Graphite Classification of Gray Cast Iron in Metallographic via A Deep Learning Approach

Wesley Huang[1*], Zhi-Yuan Su[2], Chia-Sui Wang[1], Mark Yeh[3], Jyh-Horng Chou[4]

[1] Department of Information Management, Chia Nan University of Pharmacy and Science, Taiwan
[2] Department of Multimedia and Game Development, Chia Nan University of Pharmacy and Science, Taiwan
[3] Department of Electrical Engineering, National Sun Yat-sen University, Taiwan
[4] Department of Mechanical and Computer-Aided Engineering, Feng Chia University, Taiwan
wasener@mail.cnu.edu.tw, zysu@mail.cnu.edu.tw, box1025@mail.cnu.edu.tw, mark3581923@gmail.com,
choujhtaiwan@gmail.com

Abstract

In addition to measurements of physical and mechanical properties, quality inspections also include metallographic analyses. When gray casting iron material, different manufacturing processes cause different microstructures in the material, whose metallographic images also perform large differences. The metallographic properties of gray iron can be divided into six types (from Type A to Type F). The proportion of types will influence the strength, wear resistance, and lifetime of specimens. The determination of type is usually dependent on manual judgments. In this study, two approaches were developed to analyze six metallographic types of gray casting iron. The first approach was to determine the type according to features of the detected particles in the metallographic materials by morphology algorithm. Types A, C, and F could be identified with the shape factor (SF) of gray casting iron. Then, the remained part could be identified using average grayscale values of the part-region of the metallographic material. Second approach was to identify Types A, C, and F with SF method and then identify the remaining part through the classification of the YOLO V3 deep learning algorithm. The results showed that the second approach performed more suitably in identifying the types of metallographic of gray casting iron.

Keywords: Feature recognition, Morphology, Metallographic, Gray casting iron, Deep learning networks

1 Introduction

Gray casting iron is usually manufactured for equipment bases, braking discs or drums for transportation vehicles, and the bases of mother machines. The quality of materials is very much in demand on tool life and mechanical properties. Six types of gray casting iron have been determined, according to the distribution and shape of graphite particles [1]. Type A has a random flake graphite shape in a uniform distribution. Type B has a rosette flake graphite shape. Type C has a Kish graphite shape (i.e., hyper-eutectic compositions). Type D has an undercooled flake graphite shape. Type E has an interdendritic flake graphite shape (i.e., hypo-eutectic compositions). Type F has new-born start graphite (i.e., spider-like graphite). In ideal manufacturing conditions, the proportion of type A should be over 90% in the material to provide more strength and wear resistance. For the reason mentioned above, investigating the proportion of each type has become more and more important.

Inspections of steel materials are based on identifying failures as well as general metallographic analyses. Some applications concern wear volume calculation and tool life prediction. Papa and Albuquerque used digital image inspections on gray iron and ductile casting iron to estimate the graphite-ferrite-pearlite proportions using two different approaches [2-3]. In the research of Papa, three machine learning approaches were used to train and validate the data sets, including support vector machines (SVM), Bayesian networks, and optimum-path forest-based classifiers (OPF). The performance used was x2 in McNemar's test [3]. The Otsu method was used for binarization to determine the threshold value of the image. The results showed that OPF had the best performance. In the research of Albuquerque, the neural networks included multilayer perceptron neural networks (MLP) as well as self-organizing map neural networks (SOM). Both algorithms were applied to calculate the graphite-ferrite-pearlite proportions in both gray and ductile casting iron materials. The results showed that MLP had better performance.

Recently, metallographic images analyzed by digital artificial intelligence approaches have been issued. The majority of the research has been carried out with machine learning approaches. Few issues discussed the deep learning algorithm and other approaches. The survey materials have included steel, aluminum alloys, and titanium alloys. Campbell, Murray, and Yakushina et al. discussed the grain size and volume fraction of Ti6Al4V by the feature-based analysis approach [4]. Naik, Sajid, and Kiran applied supervised machine learning methods such as Naïve Bayes (NB), k-nearest neighbor (K-NN), linear discriminant analysis (LDA), and decision tree (DT) classifiers to recognize the grain boundaries for classifying the phase of ASTM A36 steels [5]. Azimi, Britz, and Engstler et al. used general conventional neural networks and fully conventional neural networks to classify different phases of database images according to the segmentation-based classification of microstructures [6]. Kiyomura, Wang, and Ogawa et al. applied homology and Bayesian optimization of machine learning algorithms to

---

*Corresponding Author: Wesley Huang; E-mail: wasener@mail.cnu.edu.tw

DOI: 10.53106/160792642022072304023

predict the characteristic properties and microstructural optimization of globular cementite in pearlite steel [7]. Our previous research applied image morphology approaches to progress metallographic images of carbon steel and to estimate the grain size of pearlite mixture-phase steel [8]. Tsutsui, Terasaki, and Uto et al. used machine learning approaches to utilize feature extraction with texture analysis on several microstructures, including martensite, upper bainite, lower bainite, and other mixture phases, in scanning electronic microscope (SEM) images [9]. The machine learning approaches used in the research included the random forest (RF) and gradient boosting machine (GBM). Different from other machine learning or feature-based segmentation algorithms, Sarkar, Mali, and Sarkar used the genetic algorithm (GA) to classify the feature selection of seven kinds of materials [10]. Naik and Kiran integrated a type of feed-forward neural network (FFNN) named the complex-step perturbation approach (CSPA) to optimize the training model [11]. Chen, Sun, and Fu et al. applied a semi-supervised learning framework to perform metallographic image segmentation of aluminum alloy materials. A modification from U-Net (a convolution process) named semi-supervised segmentation was used to train the own image dataset (1024 pixels × 768 pixels resolution), and the results were competitive [12].

However, compared with previous literature, to the authors' best knowledge, there is no literature on gray cast iron metallographic image recognition via the deep learning approach. This research made two contributions: one was to discuss the recognition of gray cast iron using the feature-based morphology classification approach; the other was to integrate deep learning and feature-based algorithms to classify different categories with complex shapes of metallographic images.

This study combined morphology and deep learning approaches to classify different types of gray casting iron. Using the calculation criteria of the shape factors (SFs), the morphological features could be distinguished for three obvious types. The last three types could be classified with a deep learning approach. Chapter two introduces two approaches for analyzing the image features of gray graphite cast iron. Chapter three introduces the experimental details of analysis environments. Chapter four shows the results and discusses the differences between the two approaches. Finally, we give a short conclusion of this study and future works of extensible research. The contribution of this paper was to discuss the feasibility of integrating feature-based image morphology analysis with the deep learning neural network algorithm. Future research can extend more image categories of each type of gray casting iron.

2 Image Feature Analysis

2.1 Image Analysis Process

The image analysis procedure is shown in Figure 1 and includes six steps: (1) grayscaling; (2) sharpening processing; (3) binarization processing; (4) connecting component processing; (5) decision-making system; and (6) statistics [13]. Firstly, among the whole processing steps, the grayscaling process refers to the operation process of converting the existing 24 bits of red, green, and blue (RGB) color data into 8 bits of grayscale data. When the input image ignores the influence of color temperature, grayscaling the color image can prevent the increase of calculation errors due to the existence of three different primary colors. Secondly, sharpening processing includes increasing the grayscale gradient between grayscale images, reducing fuzzy and unrecognizable image features, and improving the accuracy of the image in the later binarization process. Thirdly, grayscale image binarization converts the grayscale image into a black and white binary image according to the threshold. The binarization process is as follows: Suppose the width of the image is w pixels and the height is h pixels; when the pixel is located at k, the grayscale value is marked as G_k ($0 \leq G_k \leq 255$). When G_k is equal to 0, the pixel is viewed as black; on the other hand, when G_k is equal to 255, the pixel is viewed as white. When the binarization threshold value is T, the B_k at pixel k after the binarization process can be defined as Eq. (1):

$$B_k = \begin{cases} 0, & if\ G_k \leq \mathrm{T} \\ 255, & if\ G_k > \mathrm{T} \end{cases} . \tag{1}$$

The size of the metallographic image is $w \times h$ pixels.

After the binarization process, the connected components process is carried out. The purpose of this process is to identify the features of each pixel image and connect them as a physical group. The pixel connection is applied to detect whether pixels are in the same group. The pixel connection component process is as follows: When the value of pixel k is 0 (black), four pixel values (k+1, k-1, k-w, and k+w) are also checked at the same time (i.e., four locations representing the top, bottom, left, and right of the pixel image, respectively). If the pixel image is identified as black (0), it is regarded as belonging to the same group as pixel k. The process is named four-connection component detection. For eight-connection component detection, an additional four pixel values (k-w+1, k-w-1, k+w+1, and k+w-1) are integrated into the original four-connection pixels. Meanwhile, this connection can expand out the second layer outside the periphery of pixel k as 24 channels. In this issue, considering the calculation speed factor, the eight-connection component detection approach was progressed.

After the pixel connectivity check processing, each group can be found and identified. Then, the image group analysis decision system can be carried out to determine the type of graphite carbon particles. The decision-making system is described in the next chapter. After the calculation and classification of types, the amount, area proportion, and average area of each graphite carbon cluster are counted and analyzed.

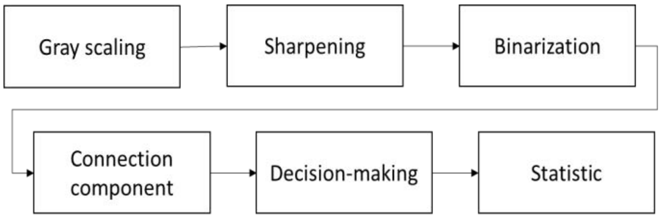


Figure 1. Image process flow chart

2.2 Feature Decision Making Systems

This study used two feature decision making approaches. One was based on only applying feature detection with two morphology methods, while the other was a hybrid approach combining the same first feature detection stage for determining type A, type C, and type F, and then applying a deep learning algorithm to determine the last three types.

The architecture of the two-stage feature decision making system (TFDMS) in this study is shown in Figure 2. After the connection component process, each connected graphite particle can be distinguished, and the shape factor (SF) can be estimated by the morphology application. The definition of SF is shown as Eq. (2):

$$\mathrm{SF} = \frac{A_{graphite}}{A_{reference-circle}} \times 100\%, \tag{2}$$

where $A_{graphite}$ is the true area of the graphite particle and $A_{reference\text{-}circle}$ refers to the area of the reference circle. Here, the reference circle is defined using the diameter from the longest length of the graphite particle.

The first stage of the determination is defined as follows: The SF of type A should be less than 40%, the SF of type C should be between 40 to 50%, and the SF of type F should be between 50 to 60%. Particles out of these ranges are ignored and remain in the image during this stage.

In the second stage, the classification is progressed in the image omitted type A, type C, and type F particles (i.e., the particles of three types are regarded as a background image and ignored). The gradient of a histogram of grayscale in each part of the region is computed as the threshold of classification. The number of regions n can be defined in Eq. (3):

$$n = 300 \times (\frac{100}{M})^2; \tag{3}$$

where M is the magnification of the microscope. By summing up the total B_i in each region, the total grayscale values in each pixel of a region can be estimated. Because the image is binarized, the total grayscale values can be expressed as $\sum B_i$,. After calculating n regions, the minimum value within all regions can be estimated as $min\,(\sum B_i)$. The maximum grayscale gradient for all regions can also be estimated as Δ_{max}, as shown in Eq. (4):

$$\Delta_{max} = \max\,[\sum B_i - \min\,(\sum B_i)]. \tag{4}$$

The type D category was defined as having a value of Δ_{max} greater than 25000, the type E category was defined as having a value of Δ_{max} between 20000 and 24500, and the type B category was defined as having a value of Δ_{max} between 18000 and 20000. Particles in the image that did not belong to type B, type D, or type E were ignored and not calculated.

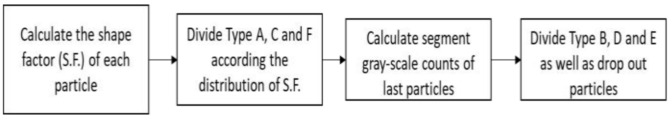


Figure 2. Architecture of the two-stage feature decision making system (TFDMS)

2.3 Deep Learning Approach

In this study, the YOLO V3 algorithm was applied for deep learning training. The training was aimed at type B, type D, and type E, instead of the second stage described in the previous section. Deep learning methods have been issued since Yann LeCun proposed a framework in 1989 [14], and Krizhevsky carried out the convolutional neural network (CNN) algorithm in 2012 [15]. The general algorithms applied in image processing include image classification and object detection and recognition. Among these algorithms, general algorithms for image classification include VGG, ResNet of Microsoft company, Inception of Google company, and DenseNet, etc. The algorithms used in target detection and recognition can be divided into two categories: the instance bounding boxes (BBox) method and the regression method. The instance BB method is used to get the main area of the desired object so as to identify and classify the selected area. Algorithms in this category include faster RCNN, R-FCN, FPN, SPP, and RES-Net. The regression classifies the image in the Bbox method into numerous categories, including YOLO and SSD, etc. In this study, YOLO V3 was applied to train and test the metallographic images.

You Only Look Once (YOLO), a famous target detection method, has been upgraded from YOLO V1 to YOLO V3. YOLO V1 is a typical one-stage object detection method in a regression algorithm [16-19]. The calculation speed is fast and the detection is efficient. The principles and ideas in the algorithm are also very simple. Firstly, an image is reshaped to a size of 448 × 448 pixels square. Since the fully connected layer is used in the network, the size of the image needs to be input into the convolution neural network (CNN) at a fixed size and then divided into S × S cells, where $S \in I \geq 0$. The objection is predicted with each location of BBox and the corresponding content.

In terms of YOLO V2, which is also named YOLO9000 [20-23], the algorithm can recognize more than 9000 types of objects. In the research, a mean average precision (mAP) of 76.8% can be achieved under a 76 FPS frame speed using the VOC2007 dataset, and an mAP of 78.6% can be achieved at 40 FPS. Compared with V1, V2 was improved and implemented an ‘anchor’ issued in faster R-CNN, whose number and shape of anchors is determined by progressing a k-means classification with all bounding boxes of all training images. In V2, five anchors are chosen, as this represents a good compromise between recall and model efficiency. The prediction of the center coordinates in the BBox is based on the offset of the anchor coordinates, and the offset is estimated by the anchor center point in V1 relative to the location of the upper left corner. In the network layer, Darknet19 is used. Different from V1, a fully convolutional network is implemented instead of the original fully connected layer, and the global average pooling layer is used instead of the last pooling layer to increase the resolution of the features.

In YOLO V3, Darknet53 is also used for the network part [24-29]. Meanwhile, each box uses multi-label classification to predict the probable category of the container within the BBox. The algorithm replaces the softmax function in V2 with a logistic regression loss and uses the binary cross-entropy loss for prediction during the training period. YOLO V3 uses a multi-scale composition method to carry out predictions, as shown in Figure 3 [30]. Similar to FPN, up-sampling and fusion approaches are applied to make predictions on feature maps with multiple scales. The improvement performance on small targets is apparent.

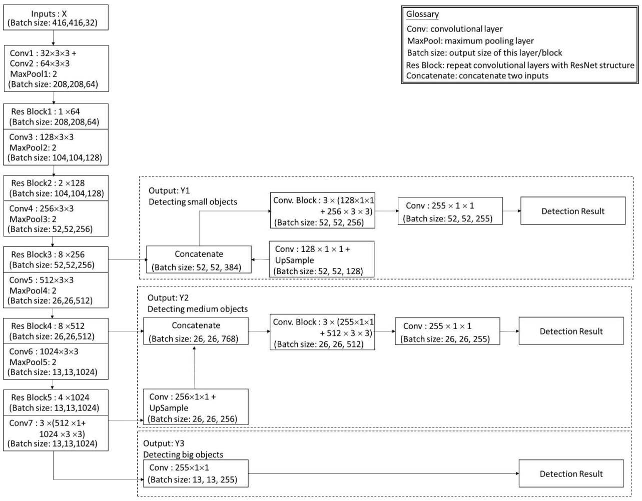


Figure 3. Architecture of the YOLO V3 algorithm

3 Experimental Details

3.1 Metallographic Image Environments

The grafy casting iron material used in this study was FC300. According to Japanese Industrial Standard (JIS) code G5501, the chemical components are: C (2.9–3.2%); Si (1.2–1.5%); Mn (0.9–1.1%); P (≤ 0.15%); and S (≤0.1%). In this study, the optical microscope was an Olympus BX51M, the metallographic camera was a Sentech STC-MC152USB, the image capture pixel was 1360 × 1024, RGB was used for image capture, and images were sampled at an optical magnification of 100 times. The captured images were output to a 22 inch 16:9 screen for display.

3.2 Programming Environments

The program languages used in this study were separated into two parts. Microsoft® Visual C# was used for image feature detection, while Python 3.6 with the TensorFlow 2.1 and DarkFlow modules were used for deep learning recognition. The main user interface was developed using Microsoft Visual C#, and the Python language was embedded using the Pyinstaller 4.5 module. The computer used for training the prepared dataset was equipped with an Intel® Core™ i7 CPU, 32 GB DDR4 RAM, and NVIDIA® GForce 1080 11GB and 4 TB SATA memory storage.

3.3 Dataset Preparation

In this study, 90 type A, 20 type B, 80 type C, 20 type B, and D mixed, 20 type D and E mixed, 30 type D, 30 type E, and 70 type F images were obtained for training. The number of metallographic images was limited to the probability of formation of the types. Each image was selected with 10 connection components as train labels and the other 2 connection components as test labels. From this, through the establishment of labels, the dataset of this study was further schemed. The LabelImg module was used for labelling the features, and image features were labeled by rectangle shapes.

4 Results and Discussion

4.1 Feature-based Analysis

Two fields of view were captured from the specimen and analyzed, and the results are shown in Figure 4. Figures 4(a) and 4(b) are the captured gray cast iron metallographic images, and Figures 4(c) and 4(d) are the analyzed metallographic images. In the figure, type A is displayed as green, type B is displayed as red, type C is displayed as yellow, type D is displayed as magenta, type E is displayed as cyan, and type D is displayed as blue. In addition, particles at the edges around the image were ignored in the calculation and shown as the original black color.

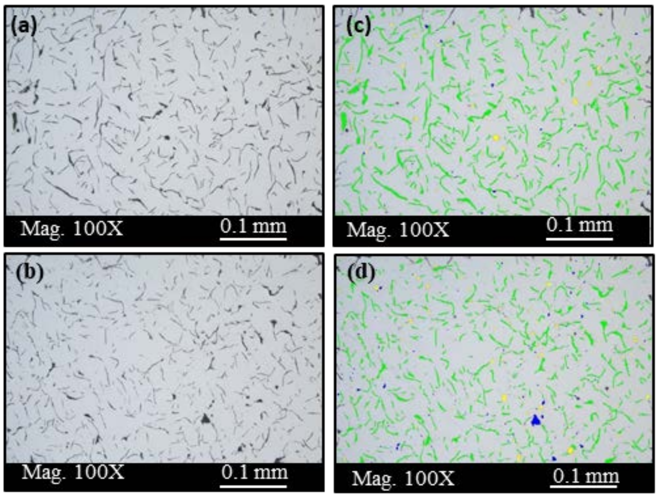


Figure 4. Original and resulting images: (a) and (b) are the captured gray cast iron metallographic images; (c) and (d) are the metallographic analyzed images.

The identification results of the two fields of view are shown in Table 1. The amount of type A was more than 90%, some were identified as type C, and others were identified as type F. As shown in Figure 3, the appearance of type A graphite was biased toward the middle of the strip and was spiky on two sides of the strip. Meanwhile, the appearance of the identified type C graphite was elongated, even if it had a bifurcated shape. The spherical-like graphite in the image was identified as type F. The results indicated the flexibility of the first stage of TFDMS. As shown in the figure, type B, type D, and type E showed zero percentage of results, indicating there was no existence of the corresponding type in metallographic.

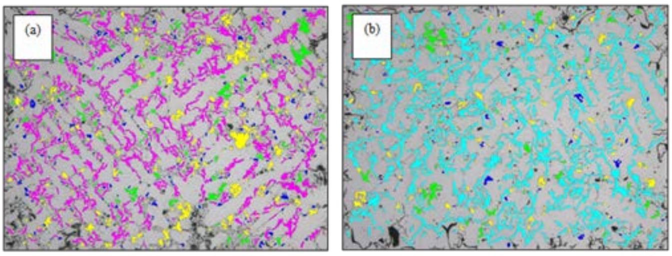
Table 1. Identification results of the two fields of view

View field	Type	1	2
Area percentage (%)	Type-A	91.00	92.19
	Type-B	0.00	0.00
	Type-C	5.13	4.18
	Type-D	0.00	0.00
	Type-E	0.00	0.00
	Type-F	3.87	3.63

Table 2 shows the classification results of stage 2 in the first approach. The proportions of type A and type C decreased as about 10-12%, and the proportions of type D and type E increased as about 80.9 and 65.5% in the image, respectively. Figure 5 shows the identification results of type D and type E of metallography. Figure 5(a) illustrates type D in magenta, and Figure 5(b) shows type E in cyan, separately.

Table 2. Results of TFDMS for two sample view fields

	Area proportion of each type						Counted field area	Particle counts
	A	B	C	D	E	F		
Units	%						um²	-
Type-D	10.70	0.00	17.53	65.53	0.00	6.25	126121	932
Type-E	5.54	0.00	8.22	0.00	80.89	5.34	156397	727



(a) type D is shown in magenta (b) type E is shown in cyan

Figure 5. Identification results of type D and type E of metallography

Three other types could be identified in TFDMS; however, the disadvantage of the approach was that the entire image was used for region segmentation and identification. Therefore, the identification result could only be shown in one singular type within three types. In other words, mixtures of type B, type D, and type E could not be performed in local region identification. Therefore, TFDMS was not suitable for the mixtures of type B, type D, and type E.

4.2 Combined with Deep Learning Approach

After the first stage of feature recognition, the remaining features were classified using a deep learning algorithm. Table 3 is the recognition results of the second stage according to the YOLO V3 algorithm. In the table, 'type' refers to the training content of different types of microstructures, including type B, type D, type E, type D, and type E mixtures, as well as type B and type D mixtures. The training dataset (D_{Train}) refers to the number of labels in the training process, the test dataset (D_{Test}) refers to the number of labels in the test process, correct numbers (CN) refers to the number of correct identifications, and correct rate (CR) refers to the correct rate in the whole recognition process, which can be expressed by Eq. (5):

$$CR = \frac{CN}{D_{Test}} \times 100\%. \quad (5)$$

Average time (t) means the average time for calculating the D_{Test} dataset.

Table 3. Recognition results of second stage proceed by the YOLO V3 algorithm

Model Name: "YOLO V3"					
Type	Train Dataset	Test DataSet	Correct Numbers	Correct Rate (%)	Average Time (s)
Symbol	TrainD	TestD	CN	CR	t
Type-B	200	40	36	90.0	0.913
Type-D	300	60	55	91.7	0.954
Type-E	300	60	53	88.3	0.986
Type-DE mixture	200	40	32	80.0	1.125
Type-BD mixture	200	40	29	72.5	1.017

The experimental results showed that when the shape of the image or microstructures was more complex, the recognition accuracy decreased, especially for mixed microstructures, for which the recognition accuracy was only up to 80.0%. The cause of the low accuracy could be the lack of sufficient numbers in the training dataset, which resulted in poor calculation convergence. In terms of the average calculation time, the recognition time for single types was shorter while the recognition time for composite types was relatively longer, indicating that the composite types still influenced the calculation time.

5 Conclusion

In this study, two approaches were issued to classify six types of gray cast iron metallography. The first approach used TFDMS, two-stage feature decision making systems. The first stage was to classify using SF and the second stage was to verify with difference grayscale in blocks. In the second approach, SF was used to distinguish at the first stage of half part, and the development of deep learning YOLO V3 was used to distinguish other types as the second stage. The conclusions are summarized below:

(1) Regarding the experimental results of TFDMS for gray cast iron shown in Table 1, type A particles could be identified clearly, and other feature particles could be distinguished.

(2) Type B, type D, and type E metallographic images could be identified in stage 2 of TFDMS. Nevertheless, the identification result could be only shown in one singular type within three types.

(3) The applicability of the second approach combined with two algorithms can distinguish the area ratio of each type. Through using this method to determine the quality of gray cast iron materials, the quality and life of the workpiece could be maintained. The accuracy of the second method was higher than that of the first method.

(4) The research was limited to the analysis of metallographic images of gray cast iron material.

(5) In our future works, more deep learning algorithms will be carried out to compare the efficiency and accuracy of each algorithm. Meanwhile, the metallographic image dataset will expand the number of images to enhance the accuracy of the algorithm.

We only discussed one deep learning approach in this study to prove the feasibility of the two mentioned approaches. In the future, more shape categories of each type will be added to the dataset to increase the accuracy of recognition. Meanwhile, more deep learning approaches will be applied to select the most suitable one for recognizing the type of gray casting iron.

References

[1] China National Standards, *Gray cast iron metallography*, GB7216-1987, 1987.

[2] J. P. Papa, R. Y. M. Nakamura, V. H. C. Albuquerque, A. X. Falcão, J. M. R. S. Tavares, Computer techniques towards the automatic characterization of graphite particles in metallographic images of industrial materials, *Expert Systems with Applications: An International Journal*, Vol. 40, No. 2, pp. 590-597, February, 2013.

[3] V. H. C. Albuquerque, A. R. Alexandria, P. C. Cortez, J. M. R. S. Tavares, Evaluation of multilayer perceptron and self-organizing map neural network topologies applied on microstructure segmentation from metallographic images, *NDT & E International*, Vol. 42, No. 7, pp. 644-651, October, 2009.

[4] A. Campbell, P. Murray, E. Yakushina, S. Marshall, W. Ion, New methods for automatic quantification of microstructural features using digital image processing, *Materials & Design*, Vol. 141, pp. 395-406, March, 2018.

[5] D. Naik, H. U. Sajid, R. Kiran, Texture-Based Metallurgical Phase Identification in Structural Steels: A Supervised Machine Learning Approach, *Metals*, Vol. 9, No. 5, Article No. 546, May, 2019.

[6] S. Azimi, D. Britz, M. Engstler, M. Fritz, F. Mücklich, Advanced Steel Microstructural Classification by Deep Learning Methods, *Scientific Reports*, Vol. 8, Article No. 2128, February, 2018.

[7] K. Kiyomura, Z.-L. Wang, T. Ogawa, Y. Adachi, Characterization and Optimization of Pearlite Microstructure Using Persistent Homology and Bayesian Optimization, *ISIJ International*, Vol. 62, No. 2, pp. 307-312, February, 2022.

[8] W. Huang, C.-S. Wang, Y.-F. Chang, C.-M. Yeh, A Digital Evaluation Approach for Nodule Size Analysis of Pearlite Mixture Phase Structure in Steel, *Metallography, Microstructure, and Analysis*, Vol. 9, No. 3, pp. 457-465, June, 2020.

[9] K. Tsutsui, H. Terasaki, K. Uto, T. Maemura, S. Hiramatsu, K. Hayashi, K. Moriguchi, S. Morito, A methodology of steel microstructure recognition using SEM images by machine learning based on textural analysis, *Materials Today Communications*, Vol. 25, Article No. 101514, December, 2020.

[10] A. Khan, S. Sarkar, K. Mali, R. Sarkar, A Genetic Algorithm Based Feature Selection Approach for Microstructural Image Classification, *Experimental Techniques*, Vol. 46, No. 2, pp. 335-347, April, 2022.

[11] D. Naik, R. Kiran, A Novel Sensitivity-based Method for Feature Selection, *Journal of Big Data*, Vol. 8, Article No. 128, October, 2021.

[12] D.-L. Chen, D.-P. Sun, J. Fu, S.-X. Liu, Semi-Supervised Learning Framework for Aluminum Alloy Metallographic Image Segmentation, *IEEE Access*, Vol. 9, pp. 30858-30867, February, 2021.

[13] L. I. Kuncheva, *Combining pattern classifiers: methods and algorithms*, Wiley-Interscience, 2004.

[14] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel, Backpropagation Applied to Handwritten Zip Code Recognition, *Neural Computation*, Vol. 1, No. 4, pp. 541-551, December, 1989.

[15] A. Krizhevsky, I. Sutskever, G. E. Hinton, Imagenet classification with deep convolutional neural networks, *Advances in Neural Information Processing Systems*, Vol. 1, Lake Tahoe, Nevada, United States, 2012, pp. 1097-1105.

[16] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, You only look once: unified, real-time object detection, *IEEE conference on computer vision and pattern recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 779-788.

[17] C. Lin, H. Xie. H. Zheng, PedJointNet: Joint Head-Shoulder and Full Body Deep Network for Pedestrian Detection, *IEEE Access*, Vol. 7, pp. 47687-47697, April, 2019.

[18] Z. Liu, S. Wang, Broken Corn Detection Based on an Adjusted YOLO with Focal Loss, *IEEE Access*, Vol. 7, pp. 68281-68289, May, 2019.

[19] L. Li, X. Chen, H. Zhou, L. Wang, Recognition and Application of Infrared Thermal Image Among Power Facilities Based on YOLO, *2019 Chinese Control And Decision Conference (CCDC)*, Nanchang, China, 2019, pp. 5939-5943.

[20] J. Redmon, A. Farhadi, YOLO9000: Better, Faster, Stronger, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 7263-7271.

[21] S. Zhang, S. Lan, Q. Bu, S. Li, YOLO based Intelligent Tracking System for Curling Sport, *2019 IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS)*, Beijing, China, 2019, pp. 371-374.

[22] X. Zhang, Z. Qiu, P. Huang, J. Hu, J. Luo, Application Research of YOLO v2 Combined with Color Identification, *2018 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC)*, Zhengzhou, China, 2018, pp. 138-1383.

[23] Z. Wang, K. Xu, S. Wu, L. Liu, L. Liu, D. Wang, Sparse-YOLO: Hardware/Software Co-Design of an FPGA Accelerator for YOLOv2, *IEEE Access*, Vol. 8, pp. 116569-116585, June, 2020.

[24] J. Redmon, A. Farhadi, YOLOV3: An Incremental Improvement, *arXiv Computer Science*, April, 2018. https://arxiv.org/abs/1804.02767

[25] E. N. Ukhwah, E. M. Yuniarno, Y. K. Suprapto, Asphalt Pavement Pothole Detection using Deep learning method based on YOLO Neural Network, *2019 International Seminar on Intelligent Technology and Its Applications (ISITIA)*, Surabaya, Indonesia, 2019, pp. 35-40.

[26] Y. Li, Y. Tian, J. Tian, F. Zhou, An Efficient Method for DPM Code Localization Based on Depthwise Separable Convolution, *IEEE Access*, Vol. 7, pp. 42014-42023, March, 2019.

[27] J. Hu, X. Gao, H. Wu, S. Gao, Detection of Workers

Without the Helments in Videos Based on YOLO V3, *2019 12th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI 2019)*, Huaqiao, Jiangsu, China, 2019, pp. 1-4.

[28] M. Jiang, T. Hai, Z. Pan, H. Wang, Y. Jia, C. Deng, Multi-Agent Deep Reinforcement Learning for Multi-Object Tracker, *IEEE Access*, Vol. 7, pp. 32400-32407, February, 2019.

[29] U. Nazir, U. K. Mian, M. U. Sohail, M. Taj, M. Uppal, Kiln-Net: A Gated Neural Network for Detection of Brick Kilns in South Asia, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, Vol. 13, pp. 3251-3262, June, 2020.

[30] CyberAI Lab, A Closer Look at YOLOV3, Qiita, July 23, 2018, from https://qiita.com/cyberailab/items/850806c08af08853bac8.

Biographies

Wesley Huang is an Assistant Professor in the General Education Center, Chia Nan University of Pharmacy and Science. His research interests include metallographic image analysis, artificial intelligence applications and feature-based image analysis algorithm. He has published articles in Applied Soft Computing, Expert Systems with Applications, and Metallography, Microstructure, and Analysis.

Zhi-Yuan Su is a Professor in the Department of Multimedia and Game Development, Chia-Nan University of Pharmacy and Science. His research focuses on network application and digital game-based learning. He has published articles in journals such as IEEE Access, International Journal of Information Management and Journal of Internet Technology.

Chia-Sui Wang is a Professor in the Department of Information Management Development, Chia-Nan University of Pharmacy and Science. Her research is focus on information education, and digital game-based learning. She has published articles in journals including Innovations in Education and Teaching International, the Electronic Library and Journal of Internet Technology.

Mark Yeh graduated from the Department of Electrical Engineering, National Sun Yat-sen University. He studies for a Master of Electrical Engineering in the National Cheng Kung University. His research focuses on image processing and video compression.

Jyh-Horng Chou is currently the Chair Professor of Mechanical and Computer-Aided Engineering Department, Feng Chia University. His research interests include intelligent systems and control, computational intelligence, and robust control and optimization. He is also a Fellow of Institution of Engineering and Technology (IET), Chinese Automatic Control Society (CACS), and Chinese Institute of Automation Engineer (CIAE).