and a pixel pitch of 0.2 μm. The database was also labeled by experienced mechanical engineering faculty as well as students using polygons to label the images and depict the microscopic raw pore morphology of aluminum alloys.

We note that the annotated area of the original pore in each image fills only a small portion of the entire original image area, and that there are large areas in the image with similar morphology to the microscopic original pore pattern of the aluminum alloy under study. Fig. 5(b) shows one of the examples. Thus, the data extracted from these regions can be used to achieve better migration learning. For the intermediate domain data as a secondary role, we extracted 227 × 227 non-overlapping images from non-labeled regions in different real SEM images. Finally, we obtained a total of 3000 patches. Fig. 1 shows some examples of these three datasets. We found that many patch features in the intermediate domain dataset have high similarity to the target domain, which motivated us to use these data for two-stage migration learning.

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202307_9(7).0014

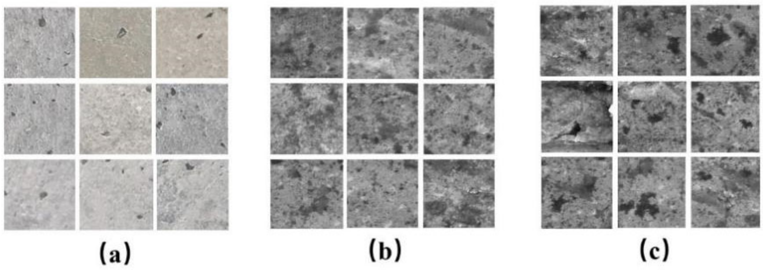

Figure 1. Picture of the dataset: (a) source domain (b) intermediate domain (c) target domain

In this study, our goal is to obtain a high-precision aluminum alloy microscopic primitive hole recognition segmentation task with limited training samples. Accordingly, we design a new Mask RCNN architecture and propose a new two-stage migration learning strategy.

The first important problem analyzed is the use of Mask-RCNN, considered as a combination of fast RCNN (class + enclosing box) and full convolutional network (FCN) for object detection. This in which due to the region of interest pooling layer in the fast region neural network may lead to misalignment of the target objects' positions in the feature map with their positions in the original image, which is solved by refining the masks using the region of interest alignment layer. The region of interest alignment is a quantization-free layer that preserves the exact spatial location by replacing quantization with bilinear interpolation, which is used to compute the floating point location values in the input.

The proposed method recognizes or segments images from a pixel perspective, but the performance of the traditional Canny algorithm in particular is significantly affected when the images contain similar features in the background. The experimental results also verify the poor actual results of Canny and there is little difference between the recognized images of other comparison networks and the proposed model by visual observation, but the proposed model provides better performance for detailed contour and edge information recognition segmentation. In fact, the features of an image should include overall level (classification), region level (object detection) and pixel level (semantic segmentation). If these three levels are combined, features can be accurately extracted to detect and segment defects in images. In this experiment, the proposed recognition segmentation network has no advantage in terms of inference speed, which may be due to some feature fusion modules used to maintain details and increase the number of parameters and computations. In fact, the features extracted by the convolutional network are not valid for all channels, which means that the channel features have some redundancy. Therefore, pruning the redundant channels can greatly reduce the number of parameters and improve the inference speed. In future work, further research will be conducted in this area.

5. CONCLUSION

Microscopic pristine pores of aluminum alloy include air holes, shrinkage holes and gas shrinkage holes, which usually appear inside the specimen during manufacturing solidification. Based on the improved Mask RCNN and migration learning strategy, a microscopic pristine hole identification and segmentation model for aluminum alloy materials is proposed. In this paper, two-stage migration learning is performed using three sets of data, where the data set contains source, intermediate, and target domains. To collect correct and detailed information, the images are cut into small pieces of 227 × 227 pixels size and the data are used in 8:1:1 ratio for training, validation and testing. For issues such as small amount of training data and noise, a pre-processing model and data enhancement method with a macroscopic image dataset of

ISSN: 2472-3703 DOI: 10.6911/WSRJ.202307_9(7).0014

concrete holes and a real aluminum alloy microscopic raw hole unlabeled SEM image dataset were used. The experimental and test results validate the effectiveness and accuracy of the proposed microscopic primitive hole identification and segmentation model for aluminum alloy materials. However, it is also necessary to achieve optimization in terms of inference time. Therefore, more optimization work on the segmentation network (Deep LinkNet) and post-processing method (CRF) not covered in this paper is needed in future research. In addition, to enrich the dataset, adversarial neural networks can be used to generate more training samples.

ACKNOWLEDGMENTS

This study was funded by : National Natural Science Foundation of China (Grant No. 52205163, 52005352), General project of Liaoning Provincial Natural Science Foundation (2022-MS-280), General Projects of Basic Scientific Research Projects in Colleges and Universities for Liaoning Provincial Education Department (LJKMZ20220932), Open Fund of Key Laboratory of Fundamental Science for National Defense of Aeronautical Digital Manufacturing Process of Shenyang Aerospace University (SHSYS202107).

REFERENCES

[1] Zhang B, Cockcroft S L, Maijer D M, et al. Casting defects in low-pressure die-cast aluminum alloy wheels[J]. Jom, 2005, 57: 36-43.

[2] Yao L. Experimental investigation and numerical modeling of microporosity formation in aluminum alloy A356[D]. University of British Columbia, 2011.

[3] Kim I H, Jeon H, Baek S C, et al. Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle[J]. Sensors, 2018, 18(6): 1881.

[4] Zhuang F, Qi Z, Duan K, et al. A comprehensive survey on transfer learning[J]. Proceedings of the IEEE, 2020, 109(1): 43-76.

[5] Liu Z, Cao Y, Wang Y, et al. Computer vision-based concrete crack detection using U-net fully convolutional networks[J]. Automation in Construction, 2019, 104: 129-139.

[6] He K, Gkioxari G, Dollár P, et al. Mask r-cnn[C]//Proceedings of the IEEE international conference on computer vision. 2017: 2961-2969.

[7] Lin X, Zhu S, Zhang J, et al. Rice Planthopper Image Classification Method Based on Transfer Learning and Mask R-CNN[J]. Transactions of the Chinese Society for Agricultural Machinery, 2019, 50(07):201-207.

[8] Chen J, Zhang D, Zeb A, et al. Identification of rice plant diseases using lightweight attention networks[J]. Expert Systems with Applications, 2021, 169: 114514.