



OPEN

Unsupervised microstructure segmentation by mimicking metallurgists' approach to pattern recognition

Hoheok Kim¹, Junya Inoue^{1,2✉} & Tadashi Kasuya²

An efficient deep learning method is presented for distinguishing microstructures of a low carbon steel. There have been numerous endeavors to reproduce the human capability of perceptually classifying different textures using machine learning methods, but this is still very challenging owing to the need for a vast labeled image dataset. In this study, we introduce an unsupervised machine learning technique based on convolutional neural networks and a superpixel algorithm for the segmentation of a low-carbon steel microstructure without the need for labeled images. The effectiveness of the method is demonstrated with optical microscopy images of steel microstructures having different patterns taken at different resolutions. In addition, several evaluation criteria for unsupervised segmentation results are investigated along with the hyperparameter optimization.

A microstructure is a small-scale internal structure of a material, which strongly affects its mechanical, chemical, and electric properties. In particular, steel alloys are known to exhibit a wide range of mechanical properties due to the formation of a wide variety of microstructures such as ferrite, pearlite, bainite, and martensite depending on the cooling process. Figure 1 depicts typical examples of microstructures observed in steel alloys. A difference in microstructure results in different mechanical properties¹. Dual-phase steel, which consists of a ferrite matrix containing hard martensitic islands and is widely used in the automobile industry, is a good example of this phenomenon. A higher yield strength can be achieved not only by increasing the volume fraction of the hard martensitic microstructure² but also by modifying the microstructural morphology³. Accordingly, characterizing the microstructures of steel is an important task for the development of advanced high-strength steels. However, this is very challenging because the microstructure usually consists of one or more phases that are not easily distinguishable. Conventionally, the phases of steel alloys have been classified by the manual analysis of light optical microscopy (LOM) or scanning electron microscopy (SEM) images⁴. However, this approach has the drawback that it requires a labor-intensive pixelwise classification performed by experienced experts. Therefore, approaches based on machine learning algorithms have attracted great attention due to their efficiency. For example, Choi et al.⁵ introduced a classification algorithm based on support vector machines (SVM)⁶ for detecting defects on the surface of a steel product. Further applications of SVM were presented by Gola et al.^{7,8}, where the microstructures of steel alloys were classified into constituent phases. Another frequently applied technique is the random forest, which is a classification algorithm composed of multiple decision trees⁹. Various studies have found that steel microstructures can be accurately classified using random-forest-based methods^{10–12}.

Nowadays, one of the most emerging microstructure classification schemes is a deep learning based approach. Deep learning is also a class of machine learning algorithms that use multiple processing layers to learn representations of the raw input¹³. These multilayers are called neural networks and are being actively applied to microstructure classification tasks. de Albuquerque et al.^{14,15} first proposed the application of artificial neural networks to classify and quantify simple nodular microstructures from cast iron images. More recently, convolutional neural networks (CNN) have been intensively used in the field of computer vision owing to their fast and efficient classification performance¹⁶. Azimi et al. reported the successful implementation of VGGNet¹⁷, which is a pretrained CNN proposed by Krizhevsky, for classifying microstructures from LOM and SEM images of a steel¹⁸. Since then, a number of studies applying CNN have been conducted such as the application of DenseNet¹⁹

¹Institute for Industrial Science, The University of Tokyo, 4-6-1 Komaba, Meguro, Tokyo 153-0041, Japan. ²Graduate School of Engineering, The University of Tokyo, 7-3-1 Hongo, Bunkyo, Tokyo 113-8656, Japan. ✉email: inoue@material.t.u-tokyo.ac.jp

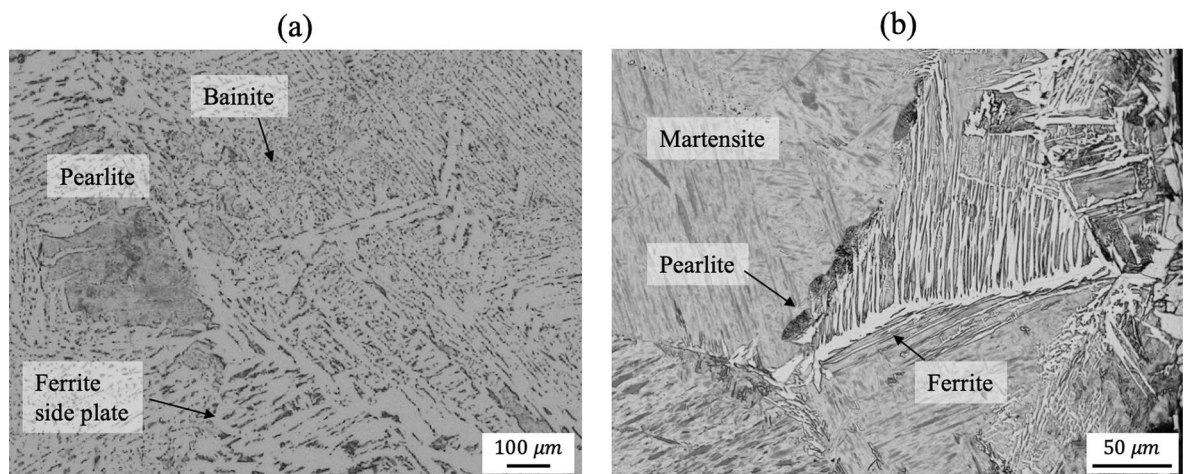


Figure 1. Typical examples of microstructures observed in steel alloys.

to detect defects in steels²⁰ and ResNet18²¹ to classify microstructures of welded steels²². It was verified that the performance of CNN-based methods is as good as that of humans.

Deep learning algorithms are trained with a vast number of labeled images so that they can learn how features are related to the target and this scenario is referred to as supervised learning. However, there are several difficulties when applying supervised learning algorithms to microstructure classification. First, microstructures are not easily distinguishable even for experienced experts, so the preparation of a large labeled dataset is extremely labor-intensive and time-consuming. Second, the size of the input image is limited when using pretrained parameters. For example, the width and height of the input image should be fixed to 224×224 pixels when implementing well-known networks such as VGGnet¹⁷, DenseNet¹⁹, and ResNet18²¹.

In contrast, human researchers perceptually distinguish different microstructures from various patterns under various illumination conditions without the need for labeled images. In view of the fact that metallurgists can identify different microstructures hidden in a single micrograph even at first glance, we design an unsupervised segmentation method for low-carbon steels by mimicking the way in which metallurgists investigate each micrograph. The algorithm is strongly motivated by the one proposed by Kanazaki²³ and is based on CNN accompanied by a superpixel algorithm. Segmentation results demonstrate that the proposed method can be used to distinguish microstructures. In addition, the quality of the segmented images is assessed using evaluation criteria for unsupervised segmentation scenarios.

Methods

The approach to identifying hidden microstructures in a single image adopted by metallurgists generally consists of three steps; at first, the whole image is roughly subdivided into many small regions of interest (ROIs) with identical contrast and texture characteristic length, then microstructural features representing each ROI are searched. Finally, ROIs are grouped into several classes that have some similarity with respect to the derived microstructural features. Accordingly, the unsupervised segmentation algorithm is designed as follows. After a single micrograph is input, it first undergoes superpixel segmentation to acquire small ROIs with identical contrast and texture characteristic length. Then, a CNN computation is carried out to derive a feature representing each ROI. The network parameters are trained so that the feature with the highest frequency in each ROI dominates the other features in the region. In addition, by applying the same CNN computation to all the superpixels, several features commonly appearing in several different ROIs are selected automatically. In other words, connected pixels with similar colors and other low-level properties, such as hue, luminance, and contrast, are grouped and assigned the same label by the superpixel computation and the spatially separated groups having similar textures are assigned the same label by CNN computation. By combining them, a group of pixels having similar features can be categorized into the same cluster and a schematic of the network is given in Fig. 2.

The following subsections provide detailed explanations of the algorithms employed in each step of the unsupervised segmentation method. First, the fundamental knowledge of the CNNs is given in order of the computation procedure. Then, an explanation of the superpixel algorithm is presented with sample images. Finally, several evaluation criteria are addressed to assess the results of segmentation without labeled images.

CNN. The structure of the CNN used in this study consists of an input layer, hidden layers, and an output layer as illustrated in Fig. 3. The input layer accepts the external image data and the output layer gives the predicted answer computed by the network. The computation is mainly conducted in hidden layers consisting of one or more convolutional filters (or kernels) that are indispensable building blocks of the CNN. Filters are typically composed of a set of learnable parameters and perform 2D filtering on input images by conducting the linear operation

$$h_n = (W_m \otimes x_n) + b_m \quad (1)$$

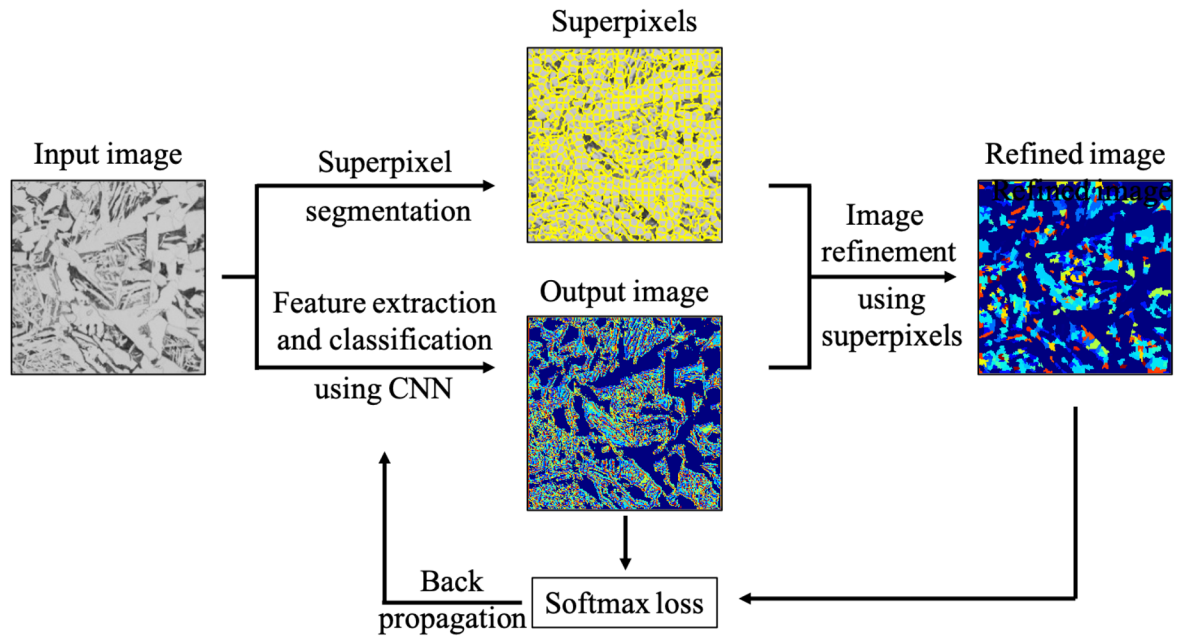


Figure 2. Schematic illustration of the applied algorithm composed of CNN and a superpixel algorithm.

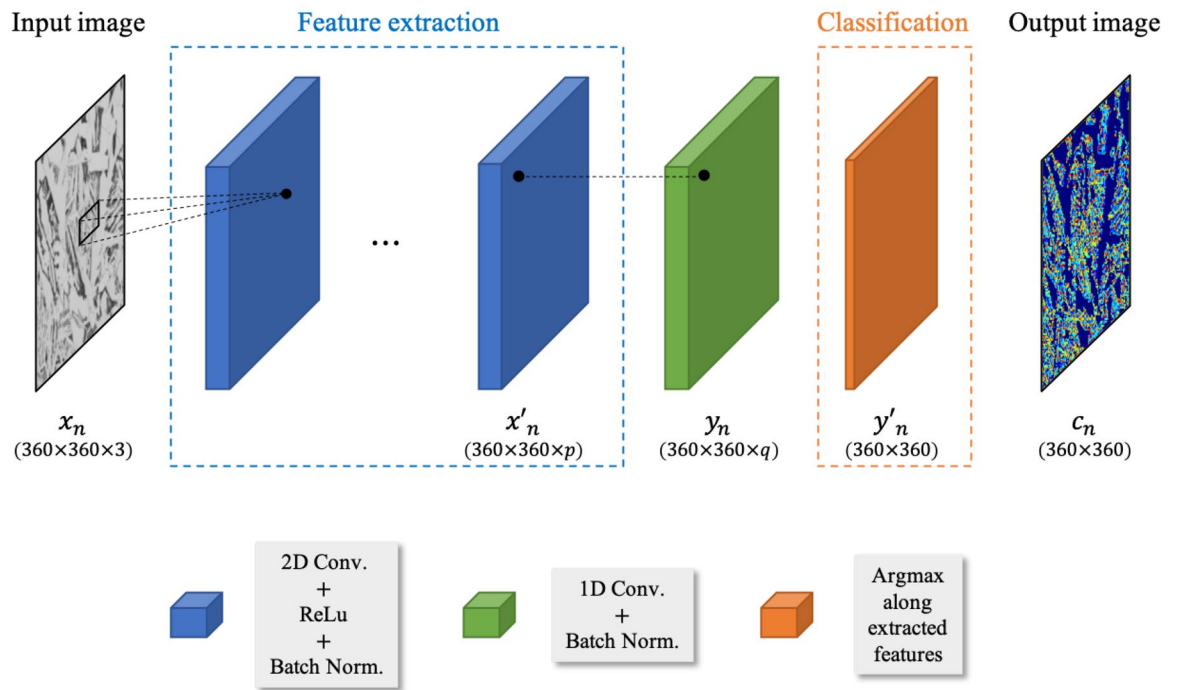


Figure 3. Architecture of the algorithm with an image size of $360 \times 360 \times 3$ and detailed information about the layers of the CNN network.

$\{x_n\}_{n=1}^N$ is a set of p -dimensional feature vectors of image pixels, where N denotes the total number of pixels in images. W_m , b_m , and h_n are respectively trainable weights of the filters, bias, and feature map obtained after convolutional operation \otimes . An activation function is followed by each convolutional layer in order to introduce nonlinearity into the neural network. One of the widely used functions is the rectified linear unit (ReLU) mathematically expressed as

$$f(x_n) = \begin{cases} x_n & \text{for } x_n > 0, \\ 0 & \text{for } x_n \leq 0. \end{cases} \quad (2)$$

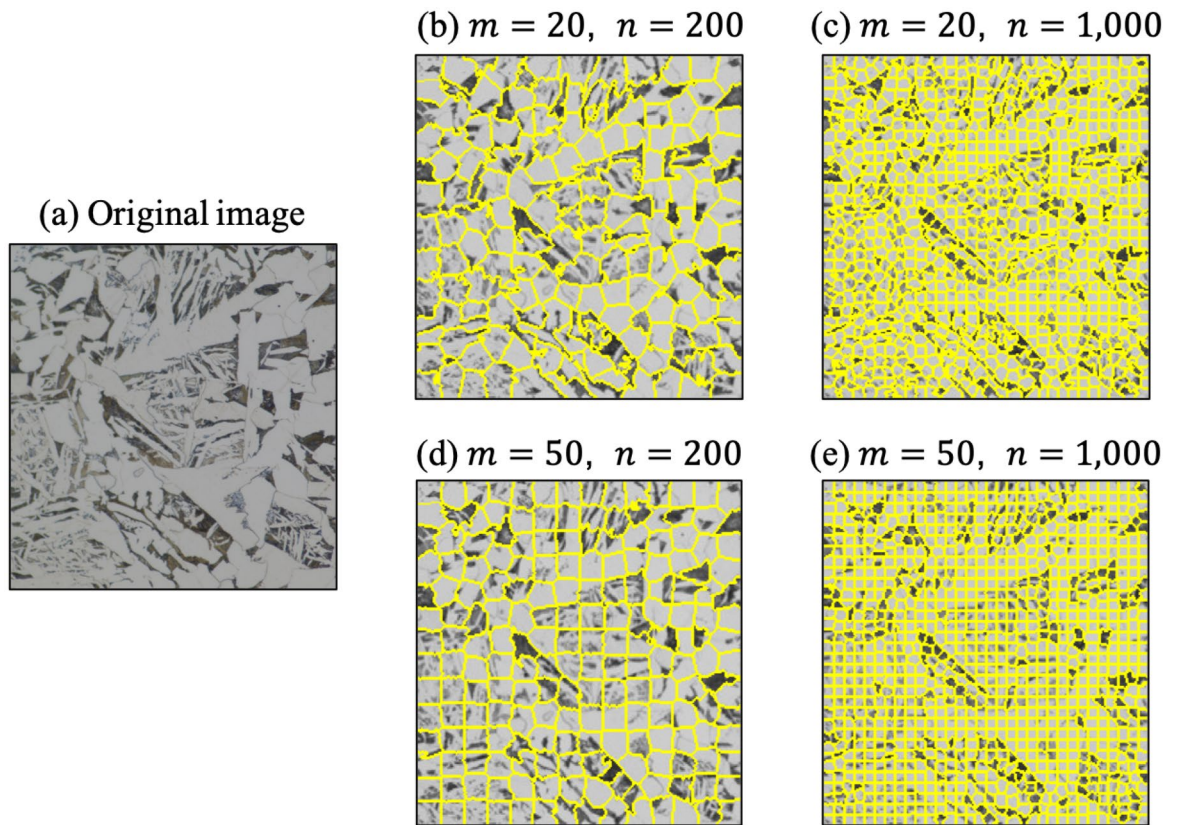


Figure 4. Effects of the compactness m and the number of superpixels n on superpixel segmentation: (a) original image, (b–e) superpixel segmentation results with various m and n .

Then, additional layers for batch normalization, which is a technique recently proposed by Ioffe and Szegedy²⁴, are connected to the activation function. The idea is to normalize the outputs of the activation function so that a subsequent convolutional layer can receive an image having zero mean and unit variance as

$$x'_n = \frac{x_n - \bar{x}_n}{\sqrt{\sigma^2(x_n) + \epsilon}}, \quad (3)$$

where \bar{x} , $\sigma^2(x)$, and ϵ are respectively the mean of x , the standard deviation of x , and a constant to provide numerical stability whose value is usually set as 1×10^{-5} . It has been reported that rescaling the image before inputting allows a faster, efficient, and more stable learning²⁵. The following layer is a linear classifier layer $\{y_n\}_{n=1}^N$ that categorizes the obtained features of each pixel into q classes. The linear relationship is applied in this study:

$$y_n = W_c x'_n + b_c, \quad (4)$$

where W_c and b_c are the weights of the 1D convolution filters and bias, respectively. After normalizing y_n so as to obtain y'_n , the argmax classification is applied to choose the features with the maximum y'_n . Finally, a 2D output image having segmentation classes c_n is obtained.

Superpixel segmentation. When a human distinguishes microstructures of steels, similar microstructures are generally grouped based on the basis of colors. For example, if two regions have the same color, then they will be allocated the same class. In addition, spatial characteristics or morphologies are considered when dividing regions. Similarly, a superpixel algorithm distinguishes different regions called superpixels, which are regions of continuous pixels having similar characteristics such as pixel intensity, by considering the color similarity and spatial proximity. Thus, it provides a convenient and compact representation of images when a human performs classification tasks. The simple linear iterative clustering (SLIC) algorithm²⁶ is introduced in this study among the various algorithms for clustering superpixels^{27–29}. There are two important hyperparameters when obtaining superpixels using this SLIC algorithm. The first is the number of superpixels, which defines the number of regions in the input image. The second is the compactness factor m , which balances the color proximity and spatial proximity. The lower the value of m , the more color proximity is emphasized. Its effect on the clustering is illustrated for a steel microstructure image in Fig. 4. Readers are recommended to refer to the literature^{26,30} for further information about the concept and implementation of the SLIC algorithm.

C	Si	Mn	P	S	Al	N	O
0.152	0.015	1.51	0.007	0.0016	0.026	18 (ppm)	28 (ppm)

Table 1. Chemical composition (wt. %) of the low-carbon steel used in this work.

Training of the CNN. Regarding the training of the CNN, the softmax function is used as a loss function, which measures the difference between the CNN output image and the refined image. The parameter tuning of the CNN is resolved using the stochastic gradient descent (SGD) optimization algorithm and backpropagation. Regarding SGD hyperparameters, it is empirically known that a learning rate of 0.1 and a momentum of 0.9 generally gives good results. Finally, the training is set to be finished after the maximum number of iterations is reached or the number of segmented regions becomes less than the number of classes set by a user.

Evaluation criteria. When evaluating a segmentation result, two approaches are generally taken. First is a supervised evaluation method which compares a segmented image with a labeled image. Even though a quantitative comparison can be easily done, it requires a manually labeled reference which is intrinsically subjective and labor intensive.

Another common alternative is an unsupervised evaluation method where the quality of a result is assessed based solely on segmented results. It enables the objective evaluation without requiring a manually labeled reference image. In this study, the performance of unsupervised segmentation with an $N \times M$ image size is evaluated using the F ^{31,32}, F' , and Q ³² metrics respectively given by Eqs. (5–7):

$$F = \frac{1}{1000(N \times M)} \sqrt{R} \sum_{i=1}^R \frac{e_i^2}{\sqrt{A_i}} \quad (5)$$

$$F' = \frac{1}{1000(N \times M)} \sqrt{\sum_{A=1}^{Max} [R(A)]^{1+1/A}} \times \sum_{i=1}^R \frac{e_i^2}{\sqrt{A_i}} \quad (6)$$

$$Q = \frac{1}{1000(N \times M)} \sqrt{R} \sum_{i=1}^R \left[\frac{e_i^2}{1 + \log A_i} + \left(\frac{R(A_i)}{A_i} \right)^2 \right]. \quad (7)$$

R , A_i , and e_i^2 are the number of segmented regions, the area of the i th region, and the average color error of the i th region, respectively. With these criteria, the segmentation result with the lowest value is preferred. The basic concept of the metrics is to assess the quality of segmented images by comparing average color error values e_i of segmented regions of input and output images. In addition to the goodness of fit of the color, terms including R are introduced in order to penalize segmentations that form too many regions. Zhang et al.'s article³³ is recommended for readers interested in a more detailed explanation about the evaluation of unsupervised image segmentation.

Implementation details. For the computational environment, a Tesla V100 NVIDIA GPU was used with the PyTorch framework and CUDA platform. The number of convolutional layer, filters in each convolution layer p and the linear classifiers q were set as 2, 100 and 50, respectively. For the SLIC superpixel algorithm, the compactness and the number of superpixels were defined as 20 and 50,000, respectively. The time taken for the unsupervised segmentation and the F value evaluation with a 1000×1000 pixel image was about 30 s and less than 1 s, respectively.

Results and discussions

Dataset. To obtain the input image data, we prepared low-carbon steel samples whose composition is presented in Table 1, which were polished and etched using 5% picral + 0.5% nital solution. Then, optical micrographs were taken at different magnifications. It was confirmed that each image consists of various microstructures such as grain boundary ferrite, ferrite side plate, pearlite, bainite, and martensite. The obtained image dataset was used as input data for training the network without modifying the image size.

Segmentation. Figure 5a shows one of the low-magnification micrographs with 1920×1440 pixels used for the input data. There are different microstructures: ferrite side plate, pearlite, and bainite. Figure 5b shows the zoomed images of the ferrite and bainite. It is clear that their morphological characteristics differ from each other in that bainite generally has finer microstructure than ferrite and includes finer carbide precipitates. However, both microstructures appear in white with small isolated dots distributed inside them. As they have the same color, traditional approaches, which rely merely on the image contrast, have encountered difficulty in distinguishing these microstructures. The segmentation result obtained using the present method is given in Fig. 5c, in which ferrite side plate, pearlite, and bainite are colored green, red, and purple, respectively. It is demonstrated that ferrite, bainite, and pearlite are well segmented. This result suggests that microstructures with similar contrasts but different morphologies can be divided on the basis of their morphological features using

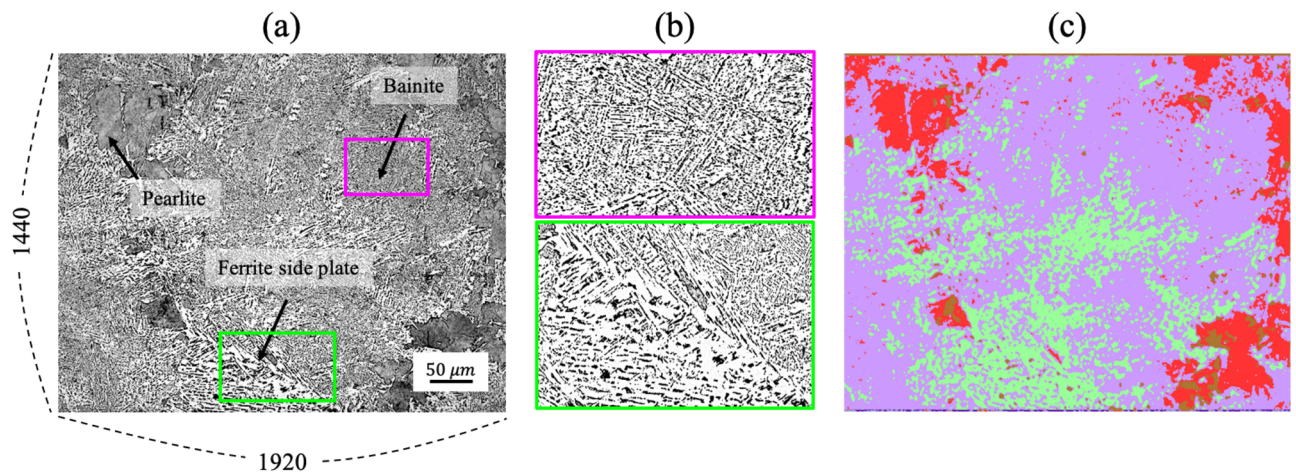


Figure 5. (a) Input microstructure image, (b) zoomed images showing bainite (upper row) and ferrite side plate (lower row) phases, and (c) segmentation result with colors indicating different classes.

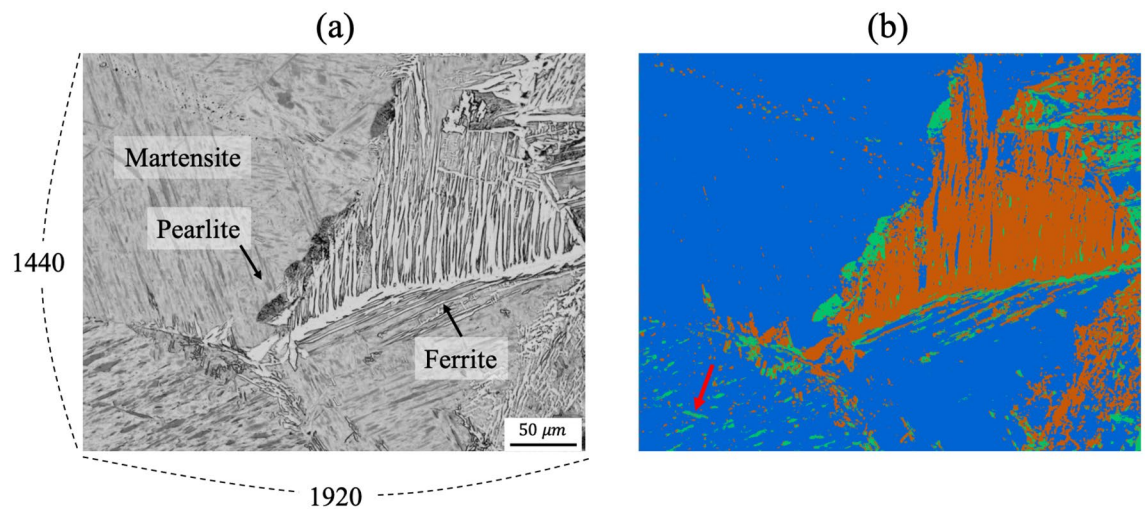


Figure 6. (a) Highly magnified input microstructure image and (b) segmented image represented by various colors indicating different classes.

the present method. It also implies that the method can be effectively applied to distinguish the microstructures in an image taken at a lower magnification.

Next, unsupervised segmentation is tested for microstructures observed at a higher magnification of 1920×1440 pixels. The input image shown in Fig. 6a consists of four different microstructures: grain boundary ferrite and needle-shaped ferrite side plate in white, pearlite in black, and martensite as the background. The result of unsupervised segmentation is shown in Fig. 6b, where ferrite, pearlite, and martensite are colored in orange, green, and blue, respectively. The result indicates that ferrite, pearlite, and martensite are well distinguished pixelwisely. However, some spots in the lower left part of Fig. 6b (marked with a red dotted circle) were recognized as a pearlite region because they are also displayed in black. Nevertheless, the overall segmentation was successful considering the complexity of the input microstructure.

Lastly, the segmentation was conducted for a micrograph with a much lower resolution. Figure 7a shows the input micrograph of a typical low-carbon microstructure with only 360×360 pixels. The microstructure is composed of grain boundary ferrite, ferrite side plate, and pearlite. Figure 7b shows the segmentation result with grain boundary ferrite in green, ferrite side plate in blue, and pearlite in light green. Unlike the previous cases, the segmented areas have rounded corners. This might be caused by the lack of information due to the low resolution. Other than that, the overall segmentation appeared to be successful even with the limited resolution.

In addition to the aforementioned segmentation results, other results for low-carbon steels with various microstructures are given in Fig. 8. They also indicate that the performance of the present method was successfully demonstrated.

Evaluation. Figure 9 shows an example of the comparison between manual and unsupervised segmentation results along with the fraction of their constituent areas. The image labeled by our experienced experts indicates

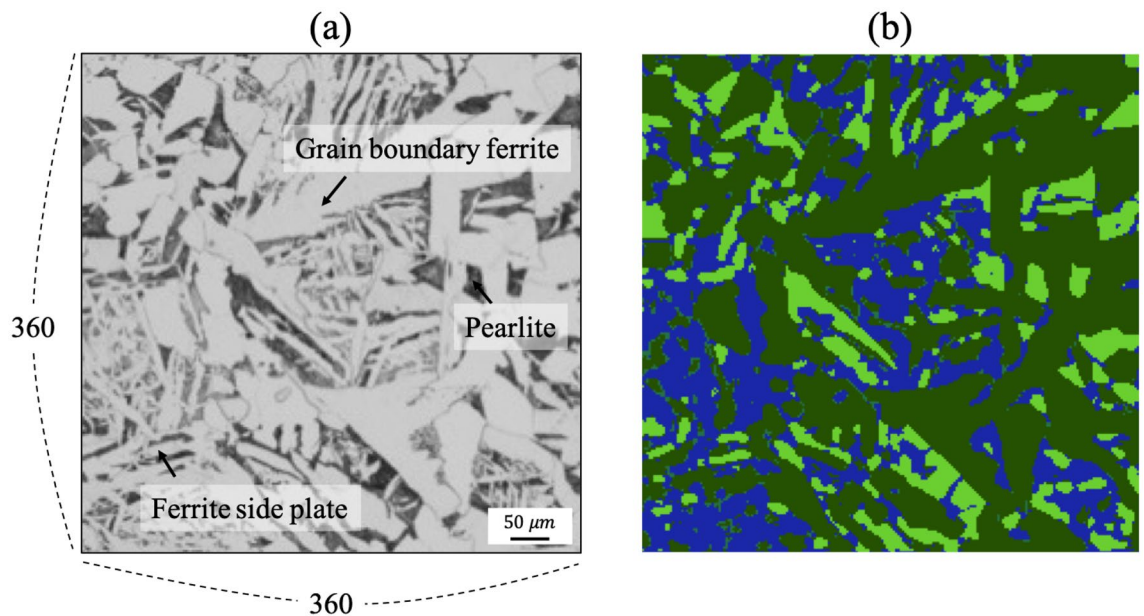


Figure 7. (a) Low resolution input microstructure image and (b) segmented image represented by various colors indicating different classes.

that the microstructure is composed of three different phases; 65% of martensite (M), 32% of ferrite (F), and 3% of pearlite (P). Likewise, the fraction of constituent regions of the segmented images obtained for three different numbers of classes are given in Fig. 9b–d. In the unsupervised segmentation results, the each region is not classified into a certain category. However, It is easily noticeable that the result consists of three classes (Fig. 9b) is consistent with the reference image in terms of the segmentation boundary. In addition, the fractions of the phases are in good agreement with the manual analysis. It is confirmed that the accuracy of the proposed unsupervised segmentation method is comparable to that of the previously reported supervised method¹⁰.

Since there can be numerous possible segmentation results for one microstructure if a hyperparameter, such as number of classes, is varied as clearly demonstrated in Fig. 9. Therefore, we need a measure to clarify which of the parameters is the most appropriate among given results. Since this deep learning method produces a different result each time owing to the randomness occurring for various reasons such as the weight initialization, each of the unsupervised evaluation criterion is averaged over ten times repetition for the evaluation of the different segmentation results. Figure 10a gives the evaluation results for Fig. 6a with increasing number of classes using the three unsupervised segmentation evaluation criteria. F and F' are lowest when there are three classes. This means that the microstructure is better to be divided into three regions. As the number of classes increases, the values also increase, meaning that quality of the result deteriorates. On the other hand, Q shows its minimum when there are four classes. In addition, it was observed that the Q with two classes is much higher than F and F' . This is because Q was designed to give a high penalty for regions with a large area having a very little variation in color³³. Therefore, when there were only two regions which inevitably include various colors, Q was higher.

Even though these criteria provide an easy way to evaluate the segmentation result, it is still difficult to compare criteria since their scales are different. In order to compare these criteria, their values were normalized to the range [0, 1] as shown in Fig. 10b. In addition, they were ranked according to their values as given in Table 2 in order to compare them with each other more easily. Note that normalization does not affect the rankings estimated by these evaluation criteria. A noticeable difference is that segmentation result with nine classes marked the third rank with Q criterion, which is not preferred with other criteria. As mentioned previously, Q disfavors large areas with various colors and prefers small areas with homogenous colors. Therefore, a larger number of classes is likely to be chosen with Q than with F and F' . Through these results, it was concluded that F is the most appropriate criterion for evaluating the unsupervised segmentation of a steel microstructure.

Hyperparameter optimization. As with the case of the number of classes, hyperparameters greatly affect the computation result. In this section, the effects of the compactness and the number of superpixels, which are essentially adjusted in superpixel segmentation processes, are investigated using the F criterion. The microstructure given in Fig. 6a is used as an input image and the number of classes are set as three. The segmentation is repeated 10 times for each pair of hyperparameters and the average F is taken. Figure 11 shows the dependence of F on the compactness and the number of superpixels. It is commonly observed that F decreases to a certain extent and then increases with increasing compactness. Since the compactness has a trade-off with the color

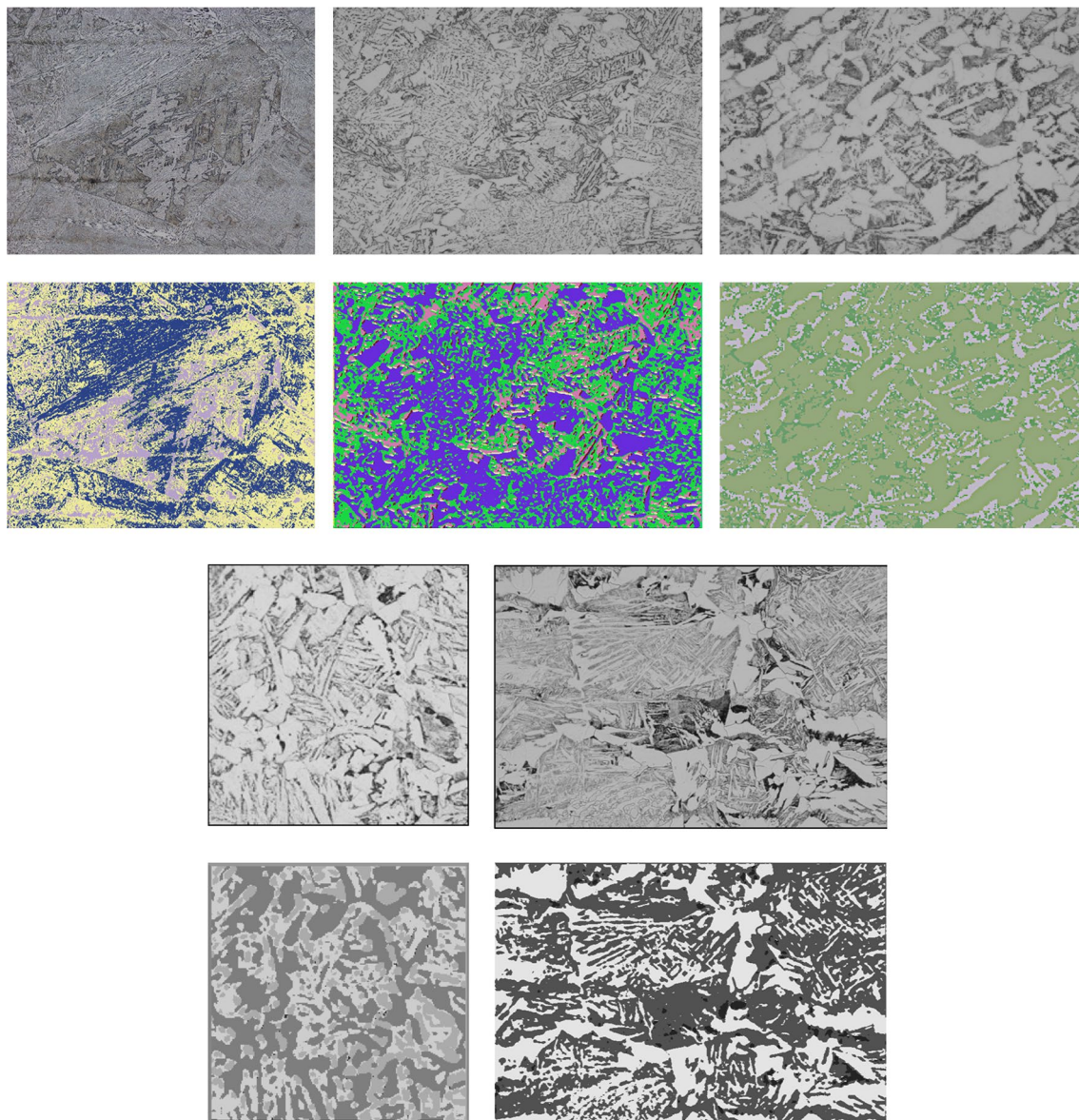


Figure 8. Examples of microstructure images (upper row) segmentation results (lower row).

similarity and spatial proximity, a very high or low compactness makes an algorithm yield superpixels based on limited information, resulting in failure to detect the region boundary³⁴. In terms of the number of superpixels, a similar tendency is shown that a very high or low number reduces the quality of segmentation; extreme values for the number of superpixels result in the incorrect segmentation of pixels^{35–37}. In conclusion, avoiding very high or too low values for the compactness and number of superpixels will provide a better segmentation result.

Conclusion

We demonstrated the segmentation of the microstructure of a low-carbon steel without labeled images using a deep learning method. Specifically, CNN and the SLIC superpixel algorithm were introduced for pixelwise segmentation. Various microstructure images of steel composed of ferrite, pearlite, bainite, and martensite are used and regions of constituent phases were well distinguished. In addition, the quality of the segmentation results was assessed on the basis of various unsupervised segmentation evaluation criteria. We found that the F criterion shows better performance than the F' and Q criteria for the segmentation of the steel microstructure. Finally, the effect of hyperparameters on the segmentation was investigated and it was found that medium values are desirable for good performance. It is concluded that the deep-learning-based approach is efficient and fast method in distinguishing various microstructural features of low-carbon steels without the need to create labeled images.

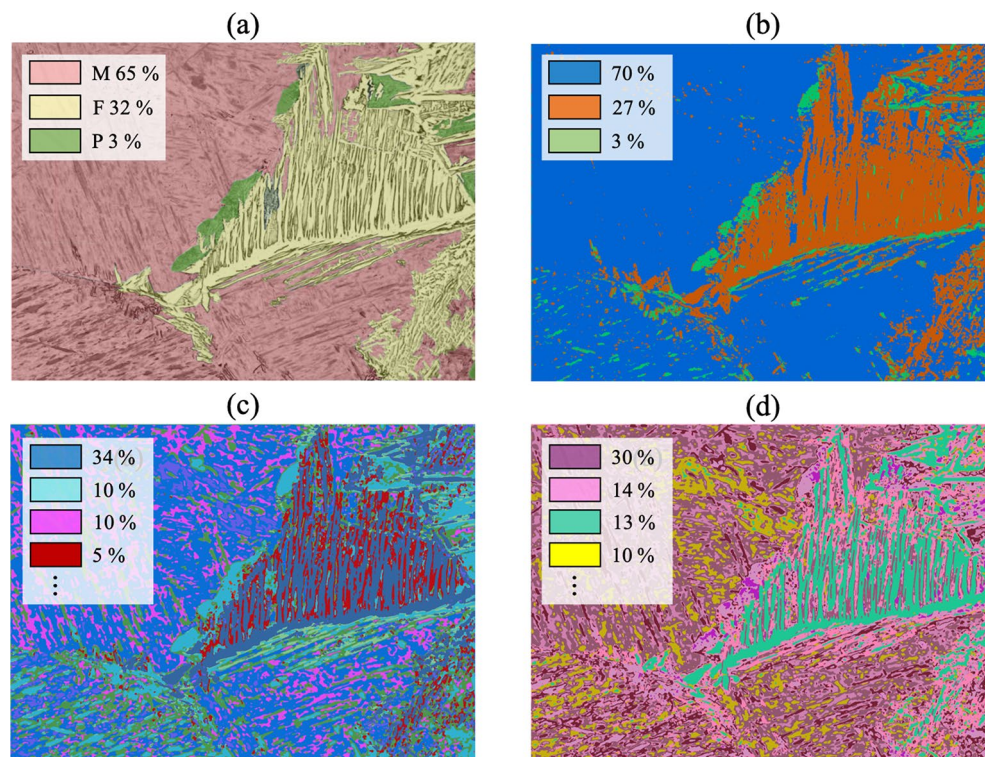


Figure 9. Comparison of (a) the labeled image consists of martensite (M), ferrite (F), and pearlite (P) phases created by manual classification and the segmented images with different hyperparameter settings for the number of classes; (b) 3 classes, (c) 10 classes, and (d) 20 classes.

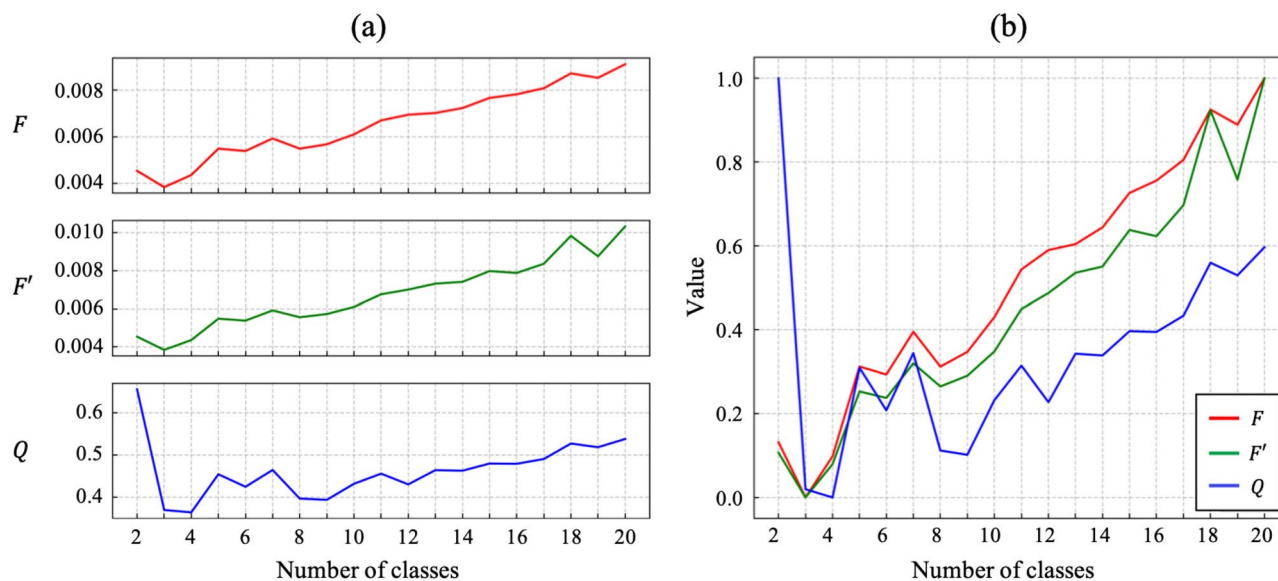


Figure 10. (a) Estimated F, F', and Q values for segmentation results with different numbers of classes and (b) their corresponding values normalized to the range [0, 1].

Rank	1st	2nd	3rd	4th	5th
F	3 (0)	4 (0.10)	2 (0.13)	6 (0.30)	5 (0.31)
F'	3 (0)	4 (0.08)	2 (0.11)	6 (0.24)	5 (0.25)
Q	4 (0)	3 (0.02)	9 (0.10)	8 (0.11)	6 (0.21)

Table 2. Numbers of classes and their values arranged in order of the rank determined from the evaluation criteria.

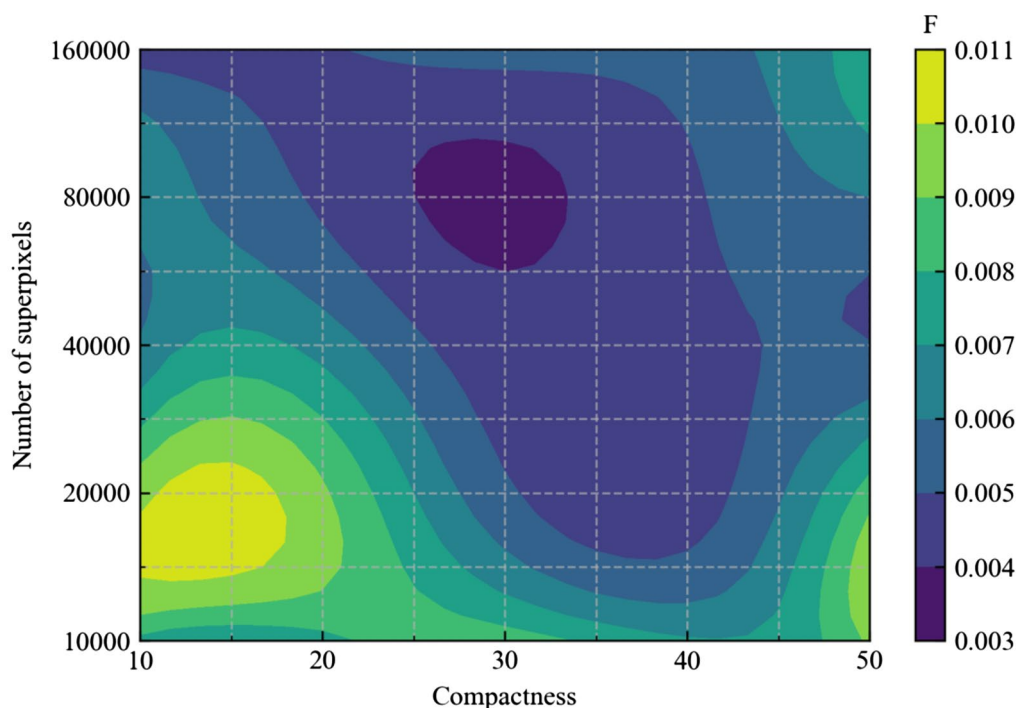


Figure 11. Contour map of F as a function of compactness values and numbers of superpixels.

Data availability

Source code of the algorithm described in this paper is available upon request.

Received: 8 July 2020; Accepted: 8 October 2020

Published online: 20 October 2020

References

- Bhadeshia, H. & Honeycombe, R. *Steels: Microstructure and Properties*. (Butterworth-Heinemann, 2017).
- Lai, Q. *et al.* Influence of martensite volume fraction and hardness on the plastic behavior of dual-phase steels: Experiments and micromechanical modeling. *Int. J. Plast.* **80**, 187–203 (2016).
- Bag, A., Ray, K. K. & Dwarakadasa, E. S. Influence of martensite content and morphology on tensile and impact properties of high-martensite dual-phase steels. *Metall. Mater. Trans. A* **30**, 1193–1202 (1999).
- Thewlis, G. Classification and quantification of microstructures in steels. *Mater. Sci. Technol.* **20**, 143–160 (2004).
- Choi, K., Koo, K. & Lee, J. Development of defect classification algorithm for POSCO rolling strip surface inspection system. in *SICE-ICASE International Joint Conference 2499–2502* (2006). <https://doi.org/10.1109/SICE.2006.314681>.
- Cortes, C. & Vapnik, V. Support-vector networks. *Mach. Learn.* **20**, 273–297 (1995).
- Gola, J. *et al.* Advanced microstructure classification by data mining methods. *Comput. Mater. Sci.* **148**, 324–335 (2018).
- Gola, J. *et al.* Objective microstructure classification by support vector machine (SVM) using a combination of morphological parameters and textural features for low carbon steels. *Comput. Mater. Sci.* **160**, 186–196 (2019).
- Breimen, L. Random forests. *Mach. Learn.* **45**, 5–32 (2001).
- Bulgarevich, D. S., Tsukamoto, S., Kasuya, T., Demura, M. & Watanabe, M. Pattern recognition with machine learning on optical microscopy images of typical metallurgical microstructures. *Sci. Rep.* **8** (2018).
- Bulgarevich, D. S., Tsukamoto, S., Kasuya, T., Demura, M. & Watanabe, M. Automatic steel labeling on certain microstructural constituents with image processing and machine learning tools. *Sci. Technol. Adv. Mater.* **20**, 532–542 (2019).
- Gupta, S., Sarkar, J., Kundu, M., Bandyopadhyay, N. R. & Ganguly, S. Automatic recognition of SEM microstructure and phases of steel using LBP and random decision forest operator. *Measurement* **151**, 107224 (2020).
- LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
- de Albuquerque, V. H. C., Cortez, P. C., de Alexandria, A. R. & Tavares, J. M. R. S. A new solution for automatic microstructures analysis from images based on a backpropagation artificial neural network. *Nondestruct. Test. Eval.* **23**, 273–283 (2008).

15. de Albuquerque, V. H. C., de Alexandria, A. R., Cortez, P. C. & Tavares, J. M. R. S. Evaluation of multilayer perceptron and self-organizing map neural network topologies applied on microstructure segmentation from metallographic images. *NDT E Int.* **42**, 644–651 (2009).
16. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324 (1998).
17. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv14091556 (2014).
18. Azimi, S. M., Britz, D., Engstler, M., Fritz, M. & Mücklich, F. Advanced steel microstructural classification by deep learning methods. *Sci. Rep.* **8** (2018).
19. Huang, G., Liu, Z., Van Der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. *IEEE Conf. Comput. Vis. Pattern Recognit. CVPR* 2261–2269 (2017). <https://doi.org/10.1109/CVPR.2017.243>.
20. Roberts, G. *et al.* Deep learning for semantic segmentation of defects in advanced STEM images of steels. *Sci. Rep.* **9** (2019).
21. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* 770–778 (2016). <https://doi.org/10.1109/CVPR.2016.90>.
22. Jang, J. *et al.* Residual neural network-based fully convolutional network for microstructure segmentation. *Sci. Technol. Weld. Join.* **25**, 282–289 (2020).
23. Kanazaki, A. Unsupervised image segmentation by backpropagation. *IEEE Int. Conf. Acoust. Speech Signal Process. ICASSP* 1543–1547 (2018) <https://doi.org/10.1109/ICASSP.2018.8462533>.
24. Ioffe, S. & Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv150203167 (2015).
25. Santurkar, S., Tsipras, D., Ilyas, A. & Madry, A. How does batch normalization help optimization? *Adv. Neural Inf. Process. Syst.* 2483–2493 (2018).
26. Achanta, R. *et al.* Slic superpixels. in *EPFL Technical Report No. 149300* (2010).
27. Ren, X. & Malik, J. Learning a classification model for segmentation. *Proc. Ninth IEEE Int. Conf. Comput. Vis.* **1**, 10–17 (2003).
28. Liu, M.-Y., Tuzel, O., Ramalingam, S. & Chellappa, R. Entropy rate superpixel segmentation. in *CVPR 2011 IEEE* 2097–2104 (2011).
29. Van den Bergh, M., Boix, X., Roig, G., de Capitani, B. & Van Gool, L. Seeds: Superpixels extracted via energy-driven sampling. *Eur. Conf. Comput. Vis.* 13–26 (2012).
30. Achanta, R. *et al.* SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans. Pattern Anal. Mach. Intell.* **34**, 2274–2282 (2012).
31. Liu, J. & Yang, Y. Multiresolution color image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **16**, 689–700 (1994).
32. Borsotti, M., Campadelli, P. & Schettini, R. Quantitative evaluation of color image segmentation results. *Pattern Recognit. Lett.* **19**, 741–747 (1998).
33. Zhang, H., Fritts, J. E. & Goldman, S. A. Image segmentation evaluation: A survey of unsupervised methods. *Comput. Vis. Image Underst.* **110**, 260–280 (2008).
34. Lv, X., Ming, D., Chen, Y. & Wang, M. Very high resolution remote sensing image classification with SEEDS-CNN and scale effect analysis for superpixel CNN classification. *Int. J. Remote Sens.* **40**, 506–531 (2019).
35. Xu, Y. *et al.* Efficient optic cup detection from intra-image learning with retinal structure priors. *Int. Conf. Med. Image Comput. Comput.-Assist. Interv.* **7510**, 58–65 (2012).
36. Zhao, Y. *et al.* Intensity and compactness enabled saliency estimation for leakage detection in diabetic and malarial retinopathy. *IEEE Trans. Med. Imaging* **36**, 51–63 (2017).
37. Jiang, J. *et al.* SuperPCA: A superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **56**, 4581–4593 (2018).

Acknowledgements

This work was supported by the Council for Science, Technology and Innovation, Cross-ministerial Strategic Innovation Promotion Program (SIP), “Structural Materials for Innovation” (funding agency: JST).

Author contributions

H.K. and J.I. wrote the main manuscript text and T.K. prepared micrographs used in the present study. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.I.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020