


Article

Automatic Recognition of Microstructures of Air-Plasma-Sprayed Thermal Barrier Coatings Using a Deep Convolutional Neural Network

Xiao Shan ¹, Tianmeng Huang ¹, Lirong Luo ², Jie Lu ¹, Huangyue Cai ^{1,*}, Junwei Zhao ³, Gang Sheng ³ and Xiaofeng Zhao ^{1,*} 

- ¹ Shanghai Key Laboratory of Advanced High-Temperature Materials and Precision Forming, School of Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China
² Engineering Research Center of Nano-Geo Materials of Ministry Education, Faculty of Materials Science and Chemistry, China University of Geosciences, Wuhan 430074, China
³ Chengdu Chengfa Taida Aviation Technology Co., Ltd., Chengdu 610000, China
* Correspondence: caihy27@sjtu.edu.cn (H.C.); xiaofengzhao@sjtu.edu.cn (X.Z.)

Abstract: Either to obtain desirable microstructures by adjusting processing parameters or to predict the properties of a thermal barrier coating (TBC) according to its microstructure, fast and reliable quantitation of the microstructure is imperative. In this research, a machine-learning-based approach—a deep convolution neural network (DCNN)—was established to accurately quantify the microstructure of air-plasma-sprayed (APS) TBCs based on 2D images. Four scanning electron microscopy (SEM) images (view field: 150 μm \times 150 μm , image size: 3072 pixel \times 3072 pixel) were taken and labeled to train the DCNN. After training, the DCNN could recognize correctly 98.5% of the pixels in the SEM images of typical APS TBCs. This study demonstrated that a small dataset of SEM images could be enough to train a DCNN, making it a powerful and feasible method for quantitatively characterizing the microstructure of APS TBCs.

Keywords: thermal barrier coatings (TBC); plasma sprayed; microstructural characterization; machine learning; convolution neural network



Citation: Shan, X.; Huang, T.; Luo, L.; Lu, J.; Cai, H.; Zhao, J.; Sheng, G.; Zhao, X. Automatic Recognition of Microstructures of Air-Plasma-Sprayed Thermal Barrier Coatings Using a Deep Convolutional Neural Network. *Coatings* **2023**, *13*, 29. <https://doi.org/10.3390/coatings13010029>

Academic Editor: Ravisankar Naraparaju

Received: 29 November 2022
Revised: 18 December 2022
Accepted: 20 December 2022
Published: 23 December 2022



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Air-plasma-sprayed (APS) thermal barrier coatings (TBCs) are widely used to protect the metals in hot sections of gas-turbine engines [1]. They have a unique and complex microstructure: impingement of molten spray particles onto substrates results in “splats”, and the successive build-up of the “splats” results in a layered microstructure [2]. Several types of microstructure defect exist in APS TBCs: cracks, pores, and unmelted particles [3,4]. Cracks can be further classified into two categories: intersplat cracks caused by imperfect contacts between splats, and intrasplat cracks caused by relaxation of the quenching stress [5]. Pores, also called globular pores in some papers, are formed by incomplete contact or partially molten particles [6].

Both for aircraft and for industrial gas-turbine engines, APS TBCs require low thermal conductivity, high strain tolerance, high fracture toughness, and high sintering resistance [7,8]. All these properties are influenced largely by the microstructure and have a big impact on the lifetime of an APS TBC [9]. Therefore, investigation of the microstructure is of great significance in the area of TBC research. Either to obtain desirable microstructures by adjusting processing parameters or to predict the properties of an APS TBC according to its microstructure, quantitatively characterizing the microstructures of APS TBCs is imperative [10], and a fast and reliable characterization method will be beneficial to the whole APS TBC research community.

Image analysis with scanning electron microscopy (SEM, Mira, TESCAN, Czech Republic) images plays a crucial role in the quantitative characterization of the microstructures

of APS TBCs, due to its feasibility and reliability [3,10,11]. Figure 1 illustrates the procedures of conventional image analysis [10,12]. First, thresholding of a grayscale SEM image (Figure 1A) is performed, which yields a binary image (Figure 1B) wherein microstructural defects (e.g., equiaxed pores and cracks) are separated from the coating material. Thresholding is effective because the gray values of microstructural features are usually lower than those of the coating material. Second, morphological filtering of the opening operation is performed on the binary image, which can separate cracks (Figure 1C) from equiaxed pores (Figure 1D). That opening operation can perform such segmentation by relying on the fact that cracks are usually thinner than equiaxed pores.

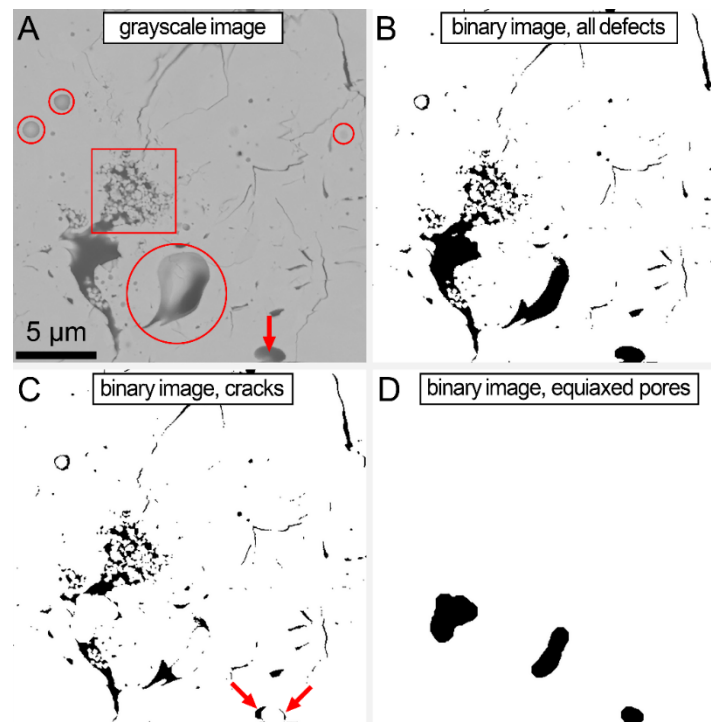


Figure 1. Images illustrating the procedures of conventional images analysis for SEM images of APS TBCs. (A) gray-scale BSE image, (B) binary image with all pores, (C) binary image with cracks and (D) binary image with equiaxed pores. The circles in (A) indicate pores whose gray values are close to the coating material; the squares indicate unmelted regions. The arrow in (A) indicates an equiaxed pore, some of which tend to be classified as cracks using an opening operation. The arrow in (C) indicate an equiaxed pore’s protrusions that were wrongly classified as cracks during opening operation.

Although the above-mentioned image analysis method has been widely used, both the thresholding and the opening operations have inherent issues. For thresholding, some defects—for example, the equiaxed pores indicated by the circles in Figure 1A—can hardly be recognized if they have gray values close to the coating material. In opening operations, the protrusions of some equiaxed pores are often classified as cracks (indicated by the arrows in Figure 1C). In addition to these issues, the conventional image analysis method for APS TBCs is based on explicit programming. Therefore, specific code must be developed if a certain type of microstructure—for example, unmelted regions—needs to be recognized, which is possibly difficult to realize.

Recently, machine learning—an approach to realizing artificial intelligence—has become increasingly important in materials science and engineering [13,14]. It has been applied in research areas such as property prediction, discovery and design of materials, characterization of materials, knowledge extraction via text mining, and molecular simulation [13]. Even in the area of TBC research, the last few years have witnessed its applications [15–21]. Previous studies have demonstrated that some machine learning

models (e.g., random forest and convolution neural network) can be used to analyze optical microscopy images [22,23]. Compared with conventional image analysis, machine learning does not need explicit code for recognizing a certain microstructure. Feeding SEM and labeled images into a proper machine learning model may realize automatic recognition. So far, however, whether machine learning can realize pixel-wise automatic microstructure recognition of SEM images of APS TBCs is still unclear. In this research, a deep convolution neural network (DCNN) model was built. Four SEM images were taken and labeled to train the model. After training, the model's performance was evaluated, and the predicted images were examined in detail, through which possible reasons for wrong recognitions were analyzed.

2. Research Methodology

The core components of the model and the training process are illustrated in Figure 2. First, input (SEM images) and true targets (labeled SEM images) are needed, because supervised learning was adopted in this research. The input is fed into the model. Next, the model makes a prediction, which is then compared with the true target. Taking the prediction and the true target as inputs, the loss function computes a loss score, which quantifies the difference between them. Based on the score, the optimizer adjusts the model's parameters in the correct direction, aiming at lowering the score in the next training loop. The above depicts one training loop. After all the training images have been used, one epoch finishes. After many epochs, the prediction might be very close to the true target.

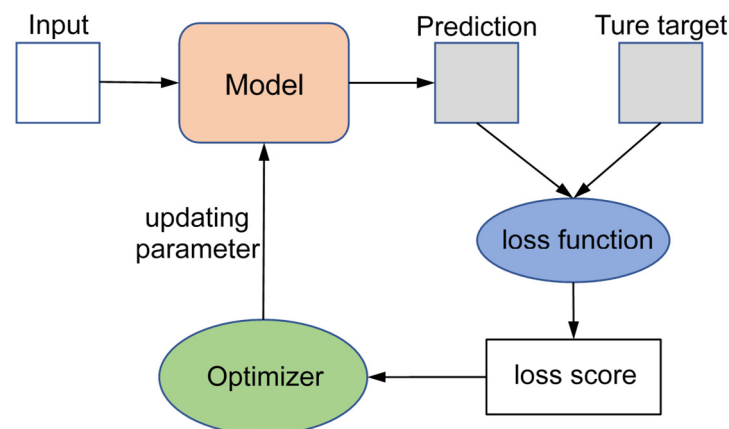


Figure 2. Schematic showing the core components of the machine learning model and the training process.

2.1. Dataset Preparation

2.1.1. Acquiring SEM Images

Four backscatter electron (BSE) SEM images were taken on the cross-section of a yttria-stabilized zirconia (YSZ) APS TBC. As can be seen in Figure 3A, this APS TBC contains not only typical APS TBC microstructures—cracks and equiaxed pores—but also unmelted regions. The view field of the BSE image is $150\ \mu\text{m} \times 150\ \mu\text{m}$, and the image size is 3072 pixels \times 3072 pixels. The numerical resolution reaches $0.046\ \mu\text{m}/\text{pixel}$, which is enough for quantitatively characterizing the microstructure of plasma-sprayed TBCs [12].

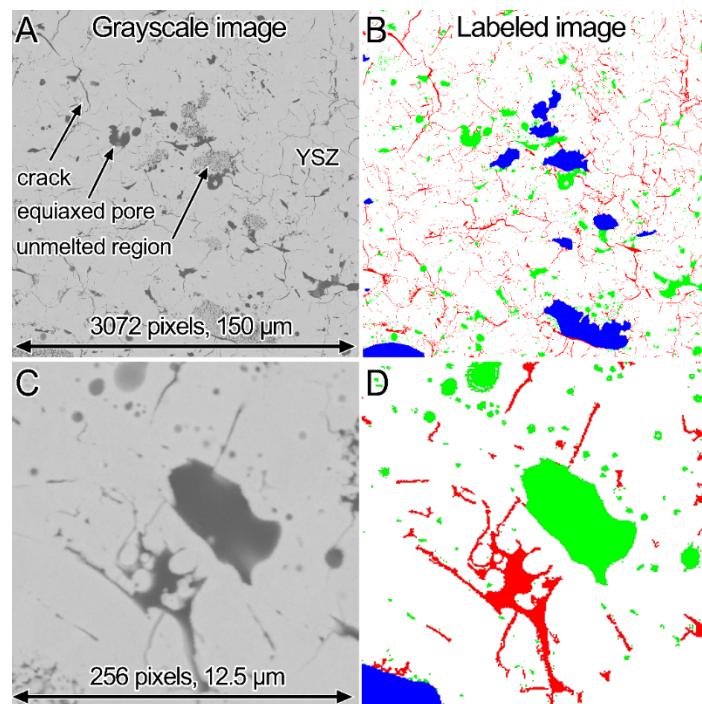


Figure 3. (A) Initial grayscale SEM image, (B) manually labeled image, (C) cropped grayscale image used for training, and (D) manually labeled image of the image in (C). Cracks, equiaxed pores, unmelted regions, and YSZ are indicated by red, green, blue, and white, respectively.

2.1.2. Labeling SEM Images

The task of the model is pixel-wise semantic segmentation—to classify each pixel in the images—so pixel-wise labeling is warranted. The four SEM images contain $3072 \times 3072 \times 4 = 37,748,736$ pixels. If all the pixels were labeled by hand, it would be a labor-heavy and time-consuming process. In addition to manual labeling, therefore, automatic labeling was also adopted. Figure 4 shows the labeling process used in this research. First, the unmelted regions were manually labeled using Photoshop software (version Photoshop CC2018). Second, non-YSZ regions (also referred to as defects in this research, including cracks, equiaxed pores, and unmelted regions) were automatically extracted from the remaining region using thresholding. However, some non-YSZ regions could not be extracted; these problematic regions were restored manually. Afterwards, isolated equiaxed pores were automatically identified based on their shapes. Next, the remaining equiaxed pores (e.g., those connected to other types of microstructures) were manually selected using Photoshop software. Finally, the labeled images were obtained. A truly labeled SEM image is shown in Figure 3B. After labeling, the area percentages of different classes of microstructures were calculated, as shown in Figure 5. One point worth stressing here is that the total area percentage of the three classes of defects is not necessarily equal to the porosity of an APS TBC, because the unmelted regions also contain YSZ (see the region indicated by the square in Figure 1A). However, the two are equal when no unmelted regions are involved.

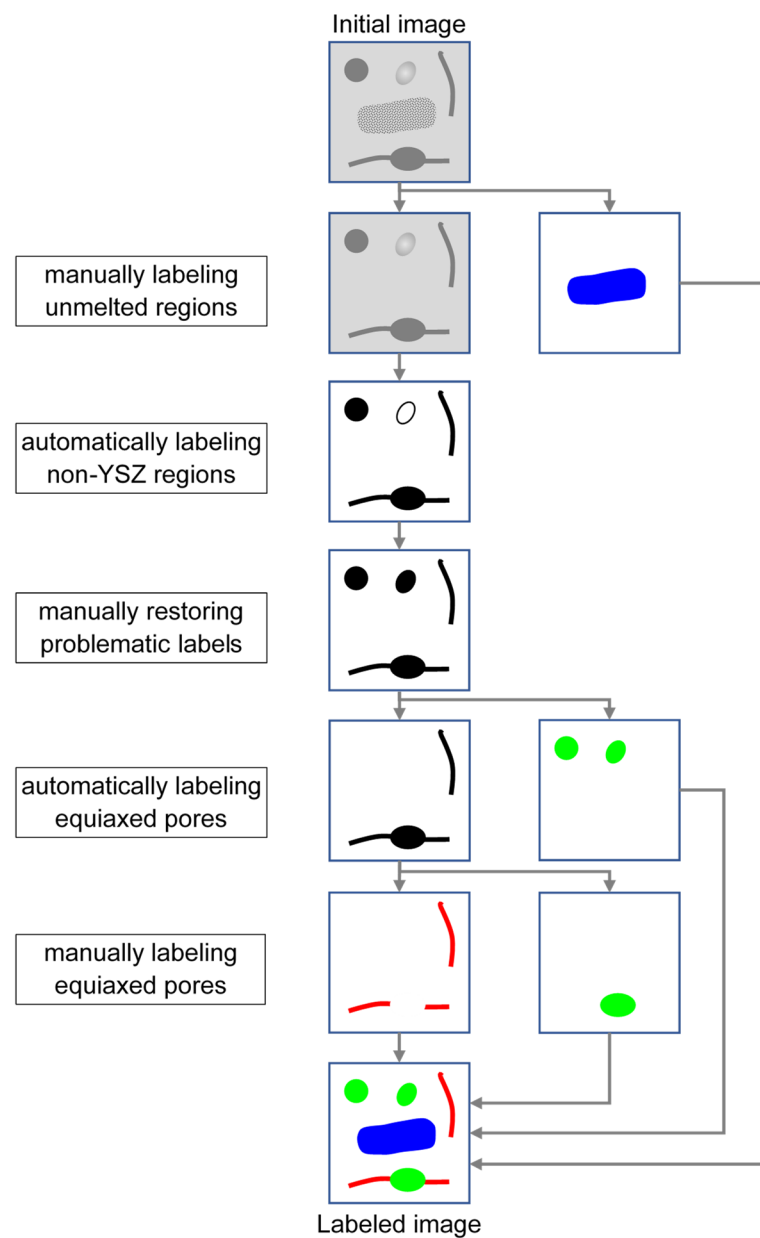


Figure 4. Schematic showing the labeling process.

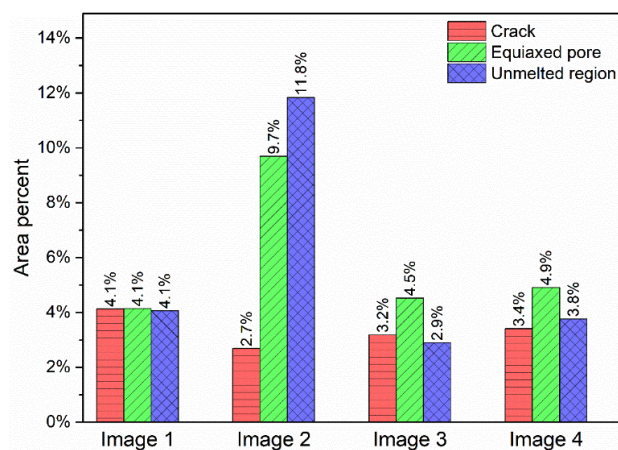


Figure 5. Area percentages of different classes of microstructures in the four SEM images.

2.1.3. Dataset Splitting

The image dataset was split into a training set, a validation set, and a test set. The training set was used to train the model. During the training process, the performance of the model on training data will always improve, but this does not mean that the model will perform well on data it has never seen. Therefore, a validation set was reserved; its purpose is to monitor how the model performs on new data. After training, the final performance of the model was evaluated using the test set. For dataset splitting, each of the initial 3072-pixel-wide images was divided into 144 256-pixel-wide images, so an image dataset containing 576 small images was obtained. The dataset was then split into three subsets according to Table 1.

Table 1. The numbers of small images in the training, validation, and testing sets.

Dataset	Image Number
Training	462
Validation	57
Test	57

2.2. Model Building

2.2.1. Workspace Setup

The DCNN model was built using Python in a Jupyter notebook. To enable fast experimentation with deep neural networks, Keras—a widely used deep learning framework—was used.

2.2.2. Model Architecture

Figure 6 shows the overall architecture of the DCNN model; its detailed specifications are presented in Table 2. The model has an encoder–decoder structure with skip connections. It is composed of 23 layers (Table 2), including convolution layers, pooling layers, up-sampling layers, and dropout layers. The convolution layers are to extract features of the images. They are composed of a certain number of kernels, whose parameters (commonly known as weights) are random values before training and need to be adjusted by model training. Pooling layers perform down-sampling operations, aiming at reducing the sizes of feature-maps to process and induce spatial-filter hierarchies. Contrary to pooling layers, up-sampling layers increase the sizes of feature-maps. Dropout layers were introduced to the model because they can reduce overfitting.

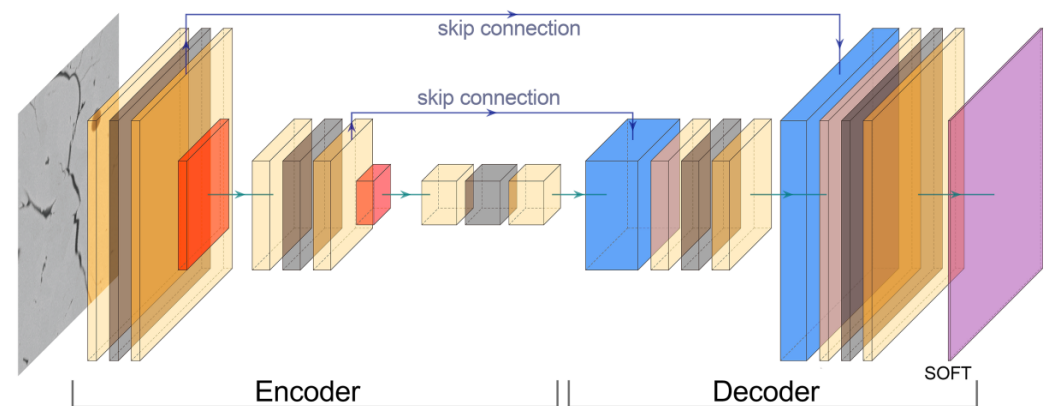


Figure 6. 3D schematic showing the overall architecture of the DCNN.

Table 2. The detailed specifications of the DCNN.

No.	Layer	Kernel	Parameter	Input Size	Output Size
1	Input	-	0	(N, 256, 256, 3)	(N, 256, 256, 3)
2	Convolution	(3, 3) × 32	896	(N, 256, 256, 3)	(N, 256, 256, 32)
3	Dropout	-	0	(N, 256, 256, 32)	(N, 256, 256, 32)
4	Convolution	(3, 3) × 32	9248	(N, 256, 256, 32)	(N, 256, 256, 32)
5	Pooling	-	0	(N, 256, 256, 32)	(N, 128, 128, 32)
6	Convolution	(3, 3) × 64	18,496	(N, 128, 128, 32)	(N, 128, 128, 64)
7	Dropout	-	0	(N, 128, 128, 64)	(N, 128, 128, 64)
8	Convolution	(3, 3) × 64	36,928	(N, 128, 128, 64)	(N, 128, 128, 64)
9	Pooling	-	0	(N, 128, 128, 64)	(N, 64, 64, 64)
10	Convolution	(3, 3) × 128	73,856	(N, 64, 64, 64)	(N, 64, 64, 128)
11	Dropout	-	0	(N, 64, 64, 128)	(N, 64, 64, 128)
12	Convolution	(3, 3) × 128	147,584	(N, 64, 64, 128)	(N, 64, 64, 128)
13	Up sampling	-	0	(N, 64, 64, 128)	(N, 128, 128, 128)
14	Concatenate	-	0	(N, 128, 128, 128) (N, 128, 128, 64)	(N, 128, 128, 192)
15	Convolution	(3, 3) × 64	110,656	(N, 128, 128, 192)	(N, 128, 128, 64)
16	Dropout	-	0	(N, 128, 128, 64)	(N, 128, 128, 64)
17	Convolution	(3, 3) × 64	36,928	(N, 128, 128, 64)	(N, 128, 128, 64)
18	Up sampling	-	0	(N, 128, 128, 64)	(N, 256, 256, 64)
19	Concatenate	-	0	(N, 256, 256, 64) (N, 256, 256, 32)	(N, 256, 256, 96)
20	Convolution	(3, 3) × 32	27,680	(N, 256, 256, 96)	(N, 256, 256, 32)
21	Dropout	-	0	(N, 256, 256, 32)	(N, 256, 256, 32)
22	Convolution	(3, 3) × 32	9248	(N, 256, 256, 32)	(N, 256, 256, 32)
23	Convolution	(1, 1) × 4	132	(N, 256, 256, 32)	(N, 256, 256, 4)
Total trainable parameters: 471,652					

2.2.3. Loss and Metrics

There are four classes of microstructures in the SEM images, and most of the pixels are YSZ, so the task of the model is an imbalanced multiclass segmentation problem. Therefore, the generalized Dice loss [24], which is effective for this kind of segmentation, was used in this research.

Metrics are functions that are used to intuitively judge a model's performance. Three of the most commonly used metrics for semantic segmentation tasks were adopted in this research: (pixel) accuracy, mean (pixel) accuracy, mean intersection over union (mean IoU). Accuracy is the percentage of pixels in an image that are classified correctly. Accuracy can be computed using the following equation:

$$\text{accuracy} = \frac{\sum_{i=1}^k N_{ii}}{\sum_{i=1}^k \sum_{j=1}^k N_{ij}} \quad (1)$$

where k is the number of classes in a certain image, N_{ij} is the number of pixels of class i predicted to be class j . Classes 1–4 represent cracks, equiaxed pores, unmelted regions, and YSZ, respectively.

Accuracy is one of the easiest to understand conceptually. However, it is not an ideal metric when class imbalance—one class dominating the image—occurs. For example, suppose that the area of the non-YSZ region (including cracks, equiaxed pores, and unmelted regions) were 5%; the accuracy could reach up to 95% if all the pixels were classified as YSZ. To compensate for this issue, mean accuracy was introduced. It is the average of the accuracies of all the classes. The accuracy for class i is the number of pixels of a class that

are predicted correctly, divided by the number of all the pixels of the class in the labeled image, whose formula is as follows:

$$\text{accuracy}_i = \frac{N_{ii}}{\sum_{j=1}^k N_{ij}} \quad (2)$$

Based on this equation, the mean accuracy can be calculated using the following formula:

$$\text{mean accuracy} = \frac{1}{k} \sum_{i=0}^k \text{accuracy}_i \quad (3)$$

where k is the number of classes in a certain image. Mean interaction-over-union (IoU) is also a popular metric when class imbalance occurs. It is the average of the IoUs of all the classes. IoU is stricter than accuracy. For class i , it is the number of pixels of a class that are predicted correctly divided by the sum of the number of all the pixels of the class in the labeled image and the number of pixels that are wrongly classified as class i . The following formula shows the way calculating it:

$$\text{IoU}_i = \frac{N_{ii}}{\sum_{j=1}^k N_{ij} + \left(\sum_{j=1}^k N_{ji} - N_{ii} \right)} \quad (4)$$

Based on this equation, the mean IoU can be calculated using the following formula:

$$\text{mean IoU} = \frac{1}{k} \sum_{i=1}^k \text{IoU}_i \quad (5)$$

2.3. Model Training

The model was trained using four graphic processing units (GPU, NVIDIA Tesla V100 32 GB) for 100 epochs, and the loss scores were recorded after each epoch.

3. Results

3.1. Evolution of the Model during the Training Process

Figure 7 shows the evolutions of the losses (the generalized Dice loss) during training. Overall, the losses changed significantly during the first 15 epochs, and then changed slowly in the following epochs. The training loss kept decreasing till the end. The validation loss stopped improving from about the 83ed epoch, indicating that the model began to overfit and would not obtain better results on never-before-seen data with more adjustment. Therefore, 100 epochs are enough for training the model. The model with the lowest loss represented the best model and was used for the following analysis.

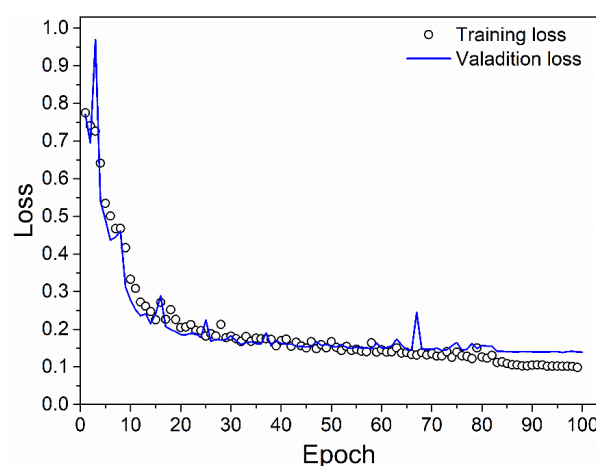


Figure 7. Training and validation losses as a function of epoch.

3.2. Model's Performance on All the Test Images

Figure 8 presents about half of the images predicted by the trained model, together with the initial grayscale images and the manually labeled images. These images give an overall visual impression of the model's performance. As can be seen in this figure, in most cases, the predicted images are very similar to the labeled images.

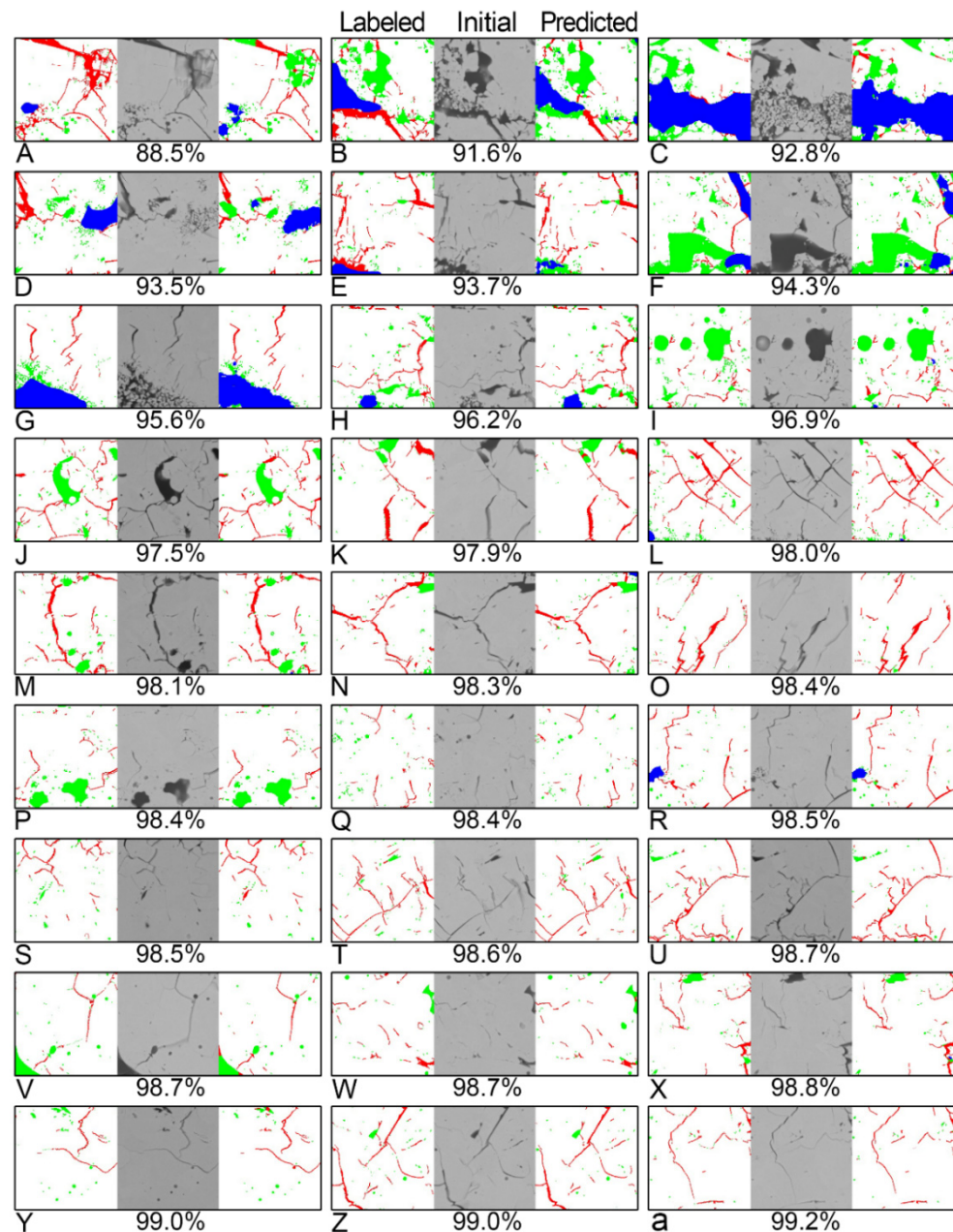


Figure 8. Images predicted by the model, together with the initial grayscale images, and the manually labeled images. These images were selected this way: all the predicted images were ranked according to accuracy, and then the even-numbered images were selected. The accuracy for each image is presented below each grayscale image. (A–Z,a) Images predicted by the model, together with the initial grayscale images, and the manually labeled images. These images were selected this way: all the predicted images were ranked according to accuracy, and then the even-numbered images were selected. The accuracy for each image is presented below each grayscale image.

Figure 9A shows the metrics of the trained model calculated based on the 57 test images. The accuracy reached up to 97.0%, meaning that only 3.0% of pixels in the grayscale

images were wrongly predicted. The mean accuracy and mean IOU dropped to 80.1% and 71.4%, respectively. The reason for this is that the accuracies of cracks, equiaxed pores, and unmelted regions were not high enough, as can be seen in Figure 9B.

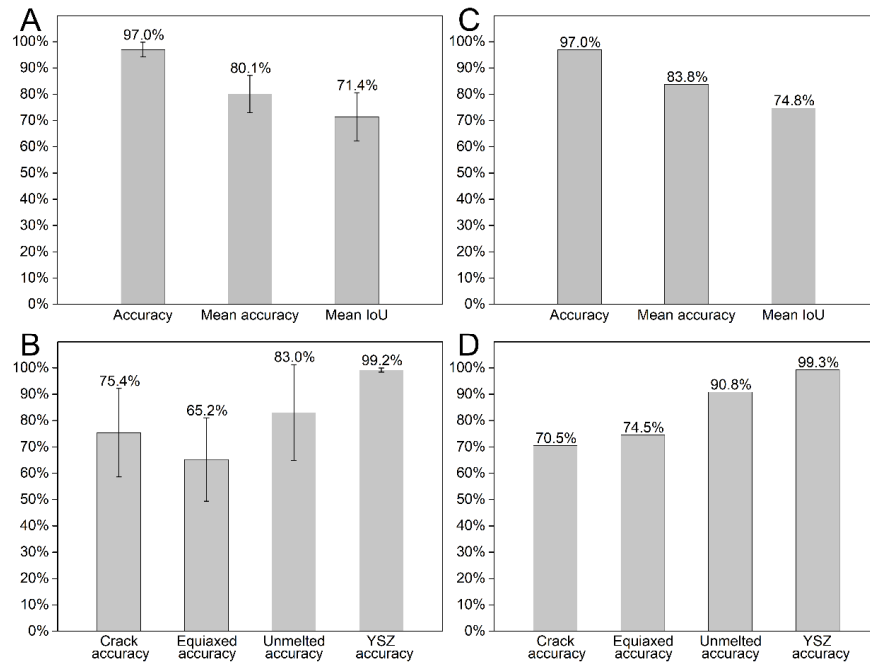


Figure 9. (A,B) Metrics of the trained model evaluated using 57 test images: (A) Overall metrics and (B) accuracy for each class of microstructure; the error bars represent standard deviations. (C,D) Metrics of the trained model evaluated using one hypothetical image containing all the test images: (A) overall metrics and (B) accuracy for each class of microstructure.

The data in Figure 9A,B were obtained based on small images ($256 \text{ pixels} \times 256 \text{ pixels}$, $12.5 \mu\text{m} \times 12.5 \mu\text{m}$), but sometimes evaluations on small images can be misleading. To illustrate, assume that there is only one test image which contains all the 57 test images. Then, the model's performance was evaluated using this large hypothetical image ($1933 \text{ pixel} \times 1933 \text{ pixel}$, $94 \mu\text{m} \times 94 \mu\text{m}$). The results are shown in Figure 9C,D. As can be seen in Figure 9C, the accuracy did not change, but both the mean accuracy and mean IoU increased slightly. In Figure 9D, note that the unmelted accuracy increases from 83.0% (Figure 9B) to 90.8%. This indicates that for the 57 test images, the images resulting in high unmelted accuracy tended to contain large areas of unmelted regions, which in turn means that a larger fraction of unmelted region pixels can be recognized if they are in larger unmelted regions. The same phenomenon occurred for equiaxed pores. The comparison between Figure 9B,D indicates that the size of the test image plays an important role in the model evaluation process, and the metrics evaluated using the test images whose sizes are closer to a usual size may be more representative.

In addition to pixel-level microstructure recognition, sometimes a trained model may also be used to calculate the area percentages of different microstructures. Figure 10 shows the true and predicted area percentages of different microstructures in the test images. It is shown that for some test images, the predicted area percentages of cracks (Figure 10A) and equiaxed pores (Figure 10B) deviate far from the true area percentages. In contrast, the predicted area percentage of the defect for each image is almost equal to the true value.

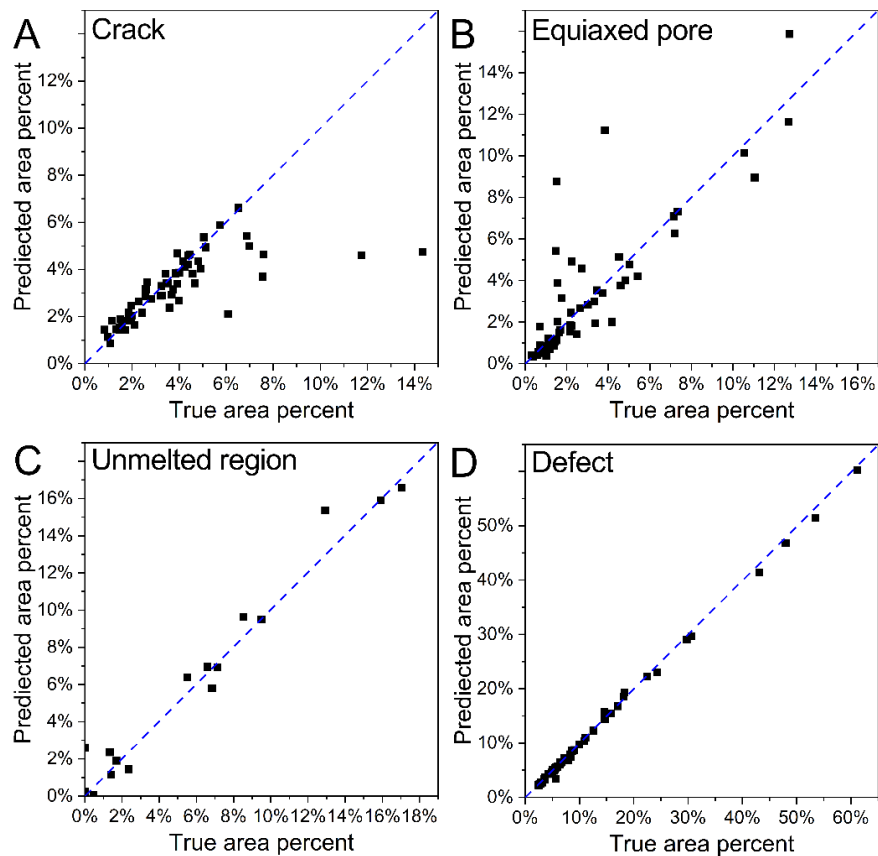


Figure 10. True and predicted area percentages of different microstructures in the test images: (A) crack, (B) equiaxed pore, (C) unmelted region, and (D) defect.

The large difference between true and predicted values for crack and equiaxed pore may have been caused by the small image size (256 pixels × 256 pixels, 12.5 μm × 12.5 μm), which is unlikely to be used in practice.

The model was also evaluated using the aforementioned hypothetical image (1933 pixel × 1933 pixel, 94 μm × 94 μm), and the results are shown in Figure 11. For this large image, the trained model exhibited good ability of estimating the area percentages of different microstructures.

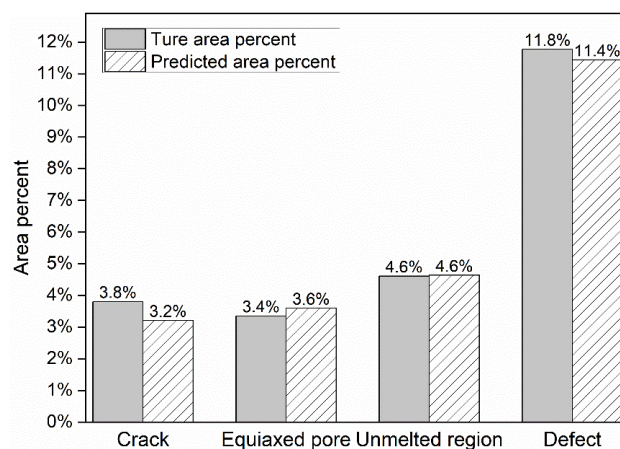


Figure 11. True and predicted area percentages for each class of microstructure in a hypothetical image containing all 57 test images.

3.3. Model's Performance on Typical APS TBC Images

It can be deduced from Figure 5 that some test images contain larger unmelted regions than usual, so the model's performance in Section 3.2 is indeed not representative enough of typical APS TBCs. To obtain more representative data, test images having area percentages of unmelted regions larger than 1% were removed, by which 39 typical APS TBC images remained and were used to evaluate the model. The results are shown in Figure 12. The accuracy reached up to 98.5%, and the crack accuracy increased to 82.6% (75.4% in Figure 9B). Note that the unmelted accuracy decreased to 0.0%, meaning that no unmelted region was successfully recognized, presumably because the features of unmelted regions become vague such areas are too small.

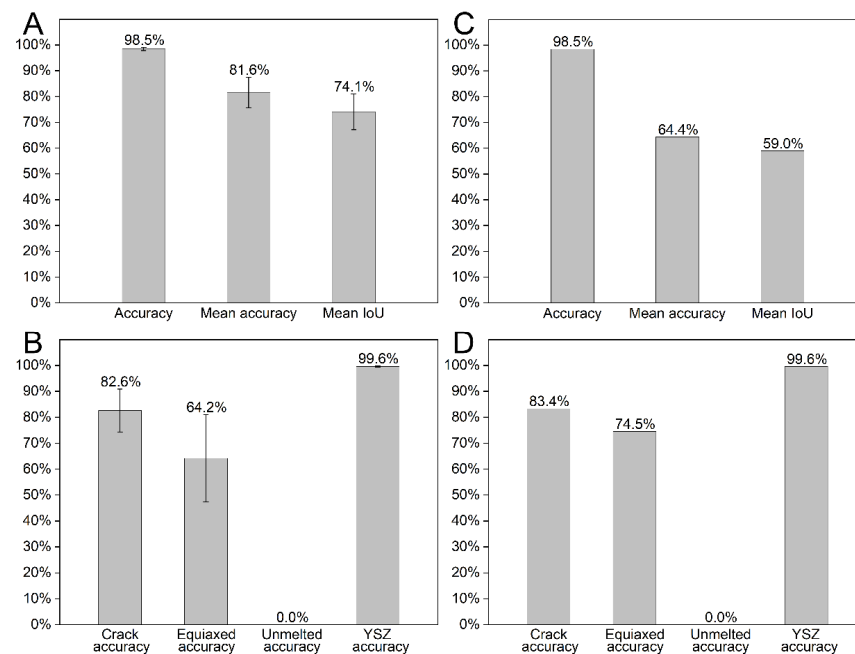


Figure 12. (A,B) Metrics of the trained model evaluated using 39 more typical APS TBC images selected from all the 57 test images: (A) Overall metrics and (B) accuracy for each class of microstructure; the error bars represent standard deviations. (C,D) Metrics of the trained model evaluated using one hypothetical image containing all 39 more typical APS TBC images: (A) overall metrics and (B) accuracy for each class of microstructure.

Figure 12C,D present the metrics of the model evaluated using one hypothetical image (1599 pixels \times 1599 pixels, 78 μm \times 78 μm) containing all 39 test images. The accuracy is the same as that in Figure 12A. However, the mean accuracy and mean IoU are significantly lower; the reason is that no unmelted region pixels were successfully recognized (Figure 12D: unmelted accuracy = 0.0%). Despite the decreases in the mean accuracy and the mean IoU, accuracies of typical microstructures (cracks, equiaxed pores, and YSZ) did not decline; in particular, the equiaxed accuracy increased from 64.2% to 74.5%, reaching the same value as that in Figure 9D.

Figure 13 shows the true and predicted area percentages of different microstructures in the 39 typical APS TBC images. Compared with Figure 10, it seems that the model exhibits better performance when evaluating the area percentages of different microstructures. Figure 14 shows the data for the large hypothetical image containing all 39 test images. For this large image, the trained model also exhibited a good ability to estimate the area percentages of different microstructures.

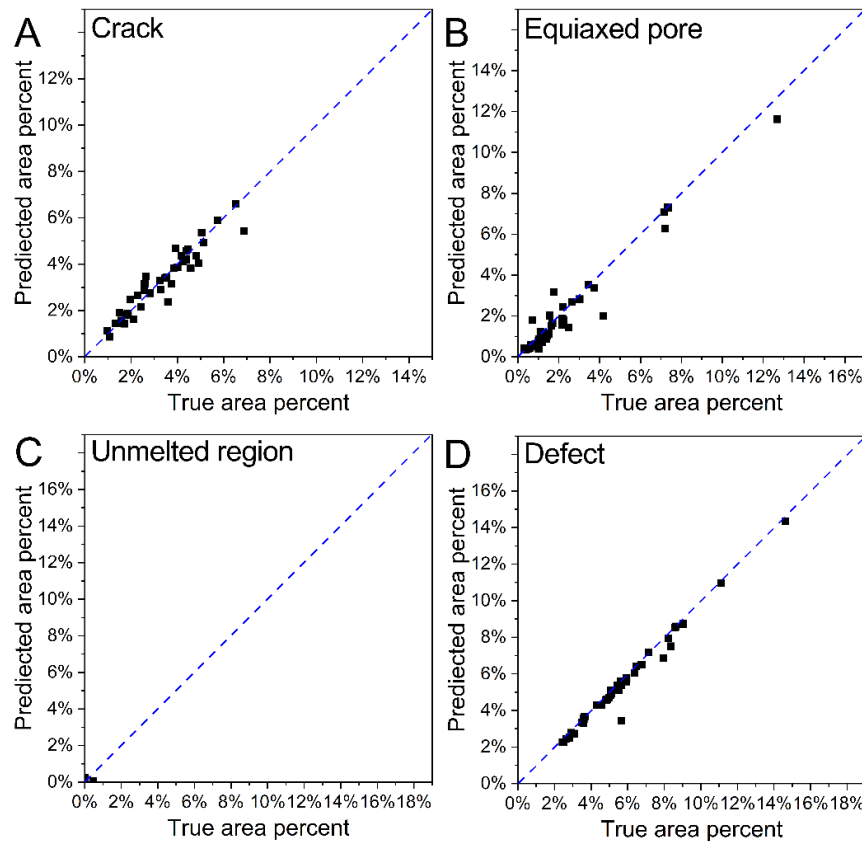


Figure 13. True and predicted area percentages of different objects in the 39 more typical APS TBC images selected from all the 57 test images: (A) crack, (B) equiaxed pore, (C) unmelted region, and (D) defect.

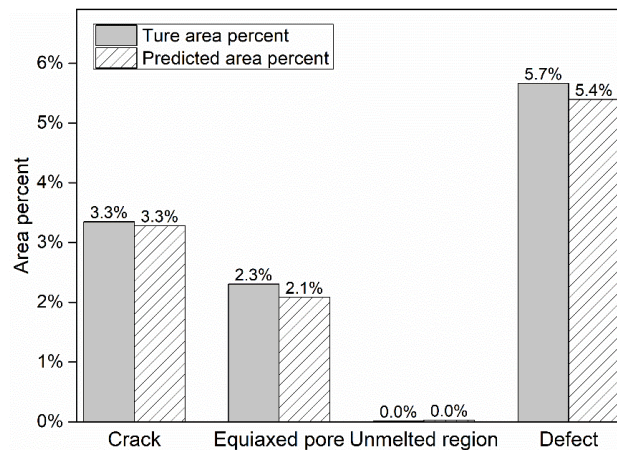


Figure 14. True and predicted area percentages for each class of microstructure in a hypothetical image containing all the 39 typical APS TBC images.

4. Discussion

Although the overall pixel accuracy of the model can reach up to 97.0%, the accuracies for some classes are still not high enough. Possible reasons are discussed in this section. Figure 15 shows the predicted amount of each class of microstructure in a hypothetical image containing all 57 test images. For YSZ, the model has very high accuracy, such that only a tiny fraction ($0.3\% + 0.2\% + 0.2\% = 0.7\%$) of YSZ was predicted as other classes. Non-negligible fractions of the non-YSZ classes can be wrongly predicted as YSZ, accounting for ~30%, ~50%, and ~60% of the wrong predictions of the cracks, equiaxed

pores, and unmelted regions, respectively. The most likely reason is that the all the three non-YSZ classes are immediately next to YSZ, but sometimes the boundaries between YSZ and non-YSZ regions are difficult to precisely locate at the pixel level.

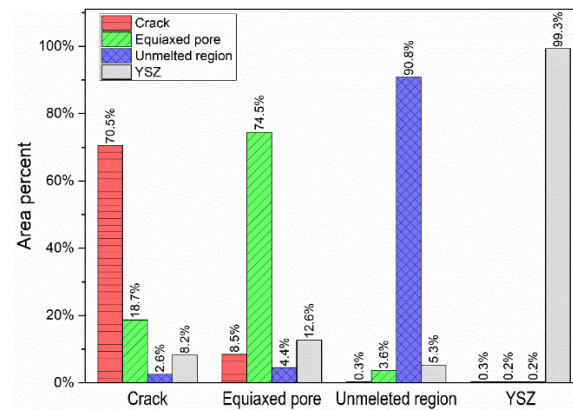


Figure 15. Predicted amount each class of microstructure in a hypothetical image containing all 57 test images.

In addition to YSZ, cracks tended to be wrongly predicted as equiaxed pores, accounting for as many as ~60% of the wrong predictions. Conversely, equiaxed pores also tended to be wrongly predicted as cracks, accounting for ~30% of the wrong predictions. A detailed examination of the predicted image revealed that there are two common situations in which misclassifications between cracks and equiaxed pores can occur. The first is that some cracks next to the unmelted regions tend to be predicted as equiaxed pores, as can be seen in Figure 16. The second occurred when cracks were linked to equiaxed pores. As can be seen in Figure 17, the arrow points to an equiaxed pore, but the model classified it as a crack.

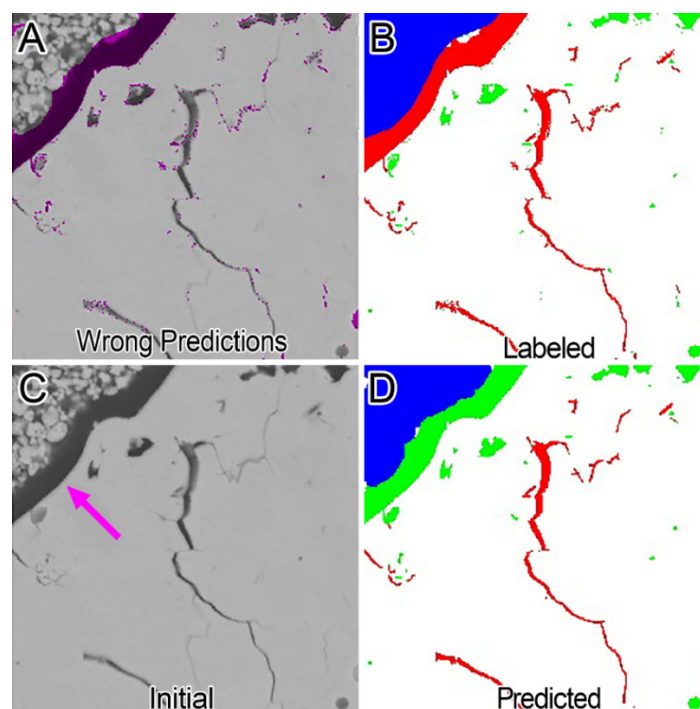


Figure 16. Test image for which cracks were wrongly predicted as equiaxed pores. (A) The grayscale image with wrong predictions are marked by purple, (B) the labeled image, (C) the initial grayscale image The arrow indicates a crack next to an unmelted region that was wrongly predicted as an equiaxed pore, and (D) the predicted image.

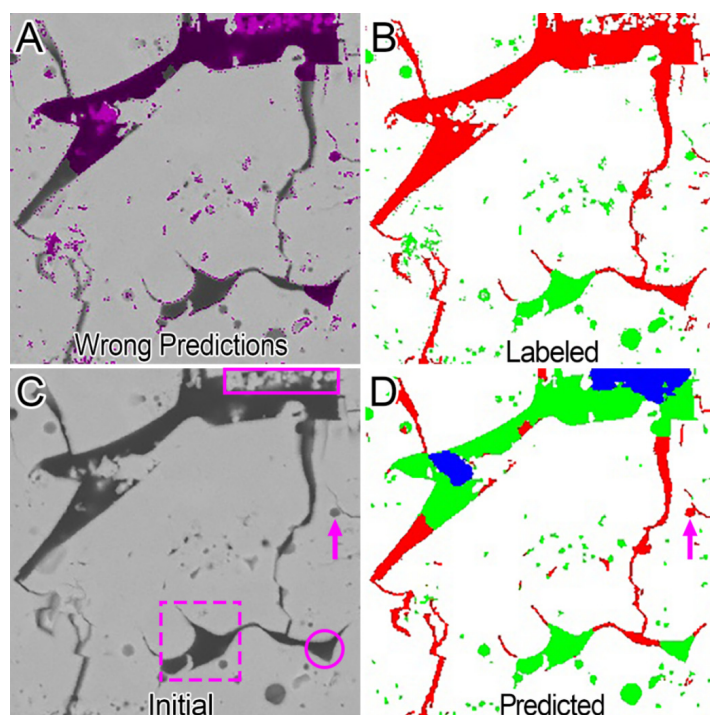


Figure 17. The lowest-accuracy (87.0%) test image (A), in which wrong predictions are marked by purple, together with the labeled image (B), the initial grayscale image (C) (violet solid rectangle indicates a region where wrong predictions occurred due to wrong label, violet dashed square indicates a region where an equiaxed pore and cracks are connected, circle indicates a region where both the wrong prediction and the label seem to be reasonable), and the predicted image (D) (arrow indicates an equiaxed pore that was wrongly predicted as a crack).

In addition to the routine wrong predictions, there were also other types of wrong predictions that are worth noting. First, although some predictions are wrong, a comparison of the predicted image with the labeled image reveals that the model seems to be more correct. For example, the region in the rectangle in Figure 17 is unmelted, and the model predicted it as unmelted (Figure 17D), but the region was labeled as cracked and YSZ (Figure 17B). Such “wrong” predictions are actually caused by labeling errors. Therefore, it is expected the model’s performance will be further improved by minimizing labeling error.

Second, sometimes both the wrong prediction and the label seem to be reasonable. For example, the defect in the circle in Figure 17 was labeled as a crack, but the model predicted it as an equiaxed pore. Connections between equiaxed pores and cracks are very common in APS TBCs. In this research, manual labeling was used to separate them. For example, the coarse part of the defect in the dashed square in Figure 17C was labeled as equiaxed pores. Therefore, predicting the defect in the circle as an equiaxed pore is also reasonable.

The above analysis shows that the misclassification between cracks and equiaxed pores is an important factor causing wrong predictions. One reason is that so far there are no widely accepted rigorous definitions of cracks and equiaxed pores, so sometimes it is hard to determine what a pixel belongs to, even for a human expert, during the image annotation process. In such situations, misclassification between crack and equiaxed pores can hardly be avoided.

In addition to image-related factors, model configuration may also play a significant role in influencing the model’s performance. The study is only a preliminary attempt to assess the possibility of utilizing a machine-learning-based model to quantitatively characterize the microstructure of APS TBCs, so only one specific model was adopted. To obtain a high-performance model, future work will investigate the influence of model configuration (e.g., model architecture and loss function) on the model’s performance. Additionally,

once a high-performance model has been obtained, a comparative study of the model and conventional image analysis method will then be conducted.

5. Summary

This research aimed to answer a key question: whether machine learning can realize pixel-wise automatic microstructure recognition of SEM images of APS TBCs. To answer this question, a machine-learning-based model—a deep convolution neural network (DCNN)—was established. Four SEM images of APS TBCs (containing not only cracks and equiaxed pores, but also unmelted regions) were taken and labeled for training the model. The results showed that the (pixel) accuracy of the trained model reached up to 97.0%—that is, 97.0% of the pixels in the SEM images can be recognized correctly; for SEM images of typical APS TBCs (containing less unmelted regions), the accuracy increased to 98.5%. Despite the high overall accuracy, the accuracies for some classes of microstructure were not high enough. Possible reasons were analyzed. It was found that wrong predictions were mainly due to the misclassification between YSZ and non-YSZ, and between cracks and equiaxed pores. The major finding of this research is that a small dataset of SEM images could be enough to train a DCNN, which may provide a powerful and feasible method for quantitatively characterizing the microstructures of APS TBCs.

Author Contributions: Conceptualization, X.S.; methodology, X.S., L.L., J.L. and H.C.; data curation, X.S.; funding acquisition, X.S.; project administration, X.S.; writing—original draft, X.S.; Formal analysis, T.H.; investigation, T.H.; validation, L.L.; resources, L.L., J.L., H.C., J.Z. and G.S.; visualization, H.C.; software, X.Z.; supervision, X.Z., writing—reviewing and editing, X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The authors would acknowledge the financial support from the National Natural Science Foundation of China (No. 52102071 and 52102072).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The full training datasets and all validation and test cases are available from the corresponding author upon reasonable request. The data are not publicly available due to the reason that at this time as the data also forms part of an ongoing study.

Acknowledgments: We would like to thank Guo Wei from Northwestern Polytechnic University for stimulating discussions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bakan, E.; Vaßen, R. Ceramic top coats of plasma-sprayed thermal barrier coatings: Materials, processes, and properties. *J. Therm. Spray Technol.* **2017**, *26*, 992–1010. [[CrossRef](#)]
2. McPherson, R. A review of microstructure and properties of plasma sprayed ceramic coatings. *Surf. Coat. Technol.* **1989**, *39–40*, 173–181. [[CrossRef](#)]
3. Heberlein, J.; Fauchais, P.; Boulos, M. *Thermal Spray Fundamentals: From Powder to Part*; Springer: New York, NY, USA, 2014.
4. Dwivedi, G.; Viswanathan, V.; Sampath, S.; Shyam, A.; Lara-Curzio, E. Fracture toughness of plasma-sprayed thermal barrier ceramics: Influence of processing, microstructure, and thermal aging. *J. Am. Ceram. Soc.* **2014**, *97*, 2736–2744. [[CrossRef](#)]
5. Kulkarni, A.; Wang, Z.; Nakamura, T.; Sampath, S.; Golland, A.; Herman, H.; Allen, J.; Ilavsky, J.; Long, G.; Frahm, J.; et al. Comprehensive microstructural characterization and predictive property modeling of plasma-sprayed zirconia coatings. *Acta Mater.* **2003**, *51*, 2457–2475. [[CrossRef](#)]
6. Wang, Z.; Kulkarni, A.; Deshpande, S.; Nakamura, T.; Herman, H. Effects of pores and interfaces on effective properties of plasma sprayed zirconia coatings. *Acta Mater.* **2003**, *51*, 5319–5334. [[CrossRef](#)]
7. Cao, X.; Vassen, R.; Stoeber, D. Ceramic materials for thermal barrier coatings. *J. Eur. Ceram. Soc.* **2004**, *24*, 1–10. [[CrossRef](#)]
8. Frommherz, M.; Scholz, A.; Oechsner, M.; Bakan, E.; Vassen, R. Gadolinium zirconate/YSZ thermal barrier coatings: Mixed-mode interfacial fracture toughness and sintering behavior. *Surf. Coat. Technol.* **2016**, *286*, 119–128. [[CrossRef](#)]
9. Clarke, D.; Oechsner, M.; Padture, N. Thermal-barrier coatings for more efficient gas-turbine engines. *MRS Bull.* **2012**, *37*, 891–898. [[CrossRef](#)]

10. Deshpande, S.; Kulkarni, A.; Sampath, S.; Herman, H. Application of image analysis for characterization of porosity in thermal spray coatings and correlation with small angle neutron scattering. *Surf. Coat. Technol.* **2004**, *187*, 6–16. [[CrossRef](#)]
11. Drexler, J.; Ortiz, A.; Padture, N. Composition effects of thermal barrier coating ceramics on their interaction with molten Ca-Mg-Al-silicate (CMAS) glass. *Acta Mater.* **2012**, *60*, 5437–5447. [[CrossRef](#)]
12. Lavigne, O.; Renollet, Y.; Poulain, M.; Rio, C.; Moretto, R.; Brannvall, P.; Wigren, J. Microstructural characterisation of plasma sprayed thermal barrier coatings by quantitative image analysis. In *Quantitative Microscopy of High Temperature Materials*; IOM Communication: London, UK, 2001; pp. 131–146.
13. Morgan, D.; Jacobs, R. Opportunities and challenges for machine learning in materials science. *Annu. Rev. Mater. Res.* **2020**, *50*, 71–103. [[CrossRef](#)]
14. Butler, K.; Davies, D.; Cartwright, H.; Isayev, O.; Walsh, A. Machine learning for molecular and materials science. *Nature* **2018**, *559*, 547–555. [[CrossRef](#)] [[PubMed](#)]
15. Chen, W.; Lu, Y.; Li, J.; Zimmerman, B. Automatic classification of microstructures in thermal barrier coating images. In Proceedings of the 2017 IEEE International Symposium on Multimedia, Taichung, Taiwan, 11–13 December 2017; pp. 99–106.
16. Lu, Y.; Chen, W.; Wang, X.; Alisworth, Z.; Tsui, M.; Al-Ghaib, H.; Zimmerman, B. Deep learning-based models for porosity measurement in thermal barrier coating images. *Int. J. Multimed. Data* **2020**, *11*, 20–35. [[CrossRef](#)]
17. Yunus, M.; Alsoufi, M. Prediction of mechanical properties of plasma sprayed thermal barrier coatings (TBCs) with genetic programming (GP). *Int. J. Eng. Trends Technol.* **2017**, *47*, 139–145. [[CrossRef](#)]
18. Ye, D.; Wang, W.; Zhou, H.; Fang, H.; Huang, J.; Li, Y.; Gong, H.; Li, Z. Characterization of thermal barrier coatings microstructural features using terahertz spectroscopy. *Surf. Coat. Technol.* **2020**, *394*, 125836. [[CrossRef](#)]
19. Ye, D.; Wang, W.; Xu, Z.; Yin, C.; Zhou, H.; Li, Y. Prediction of thermal barrier coatings microstructural features based on support vector machine optimized by cuckoo search algorithm. *Coatings* **2020**, *10*, 704. [[CrossRef](#)]
20. Ma, Z.; Zhang, W.; Luo, Z.; Sun, X.; Li, Z.; Lin, L. Ultrasonic characterization of thermal barrier coatings porosity through BP neural network optimizing Gaussian process regression algorithm. *Ultrasonics* **2020**, *100*, 105981. [[CrossRef](#)] [[PubMed](#)]
21. Padture, N.; Gell, M.; Jordan, E. Thermal barrier coatings for gas-turbine engine applications. *Science* **2002**, *296*, 280–284. [[CrossRef](#)] [[PubMed](#)]
22. 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.* **2018**, *8*, 2078. [[CrossRef](#)] [[PubMed](#)]
23. Chowdhury, A.; Kautz, E.; Yener, B.; Lewis, D. Image driven machine learning methods for microstructure recognition. *Comp. Mater. Sci.* **2016**, *123*, 176–187. [[CrossRef](#)]
24. Sudre, C.; Li, W.; Vercauteren, T.; Ourselin, S.; Cardoso, M.J. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*; Cardoso, M., Arbel, T., Carneiro, G., Syeda-Mahmood, T., Tavares, J., Moradi, M., Bradley, A., Greenspan, H., Papa, J., Madabhushi, A., et al., Eds.; Springer: Cham, Switzerland, 2017; pp. 240–248.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.