What surface characteristics truly affect thermal contact resistance -An interpretability study based on deep learning and convolutional neural networks

Authors: Man Zhou, Zhuoyan He, Peiyao Guo, I am ready to translate your academic paper. Please provide the content containing the ···tags, LaTeX formulas, and citations that require translation from Simplified Chinese to English., Ping Zhang

Date: 2025-04-11T19:08:13+00:00

Abstract

Technological advancement presents new challenges for heat dissipation. Thermal contact resistance is a critical factor in efficient thermal management. To gain a deeper understanding of thermal contact resistance, we developed a deep learning model based on convolutional neural networks for its prediction, which also serves as a reference for determining the actual contact area between two surfaces. The model utilizes the DenseNet121 architecture, and the dataset is generated according to surface fractal theory and multi-point contact mechanics. It was trained on the training set and demonstrated strong predictive performance on the testing set. Two specimen sets, produced through grinding and turning processes, provided both surface morphology and measured contact thermal resistance, which were used to evaluate the model' s accuracy, yielding comparable results. Guided Backpropagation and Class Activation Mapping were utilized for the interpretability study of the model' s visualizations. The results indicate that the contact and non-contact areas of the two surfaces each influence the prediction outcomes, thereby validating the model' s effectiveness. The surface features impacting contact thermal resistance were also directly visualized. This approach offers a new methodology to interpret the effects of surface characteristics on thermal contact resistance, advancing understanding and prediction capabilities in this field.

Full Text

Preamble

What Surface Characteristics Truly Affect Thermal Contact Resistance−An Interpretability Study Based on Deep Learning and Convolutional Neural Networks

Man Zhou, Zhuo Yan He, Pei Yao Guo, Ping Zhang*
School of Mechanical and Electrical Engineering, Guilin University of Electronic Technology, No. 1 Jinji Road, Guilin, Guangxi 541004, China

*Corresponding author, E-mail: pingzhang@guet.edu.cn

Abstract

The advancement of technology presents new challenges for heat dissipation, with thermal contact resistance (TCR) emerging as a critical factor in efficient thermal management. To gain deeper insights into TCR, we developed a deep learning model based on convolutional neural networks (CNNs) for its prediction, which simultaneously serves as a reference for determining the actual contact area between two surfaces. Our model utilizes the DenseNet121 architecture, trained on a dataset generated according to surface fractal theory and multi-point contact mechanics. The model demonstrated strong predictive performance on both training and testing sets. Two specimen sets, fabricated through grinding and turning processes, provided surface morphology data and measured contact thermal resistance for model evaluation, yielding results comparable to theoretical predictions.

Guided Backpropagation and Class Activation Mapping were employed for model interpretability and visualization. The results reveal that both contact and non-contact regions of the two surfaces influence prediction outcomes, thereby validating the model' s effectiveness. The surface features impacting contact thermal resistance were directly visualized, offering a novel methodology to interpret the effects of surface characteristics on TCR and advancing understanding and predictive capabilities in this field.

Keywords: Thermal contact resistance, Heat transfer, Convolutional neural networks, Surface topography, Interpretability, Visualization

Introduction

The state-of-the-art development of technologies such as high-performance computing [?], next-generation chips [?], and renewable energies [?] generates immense heat densities, making thermal contact resistance (TCR) a pivotal factor in scenarios requiring precise thermal management [?]. Even minimal TCR at interfaces can significantly impair computational performance or destabilize qubits [?]. By addressing TCR, researchers not only resolve immediate thermal bottlenecks but also pave the way for breakthroughs in sustainability, miniaturization, and extreme-environment technologies, thereby solidifying its position as a cornerstone of modern thermal engineering [?].

When solids come into contact, only a fraction of their apparent contact area —determined by protruding asperities—facilitates heat conduction, while the remaining gaps, typically filled with low-conductivity fluids such as air, function as thermal barriers [?]. Thermal contact resistance is defined as the temperature drop across the interface per unit heat flux. Surface characteristics have long been recognized as critical factors influencing TCR [?], with extensive experimental research demonstrating that surface roughness significantly affects thermal contact resistance [?]. However, Cui et al. [?] highlighted that even when surfaces exhibit identical roughness values, variations in processing methods can result in substantial differences in TCR. To address this limitation, they proposed incorporating additional parameters such as wavelength and surface profile height, employing diverse mathematical formulations to better characterize contact surfaces across different processing techniques.

Various studies have utilized surface fractal models to analyze contact phenomena. Ma et al. [?] conducted experimental measurements of fractal parameters, accounting for elastic, plastic, and elastoplastic deformations, and developed a theoretical prediction model for thermal contact resistance. Similarly, Sun et al. [?] extended this approach to cylindrical contact surfaces, while Chen et al. [?] enhanced their analysis by incorporating thermal stress and asperity interactions. In summary, these studies either employ surface roughness metrics such as arithmetic average roughness (Ra) and root mean square roughness (Rq), or derive fractal parameters—including fractal dimension (D) and scaling factor (G)—from fractal theory. These parameters facilitate determination of total contact area under varying mechanical states, with thermal contact resistance subsequently analyzed through heat flow theory, representing another statistical approach to surface characterization. To the best of our knowledge, theoretical models that consider complete surface topography have yet to be explored.

Directly constructing surface morphology using methods such as finite element analysis represents an alternative research approach. Dai et al. [?] reconstructed contact surfaces and conducted finite element simulations with ABAQUS to investigate the thermal contact resistance of Ti-6Al-4V, subsequently incorporating thermal expansion effects and performing thermo-mechanical coupling simulations [?]. Wang et al. [?] examined contact thermal resistance under non-uniform loading conditions, while Dong et al. [?] addressed the impact of gap conduction. Although numerical simulations based on finite element analysis can accommodate complete surface morphology, they suffer from significant drawbacks, including high computational resource demands, substantial time costs, and limited reusability, which complicate the development of simple and user-friendly general models.

In recent years, neural network-based predictive models have gained increasing traction in research [?]. Ren et al. [?] developed a predictive model for contact thermal resistance utilizing artificial neural networks, enabling accurate determination of TCR values directly from temperature distributions and contact pressure at multiple points, thus eliminating the need for complex experimental testing procedures. Moreover, Feng et al. [?] employed artificial neural networks to forecast the contact thermal resistance of copper blocks subjected to load cycling. Nonetheless, current investigations into TCR based on neural networks primarily utilize simple artificial neural networks, neglecting the potential advantages of more intricate network architectures and their inherent characteristics.

Convolutional neural networks possess powerful feature extraction capabilities, making them highly suitable for surface feature recognition. Figure 1 [Figure 1: see original paper] illustrates the overall process outlined in this article. We first generated a contact thermal resistance dataset based on surface topography using a theoretical model of contact thermal resistance. Next, we developed a deep learning prediction framework for TCR, leveraging a well-established CNN architecture. By training the model on the training set and validating it on the test set, we obtained an optimized TCR prediction model. To ensure accuracy, we further validated the model using experimentally measured surface topography data and corresponding TCR values. Finally, we employed interpretability techniques to visualize the surface features extracted by the model and explain its predictive performance.

Methods

Overview of the Methodology

Figure 1(a) presents the flowchart outlining our study' s methodology. Initially, two random surfaces were generated and used alongside other parameters as inputs for a theoretical model that calculated contact thermal resistance and actual contact area. These inputs and outputs were organized into a contact thermal resistance database. To satisfy neural network input requirements, the data underwent preprocessing. The dataset comprised surface-related features (Surf1, Surf2, Std1, Std2), environmental parameters such as contact pressure (P), and prediction targets—specifically, the actual contact area (Area) and contact thermal resistance (Rc). The data were divided into training and testing sets, followed by training an appropriate deep learning prediction model through cross-validation and hyperparameter tuning. Finally, experimental data evaluated model performance, while interpretability methods clarified the model' s decision-making process.

Figure 1(b) illustrates the neural network training process, which consists of two primary stages: forward and backward propagation. In the forward stage, input features are progressively transformed into prediction results, which are then compared to target values to calculate the loss. This loss serves as the starting point for the backward stage, during which the gradient tensor is propagated backward to update the neural network' s parameters. Figure 1(c) details the principles of guided backpropagation, which modifies the ReLU layer' s input-output behavior. During forward propagation, the output feature tensor of the ReLU layer is retained. In the backward pass, this retained input is combined with the ReLU layer' s output through element-wise operations to produce the final output tensor, which serves as the input gradient tensor for the subsequent layer. Figure 1(d) explains Class Activation Mapping principles, where the output feature tensor from forward propagation and the output gradient tensor from backward propagation are used to compute the activation map for each module.

Datasets

The data in this study derive from a theoretical computational model of thermal contact resistance [?]. Initially, fractal parameters (D) and (G) were determined, and two random surfaces of size 1024$×$1024 were generated using the Weierstrass-Mandelbrot (W-M) function [?]. These fractal parameters served as inputs to the theoretical model, which calculated the actual contact area and thermal contact resistance based on material properties—including elastic modulus, Poisson' s ratio, micro-hardness, thermal conductivity—and the contact pressure between surfaces. The generated random surfaces, along with the theoretical model' s input and output parameters, formed our dataset. The training set comprises 10,000 cases, while the testing set consists of 2,000 instances. To validate the model, experimental tests were conducted on two sets of 316 stainless steel specimens with surface roughness of approximately 0.8, prepared using grinding and turning techniques. Thermal contact resistance was measured using the steady-state method [?], and surface topography was analyzed using a surface profilometer. These experimental data were also used for model validation. For further details on dataset construction, refer to the supplement.

Selection of Descriptors

Numerous factors influence thermal contact resistance, though not all are within this study' s scope. This research primarily focuses on surface topography, making topographic data of the two surfaces the most critical descriptors. As neural networks require normalized input data, Z-score normalization was applied to surface height data, transforming it to zero mean and unit variance as follows:

$$Z_i = \frac{z_i - \mu}{\sigma}, \quad \mu = \frac{1}{N}\sum_{i=1}^{N} z_i, \quad \sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(z_i - \mu)^2}$$

where N represents the total number of data points in the surface height profile, z_i denotes the height at each point, and μ and σ are the mean and standard deviation of height, respectively. In fact, σ represents the root mean square roughness of the surface. To preserve surface topography information, the standard deviations of both surfaces were incorporated as input parameters. Consequently, a comprehensive descriptor set for surface information includes Surf1, Surf2, Std1, and Std2. To simplify the model, all surfaces were assumed to be composed of the same material (316 stainless steel) and tested under identical environmental conditions; therefore, material properties, ambient temperature, and ambient pressure were excluded from consideration. Only contact pressure was included as a parameter affecting TCR.

According to contact thermal resistance theory, the actual contact area between two surfaces ultimately dictates TCR [?]. Furthermore, the actual contact area plays crucial roles in various domains including friction, lubrication [?], and coating [?], establishing it as a more fundamental physical quantity than TCR itself. Consequently, the actual contact area was designated as an additional variable of interest. Since experimental measurement of the actual contact area is challenging and relying solely on it as a target variable complicates model validation, both the actual contact area and thermal contact resistance were designated as target variables. Accurate predictions of TCR are expected to yield correspondingly reliable estimates of the actual contact area, providing additional validation. In summary, the complete descriptor set comprises Surf1, Surf2, Std1, Std2, P, Area, and Rc, as shown in Figure 2 Figure 2: see original paper with detailed definitions provided in Table 1 .

Figure 2(a) illustrates the physical significance of the descriptors. Contact between two rough surfaces is inherently imperfect, generating contact thermal resistance that is closely associated with the magnitude of the actual contact area. Consequently, these two quantities are designated as target variables, while all information associated with the two contact surfaces (Surf1, Surf2, Std1, Std2) and contact pressure serve as input variables. Figure 2(b) shows the relationships among different features: the diagonal displays each feature' s distribution, the upper triangular section presents correlation coefficients and heatmaps, and the lower triangular section depicts paired feature distributions.

The data distribution and correlation coefficients for each descriptor are illustrated in Figure 2(b). Surface height variance ranges between 0 and 3 m, indicating comparable surface roughness. Pressure values distribute between 1 and 10 MPa, while actual contact area ranges from 0 to 0.8%, and contact thermal resistance remains below 2000 $mm^2 \cdot K/W$. The actual contact area shows high correlation coefficients with Std1, Std2, and P, indicating strong positive correlation. In contrast, contact thermal resistance exhibits more pronounced positive correlation with Std1 and Std2 but shows no significant correlation with pressure, which appears inconsistent with theoretical predictions. Examination of P' s distribution relative to Std1 and Std2 reveals that P is uniformly distributed across its range, leading to similarly uniform distribution of TCR with respect to P and obscuring any observable correlation.

Deep Learning Algorithm

Our neural network model comprises three components: an adaptor, a pre-trained CNN, and a regression network, as schematically shown in Figure 3 Figure 3: see original paper. The input surface is sequentially processed by the adaptor and pre-trained network. The regression network then takes the pre-trained network' s output along with additional parameters to perform regression and derive target parameters. The pre-trained network utilizes CNN architectures such as ResNet [**?**], DenseNet [**?**], and VGG [**?**], which have demonstrated superior performance in image processing tasks. This study employs the DenseNet121 [**?**] architecture with pre-trained parameters to accelerate training. Since image inputs typically consist of three channels but only two contact surfaces are available, the first convolutional layer' s input channels were adjusted from three to two. Furthermore, as these CNN architectures are exclusively used for surface feature extraction and typically include classification modules, we replaced the classification module with an Identity layer that outputs its inputs unchanged, ensuring only extracted features propagate to subsequent modules.

For image inputs, the standard required size is 224\$$\times 224; however, measuredsurfacedatasizeis 1024 \times 1024. Toa$
$trainednetwork, asillustratedinFigure 3(b). Theadaptoremploysstraightforwardconvolutionaloperationstore$
$trainednetwork, ensuringinputsizeisamultipleof 32 issufficienttoretainmaximuminputinformation. Whiled$
surfaces as input is feasible, it would exponentially increase training time. Consequently, after thorough consideration, we adopted the current architecture.

The regression network processes the extracted surface feature tensor and input parameters. The surface feature tensor typically exceeds one thousand elements, whereas input parameters consist of only three elements. Direct concatenation would diminish the input parameters' influence on results. Therefore, we first apply a linear layer to halve the surface feature tensor length and another linear layer to expand input parameters to a comparable length. Here, we standardize their lengths to ensure consistency. The outputs from these two linear layers are concatenated and passed through two additional linear layers to regress the final target parameters: actual contact area and contact thermal resistance.

The model was trained on a dataset comprising 10,000 samples, with performance evaluated on a separate test set of 2,000 samples. During training, the Adam optimizer was employed with a learning rate of 0.001, and mean square error (MSE) served as the loss function. A 5-fold cross-validation strategy was implemented: the training set was partitioned into five equal subsets, with four used for training and the remaining subset for validation in each iteration. This process repeated five times, ensuring each subset served as validation set once. The average validation loss across all iterations was used to assess model performance.

Interpretability of CNN Model

Convolutional Neural Networks have revolutionized feature extraction tasks, offering exceptional performance in classification and object detection. However, their complex structure often obscures the decision-making process, creating a "black box" perception. Visualization techniques provide insights into this process, making CNNs more interpretable and understandable. CNN visualization elucidates the hierarchical feature detection mechanism, where lower layers detect simple features like edges while higher layers capture more complex patterns. Activation maps and feature maps are crucial tools that highlight which input regions activate specific neurons, offering glimpses into the network' s feature detection process and aiding understanding of feature importance.

This paper employs two algorithms commonly used for image task visualization—Guided Backpropagation (GBP) and Class Activation Mapping (CAM)—to understand the surface morphology features that influence thermal contact resistance.

Guided Backpropagation
Backpropagation is the standard algorithm for training neural networks, computing the gradient of the loss function with respect to network weights by propagating errors backward from output to input layers. Guided Backpropagation, an extension of standard backpropagation, was introduced to overcome limitations in generating interpretable visualizations. Its core principle modifies the backpropagation process by retaining only positive gradients, thereby emphasizing features most relevant to specific neuron or layer activation. During forward propagation, input data is processed through the network and activations are computed for each layer. In the backward pass, guided backpropagation alters gradient flow: only positive activations pass through ReLU layers during forward propagation, and only positive gradients propagate backward through these layers, with negative gradients set to zero, as shown in Figure 1(c): $grad = ReLU(grad)$. The resulting gradient with respect to input data highlights input regions that most significantly contributed to target neuron or layer activation.

Class Activation Mapping
Class Activation Maps are a pivotal technique in deep learning that elucidates which input image regions are most influential in CNN classification decisions. By leveraging weights from the fully connected layer immediately preceding the output, CAM generates a class-specific activation map mathematically formulated as the weighted sum of feature maps from the last convolutional layer, where weights correspond to class-specific weights from the final layer:

$$CAM = W \cdot F$$

Here, W denotes the weight vector for the specific class, and F represents the feature maps. In this context, F typically represents a function of the feature tensor during forward propagation, while W generally denotes a function of the gradient tensor during backward propagation. The specific functional form is determined by the particular CAM algorithm employed, as shown in Figure 1(d). Details are listed in supplementary information. This approach offers valuable insights into the model's decision-making process by visually emphasizing image regions that most significantly contribute to predicted classes, thereby enhancing CNN interpretability across diverse applications.

Results and Discussions

As illustrated in Figure 4 Figure 4: see original paper, the training process spanned 80 epochs, with the optimal model identified at the 76th epoch, achieving an average validation loss of 0.01 during cross-validation. Model predictive performance was evaluated on both training and validation datasets. Results reveal that predicted actual contact area exhibited correlation coefficients R^2 of 0.993 and 0.893 on training and test sets, respectively (Figure 4(c)), while contact thermal resistance demonstrated correlation coefficients of 0.989 and 0.978 on training and test sets, respectively (Figure 4(d)). Relative errors were also computed, as depicted in Figure 4(b). For both prediction targets, relative errors were predominantly below 25% on the training set and below 50% on the test set, with maximum relative errors not exceeding 50% and 125%, respectively, indicating robust predictive accuracy. Regression lines for actual contact area on both datasets lay below the $x = y$ line, suggesting the model tends to underestimate actual contact area. In contrast, regression lines for contact thermal resistance were positioned above the $x = y$ line, implying overestimation of TCR.

After validating model performance, we conducted further verification using experimentally measured data. Test specimens were categorized into two groups: the first fabricated through grinding (Figure 5 Figure 5: see original paper), and the second produced via turning (Figure 5(h)). Both groups exhibit surface roughness of approximately 0.8. Based on processing characteristics, a 5$×$5 grid of local areas was selected for surface topography testing in the first group, whereas three local areas were tested from the center outward in the second group. Contact thermal resistance was measured under pressures ranging from 1 to 4.55 MPa. Surface topography data and contact pressure were utilized as model inputs to predict TCR. Test areas of upper and lower specimens were paired one-to-one, and predicted data were employed to assess TCR at specific pressures. Consequently, the first group yielded 625 datasets per pressure, while the second group produced 9 datasets. Experimental data and prediction results are presented in Figure 5. The predicted actual contact area exhibits nearly linear increase with pressure, rising from approximately 0.02% to 0.25% and 0.4% as pressure increased from 1 MPa to 10 MPa. Regarding TCR, model predictions were slightly lower for pressures below 1 MPa and slightly higher for pressures above 3 MPa, though all predictions remained within acceptable error ranges. Figures 5(d)-(g) and (k)-(n) display additional prediction results for paired surfaces. Notably, prediction errors for surfaces in Figures 5(d) and (k) were relatively minor compared to experimental results, whereas errors for surfaces in Figures 5(f) and (m) were more pronounced. Analysis revealed common characteristics: Surface 2 in Figure 5(f) and Surface 1 in Figure 5(m) exhibited overall height deviations due to local defects, which likely contributed significantly to reduced prediction accuracy.

To enhance model interpretability, we employed Guided Backpropagation and Class Activation Mapping (CAM) to visualize surfaces influencing final outcomes. Figure 6 Figure 6: see original paper, (b), (g), and (h) depict GBP results for two surface sets. To optimize visualization clarity, surfaces were manually rotated by 90°, 180°, and 270°, ensuring input data integrity. Gradients derived from GBP were further processed, producing output gradients with dimensions identical to input surfaces. These gradients were segmented into $grad_1$ and $grad_2$ along the first dimension. Positive (pos) and negative (neg) influences were defined as absolute values of gradients greater than or less than zero, respectively, with subscripts 1 and 2 corresponding to their respective gradients. At 1 MPa contact pressure, rotation significantly affects both actual contact area and corresponding TCR—specifically, increased actual contact area leads to reduced TCR. However, at contact pressures of 2 MPa or higher, rotation' s impact diminishes. For the second specimen set, rotation has negligible effects on both TCR and actual contact area, likely attributable to the manufacturing process. The rotational symmetry of textures formed by turning along the axis rendered rotation inconsequential. Summing all gradient elements reveals a positive correlation between the sum of grad elements and actual contact area, and a negative correlation with TCR. In contrast, the second specimen set exhibits no discernible trend, potentially due to minimal variations in actual contact area and TCR. Examination of gradient activation maps reveals they predominantly correspond to regions of higher or lower surface contact areas (i.e., $surf1 + surf2$), aligning with TCR theory and confirming that our model accurately captures relevant features of real surfaces.

The CAM method exhibits distinct characteristics in Figure 7 [Figure 7: see original paper]. Most notably, for surfaces, CAM visualization results focus more on lower regions of the contact surface: activation maps corresponding to lower regions of $surf1 + surf2$ have higher values and exert greater influence on results. Lower regions can be interpreted as areas where contact is less likely to occur, and non-contact areas clearly affect actual contact area—the more non-contact areas exist, the smaller the actual contact area, leading to greater thermal contact resistance. Additionally, for all CAM methods, activation maps from shallow layers are more easily understood by humans, while deep layer activation maps exhibit harder-to-interpret regions showing higher preference for certain edge regions. We offer two explanations: first, features extracted by deeper neural networks are inherently more abstract and harder to understand; second, due to CAM generation principles, deeper features typically have smaller dimensions and more channels, which can result in information loss or addition of extraneous information when converting them into CAMs of the same size as the input. Interpretation of activation maps for deeper neural network modules remains an area for further research.

Conclusion

In this study, we developed an AI-driven thermal contact resistance prediction model based on neural networks. A deep learning architecture capable of processing full surface topography data was constructed using the DenseNet121 framework to predict actual contact area and contact thermal resistance. Through 5-fold cross-validation on a training set generated from a theoretical model, the optimal prediction model was identified, achieving R^2 values of 0.893 and 0.978 for predicted actual contact area and TCR on the test set, respectively, with relative errors predominantly below 50%, indicating strong predictive accuracy. The model was further validated using experimentally measured surface topography and TCR data, with results aligning closely with predictions and demonstrating that a model trained on purely theoretical data can effectively predict real-world surface behaviors. Additionally, we conducted visualization studies using GBP and CAM. GBP activation maps revealed that the model prioritized contact regions between surfaces, with activation positions shifting accordingly when surfaces were manually rotated. At high contact pressures (\$ $2MPa), TCRremainedstableunderrotation.However, atlowerpressures($ \$1 MPa), rotational response varied between ground and turned specimens. Due to rotational symmetry of turning textures, actual contact area and TCR remained largely unaffected by rotation, whereas ground specimens exhibited significant changes. Summation of activation map data showed clear positive correlation between total activation matrix elements and actual contact area, and negative correlation with TCR, further validating model rationality.

CAM activation maps highlighted the model' s focus on regions with minimal surface contact, which are harder to engage and contribute to increased thermal resistance. Furthermore, regardless of CAM algorithm, shallow layer activation maps were consistent and interpretable, while deep layer maps were more complex and challenging to decipher—complexity that may stem from the CAM algorithm itself or the inherent nature of deep neural networks to extract abstract, high-level features.

In summary, this study effectively harnesses the robust feature extraction capabilities of convolutional neural networks to develop a contact thermal resistance prediction model utilizing complete surface morphology data, simultaneously serving as a reference for actual contact area between surfaces. Compared to finite element simulations, this model offers rapid prediction speeds and low computational resource requirements following training. Additionally, visualization methods clearly illustrate surface features influencing TCR. However, the neural network model inputs pressure and surface characteristics separately, which may hinder its ability to capture pressure' s influence on surface features. Moreover, this research lacks effective explanations for CAM visualizations generated by deep neural networks. These issues warrant further investigation. We hope this study provides a novel approach to TCR research and aids researchers in understanding the significant impact of surface morphology characteristics from a different standpoint.

Data Availability

The dataset supporting this study’ s findings is available on request from the corresponding author. The data are not publicly available due to its large size. Additional data supporting this study’ s findings, including training code, checkpoints, and test cases, are available at Joe-zhouman/TcrSurfModel.

References

1. Mahajan, R., et al., Co-Packaged Photonics For High Performance Computing: Status, Challenges And Opportunities. *Journal of Lightwave Technology*, 2022. 40(2): p. 379-392.
2. Chen, W.-Y., et al., Thermoelectric coolers for on-chip thermal management: Materials, design, and optimization. *Materials Science and Engineering: R: Reports*, 2022. 151: p. 100700.
3. Lin, J., et al., A review on recent progress, challenges and perspective of battery thermal management system. *International Journal of Heat and Mass Transfer*, 2021. 167: p. 120834.
4. Zhang, Y., et al., Simultaneous electrical and thermal rectification in a monolayer lateral heterojunction. *Science*, 2022. 378(6616): p. 169-175.
5. Chen, J., et al., Interfacial thermal resistance: Past, present, and future. *Reviews of Modern Physics*, 2022. 94(2): p. 025002.
6. Sun, D., et al., A review of thermal contact conductance research of conforming contact surfaces. *International Communications in Heat and Mass Transfer*, 2024. 159: p. 108065.
7. Murwamadala, R.D. and V.R. Veeredhi, Advances in thermal contact resistance studies. *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, 2022. 237(1): p. 201-222.
8. Wang, H., et al., Fractal analysis of the thermal contact conductance for mechanical interface. *International Journal of Heat and Mass Transfer*, 2021. 169.
9. Vu, A.T., et al., Machine learning-based predictive modeling of contact heat transfer. *International Journal of Heat and Mass Transfer*, 2021. 174.
10. Siddappa, P.G. and A. Tariq, Experimental estimation of thermal contact conductance across pressed copper–copper contacts at cryogenic-temperatures. *Applied Thermal Engineering*, 2023. 219.
11. Liu, Y., et al., Effects of contact pressure and interface temperature on thermal contact resistance between 2Cr12NiMoWV/BH137 and γ-TiAl/2Cr12NiMoWV interfaces. *Thermal Science*, 2020. 24(1 Part A): p. 313-324.
12. Pan, X.S., et al., Research Progress of Thermal Contact Resistance. *Journal of Low Temperature Physics*, 2020. 201(3-4): p. 213-253.
13. Cui, T.F., Q. Li, and Y.M. Xuan, Characterization and application of engineered regular rough surfaces in thermal contact resistance. *Applied*

Thermal Engineering, 2014. 71(1): p. 400-409.
14. Ma, C., et al., A geometrical-mechanical-thermal predictive model for thermal contact conductance in vacuum environment. *Proceedings of the Institution of Mechanical Engineers Part B-Journal of Engineering Manufacture*, 2016. 230(8): p.
15. Sun, X.G., C.X. Meng, and T.T. Duan, Fractal model of thermal contact conductance of two spherical joint surfaces considering friction coefficient. *Industrial Lubrication and Tribology*, 2022. 74(1): p. 93-101.
16. Cheng, Y., et al., Fractal model of thermal contact conductance considering thermal stress and asperity interactions. *International Journal of Heat and Mass Transfer*, 2024. 230: p. 125787.
17. Dai, Y.J., et al., A test-validated prediction model of thermal contact resistance for Ti-6Al-4V alloy. *Applied Energy*, 2018. 228: p. 1601-1617.
18. Dai, Y.J., et al., Effect of thermal expansion on thermal contact resistance prediction based on the dual-iterative thermal-mechanical coupling method. *International Journal of Heat and Mass Transfer*, 2021. 173.
19. Wang, C., et al., Effect of load application method on thermal contact resistance and uniformity of temperature distribution. *Applied Thermal Engineering*, 2023. 229: p.
20. Dong, Y., et al., Numerical study of thermal contact resistance considering spots and gap conduction effects. *Tribology International*, 2024. 193: p. 109304.
21. Zhou, T., Y.J. Zhao, and Z.H. Rao, Fundamental and estimation of thermal contact resistance between polymer matrix composites: A review. *International Journal of Heat and Mass Transfer*, 2022. 189: p. 122701.
22. Ren, X.-J., et al., Exploring thermal contact conductance between two contact solids by artificial neural network. *International Communications in Heat and Mass Transfer*, 2022. 136: p. 106182.
23. Feng, Z.Y., J.T. Yan, and Y.W. Gao, Prediction of contact resistance between copper blocks under cyclic load based on deep learning algorithm. *Aip Advances*, 2022. 12(7): p. 075009.
24. Berman, M.A.a.D.H., A multivariate Weierstrass–Mandelbrot function. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 1997. 400(1819): p. 331-350.
25. Zhang, P., Y. Xuan, and Q. Li, A high-precision instrumentation of measuring thermal contact resistance using reversible heat flux. *Experimental Thermal and Fluid Science*, 2014. 54: p. 204-211.
26. Bonnevie, E.D., et al., In Situ Studies of Cartilage Microtribology: Roles of Speed and Contact Area. *Tribology Letters*, 2011. 41(1): p. 83-95.
27. Vazirinasab, E., R. Jafari, and G. Momen, Application of superhydrophobic coatings as a corrosion barrier: A review. *Surface and Coatings Technology*, 2018. 341: p. 40-
28. He, K., et al. Deep Residual Learning for Image Recognition.
29. Huang, G., et al., Densely Connected Convolutional Networks. *arXiv e-prints*, 2016: p. arXiv:1608.06993.
30. Simonyan, K. and A. Zisserman, Very Deep Convolutional Networks for

Large-Scale Image Recognition. *arXiv e-prints*, 2014: p. arXiv:1409.1556.

Acknowledgements

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors acknowledge financial support from the National Natural Science Foundation of China (Grant No. 52076048) and Guangxi Natural Science Foundation (Grant No. 2019GXNSFFA245003).

Author Contributions

Z.M. conceived and directed the project. Z.M. and H.Z.Y. prepared datasets, constructed the training architecture, and performed training. Z.M., H.Z.Y., G.P.Y., and Z.P. performed indentation, analyzed data, and discussed results. Z.P. provided intellectual guidance, managed project finances, and reviewed progress periodically. Z.M. and H.Z.Y. co-wrote the manuscript with input from all authors.

Competing Interests

The authors declare no competing financial interests.

Supplement

Surface Fractal Modeling

For given fractal dimension D and fractal roughness G, a rough surface can be generated using the Weierstrass-Mandelbrot function [**?**]:

$$Z(x,y) = L \sum_{n=n_{\min}}^{n_{\max}} \gamma^{(D-3)n} \left\{ \cos\phi_{m,n} - \cos\left[\frac{2\pi\gamma^n}{L} \sqrt{x^2+y^2} \cos\left(\arctan\frac{y}{x} - \frac{\pi m}{M} \right) + \phi_{m,n} \right] \right\}$$

where $Z(x,y)$ represents the height of the rough surface, x and y denote lateral position coordinates, L is the measured sample length, M represents the number of superimposed ridges, γ denotes the frequency density of surface height (often equal to 1.5), n is the asperity frequency index indicating asperity size, and $\phi_{m,n}$ is the random phase ranging from 0 to 2π.

Dataset Generation Details

The dataset generation process is depicted in Figure S1. Initially, a set of fractal parameters is provided to generate a random surface according to Equation S1. These parameters, together with material properties (see Table S1) and contact pressure, are subsequently input into the theoretical model [**?**]. In this model, an initial actual contact area A_a is specified. The contact state is then determined based on A_a and the critical areas for fully plastic deformation A_{cp} and fully elastic deformation A_{ce}. If $A_a < A_{cp}$, the system is in the plastic deformation stage. If $A_{cp} \leq A_a < A_{ce}$, it transitions into the elastic-plastic deformation stage. Otherwise, all three deformation types—plastic, elastic-plastic, and elastic—occur simultaneously. Total stress is calculated based on the contact state and compared with the applied contact load. If the difference exceeds a specified tolerance (here set to 1×10^{-5}), the contact area value is adjusted, prompting repetition of the aforementioned steps. This iterative process continues until both values are within acceptable error range, at which point the area is deemed the final actual contact area and recorded. Subsequently, final thermal contact resistance is calculated based on this actual contact area and contact state.

Table S2 Material properties of 316 stainless steel at room temperature.

Property	Value
Elastic Modulus (GPa)	200
Poisson' s ratio	0.3
Micro-hardness (MPa)	2000
Thermal Conductivity (W/m · K)	16

Figure S1 Flowchart for Dataset Generation.

Experimental Details

Figure S2 (a) Integrated testing instrument for thermal adaptation coefficient and thermal contact resistance. (b) Nano 3D Optical Surface Profilometer SuperView W1.

Thermal contact resistance testing was conducted using the integrated testing instrument shown in Figure S2(a), based on the steady-state method defined in ASTM D5470-06. The testing environment was set at 20°C with a vacuum level of 10^{-3} Pa. Results are presented in Table S2. Surface morphology testing was performed using the SuperView W1 Nano 3D Optical Surface Profilometer depicted in Figure S2(b).

Table S3 Tested thermal contact resistance.

Contact Pressure (MPa)	TCR of First Specimens ($mm^2 \cdot K/W$)	TCR of Second Specimens ($mm^2 \cdot K/W$)
1	1800	1900
2	1200	1300
3	900	1000
4	750	850

CAM Method

The computational formula for Class Activation Mapping (CAM) is expressed as:

$$CAM = W \cdot F \quad \text{(S2)}$$

where F represents the feature map (CNN output) and W denotes weights typically associated with the gradient tensor during backpropagation. Various methods for computing W have resulted in different CAM approaches:

Eigen CAM [?]: $CAM = F$

Eigen Grad CAM [?]: $CAM = F \cdot \text{Element-Wise}(G)$

Grad CAM [?]: $CAM = ReLU(G \cdot F)$, where $ReLU$ means $\max(M, 0)$

Grad CAM++ [?]:

$$M_i = \sum_{j,k} F_{i,j,k}, \quad b_{i,j,k} = \frac{2G_{i,j,k}}{G_{i,j,k} + 3M_{iG_{i,j,k}} + \epsilon}, \quad W_{i,m,n} = \sum_{j,k} ReLU(G_{i,j,k}) b_{i,j,k}$$

where ϵ prevents division by zero.

X Grad CAM [?]:

$$M_i = \sum_{j,k} F_{i,j,k}, \quad b_{i,j,k} = \frac{F_{i,j,k}}{M_i + \epsilon}, \quad W_{i,m,n} = \sum_{j,k} G_{i,j,k} b_{i,j,k}$$

Layer CAM [?]: $W = ReLU(G)$

Here, F, W, and G typically possess three dimensions. The notation $\sum_{j,k}$ represents summation over the latter two dimensions (width and height), which are independent of the channel dimension. All codes are available at https://github.com/Joe-zhouman/TcrSurfModel.

References (Supplement)

1. Berman, M.A.a.D.H., A multivariate Weierstrass-Mandelbrot function. *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences*, 1997. 400(1819): p. 331-350.
2. Cheng, Y., et al., Fractal model of thermal contact conductance considering thermal stress and asperity interactions. *International Journal of Heat and Mass Transfer*, 2024. 230: p. 125787.
3. Springenberg, J.T., et al., Striving for Simplicity: The All Convolutional Net. *arXiv e-prints*, 2014: p. arXiv:1412.6806.
4. Selvaraju, R.R., et al. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization. in 2017 IEEE International Conference on Computer Vision (ICCV). 2017.

5. Chattopadhay, A., et al. Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks. in 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). 2018.
6. Fu, R., et al., Axiom-based Grad-CAM: Towards Accurate Visualization and Explanation of CNNs. *arXiv e-prints*, 2020: p. arXiv:2008.02312.
7. Jiang, P.T., et al., LayerCAM: Exploring Hierarchical Class Activation Maps for Localization. *IEEE Transactions on Image Processing*, 2021. 30: p. 5875-5888.

Note: Figure translations are in progress. See original paper for figures.