



Supervised pearlitic–ferritic steel microstructure segmentation by U-Net convolutional neural network

Mateusz Motyl¹ · Łukasz Madej¹

Received: 8 July 2022 / Revised: 1 September 2022 / Accepted: 6 September 2022 / Published online: 21 September 2022
© The Author(s) 2022

Abstract

The aim of this work is to develop an automated procedure based on machine learning capabilities for the identification of the pearlite islands within the two-phase pearlitic–ferritic steel. The input parameters for the custom implementation of a braided neural network are provided as a data set of scanning electron microscopy images of metallographic specimens. The procedures related to the processing of the data and the optimization parameters affecting the final architecture and effectiveness of the network learning stage are examined. The objective is to find the best solution to the problem of ferritic–pearlitic microstructure segmentation, allowing further processing during, e.g., 3D reconstruction of data from serial sectioning. The work examines the various quality of input data and different U-Net architectures to find the one that can identify pearlite islands with the highest precision. Two types of images acquired from secondary electron (SE) and electron backscattered diffraction (EBSD) detectors are used during the investigation. The work revealed that the developed approach offers improvements in metallographic investigations by removing the requirement for expert knowledge for the interpretation of image data prior to further characterization. It has also been proven that artificial neural networks based on the deep learning process using extensible U-Net network architectures and nonlinear learning tools can identify pearlite islands within a two-phase microstructure, while the overtraining level remains low. Convolutional neural networks do not require manual feature extraction and are able to automatically find appropriate search functions to recognize pearlite structure areas in the training process without human intervention. It was shown that the network recognizes areas of analyzed steel with satisfactory precision of 79% for EBSD and 87% for SE images.

Keywords Artificial neural networks · Supervised learning · Microstructure characterization

1 Introduction

Steels remain one of the most important and extensively used classes of materials because of their excellent mechanical properties delivered at relatively low costs. However, the mechanical properties of steels are directly related to their microstructure and morphology. Depending on the cooling rate of steels, the ferrite (F), pearlite (P), bainite (B), or martensite (M) microstructures could be formed due to the displacive and reconstructive

transformations of austenite (A) crystal structure, which are accompanied with cementite precipitation at different diffusion rates [1]. It is well known that the volume fraction, dimension, and morphology of these microstructure phases are greatly responsible for the mechanical properties of steels [2]. In this case, the correct classification of these microstructure features is crucial during the metallographic investigation to understand their role in material performance. Before most scientific and engineering investigations, microstructural features of interest have to be detected and identified according to a representation suitable for quantification done, e.g., by dedicated software. This is achieved through segmentation, the process of partitioning an image into multiple homogeneous regions or segments. Segmentation constitutes a major transition in the image analysis pipeline, replacing pixels' intensity values with region labels [3]. However, such a procedure is most of the time difficult, as the metallographic analysis

✉ Łukasz Madej
lmadej@agh.edu.pl

Mateusz Motyl
mateusz.motyl@agh.edu.pl

¹ Faculty of Metals Engineering and Industrial Computer Science, AGH University of Science and Technology, Mickiewicza 30 av., 30-059 Krakow, Poland

requires expert knowledge, which significantly extends the investigation time. Classical segmentation methods like Otsu binarization or watershed offer a lot of options for automation [4]; however, only for microstructures with clearly distinguishable phases at the image, e.g., an optical image of ferrite and martensite in dual-phase steels [5]. These approaches fail when the distinction between the features is not straightforward, like in the case of common ferritic–pearlitic steels observed under electron microscopy. To overcome that issue, the capabilities of alternative solutions based on artificial neural networks are evaluated in the current work.

Modern computer systems analyze more and more data in order to obtain useful information and formulate conclusions during, e.g., decision-making processes [6]. Despite continuous development, many systems for searching and recommending, e.g., multimedia content, rely mainly on manually prepared metadata—descriptors, category labels, and other attributes concisely representing their content. The process of creating this type of metadata is extremely tedious and time-consuming. The problem, which is related to the amount of processed data and the desired details of the description, quickly encounters limitations in terms of available human resources. It becomes necessary to generate the mentioned attributes in an automated way. However, for this to be possible, computer systems have to analyze media files at the level of their actual content. Understanding the content of images also opens the way to many applications beyond finding relevant characters or identifying patterns [7, 8]. In addition, for digital images, a number of parameters such as resolution, bit depth, data compression ratio, or color mode also have to be considered. The processing results may also be affected by the environment in which the images were taken or, finally, by the different settings of the object under study. Each of the mentioned problems may negatively influence the quality of metadata and, consequently, the efficiency of the system used for their analyses. Therefore, during the preparation of information, several image transformations are additionally performed to normalize the data [9].

One of the drivers of innovation in the field of image recognition is the continuous increase in software quality and hardware performance. A breakthrough change in this aspect turned out to be the popularization of GPUs for general applications, opening the way for the creation of models with much higher computational complexity [10]. The use of such processors also allowed the development of machine learning and deep learning approaches. The idea of using multi-layer artificial neural network (ANN) models is not new, as its beginnings can be traced back to the 1970s [11]. However, only the combination of computational capabilities and several new, seemingly minor innovations allowed a qualitative leap in this area. As a result, the practical

application of the created solutions increased in many areas, and they ceased to be exclusively academic [12].

Deep neural networks are now widely used in image and pattern processing. Examples of applications include face recognition, medical diagnosis, classification of material structure or defects, or detection of environmental objects [13]. Convolutional neural networks (CNN) are of particular use in image processing operations [14]. Wherever there is a need to focus on fragments containing specific and relatively rare events, automatic content analysis methods are an invaluable tool for sifting through increasingly large digital archives [15].

Presented progress in the CNN possibilities in the area of image processing became a motivation for their application in material science investigations, especially in the area of metallographic analysis [16]. As mentioned earlier, in this case, expert knowledge is often the only reliable approach to extracting meaningful information from the microstructure images. For single images, such a procedure is satisfactory; however, when a large number of images are obtained, for example, during the serial sectioning procedure, manual feature extraction becomes extremely laborious [17]. One of the first successful applications of a neural network approach to cast iron microstructure processing is presented [18]. Segmentation of steel microstructures with specific morphological features was also investigated by other approaches based on, e.g., ResNet [19] or VGGNet [20]. An interesting approach to developing an artificial neural network segmentation tool for microstructures based on a numerically generated training data set was also proposed [21]. The U-Net architecture is recently often used for the segmentation of different types of microstructures, e.g., martensite–ferrite in dual-phase steels [22], martensite–austenite in bainitic steels [23]. Authors of [24] applied the same approach for ferrite, bainite and martensite segmentation in low carbon steel. An interesting systematic review of the segmentation of metallographic images is presented in [25].

These approaches proved that the neural networks could be used to support metallographic investigation when a large number of data from, e.g., mentioned serial sectioning or computed tomography analysis [26], needs to be processed.

Application of a serial sectioning procedure for imaging of large 3D microstructure volume is not often used, as it is usually based on time-consuming and very precise manual labor. However, at the same time, such an approach provides a significant amount of information about microstructure morphology. Additionally, if combined with the electron backscattered diffraction (EBSD) investigation, obtained texture information could be used as input data for advanced microscale modeling approaches [27].

However, EBSD-based approach encounters difficulties when two-component materials of similar crystallographic structure are considered like investigated within the work

ferritic–pearlitic steels. In this case, the EBSD investigation cannot provide information for separation between the two constituents as described in [17]. The alternative solution that was used to overcome this limitation is based on a combination of information from the EBSD investigation with a series of 2D scanning electron microscopy (SEM/SE2) images of a ferrite–pearlite microstructure to clearly identify the morphology of both constituents. Unfortunately, there is no robust approach that can perform such identification automatically in a reliable manner. The identification of pearlite islands has to be done manually [17], which becomes tedious, especially for a large number of 2D images. Therefore, as mentioned, the evaluation of the capabilities of deep neural networks in image processing is explored in current research. The general idea and methodology of the work are based on three stages, i.e., data preparation, data processing, and model optimization tasks, as presented in Fig. 1.

The work, in particular, examines the various quality of input data and different U-Net architectures to find the one that can identify pearlite islands with the highest precision. Two types of images acquired from secondary electron (SE)

and electron backscattered diffraction (EBSD) detectors are analyzed during the investigation.

2 Data pre-processing

Details of the image acquisition process during the EBSD-based serial sectioning procedure are described in [17], while only major steps are summarized below for the clarity of this research. The ferritic–pearlitic steel sample (0.12%C, 0.35%Mn, 0.15%Si, 0.03%P, 0.03%S) was polished with 3 and 1 μm diamond suspension, and the final step with colloidal silica. During the serial sectioning procedure, the polishing times were adjusted to remove approx. 3 μm layer of the material. After each polishing operation, the sample was subjected to metallographic investigation within the scanning electron microscopy. The high-resolution secondary electron image was taken first, and then the EBSD map was acquired. EBSD data were further used to calculate the image quality (IQ) maps that can be used to extract information on the pearlite morphology (Fig. 2). Both the SE

Fig. 1 Operations in subsequent stages of ANN model development

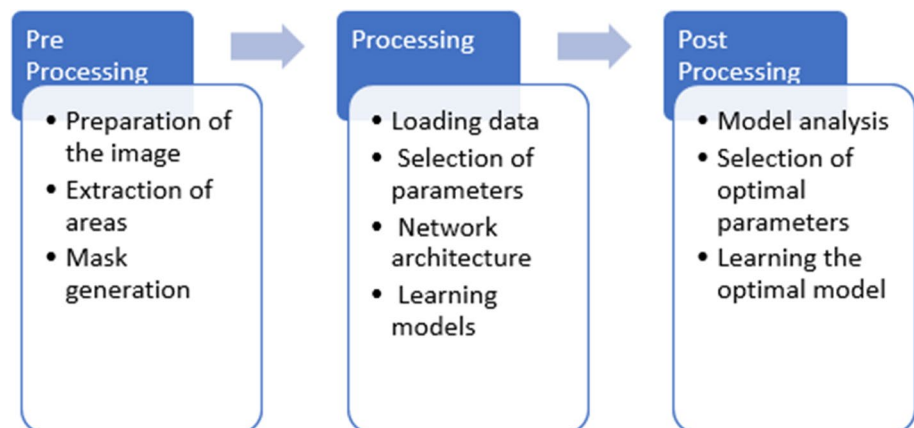
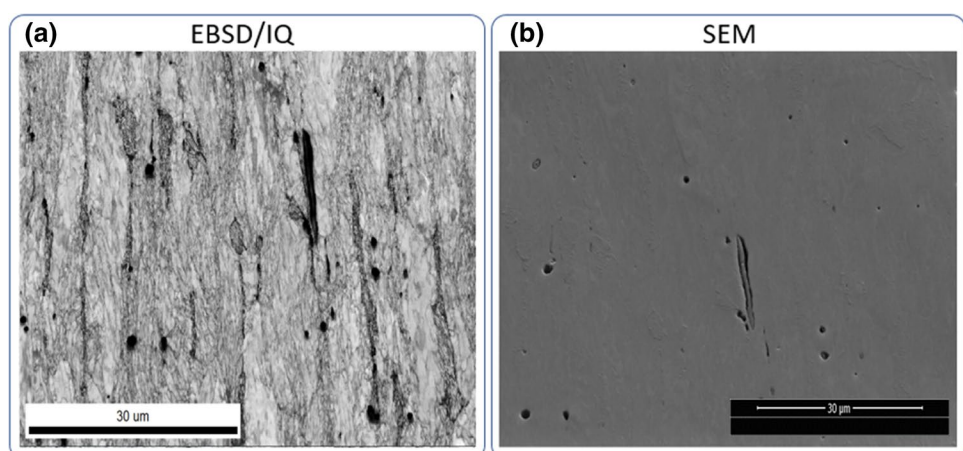


Fig. 2 Examples of images of ferritic–pearlitic steel acquired from the **a** EBSD/IQ, **b** SEM/SE detectors



and EBSD/IQ images were used to evaluate the influence of different types of input data on the ANN learning process.

The preparation of an appropriate training dataset is a crucial step in performing a proper neural network learning process. Even the most advanced network architecture models will not produce measurable results with poor quality input data. In the case of image segmentation to extract the morphology of the pearlite islands from the ferrite matrix, the most important is the proper preparation of the binary mask, which represents the shape and location of the pearlite. The preparation of the training dataset was divided into several steps.

For the preliminary dataset preparation stage, the laborious manual extraction of pearlite island areas was performed from both EBSD/IQ and SEM/SE images obtained

during the serial sectioning procedure (Fig. 3). The manually marked contours were subjected to the in-house filling algorithm as presented in Fig. 4 to provide features masks required for further processing.

The initial data set contained 2×38 microstructural images with resolutions 1536×1103 px and 1280×1024 px, respectively. Then, the data augmentation techniques were applied to increase the training dataset. Each image was divided into smaller areas with sizes 256×256 , 512×512 , and 1024×1024 px. The set of rotations with respect to the $\pm 90^\circ$ and 180° was then applied. Finally, each of the rotated images was also symmetrically reflected relative to the X -axis, Y -axis, and both the X -, Y -axis simultaneously. The augmentation procedure delivered the training dataset with 2720, 1360, and 340 images with respect to their

Fig. 3 Process of manual contours extraction procedure to identify pearlite islands for **a** SE, **b** EBSD/IQ detectors. The filled contours represent markers from the serial sectioning procedure

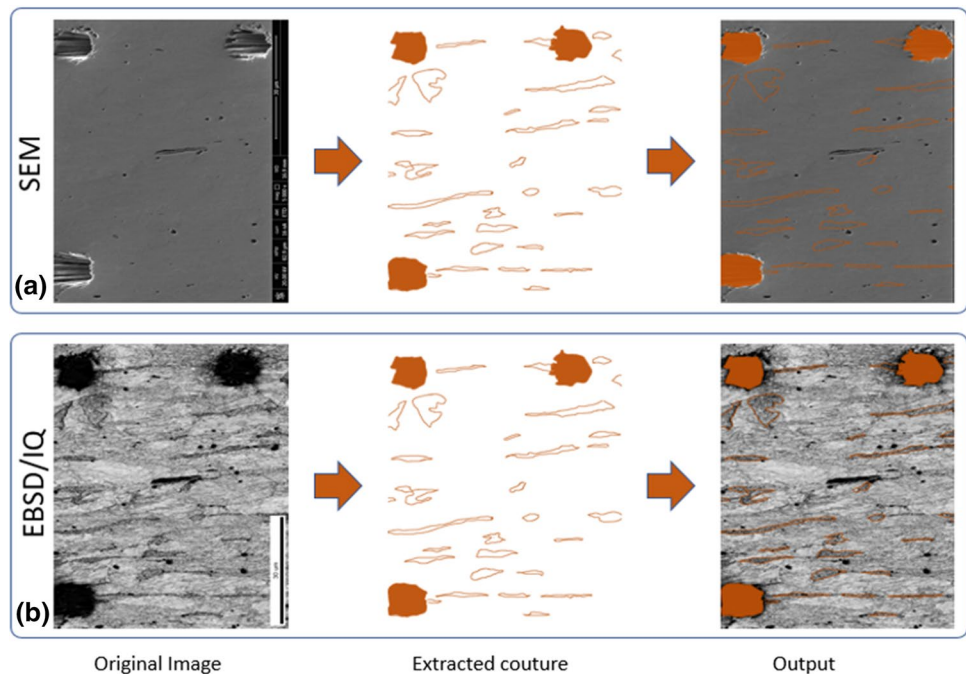
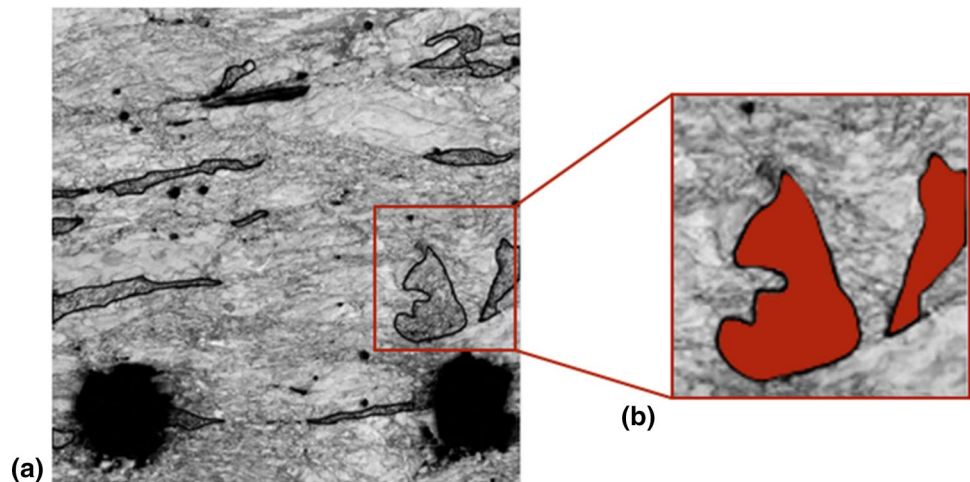


Fig. 4 Illustration of **a** manual identification process of the pearlite islands, **b** filling the pearlite contour to create the mask



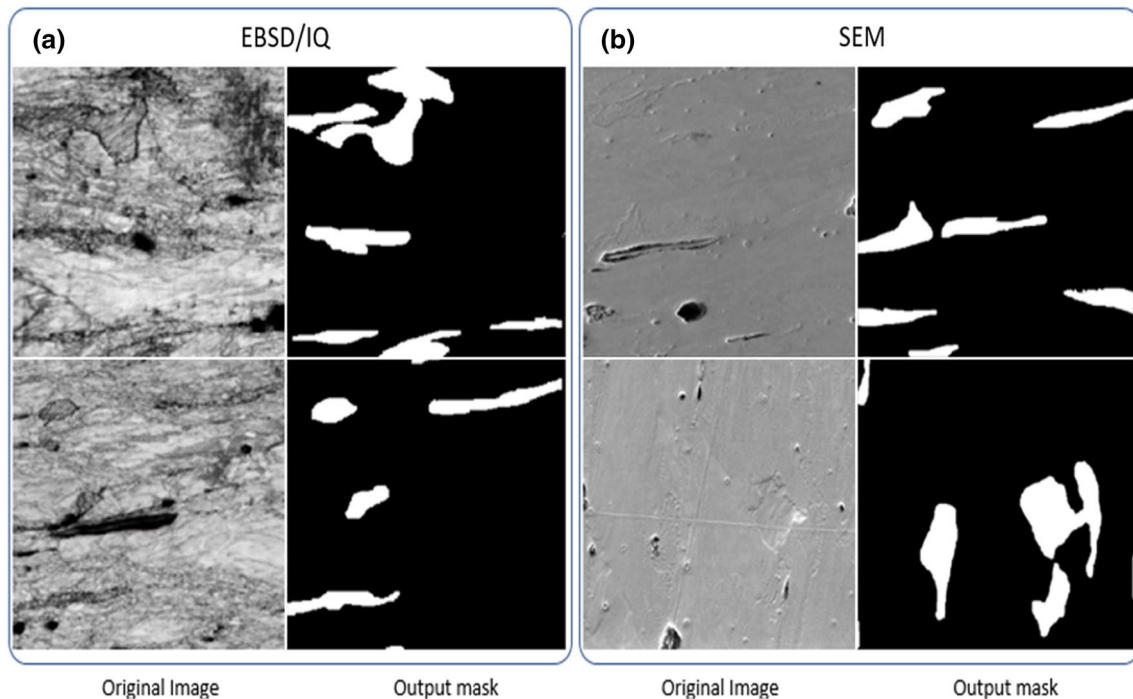


Fig. 5 Examples of ground truth masks generated from **a** EBSD/IQ, **b** SE images

Table 1 Data set size

No.	Image size	Input file extension	Data set size images
1.1	256×256	SE	2720
1.2	256×256	EBSD/IQ	2720
2.1	512×512	SE	1360
2.2	512×512	EBSD/IQ	1360
3.1	1024×1024	SE	340
3.2	1024×1024	EBSD/IQ	340

resolution. The final stage of the image preparation involved binarization operation to provide the ground truth masks for the neural network. Examples of the final binary images in the form of the ground truth masks processed from SE and EBSD/IQ images are shown in Fig. 5.

Six input data sets were developed according to the above-presented procedure for the initial analysis, as shown in Table 1.

To perform supervised learning of a neural network, it is necessary to divide the data into at least two parts: a learning set and a validation set (Fig. 6). Very often, the third subset of cases, the test set, is also separated from the collection, which is used for the final evaluation of the network quality (generalization ability). The main problem with the separation of subsets is the need to have each set of learning cases

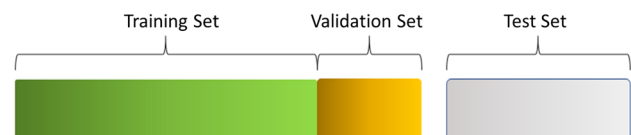


Fig. 6 Concept of the data division between training, validation, and test data sets

representative of the entire set. If unique cases are separated from the learning set, the model will not be able to correctly predict their properties. On the other hand, if standard cases are selected with very close or nearly identical counterparts in the learning set, the whole network quality assessment procedure will be ineffective. Even the overlearned model will get very good predictions during validation and testing. A learning group is a set of data that is used to train the ANN. The model learns to classify the features appropriately and builds some dependencies based on this data. Therefore, it predicts possible outcomes and makes decisions based on the given data. A validation group is a dataset used to run an unloaded test of the trained model. It is important that the data in the validation set is not previously used for model learning, as it will then not be suitable for unbiased, unencumbered testing. Finally, a test group becomes important once the model and hyperparameters are defined. The set is used to investigate the ANN predictions with the data not used previously during the

investigation. It is crucial that this data set has not been used before for learning or model validation.

The use of independent learning, validation, and testing groups is essential when a model contains many parameters and potential hyperparameters. To select the best possible model and the best possible data, one has to be very careful not to overfit the model. The more models, hyperparameters, and other options are analyzed, the larger the validation set is also required [28].

Therefore, from the developed set of images, training, validation, and testing data sets were extracted as presented in Table 2 from Chapter 3. The latter was not considered during the training process in the processing stage.

3 Processing

The processing stage is based on the concept of the U-Net convolutional deep learning neural network [28]. In the processing stage, the dataset was divided, and the parameters for the neural network model were tested to select the most promising architecture. The change of activation function, dropout rate, and size of each convolution layer were evaluated during the research. The training process setup parameters that were analyzed are gathered in Table 2. The results from the training process are evaluated based on the accuracy metrics (ACC), which represent the accuracy of the binary mask obtained in the inference process [29]. A confusion matrix provides a summary of the predictive results in a classification problem. Correct and incorrect predictions are summarized in a table with their values and are broken down by each class (Fig. 7).

The ACC is defined as follows:

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

While the loss function, which is a Binary Cross-entropy (LOSS) has the form:

$$LOSS = -\frac{1}{N} \sum_{i=0}^N y_i \log \hat{y}_i + (1 - y_i) \log (1 - \hat{y}_i), \tag{2}$$

where \hat{y}_i the i -th scalar value in the model output, y_i the target value, N the number of scalar values in the model output.

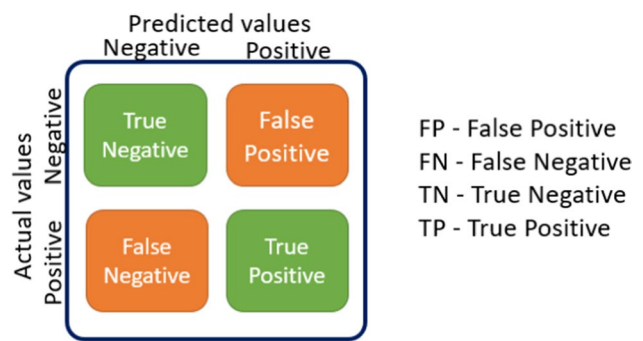


Fig. 7 The concept of the correct and incorrect predictions

The U-Net CNN structure consists of contracting and expansive paths, as shown in Fig. 8. The contracting path starts as a typical convolutional network architecture [30, 31]. It consists of the repeated application of two 3×3 convolutions (unpadded convolutions), each followed by a batch normalization layer and a rectified linear unit (ReLU). A 2×2 max-pooling operation with a stride of 2 for down-sampling, and after each down-sampling step, the number of feature channels is doubled. In contrast to the contracting path, the expansive path is composed of an up-sampling of the feature map followed by a 2×2 convolution (i.e., “up-convolution”) that halved the number of feature channels. Then a concatenation with the correspondingly cropped feature map from the contracting path is used, and two 3×3 convolutions, each also followed by batch normalization and ReLU. The cropping is necessary due to the loss of border pixels in each convolution. At the final layers, 4×4 , 2×2 and 1×1 convolutions are used to map each 8-component feature vector to the desired number of classes and to map the feature to the required value.

Therefore, many image features are used in the contracting and expansive paths to reconstruct a new image of the same size as the input one. The final developed implementation of our U-Net is shown in Fig. 8. The batch normalization, ReLU, and dropout approach with sigmoid activation in the last layer, which is different from the classical U-Net [28], are used in the current research.

Table 2 Training process setup parameters

No.	Image size	Training, validation and test data set size (%)	Epoch	Batch size (%)
Scenario 1.1	256 × 256	[60/20/20][70/15/15][80/10/10]	[30][50][70]	[5][10]
Scenario 1.2	256 × 256	[60/20/20][70/15/15][80/10/10]	[30][50][70]	[5][10]
Scenario 2.1	512 × 512	[60/20/20][70/15/15][80/10/10]	[30][50][70]	[5][10]
Scenario 2.2	512 × 512	[60/20/20][70/15/15][80/10/10]	[30][50][70]	[5][10]
Scenario 3.1	1024 × 1024	[60/20/20][70/15/15][80/10/10]	[30][50][70]	[5][10]
Scenario 3.2	1024 × 1024	[60/20/20][70/15/15][80/10/10]	[30][50][70]	[5][10]

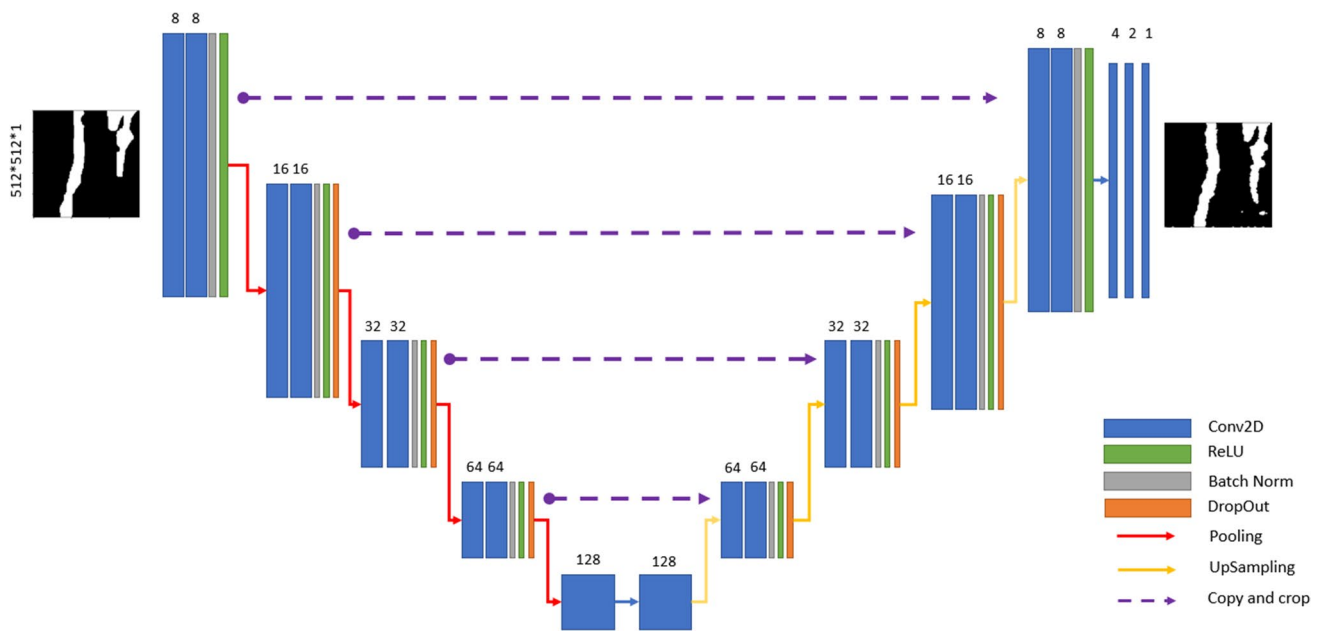


Fig. 8 The developed U-Net network architecture for microstructure image segmentation

The U-Net was developed in the Keras framework, which is a high-level neural network API (Application Programming Interface). Keras acts as an interface for the TensorFlow library [30, 32]. In total, the U-Net network is constructed from 19 convolutional layers. The input tile size so that all 2×2 max-pooling operations are applied to a layer with an even x - and y -sizes was selected.

Examples of training data results from the developed model according to scenarios from Table 2 are presented in Figs. 9, 10 and 11.

The best training test results without optimization were obtained for 512×512 px resolution, as shown in Table 3 and Fig. 12.

4 Post-processing

Finally, the post-processing operation is focused on the main elements that influence the final ANN setup under study. Parameter optimization provides an opportunity to distinguish important variables from information that does not contribute much to the network output, which will eventually be discarded. This stage is based on the parameters used in Scenario 2.1 and 2.2. As mentioned, they have a favorable size of the input data set compared to the resolution 1024×1024 px. It should also be mentioned that for Scenario 1.1 and 1.2, the ACC values are also satisfactory; however, in this case, a large part of the dataset contained masks without or with a small number of perlite areas.

When a deep neural network ends up going through a training batch, where it propagates the inputs through the layers, it needs a mechanism to decide how it will use the predicted results against the known values to adjust the parameters of the neural network. These parameters are commonly known as the weights and biases of the nodes within the hidden layers. Optimizers are classes or methods used to change the attributes of the deep learning model, such as weights and learning rate, to reduce the losses. Optimizers are necessary for the model to improve training speed and performance [30]. TensorFlow and Keras libraries mainly support nine optimizer classes, consisting of algorithms like Adadelta, Adam, RMSprop, and more [33]. Multiple optimizers have been investigated in this study. The learning rate has been set to 0.0001, batch size to 10 and epochs to 70.

The best-obtained results are shown in Table 4, and examples of identified perlite islands are in Fig. 13.

As seen from Table 4, in the best case of the post-processing stage, a decrease in classification error on the test dataset of more than 20% was achieved compared to the classification without prior parameter modifications.

Based on the qualitative comparison from Fig. 13, it can be stated that the developed U-Net architecture achieves an acceptable segmentation level of the two-phase ferritic–pearlitic microstructures. Additionally, from Table 4, it is visible that the neural network performs slightly better with the SE image set than with the EBSD/IQ image sets. It can also be summarized that the U-Net model for the first data set (SE) achieved an average ACC for Adam (adaptive moment estimation) algorithm of 87%, which is

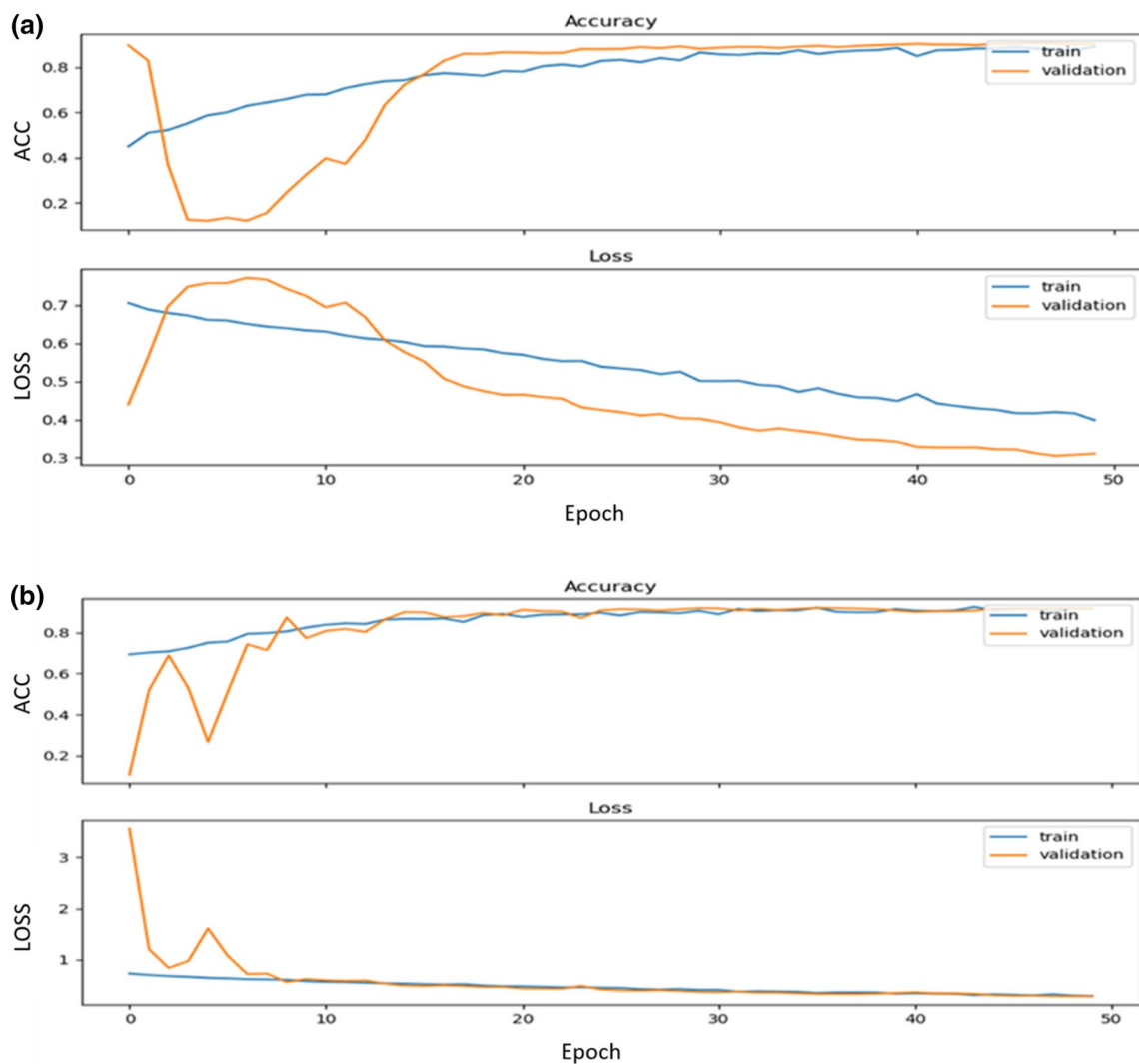


Fig. 9 Evolution of the ACC and LOSS indicators during the training process for **a** Scenario 1.1, **b** Scenario 1.2

better than the second-best algorithm AdaMax (extension to the Adam) with 81%. The second data set EBSD/IQ achieved an average ACC of 85.% for RMSprop, which is significantly better than the second-best algorithm with 75% for Adam in this case.

For quantitative image comparison, the SSIM (Structural Similarity Index) and PSNR (Peak Signal-to-Noise Ratio) [29] indicators were used. The PSNR is used to calculate the ratio between the maximum possible signal power and the power of the distorting noise, which affects the quality of its representation. This ratio between two images is computed in decibel form. The PSNR is usually calculated as the logarithm term of the decibel scale because the signals have a very wide dynamic range. This dynamic range varies between the largest and the smallest possible values, which are changeable by their quality. The PSNR is expressed as:

$$PSNR = 10 \log_{10} \left(\frac{P_v}{MSE} \right), \tag{3}$$

where P_v maximum possible pixel value of the image, MSE mean square root error between two images.

The SSIM index is a quality measurement metric calculated based on the computation of three major components termed luminance, contrast, and structural or correlation term [34]:

$$SSIM(x, y) = [l(x, y)]^\alpha [c(x, y)]^\beta [s(x, y)]^\gamma, \tag{4}$$

where l luminance (used to compare the brightness between two images), c contrast (used to differ the ranges between the brightest and darkest region of two images), s structure (used to compare the local luminance pattern between two

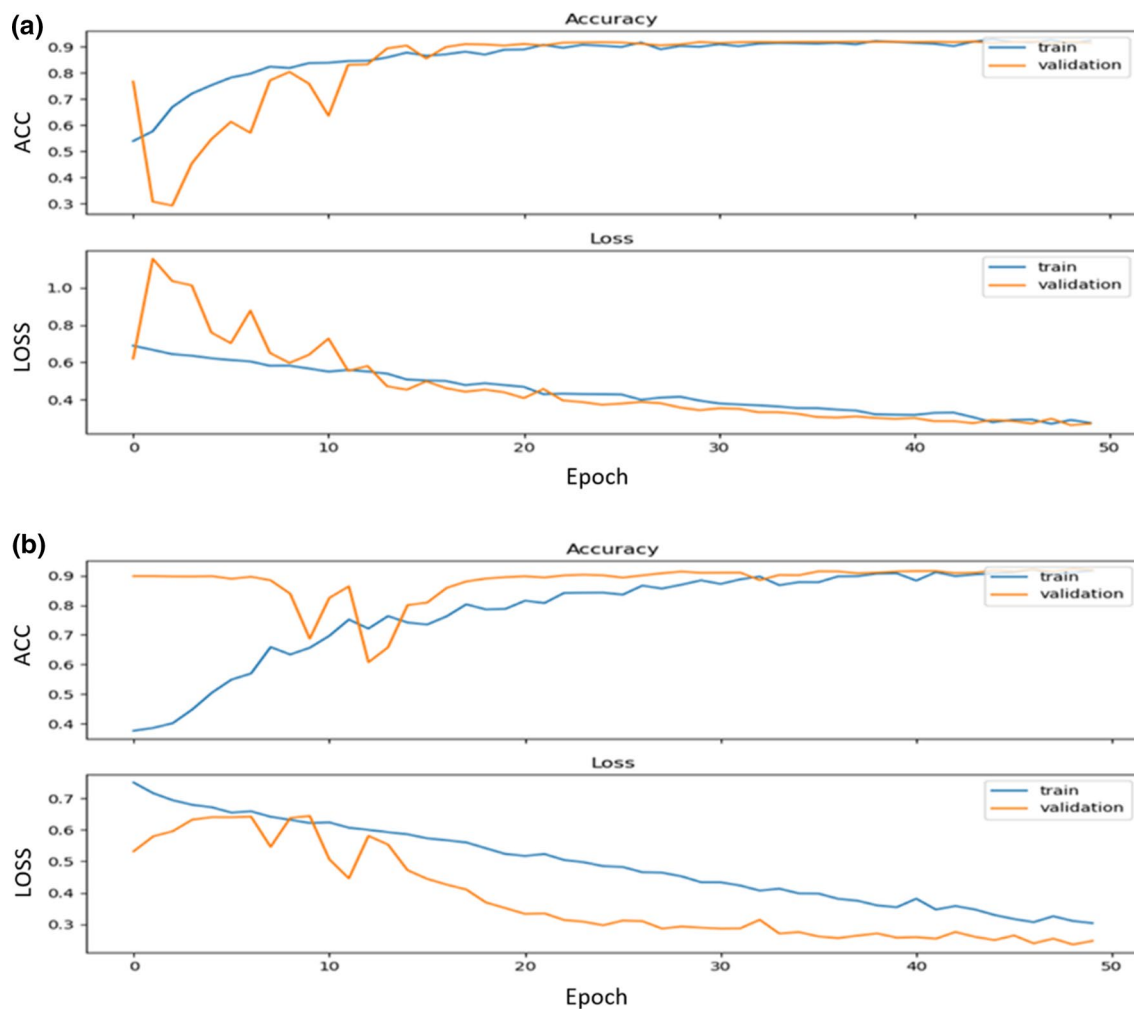


Fig. 10 Evolution of the ACC and LOSS indicators during the training process for **a** Scenario 2.1, **b** Scenario 2.2

images to find the similarity and dissimilarity of the images), α , β and γ are positive constants [33, 35].

The final validation of the predictions of the developed CNN setup is presented in a qualitative and quantitative manner in Fig. 14 and Table 5, respectively.

Regarding Table 5, one can conclude that the model has a better individual recognition performance for EBSD/IQ images but also more false-positive samples. For the SE images, results are mostly in a high level of accuracy.

Finally, it should be mentioned that thanks to data augmentation that was employed during the study, the developed ANN model needs only a few annotated images and a reasonable training time of 2 h on NVIDIA RTX 2080 (8 GB) to provide reliable results.

5 Conclusions

Based on the research, it can be stated that designed and developed deep convolutional neural network model enables successful automated microstructure segmentation of complex two-phase ferritic–pearlitic microstructures. During the research, the influence of the image acquisition technique on network predictions was also evaluated. Images from the SE and EBSD/IQ detectors were tested with various network setups. It was also identified that the image pre-processing stage is crucial. Evaluating different parameters, especially data augmentation setup and the image size that affect the network training procedure,

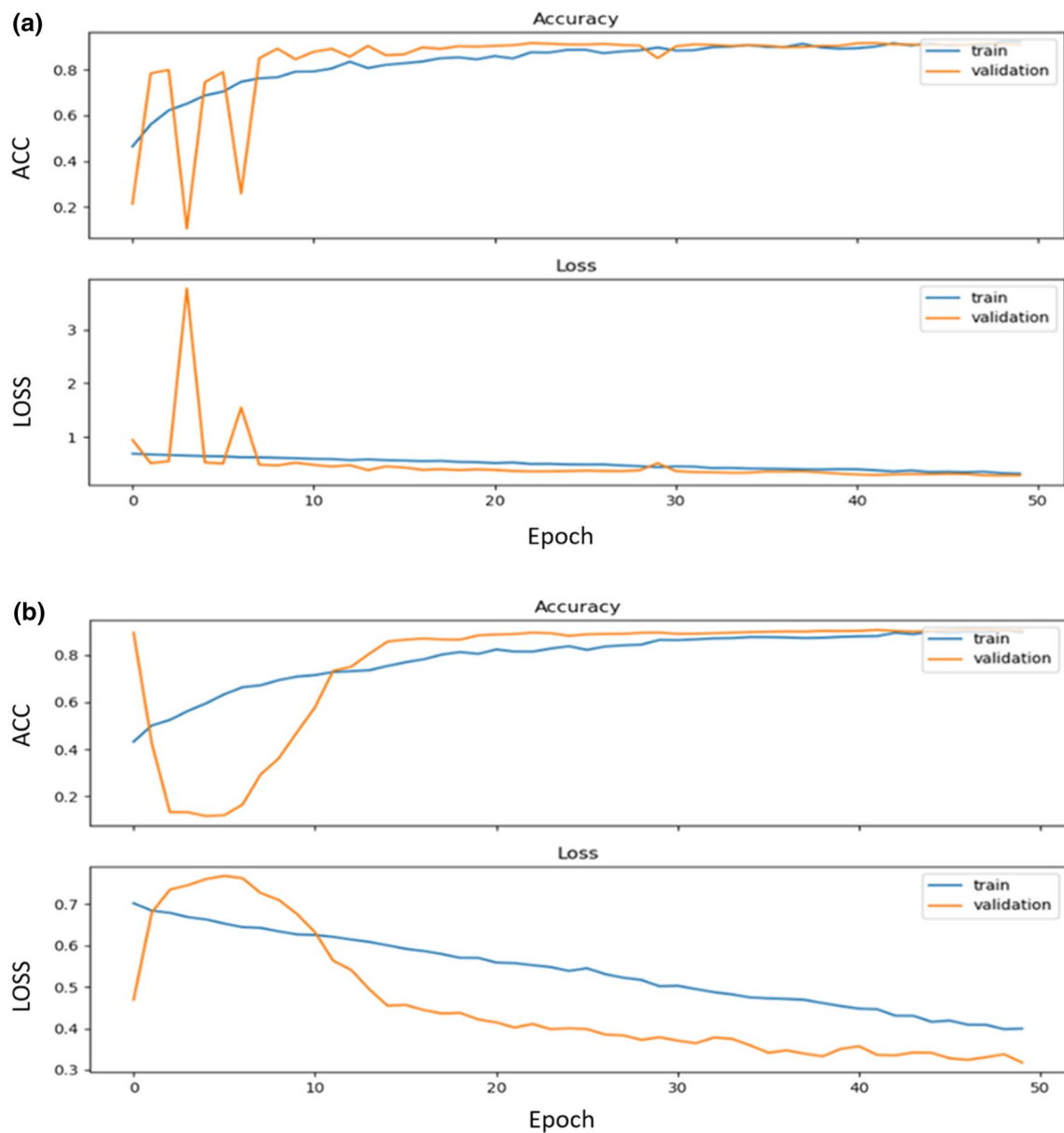


Fig. 11 Evolution of the ACC and LOSS indicators during the training process for a Scenario 3.1, b Scenario 3.2

Table 3 Processing test results for different scenarios from Table 2

Resolution and image extension no.	ACC	LOSS
Scenario 1.1	0.760	0.460
Scenario 1.2	0.880	0.130
Scenario 2.1	0.910	0.370
Scenario 2.2	0.912	0.280
Scenario 3.1	0.870	0.090
Scenario 3.2	0.840	0.300

are paramount. As a result, both datasets were augmented to multiply the data and obtain a target resolution of 512×512 px, as the smaller or larger image size resulted in poor outcomes for all transfer learning models.

The work revealed that the presented approach offers improvements in the metallographic investigations by removing the requirement for expert knowledge for the interpretation of image data prior to further characterization. It has also been proven that artificial neural networks based on the deep learning process using extensible U-Net network architectures and nonlinear learning tools can identify pearlite islands within a two-phase microstructure while the overtraining level remains low. Convolutional

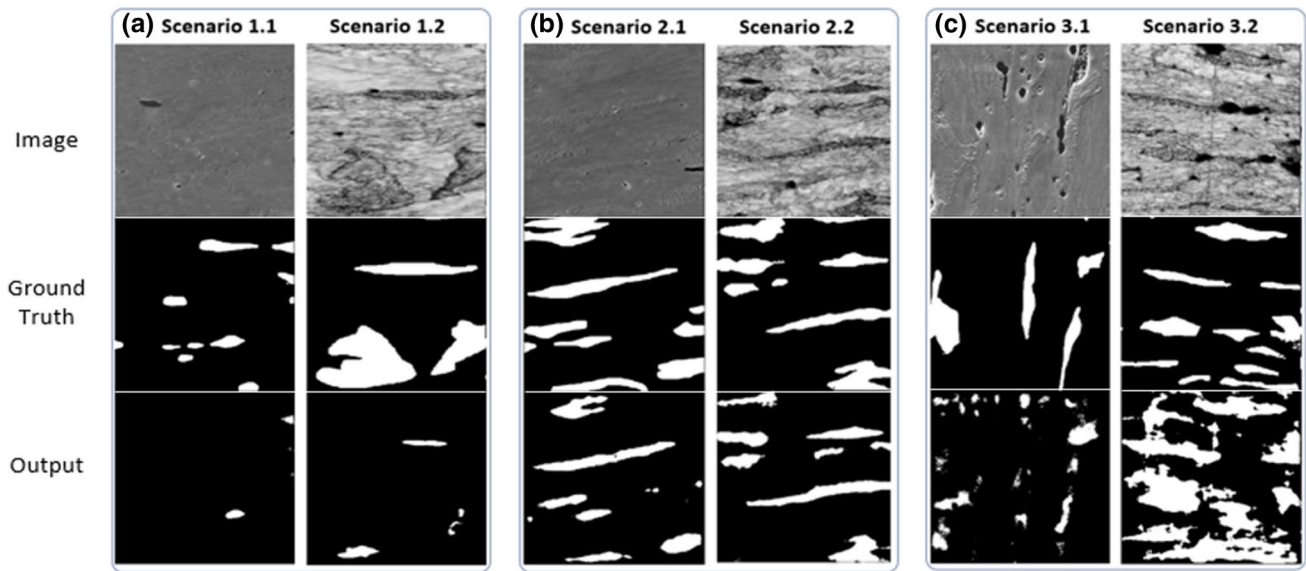


Fig. 12 Image processing output result for different resolutions and image source without optimization **a** for resolution 256×256 px, **b** for resolution 512×512 px, **c** for resolution 1024×1024 px

Table 4 Post-processing for SE and EBSD/IQ

SE					EBSD/IQ				
Alg	Train		Validation		Alg	Train		Validation	
	LOSS	ACC	LOSS	ACC		LOSS	ACC	LOSS	ACC
RMSprop	0.4150	0.8087	0.1499	0.9258	RMSprop	0.2650	0.8591	0.2799	0.9181
Adam	0.2018	0.8761	0.1310	0.9026	Adam	0.3088	0.7510	0.1910	0.9120
Adamax	0.3184	0.8189	0.3385	0.8862	Adamax	0.1438	0.7409	0.1385	0.8338

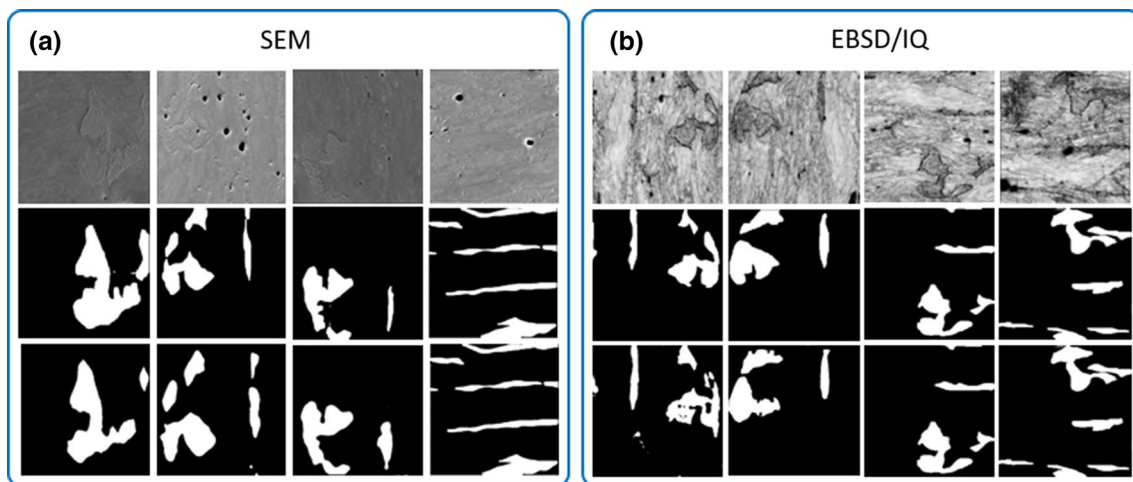


Fig. 13 Example of results from CNN identification of pearlite islands for **a** SE, **b** EBSD/IQ input data. From the top: original image, ground truth, and predicted image

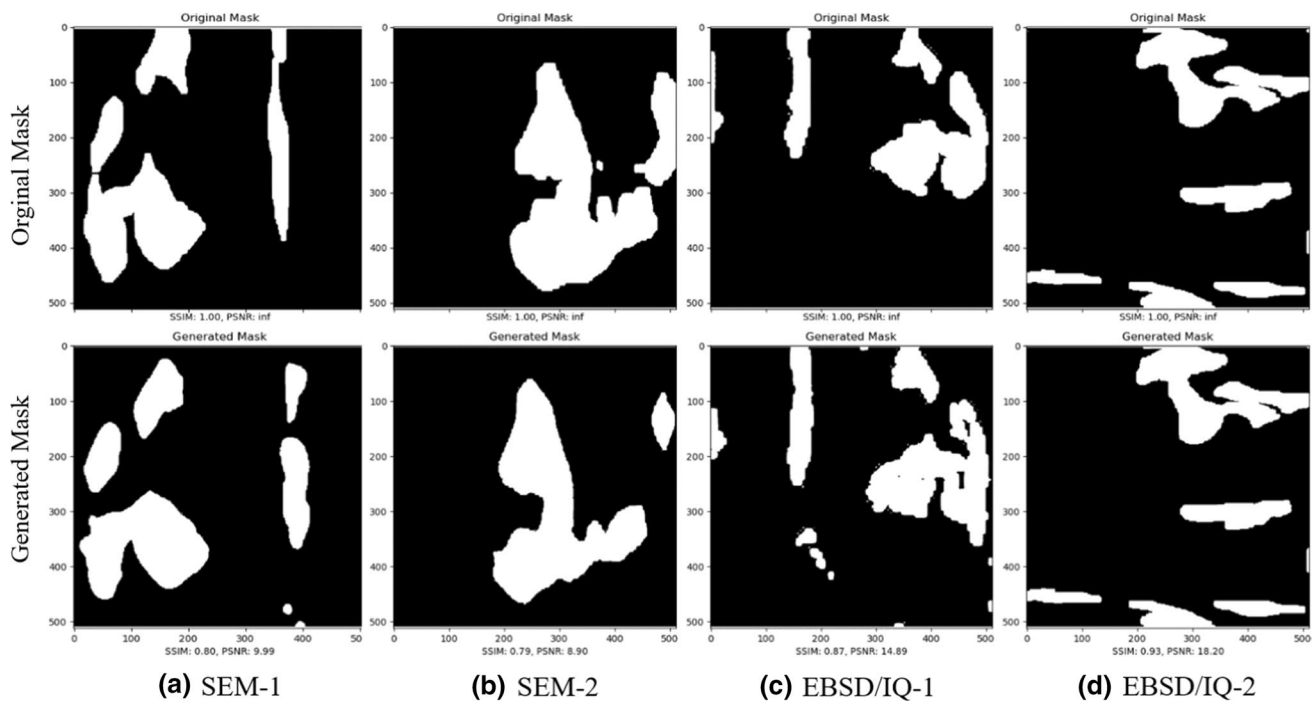


Fig. 14 Examples of a, b SE, c, d EBSD/IQ U-Net setup extractions

Table 5 Quantitative comparison of the developed U-Net performance with respect to two different input types

Figure 14	(a) SEM-1		(b) SEM-2		(c) EBSD/IQ-1		(d) EBSD/IQ-2	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
	0.80	9.99	0.79	8.90	0.87	14.89	0.93	18.20

neural networks do not require manual feature extraction and are able to automatically find appropriate search functions to recognize pearlite structure areas in the training process without human intervention. It was shown that the network recognizes areas of analyzed steel with satisfactory precision of 79% for EBSD and 87% for SE images.

With that, the developed model could be used to support automated three-dimensional reconstruction operations of two-dimensional microstructure images acquired during the serial sectioning procedure.

Acknowledgements Financial assistance of the National Science Centre Project No. 2017/27/B/ST8/00373 is acknowledged.

Data availability The raw/processed data required to reproduce these findings cannot be shared at this time as the data also forms part of an ongoing study.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

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/>.

References

1. Raabe D, Sun B, Kwiatkowski Da Silva A, Gault B, Yen H-W, Sedighiani K, Sukumar PT, Filho IRS, Katnagallu S, Jäggle E, Kürnsteiner P, Kusampudi N, Stephenson L, Herbig M, Liebscher

- CH, Springer H, Zaefferer S, Shah V, Wong S-L, Baron C, Diehl M, Roters F, Ponge D. Current challenges and opportunities in microstructure-related properties of advanced high-strength steels. *Metall Mater Trans A*. 2020;51:5517–86.
2. Adamczyk-Cieślak B, Korallnik M, Kuziak R, Majchrowicz K, Mizera J. Studies of bainitic steel for rail applications based on carbide-free, low-alloy steel. *Metall Mater Trans A*. 2021;52:5429–42.
 3. Liu X. Microstructural characterisation of pearlitic and complex phase steels using image analysis method, Xi Liu, PhD thesis, Birmingham University; 2014.
 4. Roskosz S, Chrapoński J, Madej L. Application of systematic scanning and variance analysis method to evaluation of pores arrangement in sintered steel. *Measurements*. 2021;168: 108325.
 5. Banerjee S, Ghosh SK, Datta S, Saha KS. Segmentation of dual phase steel micrograph: an automated approach. *Measurement*. 2013;46:2435–40.
 6. Bhadeshia HKDH. Neural networks in materials science. *ISIJ Int*. 1999;39:966–79.
 7. Gurney K. An introduction to neural networks. New York: UCL Press; 1997.
 8. Widrow B, Lehr MA. 30 years of adaptive neural networks: perceptron, madaline, and backpropagation. *Conf Proc IEEE*. 1990;78:1415–42.
 9. Fukushima K. Neocognitron: a self organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern*. 1980;36:193–202.
 10. Cireşan D. High-performance neural networks for visual object classification. www.arxiv.org; 2011.
 11. Cireşan D. Multi-column deep neural network for traffic sign classification. *Neural Netw*. 2012;32:333–8.
 12. Bozinovski S. Teaching space: a representation concept for adaptive pattern classification. COINS Technical Report; 1981. p. 81–2.
 13. Bengio IGY, Courville A. Deep learning. New York: MIT Press; 2017.
 14. Russell S, Norvig P. Artificial intelligence, a modern approach. 2nd ed. New York: Prentice Hall; 2003.
 15. Alpaydin E. Introduction to machine learning. New York: MIT Press; 2020.
 16. Kim H, Inoue J, Kasuya T. Unsupervised microstructure segmentation by mimicking metallurgists' approach to pattern recognition. *Sci Rep*. 2020;10:17835.
 17. Sitko M, Mojżeszko M, Rychłowski Ł, Cios G, Bała P, Muszka K, Madej L. Numerical procedure of three-dimensional reconstruction of ferrite–pearlite microstructure data from SEM/EBSD serial sectioning. *Proc Manuf*. 2020;47:1217–22.
 18. de Albuquerque VHC, Cortez PC, de Alexandria AR, Tavares JMRS. A new solution for automatic microstructures analysis from images based on a backpropagation artificial neural network. *Nondestruct Test Eval*. 2008;23:273–83.
 19. Wang S, Xia X, Ye L, Yang B. Automatic detection and classification of steel surface defect using deep convolutional neural networks. *Metals*. 2021;11:388.
 20. Abu M, Amir A, Lean YH, Zahri NAH, Azemi SA. The performance analysis of transfer learning for steel defect detection by using deep learning. *J Phys Conf Ser*. 2021;1755:1.
 21. Yeom J, Stan T, Hong S, Voorhees PW. Segmentation of experimental datasets via convolutional neural networks trained on phase field simulations. *Acta Mater*. 2021;214:1.
 22. Ostormujof TM, Purushottam Raj Purohit RRP, Breumier S, Gey N, Salib M, Germain L. Deep Learning for automated phase segmentation in EBSD maps. A case study in dual phase steel microstructures. *Mater Charact*. 2022;184:111638.
 23. Ackermann M, Iren D, Wesselmecking S, Shetty D, Krupp U. Automated segmentation of martensite-austenite islands in bainitic steel. *Mater Charact*. 2022;191:112091.
 24. Breumier S, Ostormujof TM, Frincu B, Gey N, Couturier A, Loukachenko N, Abaperea PE, Germain L. Leveraging EBSD data by deep learning for bainite, ferrite and martensite segmentation. *Mater Charact*. 2022;186:111805.
 25. Luengo J, Moreno R, Sevillano I, Charde D, Peláez-Vegas A, Fernández-Moreno M, Mesejo P, Herrera F. A tutorial on the segmentation of metallographic images: taxonomy, new Metal-DAM dataset, deep learning-based ensemble model, experimental analysis and challenges. *Inf Fus*. 2022;78:232–53.
 26. Tian W, Cheng X, Liu Q, Yu C, Gao F, Chi Y. Meso-structure segmentation of concrete CT image based on mask and regional convolution neural network. *Mater Des*. 2021;208:1.
 27. Madej L. Virtual microstructures in application to metals engineering—a review. *Arch Civ Mech Eng*. 2017;17:839–54.
 28. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: *Conference of proceeding of the medical image computing and computer-assisted intervention, Munich*; 2015. p. 234–41.
 29. Sara U, Akter M, Uddin MS. Image quality assessment through FSIM, SSIM, MSE and PSNR, a comparative study. *J Comput Commun*. 2019;7:8–18.
 30. Gulli A, Pal S. Deep learning with Keras. New York: Packt Publishing; 2017.
 31. Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. www.arxiv.org; 2014.
 32. Abu M, Amir A, Lean YH, Zahri NAH, Azemi SA. The performance analysis of transfer learning for steel defect detection by using deep learning. *J Phys Conf Ser*. 2020;1755:1.
 33. Ruder S. An overview of gradient descent optimization algorithms. www.arxiv.org; 2016.
 34. Wang Z, Simoncelli EP, Bovik AC. Multiscale structural similarity for image quality assessment. In: *Conference proceeding of the asilomar conference on signals, systems and computers, vol. 2*; 2004. p. 1398–402.
 35. Kumar R, Moyal V. Visual image quality assessment technique using FSIM. *Int J Comput Appl Technol Res*. 2013;2:250–4. <https://doi.org/10.7753/IJCATR0203.1008>.

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