



Original article

Thermal conductivity prediction of Al₂O₃-doped tetragonal YSZ coatings using deep learningQiaochuan Chen^a, Sifan Han^a, Xuemei Song^{d,e}, Yi Zeng^{d,e}, Yuexing Han^{a,b,c,*}^a School of Computer Engineering and Science, Shanghai University, 99 Shangda Road, Shanghai 200444, China^b Zhejiang Laboratory, Hangzhou 311100, China^c Key Laboratory of Silicate Cultural Relics Conservation (Shanghai University), Ministry of Education, China^d The State Key Lab of High Performance Ceramics and Superfine Microstructure, Shanghai Institute of Ceramics, Chinese Academy of Sciences, Shanghai 200050, China^e Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences, Beijing 100049, China

ARTICLE INFO

Keywords:

Thermal conductivity
Deep learning
Computer vision
Al₂O₃-doped t-YSZ coatings
Electronic structure

ABSTRACT

Thermal barrier coatings, including Al₂O₃-doped tetragonal yttria-stabilized zirconia (t-YSZ) coatings are vital in diverse applications. Thermal conductivity is a key property, but prediction often requires some complex experiments. In this study, a novel dual-structure feature extraction coupled with a multi-scale attention fusion network (RCFNet) is introduced for accurate thermal conductivity prediction using electron backscattered diffraction (EBSD) microstructure images. The method requires minimal data, and the result shows high accuracy with a Mean Square Error (MSE) of 0.003 W(m·K)⁻¹ and an R-squared (R²) value of 0.81. Furthermore, the model reveals specific correlations between microstructural characteristics and thermal conductivity, facilitating rapid structure selection for specific thermal conductivity.

1. Introduction

Thermal barrier coating materials have a wide range of applications, covering diverse fields from aero-engines to everyday household items. They make a significant contribution to the aerospace industry [1], providing excellent heat and corrosion resistance to engines, as well as effective lightweighting that drives the continuous advancement of aerospace technology. Whether in high-tech engineering or everyday life, thermal barrier coating materials play an important role, providing opportunities for sustainable development and innovation in a wide range of industries. Material images can provide a wealth of information, including material texture and microstructure, which can be directly related to material properties [2]. Therefore, in numerous applications involving coating materials with high-performance demands, investigating the image-to-property relationship becomes crucial. Exploring the image-to-property relationship can provide meaningful insights into the correlation between a material's microstructure and its corresponding performance. These findings can, in turn, lead to the development of high-performance coating materials. An in-depth analysis of the relationship between an image and its properties can strongly support and guide the innovation and development of thermal barrier coating materials.

Predicting thermal conductivity in thermal barrier coating materials is challenging for several reasons. The microstructures of thermal

barrier coating materials are complex, with numerous detailed features and interactions. Accurately exploring the correlation function between microstructure and thermal conductivity in a large structural space is a complex task. The microstructure of thermal barrier coating materials is complicated, involving multiple scales and complex shapes. It is therefore difficult to find effective features fitting for modeling and corresponding them to thermal conductivity. Additionally, the thermal conductivity of coating material is often influenced by a combination of factors, including interfacial thermal resistance, particle shape, and distribution. The interaction of these factors increases the complexity of predicting thermal conductivity, while the properties of materials are closely linked to their structures [3,4]. Even small structural changes can significantly impact their properties [5–7]. Non-homogeneous materials present an obstacle in predicting thermal conductivity due to the complexity of heat conduction paths. Moreover, data scarcity is a prevalent hurdle encountered in predicting the thermal conductivity of coating materials in the field of materials. Preparing and testing materials to obtain an ample amount of precise experimental data is a demanding and expensive task. Lack of data may result in inadequate model training, which can affect the accuracy and reliability of predictions. Studying the image-to-property relationship for predicting the thermal conductivity of thermal barrier coating materials is hard due

* Corresponding author at: School of Computer Engineering and Science, Shanghai University, 99 Shangda Road, Shanghai 200444, China.

E-mail address: han_yx@shu.edu.cn (Y. Han).

<https://doi.org/10.1016/j.jeurceramsoc.2024.04.057>

Received 24 October 2023; Received in revised form 12 December 2023; Accepted 22 April 2024

Available online 24 April 2024

0955-2219/© 2024 Published by Elsevier Ltd.

to factors such as complex microstructure, multifactorial impacts, and data shortage.

Global and local features are important when exploring the correlation between the microstructure of material images and material properties. Convolutional Neural Networks (CNNs) perform well in image classification tasks [8]. Transformer [9] was first used for sequence-to-sequence predictive modeling in Natural Language Processing (NLP) tasks, ViTDosovitskiy et al. [10] in 2020 segmented each image into patches with positional embeddings. Sequences of tokens are constructed and a cascaded transformer block is applied to extract parameterized vectors as visual representations to model global semantic information through complex spatial transformations and long-range feature dependencies.

However, CNNs are usually limited by the fixed receptive field size of the convolutional kernel, leading to difficulties in capturing global features when processing data such as material images with highly complex structures and multi-scale features. Although CNNs are able to capture local information and form abstract features in images through convolutional operations to a certain extent, the limitations of their receptive fields make them inadequate for characterizing large-scale, global information.

On the contrary, the transformer model is able to model the global relevance of sequence data using its unique Self-Attention mechanism, which gives it an advantage in modeling long-range feature dependencies and global semantic information. However, since the transformer model focuses more on modeling global information, it may be relatively insufficient for capturing local information. Its self-attention mechanism, although capable of handling dependencies between different positions in a sequence, is not sufficiently attentive to the local details of fine microstructures in image data.

Considering the limitation that existing methods cannot extract both global and local features at the same time, it is especially necessary to create new model architectures. We propose a high-performance method called RCFNet to overcome the difficulty of image-to-property conversion, specifically for thermal conductivity analysis of thermal barrier coating materials. Our method utilizes a dual-network parallel feature extraction structure coupled with a multi-scale attention fusion module (Merge). It fully integrates the advantages of CNNs and transformers, adopts ResNet-50 with fewer parameters as the backbone of the global feature extraction module, and introduces the Multi-head Self-attention (MHSA) mechanism to enhance global information extraction. Meanwhile, we adopt CNN-based ConvNext as the backbone of the local feature extraction module, and this parallel framework of local and global feature modules can efficiently capture both local spatial contextual features and global semantic information representation at different scales simultaneously. This fusion algorithm can effectively fuse feature information from different scales of local and global information to explore the essential features in the material images and extract complex microstructures and factors affecting the thermal conductivity of the material. With a small dataset, it accurately predicts the thermal conductivity of thermal barrier coating materials and captures relevant microstructural regions that are most meaningful to corresponding thermal conductivity. The proposed network structure and multi-scale attentional fusion module enhance the capability of the network to capture crucial image features, resulting in improved performance and accuracy for the regression task.

In deep learning interpretability research, the Grad-CAM technique [11] has been proposed and used to visualize the operation and decision-making process of deep learning CNNs models. Inspired by this, we also use the Grad-CAM technique to automatically identify important microstructural regions that are highly related to material properties. This provides a deeper understanding of how microstructure affects material performance, facilitating optimization of coating material design and performance prediction. This method contributes to a clearer and more intuitive correlation between microstructure and material performance, improving our understanding of microstructure–property relationships.

To overcome the challenge of limited sample size during model training, we employ Han et al.'s HP-VAE GAN generation model for data augmentation [12]. This model progressively generates new images, transitioning from coarse to fine-grained and from low-resolution to high-resolution. The augmentation expands the original training dataset, thereby improving the performance of the model.

The primary contributions of this paper are:

1. To extract both global and local features from material images, we propose a novel dual-structure feature extraction and multi-scale attention fusion network and achieve very competitive results on the task of predicting thermal conductivity from images of thermal barrier coatings.
2. We propose a multi-scale attention fusion module that can hierarchically enhance, fine-tune, and fuse global and local features. Based on this design, we make full use of limited data, minimize information loss, and optimize the use of limited data.
3. Our algorithm introduces the Grad-CAM technique, which provides interpretability of the relationship between microstructure and performance of material images, which may be useful for inference of structural design based on performance.

Fig. 1 shows the complete process for predicting thermal conductivity utilizing EBSD maps of the thermal barrier coating materials.

2. Related work

The significance of the thermal conductivity of coating materials has led to extensive investigations by both domestic and foreign researchers in its determination. The main technique currently employed is an analytical model based on effective medium theory (EMT) [13]. This model considers the composite material as macroscopically homogeneous [14]. For predicting the effective thermal conductivity of non-homogeneous materials, the Bruggeman method is commonly employed [15]. The Maxwell model is another EMT model that is relatively simple and assumes that heat transfer occurs separately in the matrix and particle filler [16]. This method has been utilized to predict the thermal conductivity for the low volume fraction of spherical particle fillers. Certain updated Maxwell models, such as the one introduced by Hasselman and Johnson (1987) [17], which consider the effect of interfacial thermal resistance, can be used for composites that contain a low volume fraction of spherical, cylindrical, or flat-plate-shaped particles. An alternative approach to determining thermal conductivity involves conducting experimental assessments of specific heat, thermal diffusivity, and density. The resulting thermal conductivity is then calculated as the product of these three factors [18]. However, this technique is laborious, time-consuming, and highly inefficient, making it unsuitable for the industry's need to identify a specific material with thermal conductivity within a specific timeframe.

Recently, some researchers have introduced machine learning (ML) to predict the thermal conductivity of materials. Sargam et al. applied the fractional hot-deck imputation (FHDI) method to fill missing data in a dataset developed from published literature [19]. They evaluated the prediction performance of various ML algorithms and attempted to optimize prediction accuracy by tuning the hyperparameters of the ML algorithm. Finally, they used a robust multilayer perceptron machine model with high accuracy for prediction. Juneja et al. also investigated prediction models using other methods [20]. The authors calculated the lattice thermal conductivity of 120 dynamically stable and nonmetallic compounds. After analyzing the property maps generated from this dataset, which consisted of tabular data encompassing information on 195 compounds, including elemental compositions, lattice parameters, and properties associated with thermal conductivity, they used several descriptors, including the maximum phonon frequency, integrated Grüneisen parameter up to 3 THz, average atomic mass, and volume of the unit cell, to develop an ML model based on Gaussian regression process for prediction. Ning et al. used principal component analysis to

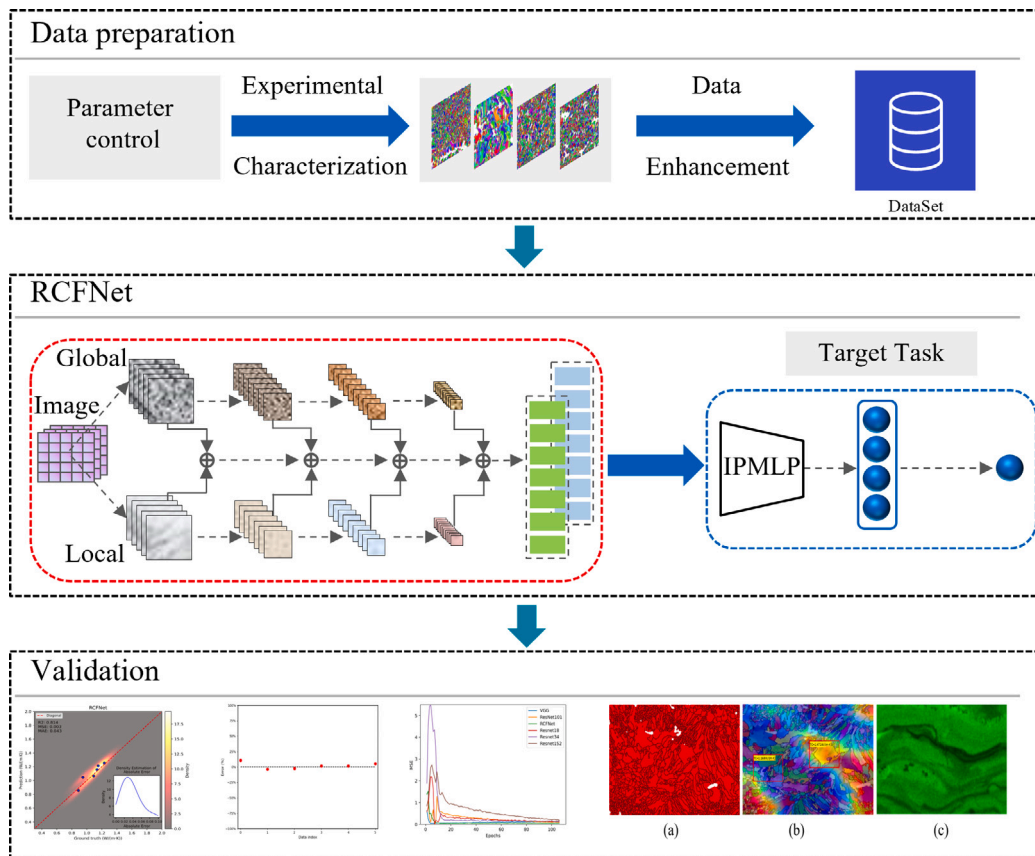


Fig. 1. Implementation of deep learning for predicting the thermal conductivity in thermal barrier coating materials.

determine the pressure [21], temperature, and porosity as the eigenvalues of the thermal conductivity of underground tar-rich coal seams. They used the Support Vector Regressor (SVR) model for prediction and achieved excellent results.

The inputs of the machine learning methods consist of physically based descriptors [22], implying that feature extraction is a manual process. Similar to the analytical methods, the final accuracy of the model heavily relies on hand-selected physically based microstructural descriptors. The interplay between the structure of materials and their properties is an intricate phenomenon that involves a significant amount of physical and chemical knowledge. Therefore, the researchers must possess a certain degree of a priori knowledge. Another limiting factor is that the material microstructure has a vast number of key features affecting the measured performance, turning the feature extraction of machine learning into an almost impossible task. Consequently, the machine learning approach has severe limitations. Since machine learning methods predictive performance is limited by feature extraction, researchers have incorporated deep learning into some aspects of material image analysis with the popularity of deep learning. Chen et al. proposed a multi-dimensional feature fusion method to address challenges in material images [23], such as complex textures, indistinct boundaries, and low contrast, leveraging datasets including spheroidite, wood, pearlite1, pearlite2, and ceram. Each category consists of 4 to 6 images, making up the training and test sets with a ratio of 3:1 or 2:1. This method improves network training with limited yet available annotated samples. Wu et al. used convolutional neural networks to predict material permeability from images by generating porous medium samples [24]. Similarly, Heidenreich et al. used microstructural information from material images to predict the effective yield surface of porous media based on images of their unit cell [25]. The dataset primarily consisted of extensive image sets, each possessing a 100×100 pixel resolution. Goodall and

Lee established a mapping between the effective mechanical properties of heterogeneous materials and mesoscale structures [26], utilizing image modeling and deep learning, which relied on datasets, including OQMD (256,620 materials), MP (43,921 non-metallic materials), and EX (3895 materials), to predict the effective mechanical properties of these materials. Utilizing 408 high-resolution images and tabular data on the chemical compositions of precipitates under varying aging conditions, Han et al. introduced deep learning to analyze the particle size evolution of precipitates in 2205 duplex stainless steel under high-temperature conditions [27], providing precise size data. This approach reduces subjective factors and contributes to a better understanding of the microstructural evolution and precipitate behavior in stainless steel.

Moreover, several researchers have developed deep learning models to predict material performance based on image data. For example, Xiang et al. employed a deep learning-aided prediction approach [28], integrating textual data encompassing chemical composition, trace element content, and testing conditions, along with 4 scanning electron microscope (SEM) images per sample, totaling 2608 images, to predict the creep rupture time of Fe-Cr-Ni heat-resistant alloys. In 2018, Han et al. used the quartet structure generation set (QSGS) to generate composite material structure and apply Lattice Boltzmann method (LBM) to calculate the effective thermal conductivity [29]. The authors assumed the reliability of LBM-derived results. Following this, they constructed a dataset based on this information and used a convolutional neural network model to predict the effective thermal conductivity of composite materials. In 2019, the particle packing method and the quartet structure generation set were utilized to capture 3D microstructures of composite materials, which resulted in corresponding 2D cross section images. CNNs were then introduced to predict effective thermal conductivity. In 2022, Gong et al. employed transfer learning techniques [30], they applied the ResNet101 network to predict the effective thermal conductivity using a dataset of UO_2/BeO composite material created through finite element methods. The results

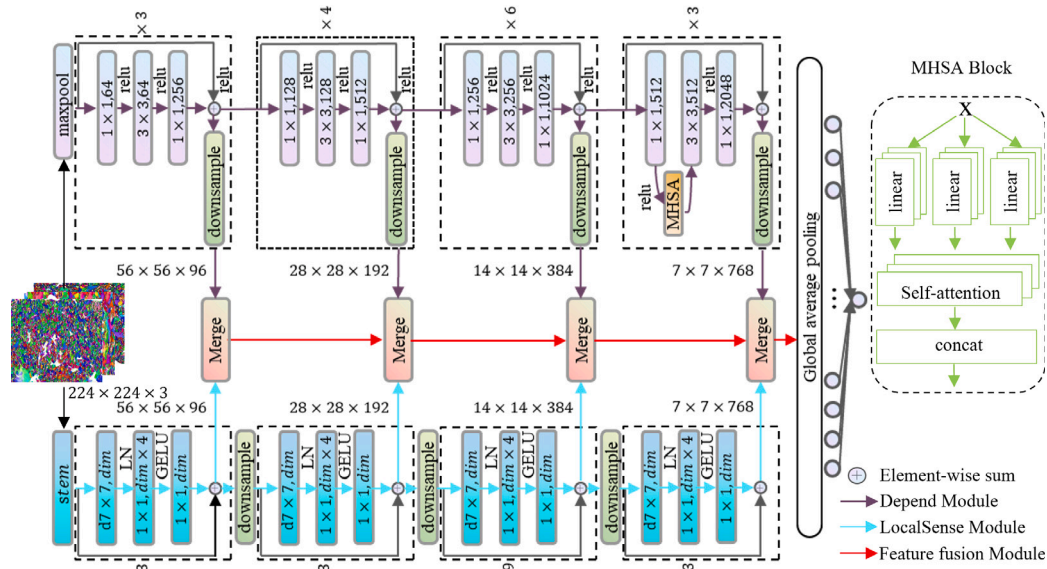


Fig. 2. Structure of RCFNet. The LocalSense module is used to extract local information from microstructure images, the Depend module is utilized for extracting global information, and the Merge module is employed for feature fusion.

showed that the ResNet101 network, characterized by its residual structure, exhibited superior predictive performance compared to conventional CNNs architectures. In 2023, Du et al. employed an enhanced AlexNet network to predict the effective thermal conductivity of sintered silver [31]. This was carried out based on a dataset encompassing authentic microscopic structures of 6156 samples, along with corresponding reliable finite element simulation-derived effective thermal conductivity.

Currently, research regarding the utilization of deep learning methods for predicting the thermal conductivity of materials is mainly conducted on numerous sample data generated by QSGS or the particle packing method. It is rarely experimented on real data obtained through actual experiments. However, the data obtained through actual experiments have increased complexity and diverse distribution, which raises the difficulty of prediction. Furthermore, material data acquisition involves highly resource-intensive experiments, often, only a small number of real samples can be gathered, further increasing the difficulty of the regression task. To address this problem of a small dataset, deep learning typically utilizes data augmentation, such as with generative models like the GAN series [32]. This approach generates many virtual samples based on the existing samples. Nonetheless, due to its need for a large number of initial samples, this method is not applicable in the material domain, where a high amount of initial samples is necessary. Song et al. [33] proposed a fusion spatial attentions method for few-shot learning from the perspective of reducing information loss, which performs fusion spatial attentions in both the original image space and the embedding space. The proposed SOD module in the image space extracts the saliency map of an image. The proposed Ada-P module introduces a meta-learner to adaptively fuse local features of the feature maps for each individual embedding. Although this approach achieves very competitive results on widely used benchmark datasets (e.g., miniImageNet, tieredImageNet, and CUB), it lacks the initial extraction of more comprehensive and useful feature information. Therefore, it may be less applicable to tasks that require comprehensive global and local features of an image. Our proposed RCFNet model has two key advantages: first, it can simultaneously extract global and local features to cope with the diversity and complexity of data distribution. Second, our multi-scale attention fusion block can enhance features at multiple levels and focus on key features across scales, thus improving the efficiency of key feature utilization. This leads to excellent performance on small sample data.

3. Method

3.1. RCFNet

Combining the local and long-range dependence information in EBSD images is crucial to leverage valuable information and improve the accuracy of the thermal conductivity prediction model for coating materials. The overall structure of RCFNet is shown in Fig. 2. The LocalSense Module is used to extract local region information from images, while the Depend Module extracts long-range dependencies information between local information. Both modules comprise four stages that extract features at different scales.

The process happens through two parallel paths that are divided into four stages specifically designed for feature extraction at different scales. The LocalSense Module normalizes the input data's features and transforms them from three channels to 96 channels through the stem layer. The Depend Module employs a MHSA to establish associations and capture long-range dependencies between the features extracted in the first three stages. In MHSA, the head is a key component unit within it, and each head is responsible for learning features in a different representation space. It is difficult for a single head to capture all the information in the feature space. The role of each head is mainly manifested in two aspects: first, it extends the model's ability to pay attention to different locations; second, it provides multiple representation subspaces for the attention layer, and different subspaces represent diversified information features. Our model employs four heads, each independently computing features during self-attention computations, thereby enhancing the model's understanding of the feature space. Afterward, they are combined and subjected to a linear transformation to discover relationships between different regions within the image.

This two-branch structure allows features from both local and global levels of the image to be obtained in parallel without interfering with each other. The Merge Module combines information from the two branches and transfers it to the Merge Module in the subsequent stage. For each scale of the Merge Module, the following operations are performed on the features from the LocalSense Module and the Depend Module respectively: a combination of spatial and channel attention (scSE) mechanisms [34], to enhance meaningful features and suppress useless features, followed by a fine-grained adjustment through the spatial attention (SA) mechanism and the channel attention (CA) mechanism respectively. Finally, the fused features undergo linear

Table 1
Specific RCFNet model parameters.

Input	224 × 224 × 3								
Stage	Output	LocalSense	Merge				Depend		
Pool	56 × 56,64	–	–				3 × 3,sd 2		
Stem	56 × 56,96	4 × 4,96,sd 4	–				–		
1	56 × 56,96	7 × 7,96 gp 96 1 × 1,384 1 × 1,96	×3	scSE	SA	CA	scSE	1 × 1,64	×3
				CA				3 × 3,64	
								sd 1	
								1 × 1,256	
								1 × 1,96	
						sd 1			
2	28 × 28,192	7 × 7,192 gp 192 1 × 1,768 1 × 1,192	×3	scSE	SA	CA	scSE	1 × 1,128	×4
				CA				3 × 3,128	
								sd 2	
								1 × 1,512	
								1 × 1,192	
						sd 1			
3	14 × 14,384	7 × 7,384 gp 384 1 × 1,1536 1 × 1,384	×9	scSE	SA	CA	scSE	1 × 1,256	×6
				CA				3 × 3,256	
								sd 2	
								1 × 1,1024	
								1 × 1,384	
						sd 1			
4	7 × 7,768	7 × 7,768 gp 768 1 × 1,3072 1 × 1,768	×3	scSE	SA	CA	scSE	1 × 1,512	×3
				CA				MHSA	
								w,h14	
								head 4	
								1 × 1,2048	
Average pool, 1-d fc									

*gp = groups, CA = Channel Attention, SA = Spatial Attention and sd = stride.

transformation through global average pooling and normalization to obtain the thermal conductivity. The exact parameters, including the model architecture, input data, layer details, and details about the convolutional kernel sizes and strides used, are shown in Table 1.

3.2. Depend module

Typically, the internal structure and composition of a material determine its performance. This information should not only come from a localized image region but also consider the interdependence between various regions comprehensively, as indicated in Ref. [35]. Therefore, it is crucial to obtain the necessary dependence information between different parts of the image. In this context, the MHSA is utilized to address dependencies between local parts of the image. This mechanism can take into account the interactions between different locations when computing the attention weights. Through computing attention weights on multiple heads, each head can focus on a different feature subspace, thereby capturing global dependencies at different scales and semantic levels. The multi-head mechanism can synthesize multiple perspectives and semantic information, resulting in a more comprehensive understanding of the global structure and associative relationships in an image. Time complexity measures the relationship between the running time of an algorithm and the size of the problem. Specifically, it represents the relationship between the number of basic operations required for the execution of the algorithm and the size of the inputs. Eq. (1) provides the time complexity of the multi-head self-attention mechanism.

$$\Omega(MHSA) = 4hwC^2 + 2(hw)^2C \quad (1)$$

where h represents the height of the feature map, w represents the width of the feature map, and C represents the number of channels in the feature map. The four feature extraction stages of the Depend

Module all utilize residual structures, and in the final stage, a multi-head self-attention mechanism is employed. This process can be defined by the Eq. (2):

$$D_i = f^{1 \times 1} (MHSA (LN (f^{1 \times 1} (D_{i-1})))) + D_{i-1} \quad (2)$$

where D_i represents the features obtained from the i th stage, $f^{1 \times 1}$ denotes the convolution operation with a kernel size of 1, and LN signifies layer normalization. Finally, the interdependent information extracted from various parts of the image is incorporated into the Merge module for fusion.

3.3. LocalSense module

In material images, it is common to find regions with distinct characteristics or properties. For instance, these regions may differ in terms of organizational structures, shapes, or components. Analyzing these local features can capture the differences between the regions, highlight their importance in the material image, and draw focus toward particular regions. To extract specific local information of the material image, we utilize Depthwise convolution within each sub-module of the LocalSense Module, to significantly reduce computing and storage resources demand, all while maintaining a certain level of model performance. This convolution processes each channel of the input feature map with its respective convolution kernel separately, to obtain the feature map of the corresponding channel. These output feature maps are merged by channel into the final output feature maps; the local area features in the image then are derived by employing the smoother and unsaturated activation function GELU and linear transformation. The whole procedure can be expressed using Eq. (3).

$$L_i = f^{1 \times 1} (LN (f^{d \times 7} (L_{i-1}))) + L_{i-1} \quad (3)$$

where $f^{d \times 7}$ represents the deep convolution with a kernel size of 7, and L_i represents the features obtained from the i th stage. Ultimately,

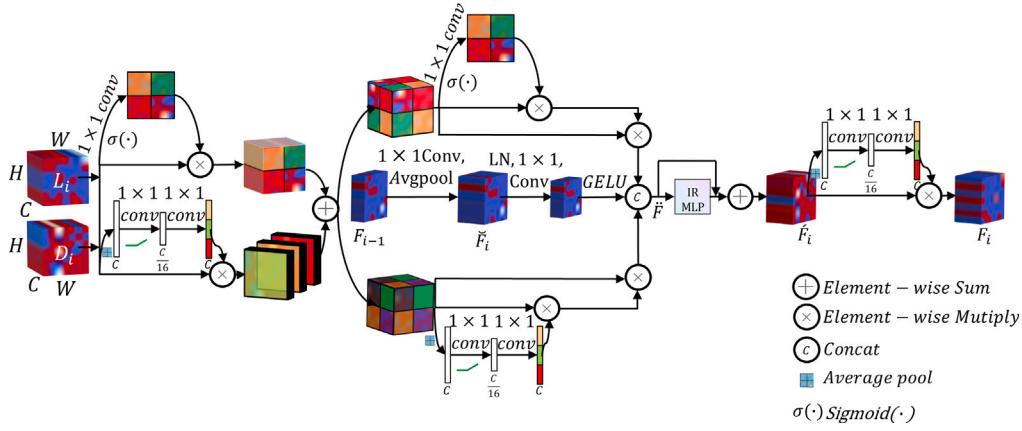


Fig. 3. Merge module.

the localized features of specific regions within the image are fed into the Merge module for fusion.

3.4. Merge

At each stage, the fusion module Merge can adaptively fuse features from the Depend Module and LocalSense Module, as well as semantic information from the previous Merge module, as shown in Fig. 3. Here, L_i represents the information on the local focus of the image at the i th stage obtained from the LocalSense Module, D_i represents the information regarding long-range dependencies between localizations in the image at the i th stage obtained from the Depend Module, and F_i represents the information obtained from the fusion module at the i th stage.

The merge module follows a sequence of feature enhancement, fine-grained adjustment, and further feature refinement. In particular, the scSE attention mechanism enhances the discriminative and expressive ability of the features from the Depend Module and LocalSense Module, respectively, ensuring an excellent distinction between categories or semantic information representation, making use of useful information in the image, resulting in a rich and a helpful feature representation. The channel attention mechanism is used to further refine the features from the Depend Module that have been processed by the scSE attention mechanism. Since the focus is on the interdependence between various regions in the image, the fine-grained adjustment helps the network to pay more attention to the important channels' positions in a targeted way. Likewise, the features from the LocalSense Module that undergo processing from the scSE attention mechanism to concentrate on specific localized regions in the image, are further adjusted through the spatial attention mechanism at a finer level, which allows the network to emphasize and employ the vital spatial location information. Lastly, the two types of features mentioned above are concatenated with the results from the previous merge module, and subjected to a linear transformation and residual connection to obtain the fused outcome. Subsequently, a channel attention mechanism is applied for further refinement. This process can be represented by the following steps:

First, the feature maps from both modules are modeled with scSE attention to capture the spatial relationship between them, as well as the channel relationship, simultaneously, as defined in the Eqs. (4) and (5):

$$\hat{D}_i = \text{scSE}(D_i) \otimes D_i \quad (4)$$

$$\hat{L}_i = \text{scSE}(L_i) \otimes L_i \quad (5)$$

where \otimes represents element-wise multiplication, \hat{D}_i and \hat{L}_i indicate the outputs obtained by the scSE attention mechanism. The method to

process \hat{D}_i via the channel attention mechanism and \hat{L}_i through the spatial attention mechanism can be expressed as Eqs. (6) and (7):

$$\tilde{D}_i = CA(\hat{D}_i) \otimes \hat{D}_i \quad (6)$$

$$\tilde{L}_i = SA(\hat{L}_i) \otimes \hat{L}_i \quad (7)$$

where \tilde{D}_i and \tilde{L}_i represent the results obtained through the Channel Attention mechanism and the Spatial Attention mechanism, respectively. We downsample the output feature map F_{i-1} from the previous stage of the Merge module to facilitate fusion with \tilde{D}_i and \tilde{L}_i as inputs to the Merge fusion module, as defined in the Eq. (8):

$$\check{F}_i = \frac{1}{|\mathcal{R}_{ij}|} \sum_{(p,q) \in \mathcal{R}_{ij}} (F_{i-1})_{pq} \quad (8)$$

where \check{F}_i represents the downsampling result from the previous Merge module, and $|\mathcal{R}_{ij}|$ denotes the number of elements in rectangular region \mathcal{R} . We fuse \tilde{D}_i , \tilde{L}_i , and \check{F}_i to fully leverage multi-directional information, as defined in the Eq. (9):

$$\ddot{F}_i = f^{3 \times 3}(\tilde{D}_i \oplus \tilde{L}_i \oplus \check{F}_i) \quad (9)$$

Here, \oplus denotes concatenation along the dimension 1, and \ddot{F}_i represents the fusion result from Depend Module, LocalSense Module, and the downsampling output of the previous merge module. After processing the \ddot{F}_i with the *IRMLP* module, we use the channel attention mechanism to further refine the features and obtain the final result, as presented in Eq. (10):

$$F_i = CA(\text{IRMLP}(\ddot{F}_i) + \ddot{F}_i) \quad (10)$$

the *IRMLP* module involves convolution and downsampling operations, with F_i representing the outcome obtained from the merge module.

Following each iteration of the model on the training set, the validation set, uninvolved in training, is used to test the model's generalization ability by computing the MSE to measure the degree of deviation between predictive and true values. Ultimately, the model with the lowest MSE value on the validation set is chosen as our final model. Then, we evaluate the MSE and the R^2 metrics on the test dataset to signify the degree of deviation between predictive and true values, as well as to measure the fit of the model, as presented in Eqs. (11) and (12), respectively. The R^2 provides a measure of how well the model fits the data, with a value closer to 1 indicating a better fit.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (11)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (12)$$

In the equations, n represents the number of samples, y_i represents the true value of the i th sample, \hat{y}_i represents the predictive value by the model for the i th sample, and $\bar{\hat{y}}$ denotes the mean.

The network takes the EBSD images as the input. The images size are uniformly set to 3-channel, and the width and height are set to 224×224 pixels. The output of the network is the predictive thermal conductivity. During training, a batch size of 32 is used, and the training is conducted for 5000 epochs. To prevent overfitting, regularization terms [36], Drop_path, and an early stopping strategy are applied. The initial learning rate is set to 1×10^{-6} , and a cosine annealing schedule is employed to decay the learning rate gradually as the training progresses [37]. This approach effectively alleviates the issue of oscillations in the output results. The network utilizes the SGD optimizer to update the network's weight parameters based on the loss function values [38].

3.5. Correlation analysis

One drawback of deep learning is its lack of interpretability. However, in the field of neural network interpretability research, Grad-CAM is a frequently used method for visually explaining the autonomous feature selection of deep learning networks. The microscopic structure of a material is directly related to its intrinsic properties such as thermal conductivity. In this study, based on the trained RCFNet model, we use the Grad-CAM method to construct visualizations of features, identifying the key local microscopic structures most closely related to the thermal conductivity of coating materials. We obtain the output of the last convolutional layer of the feature fusion algorithm, denoted as $f_{Merge4-map}$. The Grad-CAM method executes backpropagation based on the loss derived from the final predictive values. This process attains gradient information about $f_{Merge4-map}$, denoted as $Grad_{Merge4-map}$. Computing the channel-wise mean of $Grad_{Merge4-map}$ facilitates obtaining weights for each channel of the original feature $f_{Merge4-map}$. These weights are then weighted sum, followed by the ReLU activation to generate the final class activation map. This map visually represents the structures corresponding to the performance, providing a visualization of the structures most relevant to the predictive performance.

4. Experimental framework and results

4.1. Datasets

In this work, YSZ coatings doped with different amounts of Al_2O_3 are prepared by plasma spraying, and the phase compositions of the coatings are identified by X-ray diffraction (XRD) (D8 ADVANCE, Bruker, Germany) and EBSD (NordlysNano, Oxford Instruments, UK). In total, 22 images of the microstructure of Al_2O_3 -doped t-YSZ are obtained. By experimental and computational methods, the laser flash technique (TD-79 A, SIC, China) is used to determine the thermal diffusivity (α) at room temperature. Differential scanning calorimetry is used to measure the specific heat (C_p). The Archimedes principle is used to determine the density (ρ) of each sample. Finally, the thermal conductivity (k) is then calculated using the equation $k = \alpha \times \rho \times C_p$.

From these 22 microstructure images, we randomly divide them into three subsets: 12 images for the training set, 4 images for validation, and 6 images for testing purposes. Then, we use Improved-HP-VAE-GAN to augment the training set. This method is based on a single image for training, which requires less data, and it incrementally generates images from coarse to fine, and from low-resolution to high-resolution. Ultimately, we augment our training set from the original 12 data samples to 12 000 data samples, which further improves the performance of the model.

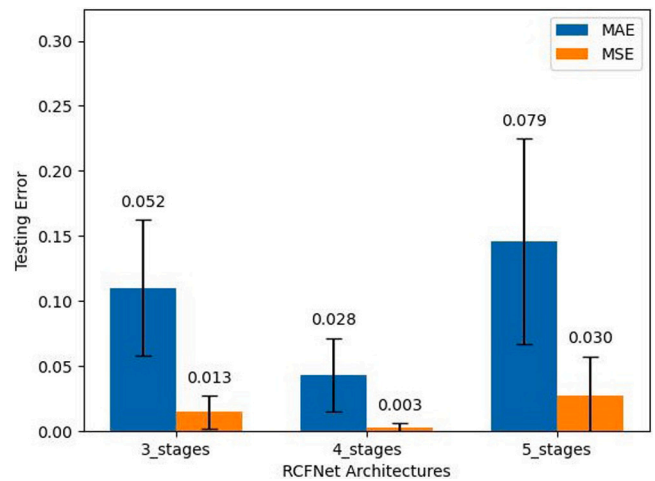


Fig. 4. Comparing the performance of RCFNet at different network depths: 3_stages, 4_stages, and 5_stages - using EBSD images that are $224 \times 224 \times 3$ in size as input.

4.2. Experimental design

In this section, we present our experimental framework in several subsections. Section 4.3 compares the performance of our model at different depths to determine the optimal configuration. Section 4.4 evaluates the accuracy of the RCFNet model in predicting thermal conductivity using metrics such as MSE and R^2 . In Section 4.5, we perform ablation experiments to isolate the effect of individual modules, such as the Depend and LocalSense modules, on the overall model performance. Section 4.6 provides a comparative analysis between our method and widely used deep learning models such as VGG, ResNet-101, ResNet-152, and ResNet-18.

Moreover, Section 4.7 compares our approach to traditional machine learning methods. Finally, in Section 4.8, we use our method to identify the key regions of the EBSD images and verify the correctness of the automatic model identification by analyzing the phase distribution map and the energy spectrum surface distribution.

4.3. Comparison of thermal conductivity prediction accuracy

The performance of a deep learning model depends heavily on the features it extracts. In general, increasing the depth of a deep learning model enables it to capture the underlying features of the data better, thereby improving its overall performance. However, the final performance of the model does not simply increase with its depth because increasing the depth of the model also increases the number of parameters. This phenomenon can lead to overfitting issues on the training data owing to the inadequate training samples where the model performs well on the training dataset but generalizes poorly on the test dataset [39].

Our research evaluates the generalization performance of RCFNet on the test dataset at different depths, as shown in Fig. 4. This figure illustrates the performance of the RCFNet network across various depths. Mean Absolute Error (MAE) and MSE are used to measure the deviation between actual and predictive values. MAE is a standard metric used in regression models that represents the average absolute error between predictions and actual values. The RCFNet network obtains the highest performance on the test set, at a depth of 4_stages, with corresponding MAE and MSE values of 0.043 and 0.003, respectively. At a depth of 4_stages, the MAE value decreases by 0.067 and the MSE value decreases by 0.012, compared to the depth of 3_stages. At the same time, the MAE value decreases by 0.103 and the MSE value decreases by 0.024, in comparison to the depth of 5_stages.

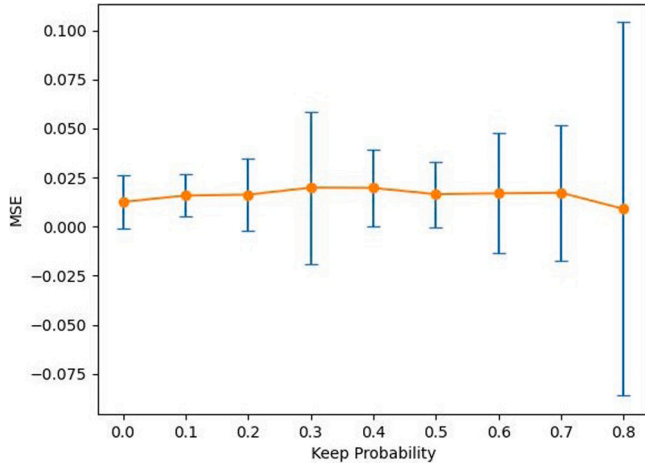


Fig. 5. Comparing the performance of the RCFNet algorithm using various keep probability values.

The performance of the model exceeds that of both 3_stages and 5_stages depths when the depth is set at 4_stages. This is because a deeper network can extract more abstract information from the input data. Consequently, it leads to improved model performance. However, as the network depth increases, overfitting can occur, bounded by the dataset's size. When the depth is set to 5_stages, the model overfits; In the forthcoming work, we will select the network depth as 4 stages.

Moreover, the data utilized in estimating the thermal conductivity of thermal barrier coating materials with deep learning solely comes from actual experimental measurements resulting in a notably limited dataset owing to various constraints. To prevent overfitting, we introduce a regularization term when training the model. Fig. 5 illustrates the model performance's variability with different Keep Probability values during training.

From Fig. 5, it can be observed that various Keep Probability values do not have a significant impact on the final performance of the model. As Keep Probability changes, there is a slight fluctuation in MSE, but the magnitude of this fluctuation is minimal. This indicates that the model's performance is not highly sensitive to the Keep Probability. This might be attributed to the fact that we have already employed normalization, which has helped mitigate overfitting to some extent. Ultimately, the depth of the RCFNet model is determined as 4_stages, with a Keep Probability value of 0.00003, which offers slightly improved performance comparatively.

4.4. Performance of RCFNet

The effectiveness of RCFNet in accurately predicting the thermal conductivity of thermal barrier coating materials using EBSD images is shown in Fig. 6. The red line represents a perfect alignment between actual and predictive values, where points closer to the red dashed line indicates a higher degree of proximity between predictive and actual values. The graph reveals that the model's predictive values cluster around the vicinity of the actual values. On the test set, the MSE, MAE, and R^2 are 0.003, 0.043, and 0.814, respectively. These results confirm that our model accurately predicts the thermal conductivity based on the EBSD images of thermal barrier coating materials.

In this context, the absolute error is defined as the difference between the predictive value and the actual value, formulated as Eq. (13). Here, k_e represents the absolute error between the actual and predictive results of thermal conductivity. The k_{RCFNet} represents the predicted values by the model, and the k_{true} represents the true values. We employ Gaussian kernel density estimation to obtain the kernel density curve of the absolute error, as shown in the inset of Fig. 6. It is clear that the

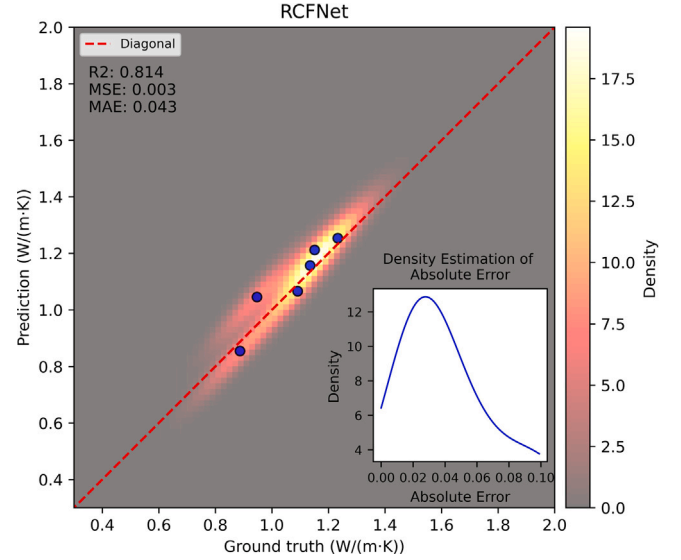


Fig. 6. Comparison between the predictive values and actual values using the RCFNet model. Points falling on the red dashed line represent a perfect alignment between predictive and actual values. The kernel density curve is depicted in the inset graph.

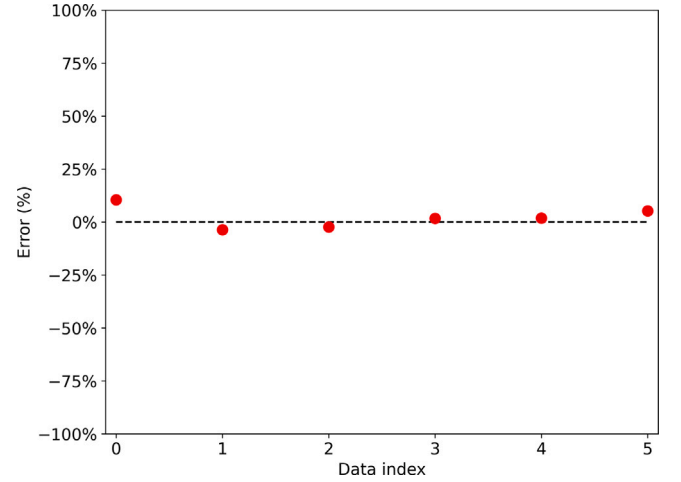


Fig. 7. The deviation ratio of RCFNet on the test set, where points on the black dashed line indicates perfect alignment between predictive and actual values, with closer proximity to the dashed line indicating smaller deviations.

absolute error is concentrated around $0.003 \text{ W (m K)}^{-1}$, affirming the accuracy of our model in predicting the thermal conductivity of thermal barrier coating materials.

$$k_e = k_{RCFNet} - k_{true} \quad (13)$$

To quantify how much the deviation between the predictive value and the actual value of the model, we illustrate the offset of the model's predictive values from the actual values in Fig. 7, as defined as Eq. (14):

$$k_{error} = \frac{k_{RCFNet} - k_{true}}{k_{true}} \quad (14)$$

k_{error} is defined as the ratio of the difference between the predictive values of the model and the actual values in the test set.

4.5. Ablation experiment

We conduct ablation experiments on various components of RCFNet using EBSD image data of the thermal barrier coating materials with

Table 2Algorithm's MSE, MAE, and R^2 scores with different modules.

	MSE	MAE	R^2
Depend module (Resnet50)	0.011	0.087	0.259
+mhsa	0.008	0.066	0.411
+LocalSense module (+Merge)	0.003	0.043	0.814

accurately measured thermal conductivity obtained from our actual experiments. We evaluate the performance of different parts of the algorithm separately.

Starting with the Depend Module, we begin by using only the pre-trained ResNet50 model for transfer learning. Then, with the addition of the multi-head self-attention mechanism, MSE and MAE are reduced by 0.003 and 0.021, respectively. This demonstrates the importance of considering the correlations among various local information in EBSD image data. Building upon this foundation, we introduce the LocalSense Module to focus on crucial local information of the model. The information from the LocalSense Module is then fused with that of the Depend Module by a merge model, resulting in our RCFNet model. This composite model exhibits further improvements across all metrics, ultimately achieving MSE and MAE of 0.003 and 0.043, respectively, and an R^2 value of 0.814. The model's performance has improved by 98.1% compared to before feature fusion. The detailed results are shown in Table 2.

4.6. Comparison between RCFNet and other similar algorithms

Compared with other similar methods, our work faces a significant challenge due to the scarcity of annotated experimental data from EBSD images of thermal barrier coating materials. The performance of deep learning models is closely tied to the amount of data available. However, to address the issue of limited data, our model employs multi-network feature extraction and multi-scale attention fusion mechanisms. These mechanisms greatly enhance important features while suppressing noise features, thus reducing our reliance on data.

To validate the effectiveness of our algorithm, we also test other models, such as VGG, ResNet-101, ResNet-152, etc., that achieved good performance in similar work on predicting the thermal conductivity of materials based on the dataset obtained from our experiments. Each model is tuned to the optimal hyperparameters such as learning rate, batch size, dropout path rate, and selecting the most appropriate optimizer to achieve optimal performance. The parameter optimization information for each model is shown in Table 3. To make a fair assessment, we employ the MSE metric as a quantitative measure to compare performance among the models. Fig. 8 illustrates the change in MSE values on the validation set over the course of training for these six models.

It can be observed that among the ResNet-based models where only the input and output layers are trained while keeping the middle layers frozen, ResNet-101 performs the best, fluctuating around an MSE value of approximately 0.1. The VGG network exhibits fluctuations in MSE values around 0.13. In contrast, our RCFNet converges to an MSE value around 0.05. In other words, it is clear that the other models do not perform well on our dataset for the task of predicting thermal conductivity. The performance of deep learning models is heavily influenced by both the quantity of data and the quality of features extracted by the model. VGG and alternative ResNet-based methods are evidently constrained by the dataset size and their own structural limitations in fully exploiting image data features. When dealing with a particularly small dataset, they struggle to efficiently utilize the information within the data, leading to higher data requirements. In contrast, our model can effectively leverage the limited data to establish a robust relationship between material structure and performance, resulting in exceptional performance.

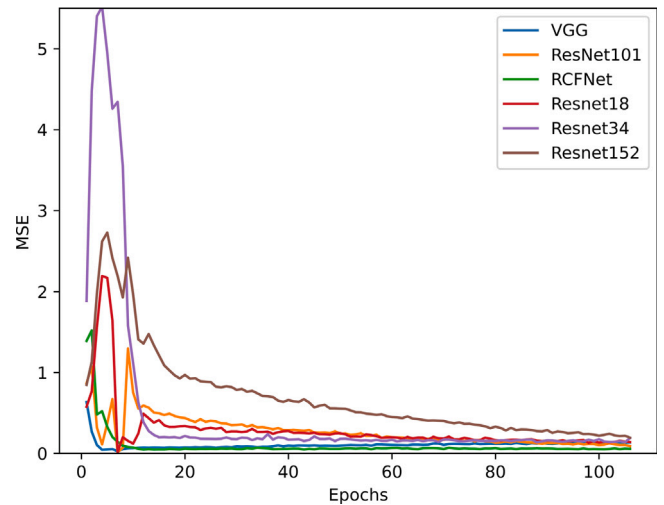


Fig. 8. Comparison of the performance of RCFNet, VGG, and ResNet series network models on the dataset used in this study. The horizontal axis represents the number of network iterations, while the vertical axis represents the MSE values.

4.7. Comparison between RCFNet and machine learning models

We extract five key features most relevant to thermal conductivity from the image data of thermal barrier coating materials: grain size distribution, aspect ratio distribution, grain boundary density, average grain size, and resolution. Based on these features, we carry out a regression task to predict the thermal conductivity using traditional machine learning models: Gradient Boosting Regressor (GBTR), K-Nearest Neighbors, Multi-Layer Perceptron (MLP), Random Forest, Ridge, and SVR. The performance of these models on the test set is presented in Fig. 9. Among these models, K-Nearest Neighbors exhibited the best performance, with an MSE of 0.019, MAE of 0.125, and an R^2 of -0.318 . The negative R^2 indicates its performance slightly lags behind the average value of the test set, and it significantly trails the deep learning RCFNet model trained and tested on the same dataset. Table 4 outlines the performance of each method on the test set. This underscores that our deep learning model, RCFNet, demonstrates better generalization performance in predicting the thermal conductivity of thermal barrier coating materials compared to machine learning models utilizing the five physical descriptors.

This can be explained from two perspectives. Firstly, relative to the excellent feature extraction capabilities of deep learning RCFNet, using a set of five physical descriptors for predicting the thermal conductivity of thermal barrier coating materials is not comprehensive. These descriptors are insufficient to capture the intricate relationship between material image structure and corresponding performance. Another reason is that model performance is greatly influenced by the quantity of data. Although traditional machine learning models are much smaller compared to the deep learning RCFNet model, their performance suffers due to the simplicity of their input data. As a result, they require a larger quantity of well-annotated data for training.

4.8. Capture and validation of key feature extraction by RCFNet

We randomly select a data sample from the test set and utilize the RCFNet model to capture its key regions. For a more comprehensive analysis, we additionally obtain the phase distribution map and energy spectrum surface distribution of the aluminum (Al) element within that region. These pieces of information are presented in Fig. 10. As shown in Fig. 10(b), it can be seen that the thermal conductivity of the left-marked area is $1.1680 \text{ W (m K)}^{-1}$, while the thermal conductivity of the right-marked area is $1.471 \text{ W (m K)}^{-1}$. This suggests that the thermal

Table 3
Deep learning model comparison parameters.

Models	Stages/Layers	Dropout path rate	Loss function	Max epochs	Batch size	Optimizer	Learning rate
RCFNet	4	0.3	MSE	5000	32	SGD	1e–6
VGG	6	0.1	MSE	5000	64	RMSprop	1e–4
ResNet101	4	0	MSE	5000	128	SGD	1e–4
ResNet18	4	0.3	MSE	5000	32	SGD	1e–7
ResNet34	4	0.1	MSE	5000	64	SGD	1e–6
ResNet152	4	0.1	MSE	5000	32	SGD	1e–5

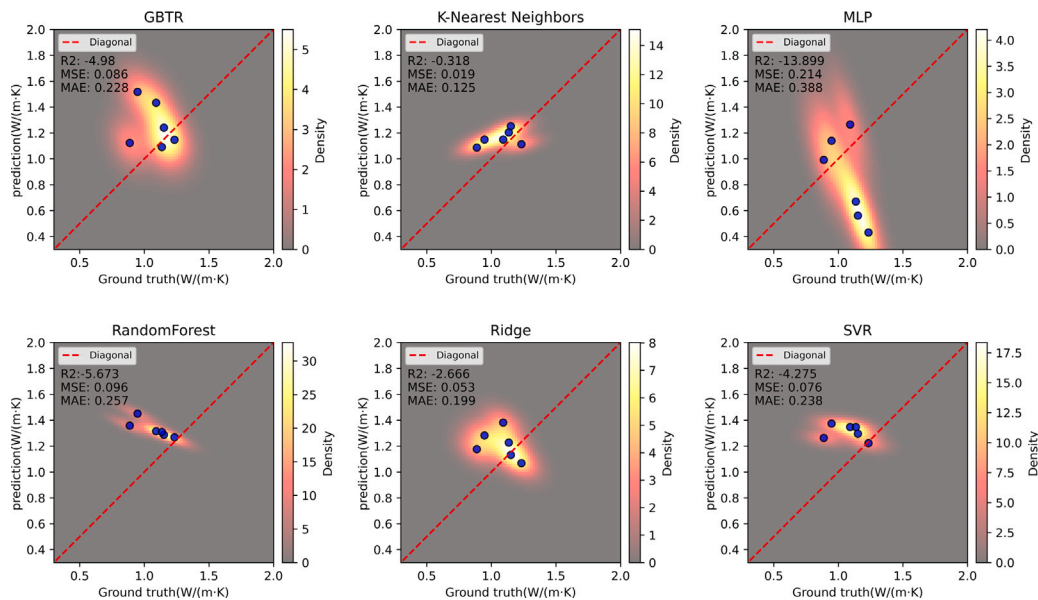


Fig. 9. The performance of conventional machine learning models on the dataset utilized in this research.

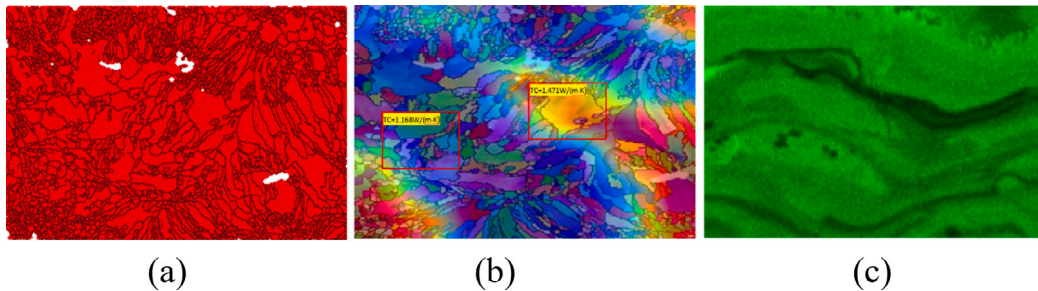


Fig. 10. (a) phase distribution map, (b) EBSD image, and (c) energy spectrum surface distribution map of the Al element.

Table 4
A comparison between RCFNet and machine learning methods, using MSE, MAE, and R² scores.

Model	MSE	MAE	R ²
GBTR	0.086	0.228	–4.980
K-Nearest neighbors	0.019	0.125	–0.318
MLP	0.214	0.388	–13.899
RandomForest	0.096	0.257	–5.673
Ridge	0.053	0.199	–2.666
SVR	0.076	0.238	–4.275
RCFNet (Depth = 4)	0.003	0.043	0.814

conductivity of the region marked on the left is significantly lower than that of the region marked on the right, which may be due to differences in material composition.

The sample is a YSZ coating doped with Al₂O₃. The entry of Al³⁺ into the YSZ lattice forms the Al₂O₃-YSZ solid solution, and its phase

structure is the tetragonal YSZ phase. Fig. 10(a) indicates that the coating is composed solely of the tetragonal YSZ phase without other impurities. During the Al₂O₃ doping process in YSZ, the substitutional atomic defect represented by Al'_{Zr} arises due to the replacement of Zr⁴⁺ by Al³⁺. Additionally, some Al³⁺ occupy interstitial positions within the ZrO₂ lattice, forming interstitial defects denoted as Al^{••}_i. This process is accompanied by the generation of oxygen vacancies V_O^{••}. Defects such as Al'_{Zr}, Al^{••}_i, and V_O^{••} enhance phonon scattering within the lattice, thereby reducing the phonon mean free path and subsequently leading to a decrease in the thermal conductivity of the coating. For the corresponding region indicated in Fig. 10(b). We collect its energy spectrum results. The spectral surface distribution of the Al element, shown in Fig. 10(c), reveals that the left-marked box area exhibits a higher enrichment of Al elements compared to the right-marked box area. This indicates that the left-marked area has a greater concentration of aluminum dopants in YSZ, leading to a larger number of defects. Consequently, this results in a lower thermal conductivity.

5. Conclusions

In this study, several Al_2O_3 -doped YSZ coatings are fabricated using APS. Their images of microstructures are obtained by EBSD. We propose a novel dual-structure feature extraction and multi-scale attention fusion network that can extract global and local features of an image simultaneously without interference. In addition, we propose a multi-scale attention fusion module that enables hierarchical enhancement, fine-grained tuning, and fusion of global and local features. This strategy optimizes the use of limited data while reducing information loss. In addition, our algorithm integrates the Grad-CAM technique to visualize the relationship between material microstructure and properties, thereby enhancing interpretability. It can also contribute to reverse engineering inference by inferring relevant structure from performance metrics.

Experimental result reveals significant advancements achieved by our method in predicting the thermal conductivity of coating. More precisely, our method attains an MSE of $0.003 \text{ W (m K)}^{-1}$ and an R^2 of 0.81 using the test dataset. These confirm the high precision and dependability of RCFNet, verifying it as an excellent method for predicting thermal conductivity in coating materials. We are convinced that the successful application of RCFNet, combined with a deeper understanding of coating material microstructures and performance, will yield valuable insights and inspiration for more research and applications in the field of coating material science. Lastly, our research provides new opportunities for expanding the potential uses of deep learning in the field of material science. In the future, we hope to show some SEM images and conduct in-depth analyses to observe potential artifacts within them, with a specific focus on aspects like shadows and grayscale biases.

CRedit authorship contribution statement

Qiaochuan Chen: Conceptualization, Reviewing and editing. **Sifan Han:** Software, Writing – original draft. **Xuemei Song:** Data curation, Visualization, Investigation. **Yi Zeng:** Formal analysis. **Yuexing Han:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by National Key Research and Development Program of China (2022YFB3707800), National Natural Science Foundation of China (Grant No. 52273228, 51971127), Natural Science Foundation of Shanghai, China (Grant No. 20ZR1419000), Key Research Project of Zhejiang Laboratory, China (No. 2021PE0AC02), Key Program of Science and Technology of Yunnan Province, China (202302AB080022), Shanghai Science and Technology Young Talents Sailing Program, China (23YF1412900).

References

- [1] T.N. Rhys-Jones, *Surf. Coat. Technol.* 43 (1990) 402.
- [2] N. Marzari, A. Ferretti, C. Wolverton, *Nat. Mater.* 20 (6) (2021) 736.
- [3] J. Fish, G.J. Wagner, S. Ketten, *Nat. Mater.* 20 (6) (2021) 774.
- [4] J.M. Rickman, T. Lookman, S.V. Kalinin, *Acta Mater.* 168 (2019) 473.
- [5] A.S. Namini, A. Motallebzadeh, B. Nayeibi, M.S. Asl, M. Azadbeh, *Mater. Chem. Phys.* 223 (2019) 789.
- [6] S. Adithya, K. Sachin, U.S. Meda, et al., *IOP Conference Series: Earth and Environmental Science*, Vol. 822, IOP Publishing, 2021, 012005.
- [7] N.A. Jaya, L. Yun-Ming, H. Cheng-Yong, M.M.A.B. Abdullah, K. Hussin, *Constr. Build. Mater.* 247 (2020) 118641.
- [8] S. Koitka, C.M. Friedrich, CLEF (Working Notes), Citeseer, 2016, pp. 304–317.
- [9] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, *Adv. Neural Inf. Process. Syst.* 30 (2017).
- [10] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, et al., 2020, arXiv preprint arXiv:2010.11929.
- [11] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, *Proceedings of the IEEE International Conference on Computer Vision*, 2017, pp. 618–626.
- [12] Y. Han, Y. Liu, Q. Chen, *Comput. Mater. Sci.* 226 (2023) 112250.
- [13] S.A. Hosseini, S. Khanniche, P.A. Greaney, G. Romano, *Int. J. Heat Mass Transfer* 183 (2022) 122040.
- [14] K. Pietrak, T.S. Wiśniewski, *J. Power Technol.* 95 (2015) 1.
- [15] Z. Zhao, H. Zhang, G. Zou, H. Ren, W. Zhuang, L. Liu, Y.N. Zhou, *J. Electron. Mater.* 48 (2019) 2811.
- [16] J.C. Maxwell, *Oxford 314* (1881) 1873.
- [17] D. Hasselman, L.F. Johnson, *J. Compos. Mater.* 21 (6) (1987) 508.
- [18] X. Song, Y. Xu, Y. Ding, J. Zhang, X. Guo, C. Jiang, W. Zheng, Y. Zeng, *Appl. Surf. Sci.* 542 (2021) 148553.
- [19] Y. Sargam, K. Wang, I.H. Cho, *J. Build. Eng.* 34 (2021) 101956.
- [20] R. Juneja, G. Yumnam, S. Satsangi, A.K. Singh, *Chem. Mater.* 31 (14) (2019) 5145.
- [21] X. Ning, L.D. Chang'an Wang, T. Zhu, X. Xue, D. Che, *International Conference on Applied Energy*, Vol. 19, 2021.
- [22] B.V. Goodarzi, A.R. Bahramian, *J. Therm. Anal. Calorim.* 147 (11) (2022) 6227.
- [23] Q. Chen, H. Wei, B. Wang, L. Ruan, Y. Han, *Mater. Today Commun.* 35 (2023) 105941.
- [24] J. Wu, X. Yin, H. Xiao, *Sci. Bull.* 63 (18) (2018) 1215.
- [25] J.N. Heidenreich, M.B. Gorji, D. Mohr, *Int. J. Plast.* 163 (2023) 103506.
- [26] R.E. Goodall, A.A. Lee, *Nat. Commun.* 11 (1) (2020) 6280.
- [27] Y. Han, R. Chi, Q. Chen, B. Wang, W. Liu, Y. He, *J. Mater. Res. Technol.* (2023).
- [28] S. Xiang, X. Chen, Z. Fan, T. Chen, X. Lian, *J. Mater. Res. Technol.* 18 (2022) 268.
- [29] H. Wei, S. Zhao, Q. Rong, H. Bao, *Int. J. Heat Mass Transfer* 127 (2018) 908.
- [30] Z. Gong, Z. Xu, J. Hu, B. Yan, X. Ding, J. Sun, P. Zhang, J. Deng, *Acta Mater.* 240 (2022) 118352.
- [31] C. Du, G. Zou, A. Zhanwen, B. Lu, B. Feng, J. Huo, Y. Xiao, Y. Jiang, L. Liu, *Int. J. Heat Mass Transfer* 201 (2023) 123654.
- [32] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, Y. Bengio, *Commun. ACM* 63 (11) (2020) 139.
- [33] H. Song, B. Deng, M. Pound, E. Özcan, I. Triguero, *Inf. Fusion* 81 (2022) 187.
- [34] A.G. Roy, N. Navab, C. Wachinger, *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16–20, 2018, Proceedings, Part I*, Springer, 2018, pp. 421–429.
- [35] S.G. Louie, Y.-H. Chan, F.H. da Jornada, Z. Li, D.Y. Qiu, *Nature Mater.* 20 (6) (2021) 728.
- [36] X. Ying, *Journal of Physics: Conference Series*, Vol. 1168, IOP Publishing, 2019, 022022.
- [37] I. Loshchilov, F. Hutter, 2016, arXiv preprint arXiv:1608.03983.
- [38] H. Robbins, S. Monro, *Ann. Math. Stat.* (1951) 400–407.
- [39] H. Zhang, L. Zhang, Y. Jiang, 2019 11th International Conference on Wireless Communications and Signal Processing, WCSP, IEEE, 2019, pp. 1–6.