ORIGINAL ARTICLE

Deep learning–based inline monitoring approach of mold coating thickness for Al-Si alloy permanent mold casting

Fangtian Deng[1] · Xingyu Rui[2] · Shuang Lu[1] · Zhang Liu[2] · Haoran Sun[2] · Wolfram Volk[1,2]

Abstract

In the permanent mold casting process, the distribution of mold coating thickness is a significant variable with respect to the coating's thermal resistance, as it strongly influences the mechanical properties of cast parts and the thermal erosion of expensive molds. However, efficient online coating thickness measurement is challenging due to the high working temperatures of the molds. To address this, we propose an indirect monitoring concept based on the analysis of the as-cast surface corresponding to the coated area. Our previous research proves linear correlations between the as-cast surface roughness parameter known as arithmetical mean height (Sa) and the coating thickness for various coating materials. Based on these correlations, we can derive the coating thickness from the analysis of the corresponding as-cast surface. In this work, we introduce a method to quickly evaluate the as-cast surface roughness by analyzing optical images with a deep-learning model. We tested six different models due to their high accuracies on ImageNet: Vision Transformer (ViT), Multi-Axis Vision Transformer (MaxViT), EfficientNetV2-S/M, MobileNetV3, Densely Connected Convolutional Networks (DenseNet), and Wide Residual Networks (Wide ResNet). The results show that the Wide ResNet50-2 model achieves the lowest mean absolute error (MAE) value of 1.060 μm and the highest R-squared (R^2) value of 0.918, and EfficientNetV2-M reaches the highest prediction accuracy of 98.39% on the test set. The absolute error of the surface roughness prediction remains well within an acceptable tolerance of ca. 2 μm for the top three models. The findings presented in this paper hold significant importance for the development of an affordable and efficient online method to evaluate mold coating thickness. In future work, we plan to enrich the sample dataset to further enhance the stability of prediction accuracy.

Keywords Permanent mold coating · Coating thickness · As-cast surface roughness · Deep learning

1 Introduction

Mold coating refers to a material applied to the surface of a foundry mold to enhance the overall quality of the resulting casting. Additionally, the coating acts as a protective barrier between the mold and the molten metal, preventing the metal from adhering to the mold. This intervention ensures a smooth, refined finish on the cast part, thereby extending the operational lifespan of valuable molds. [1, 2] Typically, the mold coating material consists of a refractory binder, known for its remarkable ability to withstand high temperatures without breaking down or degrading. This binder is

Fangtian Deng and Xingyu Rui are both corresponding authors, they contributed equally to this work.

✉ Fangtian Deng
  fangtian.deng@igcv.fraunhofer.de

✉ Xingyu Rui
  ge79jos@mytum.de

  Shuang Lu
  shuang.lu@igcv.fraunhofer.de

  Zhang Liu
  zhang.liu@tum.de

  Haoran Sun
  haoran.sun@tum.de

  Wolfram Volk
  wolfram.volk@igcv.fraunhofer.de

1 Composite and Processing Technology IGCV,
   Fraunhofer Institute for Casting, Lichtenberg Street 15,
   85748 Garching Near Munich, Germany

2 Technical University of Munich, Walther-Meißner-Straße 4,
   85748 Garching Near Munich, Germany

combined with other elements such as ceramic powders, graphite, and organic additives. [3] These components contribute to an insulating effect during the solidification of the cast part in the mold. The thermal resistance of the coating strongly depends on the coating status, specifically its thickness, with a fixed coating material and process. Linear correlations between the thermal resistance of mold coatings and the coating thickness have been established for ceramic and graphite coatings. [4] This thermal resistance can significantly impact the temperature field and cooling rate in both the cast part and mold surface. [2, 5] Overall, the status of the foundry mold coating plays a crucial role in ensuring the quality of cast parts and the operational longevity of expensive casting molds in the foundry industry.

However, up to this point, there have been no efficient inline monitoring methods for assessing the coating status on the mold surface, primarily due to the high working temperatures involved in foundry molds. The challenges and reasons behind this limitation are extensively discussed in our previous work. [2]

To tackle this issue, our focus is on developing an indirect measurement concept for coating thickness, relying on the analysis of the as-cast surface quality. Furthermore, in our previous research, we found close correlations between the as-cast surface roughness parameter, arithmetical mean height Sa, and the mold coating thickness for both coating materials, HA Eco 182 and Cillonlin AL 286. These near-line correlations were validated when the coating thicknesses exceeded 50 μm and 100 μm, respectively, as illustrated in Fig. 1 [2, 5].

The increasing roughness of the as-cast surface can be attributed to the rising surface roughness of the coated mold as the coating thickness increases. When the mold coating

is sprayed onto the mold, ceramic powders are randomly distributed in limited areas. This random distribution of powders on the coated area leads to an elevation in surface roughness within a particular range, specifically between 100 and 400 μm of coating thickness. Although the surface roughness of the coating cannot be entirely mirrored by the corresponding as-cast surface due to the surface tension of the Al alloy melt, the similar trends observed in coating and as-cast surface roughness suggest that the coating surface plays a significant role in determining the characteristics of the as-cast surface due to their close physical contact.

Additionally, in Fig. 1, HA Eco 182 exhibits a notably more consistent increase in the surface roughness of the coated mold compared to Cillonlin AL 286. This can be attributed to the narrower particle distribution of HA Eco 182, as illustrated in Fig. 2. Despite the substantial difference in the roughness of the coated mold, the slopes of increase and the absolute values of the as-cast surface roughness remain similar for these two distinct coating materials.

Building upon these linear correlations between as-cast surface roughness and coating thickness, we introduce an indirect inline monitoring concept for mold coating thickness based on the measurement of as-cast surface roughness. However, employing a laser-confocal-microscope for this measurement is often slow and sensitive to the cleanliness of the sample surface. In our case, the laser microscope VK-X100 requires approximately 5 min for an area of 3 mm × 4 mm, making it unsuitable for achieving a 100% inline monitoring rate in a casting production line.

To address this challenge, we propose the implementation of an optical sensor embedded with deep-learning algorithms to assess as-cast surface roughness with an acceptable level of accuracy. Our target speed for the prototype is less
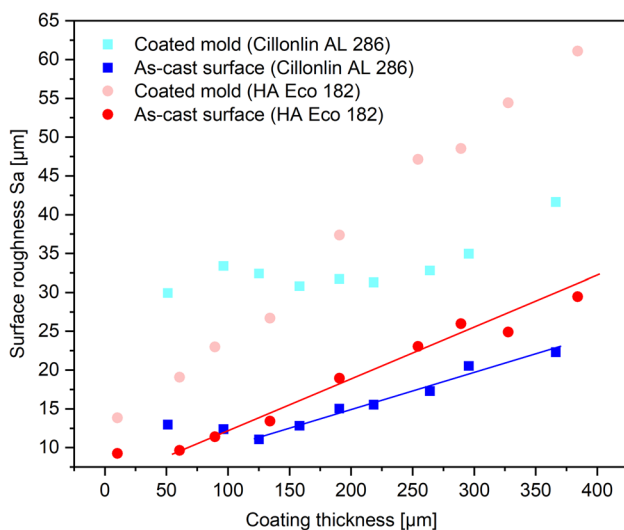

Fig. 1 Correlation between surface roughness Sa of the coated mold and as-cast surface for Cillonlin AL 286 and HA Eco 182
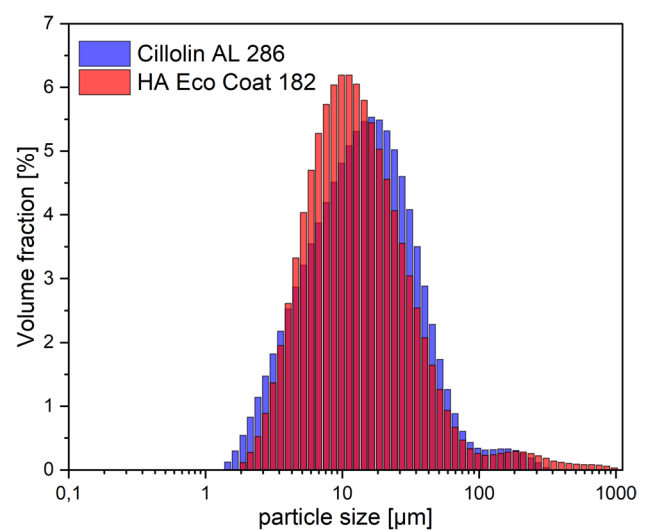

Fig. 2 Particle size distribution of Cillonlin AL 286 and HA Eco Coat 182

than 2 s for this area. This approach enables the indirect inline monitoring of mold thickness distribution without direct contact with the hot mold. Looking ahead, the standard for coating status can be defined using feedback data from cast part quality and the working time of casting molds.

2 State-of-the-art online methodologies for surface roughness evaluation

For the assessment of product surface roughness, numerous researchers have explored online methodologies, including computer vision and machine learning techniques, to classify or predict surface roughness. However, limited research has been conducted on evaluating as-cast surface quality using computer vision. A notable study of Tsai et al. introduced a machine vision system to classify the roughness of as-cast surfaces into nine classes based on Rz values (ranging from 6.3 to 400 μm) for both gray-level and binary images. Although achieving 100% classification accuracy, this system could not precisely predict roughness values [6].

In contrast, machined surface evaluation has a more extensive research history due to its significance in end-quality assessment. Earlier works by Luk et al. and Patel et al. used optical parameters and histogram analysis to estimate roughness for tool-steel samples [7, 8]. Younis et al. introduced an inline roughness measurement method based on the relationship between surface roughness (Ra) and gray level coefficients [9]. Yi et al. developed a grind surface roughness measurement using image sharpness evaluation under varying illumination [10].

As technology advanced, online roughness prediction shifted towards deep learning methods. Hu et al. constructed the "optical image-surface roughness" data set for the surface roughness prediction of the aircraft surface after coating removal. [11] A specific CNN, the "SSEResNet" (SEResNet:

Squeeze-and-Excitation Residual Networks), for regression prediction of surface roughness (Ra) was also proposed in their work. This ResNet model reaches the test MAE (mean absolute error) value of 0.245 μm. Routray et al. introduced a deep fusion network using both visual and tactile inputs for surface roughness classification. [12] Yang et al. employed an ANN with various cutting process parameters. [13]

To date, the application feasibility and the performance of the state-of-the-art deep learning methods, such as ViT, MaxViT, EfficientNet, MobileNet, DenseNet, and ResNet, have not been adequately investigated for the prediction of as-cast surface roughness parameter Sa. Here, we chose six deep-learning models with top performance on the public ImageNet dataset for recent years.

3 Experiment

3.1 Design of dataset

3.1.1 Data collection

Figure 3 illustrates an overview of our experimental process. In this study, different coating thicknesses were generated through spraying with a conventional spray can, ranging from approximately 50 μm to around 400 μm. For each coating process, two cast parts were produced. The second cast parts were selected for the analysis of surface roughness.

Before conducting the surface roughness measurement, target sampling areas were designated using a laser marking device. The geometry of these sampling areas was set as a rectangle marked with a black border, measuring 7 mm in width and 5 mm in height. Sa (the arithmetical mean height of the surface) was employed as the metric to evaluate the surface roughness of irregular surface textures. The Sa value of the marked area on the as-cast surfaces was measured
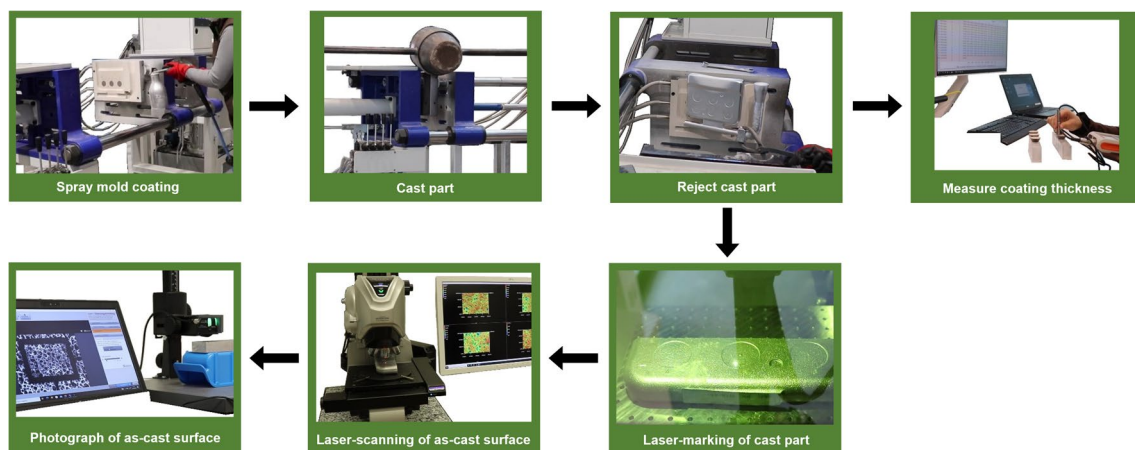

Fig. 3 Experimental construction of dataset

using a laser scanning microscope, VK-X100, produced by Keyence Corporation.

To capture the optical images of the target area, we used an industrial camera, IV2-G500MA manufactured by Keyence Corporation. The original optical image resolution is 640×480. In our work, a total of 326 optical images were collected, and the areal roughness Sa of these corresponding marked areas was measured. Among these data, 204 images are used as training set, 60 images as validation set, and 62 images as test set, respectively.

3.1.2 Data preparation

Figure 4 shows the complete process of dataset preparation for the prediction model. Firstly, the target sample area in each original image is automatically detected and cropped using the "Template Matching" method through OpenCV. Subsequently, our dataset is enriched by combining offline and online data augmentation to fully utilize the collected target images.

Secondly, in offline data augmentation, each cropped image is flipped horizontally, vertically, and horizontally and vertically via OpenCV. This operation not only preserves the image features but also mitigates the positional impact of these features. Thirdly, each image is resized to specific resolutions using bilinear interpolation mode, to ensure consistency during model training. In this case, two different resolutions, 372×270 and 224×224, are tested on all selected models. Fourthly, for online data augmentation, the ColorJitter and Gaussian Blur transforms are implemented

with Torchvision [14]. The ColorJitter transform randomly changes the brightness, contrast, saturation, and hue of images in user-defined ranges. This method can enhance the model's stability by accommodating various lighting conditions, color variations, and other factors affecting object appearance in real-world scenarios. The Gaussian Blur transform introduces blurring to an image with a Gaussian filter, enhancing the model's robustness against the variations in image clarity.

Lastly, the pixel values of the images are normalized to a standardized range. This normalization process ensures a consistent scale of input data to stabilize the training process.

3.1.3 Selected deep-learning models

In our work, six models are selected to train on our prepared training set, i.e., Vision Transformer (ViT), Multi-Axis Vision Transformer (MaxViT), EfficientNetV2-S/M, MobileNetV3, Densely Connected Convolutional Networks (DenseNet), and Wide Residual Networks (Wide ResNet). These models were chosen based on their high accuracies on ImageNet. [15]

- Vision Transformer (ViT), proposed by Dosovitskiy et al., successfully applies the transformer architecture to computer vision, demonstrating its effectiveness in image recognition tasks. [16] While achieving a top 1 accuracy of 77.91% on ImageNet-1 k when pre-trained, ViT surpasses state-of-the-art performance when pre-trained on
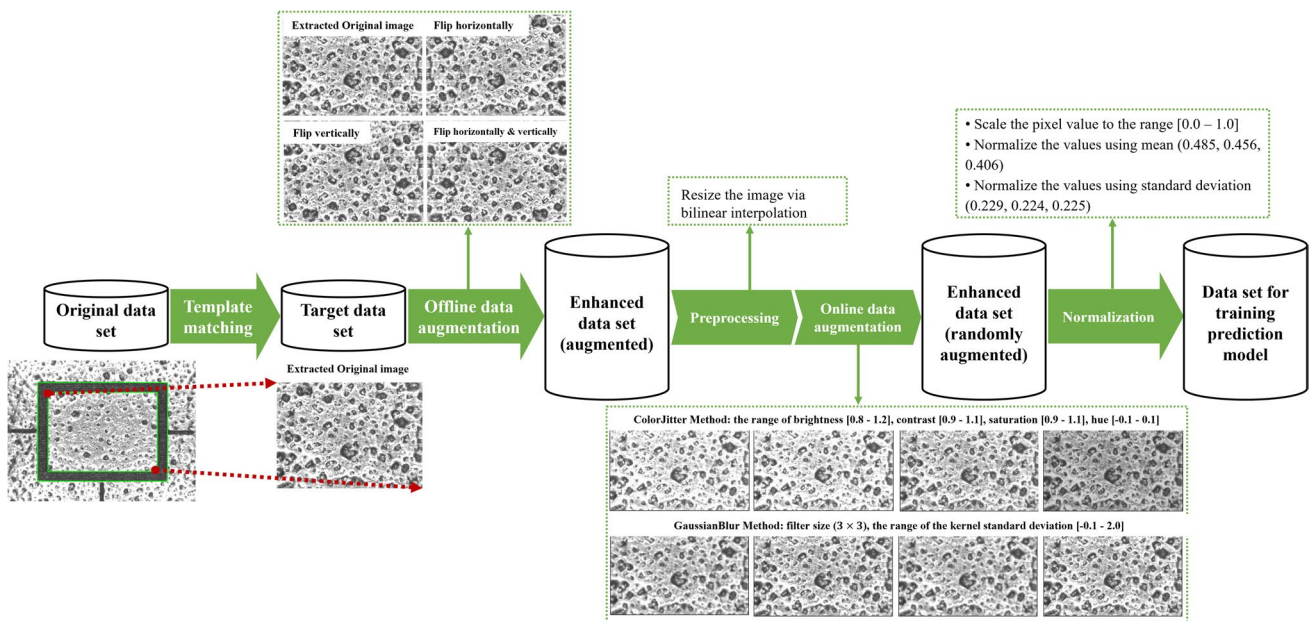

Fig. 4 Dataset preparation for prediction model

large datasets such as ImageNet-21 k or JFT-300 M, with the best model (ViT-H/14) reaching a top 1 accuracy of 88.55% on ImageNet.

- Multi-Axis Vision Transformer (MaxViT), introduced by Tu et al., presents a hybrid architecture that integrates convolutional layers with the attention mechanisms of Vision Transformers [17]. This integration enhances computational efficiency and increases accuracy compared to the "pure" Vision Transformer and Swin Transformer. MaxViT achieves a top 1 accuracy of 86.5% on ImageNet-1 k without additional pre-training data and a higher top 1 accuracy of 88.7% on ImageNet-1 k when pre-trained on ImageNet-21 k.

- EfficientNetV2, designed by Tan and Le, is one of the most recent convolutional neural network architectures that aim to achieve high accuracy with low computational complexity. [18] It introduces Compound Scaling, Fused Inverted Residual block, Neural Architecture Search (NAS), and Progressive Learning with adaptive regularization. EfficientNetV2-S and EfficientNetV2-M achieved top 1 accuracies of 83.9% and 85.1% on ImageNet. These two variants are chosen in our work based on their test performance on our datasets.

- MobileNetV3 introduced by Howard et al. is the most accurate and efficient model in the MobileNet series. [19] It is designed for running deep neural networks on mobile and embedded devices with limited computational resources. In our work, the variant MobileNetV3-Large is chosen, achieving state-of-the-art accuracy on ImageNet with a parameter count of 5.4 million and a top-1 accuracy of 75.2%.

- Wide Residual Networks (Wide ResNet), proposed by Zagoruyko and Komodakis, improves the performance by increasing the depth (adding more layers) of the neural networks based on the Deep Residual Networks (ResNet). [20, 21] In our work, the variant Wide ResNet50-2 is chosen, which closely resembles ResNet50. The difference between them is that Wide ResNet50-2 has a doubled number of channels than ResNet50 in the inner 3×3 convolution. The number of channels in outer 1×1 convolutions remains the same. Wide ResNet50-2 achieves 78.1% top 1 accuracy on ImageNet.

- Densely Connected Convolutional Networks (DenseNet), introduced by Huang and Zhuang Liu et al., innovates with its densely connected architecture. [22] In DenseNet, each layer is connected to every subsequent layer (with the same feature-map size) in a feedforward fashion. According to this paper, DenseNet requires significantly fewer parameters and computational resources to deliver comparable results, compared to the prediction accuracy of Residual Networks (ResNet). In our work, DenseNet-121 is selected.

3.2 Application of the pre-trained deep-learning models

Deep learning heavily relies on extensive training data, as it requires a substantial volume of data to capture the underlying patterns presented in the data. [23] However, the demand of a large amount of data can pose challenges in various fields due to the high cost or extended time required for data acquisition and labeling. Transfer learning in deep learning serves as a solution to this problem.

Through the pretraining on large-scale data sets, the pre-trained deep models can capture rich hierarchical and generic features. Utilizing these learned features as the starting point for training, transfer learning enables effective feature extraction for new tasks. This becomes particularly valuable when the amount of available training data is limited. This approach is instrumental in making deep learning techniques more accessible and applicable in scenarios where acquiring large amounts of labeled data is challenging or impractical.

In this study, transfer learning of advanced deep learning models is conducted utilizing the torchvision.models library. As shown in Fig. 5, the pre-trained parameters are obtained by training the model on the ImageNet1k dataset, following with the image transformation methods defined in the training recipe by Vryniotis. [24] ImageNet1k is a public dataset constructed by Deng et al., comprising 1,281,167 training images, 50,000 validation images, and 100,000 test images. [15]

The classification models pre-trained on ImageNet1k all consist of two blocks. The first block can be considered a feature extractor, generating a "feature vector" after its last operation. The second block functions as a classifier, taking the "feature vector" as input and computing the final output of the network. This transfer learning approach leverages the knowledge gained from the extensive ImageNet1k dataset to initialize the models for the specific task at hand, providing a valuable starting point for further training on the target dataset.

4 Experiment and results

4.1 Experimental environment and parameter settings

In this work, various Python libraries are utilized, including numpy, sys, sklearn, time, matplotlib, os, PIL, random, PyTorch, and Torchvision. The software environment details are outlined in Table 1. It is widely acknowledged that the hyperparameter settings and training strategies significantly
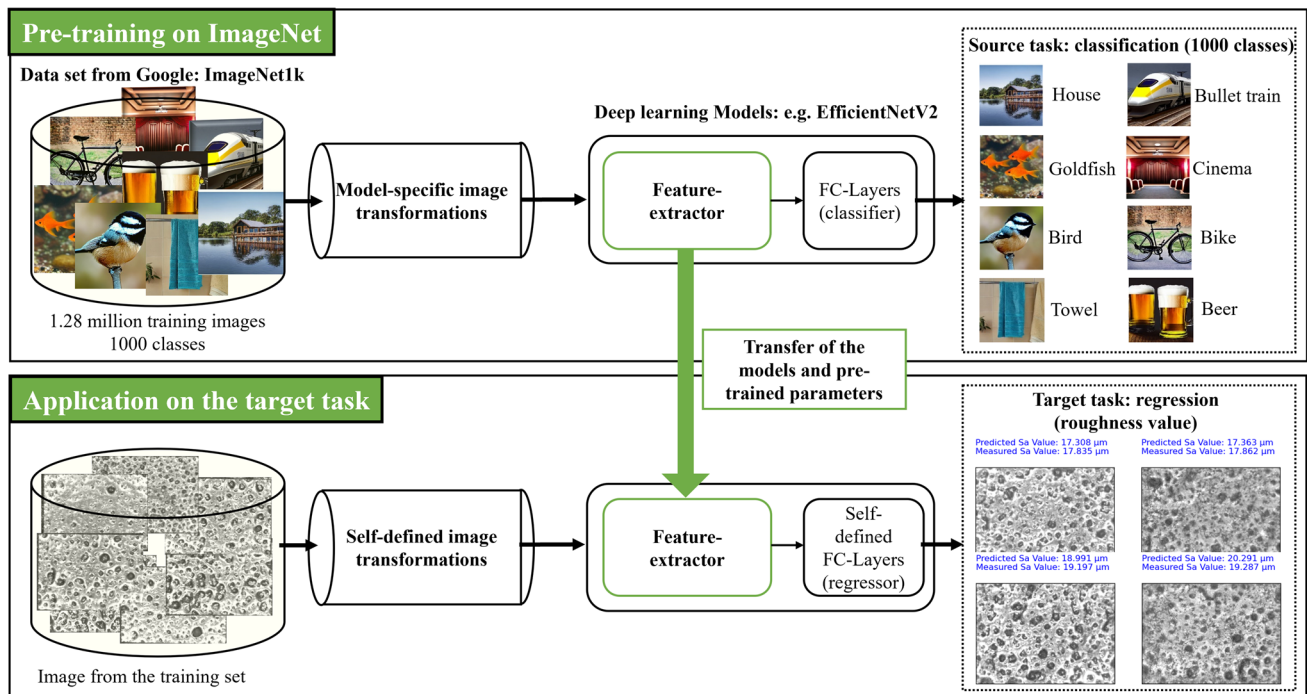
Fig. 5 Transfer learning of the pre-trained deep-learning models

Table 1 Hyperparameter settings

Batch size	8
Number of training epoch	70
Initial learning rate	0.001
Optimizer	Adam (adaptive moment estimation)
β_1 in Adam	0.9
β_2 in Adam	0.99
ε in Adam	10^{-8}
Learning rate decay method	stepLR
γ Decay factor of the leaning rate	0.5
Step size of learning rate decay	10

Table 2 Software environment

Operating system	Windows 10 (64-Bit)
Graphics card	NVIDIA RTX A6000
PyTorch version	1.13.0
CUDA version	11.6

impact the training of deep learning models. The hyperparameter settings in this work are determined based on conventions and several experiments, and they are listed in Table 2.

4.2 Evaluation metrics of the prediction results

In comparing the performances of the three models, five evaluation metrics are selected. These metrics include the mean square error (MSE) loss function, root mean square error (RMSE), mean absolute error (MAE), coefficient of determination (R^2), and the ratio of samples within 20% relative error tolerance (PR20RET). These metrics are listed in Table 3.

The mean square error (MSE) loss function is specifically well-suited for regression tasks and is employed as the evaluation metric during the training process. The other metrics are utilized for assessing the prediction results.

4.3 Experimental data of five selected models

The experimental data in Table 4 present the values of the five aforementioned evaluation metrics for the selected models using input images with two resolutions. For the models ViT-B/16 and MaxViT-T, the input images need to be resized to the resolution of 224×224 to utilize the pre-trained position embeddings of the ViT models available in the Torchvision.models package. Other models exhibit different impacts of input image resolution.

Notably, for EfficientNetV2-M and DenseNet-121, the smaller resolution of 224×224 has a positive influence on the experimental results. After comparison, the top

Table 3 Evaluation metrics of model performance

Evaluation metrics	Mathematical expression
Mean square error loss function (MSE)	$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2$
Root mean square error (RMSE)	$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}$
Mean absolute error (MAE)	$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \widehat{y_i}\right|$
Coefficient of determination (R^2)	$R2 = 1 - \dfrac{\sum_{i=1}^{n}(y_i - \widehat{y_i})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}$
Ratio of the predicted samples within 20% relative error tolerance (PR20RET)	$PR20RET = \frac{m}{n} \times 100\%$

n the number of the samples

y_i the ground truth value of the samples, here is Sa value

$\widehat{y_i}$ the predicted value of the i-th sample

$\overline{y}$ the average value of y_i

m the number of the samples with the prediction results within 20% relative error tolerance

Table 4 Experimental results

Model	Resolution	MSE /μm	RMSE /μm	MAE /μm	R2	PR20RET
ViT-B/16	224 × 224	2.551	1.597	1.315	0.891	93.55% (58/62)
MaxViT-T	224 × 224	2.221	1.490	1.168	0.906	88.71% (55/62)
EfficientNetV2-M	**224 × 224**	**2.089**	**1.445**	**1.153**	**0.911**	**98.39% (61/62)**
	372 × 270	2.501	1.581	1.305	0.894	90.32% (56/62)
EfficientNetV2-S	224 × 224	2.629	1.621	1.270	0.888	90.32% (56/62)
	372 × 270	**2.027**	**1.424**	**1.122**	**0.914**	**95.16% (59/62)**
MobileNetV3	224 × 224	2.965	1.722	1.282	0.874	90.32% (56/62)
	372 × 270	2.456	1.567	1.214	0.896	91.94% (57/62)
Wide ResNet50-2	224 × 224	2.387	1.545	1.210	0.898	93.55% (58/62)
	372 × 270	**1.923**	**1.387**	**1.060**	**0.918**	**91.94% (57/62)**
DenseNet-121	224 × 224	2.097	1.448	1.182	0.911	91.94% (57/62)
	372 × 270	2.839	1.685	1.278	0.879	88.71% (55/62)

three-performing models are Wide ResNet50-2 (trained with image resolution 372×270), EfficientNetV2-S (trained with image resolution 372×270), and EfficientNetV2-M (trained with image resolution 224×224), which are highlighted as bold text in Table 4. Among them, Wide ResNet50-2 achieves the highest R2 value of 0.918 and the lowest MAE value of 1.060 μm. Additionally, the PR20RET value of EfficientNetV2-M exceeds 98%.

4.4 Visualization of the prediction results of the top three best-performance models

The MSE Loss-epoch progressions depicted in Fig. 6 illustrate that the training losses of these three models fall below 5 μm^2 by the 10th epoch, indicating a relatively fast convergence speed. Simultaneously, their corresponding validation losses tend to stabilize at around 5 μm2 with slight fluctuations until the end of the training process.

This pattern suggests that the models converge quickly during the initial epochs of training, reaching a satisfactory level of performance. The stabilization of validation losses indicates that the models are not obviously overfitting and are generalizing well to the validation data. The observed behavior is indicative of effective training and a balanced model performance.

Figure 7 illustrates that the absolute error distributions of the predicted Sa values are homogeneous. Particularly noteworthy is that the error value remains less than 2 μm when the Sa value is above 23 μm. This indicates that the absolute error value does not increase as the surface roughness Sa increases, ensuring a satisfactory evaluation result of the rough as-cast surface when the corresponding mold coating becomes thick.
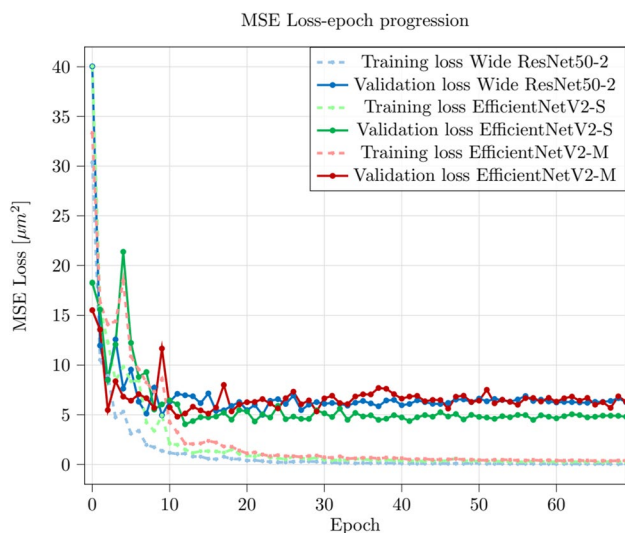
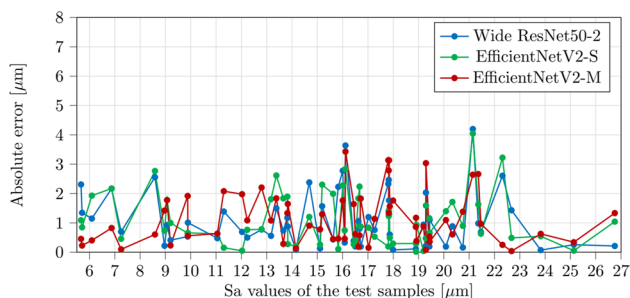**Fig. 6** The training and validation loss of the 3 top-performance models

Fig. 7 Absolute error of predicted and measured surface roughness values of top 3 models

The observed pattern in the absolute error distribution suggests that the model maintains a consistent level of accuracy across different ranges of surface roughness Sa. This is crucial for reliable evaluation, especially in scenarios where the mold coating thickness varies, and the resulting surface roughness spans a broad range.

5 Conclusion

In this paper, the prediction of as-cast surface roughness Sa based on optical images using deep-learning methods is investigated. The dataset construction involves capturing grayscale images with a resolution of 640×480 pixels using an affordable industrial camera. Subsequently, a laser confocal microscope is employed to evaluate the surface roughness parameter Sa, which serves as the labeling for these optical images. Online and offline data augmentation methods are applied during data preparation to enhance the

dataset. Based on the experimental results and comparative analysis, the conclusions of the end-to-end as-cast surface roughness prediction can be summarized as follows:

- EfficientNetV2-M achieves the highest ratio of predicted samples within 20% relative error tolerance (PR20RET) of 98.39% on the test set.
- Wide ResNet50-2 and EfficientNetV2-S achieve mean absolute error (MAE) values of 1.060 μm and 1.122 μm and R^2 values of 0.918 and 0.914, respectively, on the test set.
- For these three models, the absolute error value of the predicted Sa values maintains good tolerance across the entire range without exhibiting an increasing tendency.

Among the top three models, Wide ResNet50-2 stands out as it achieves a better prediction ratio for a larger portion of the test set due to the highest MAE value and R^2 value. Moreover, these three models demonstrate the capability to provide satisfactory prediction results of surface roughness directly from optical images after training on a limited dataset.

In a broader context, the linear correlations between mold coating thickness and as-cast areal surface roughness Sa, as proven in previous work, can be leveraged. This allows for the efficient and cost-effective derivation of corresponding mold coating thickness based on the prediction results of as-cast surface roughness.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Davis JR (1995) Selection of material for permanent molds, ASM specialty handbook: tool materials. (ASM International)

2. Deng F, Attaluri M, Klan S, Volk W (2022) An indirect evaluation method of mold coating thickness in AlSi alloy permanent mold casting production. Inter Metalcast 2072–2084. https://doi.org/10.1007/s40962-022-00920-8

3. Nwaogu UC, Tiedje NS (2011) Foundry coating technology: a review. Mater Sci Appl 2(8):1143–1160. https://doi.org/10.4236/msa.2011.28155

4. Hamasaiid A, Dargusch MS, Davidson C, Tovar S, Loulou T, Rezaï-Aria F, Dour G, (2007) Effect of mold coating materials and thickness on heat transfer in permanent mold casting of aluminium alloys. Metall Mater Trans A 38(6):1303–1316. https://doi.org/10.1007/s11661-007-9145-2

5. Deng F, Attaluri M, Klan S, Volk W (2023) Study on the influence of mold coating thickness on the thermal analysis, As-cast surface roughness, and microstructure in aluminum alloy permanent mold casting. 2nd Inceight Casting C8 Conference, Fraunhofer LBF, Darmstadt. 6–8

6. Tsai DM, Tseng CF (1999) Surface roughness classification for castings. Pattern Recog 32(3):389–405. https://doi.org/10.1016/S0031-3203(98)00077-6

7. Luk F, Huynh V, North W (1989) Measurement of surface roughness by a machine vision system. J Phys E: Sci Inst 22(12):977. https://doi.org/10.1088/0022-3735/22/12/001

8. Patel D, Mysore K, Thakkar K (2019) Noncontact surface roughness assessment using machine vision system. In: G. S. V. L. Narasimham, A. Veeresh Babu, S. Sreenatha Reddy & R. Dhanasekaran, Hrsg. Recent Trends in Mechanical Engineering - Select Proceedings of ICIME 2019. Singapore: Springer Nature Singapore Pte Ltd. pp 567–577. https://doi.org/10.1007/978-981-15-1124-0

9. Younis M (1998) Online surface roughness measurements using image processing towards an adaptive control. Comput Ind Eng 35(1–2):49–52. https://doi.org/10.1016/S0360-8352(98)00017-5

10. Yi H, Liu J, Lu E, Ao P (2016) Measuring grinding surface roughness based on the sharpness evaluation of colour images. Meas Sci Technol 27(2):025404. https://doi.org/10.1088/0957-0233/27/2/025404

11. Hu Q, Xu H, Chang Y (2022) Surface roughness prediction of aircraft after coating removal based on optical image and deep learning. Sci Rep 12:19407. https://doi.org/10.1038/s41598-022-24125-5

12. Routray PK, Kanade AS, Bhanushali J, Muniyandi M (2022) VisTa-Net: attention guided deep fusion for surface roughness classification. Preprint at https://arxiv.org/abs/2209.08516. Accessed 01 Jan 2023

13. Yang S, Natarajan U, Sekar M, Palani S (2010) Prediction of surface roughness in turning operations by computer vision using neural network trained by differential evolution algorithm. Int J Adv Manuf Technol 51:965–971

14. Contributors, T. (2017) Torchvision. URL: https://pytorch.org/vision/stable/index.html (visited on 10/30/2023). Accessed 01 Jan 2023

15. Deng J, Dong W, Socher R, Li LJ, Li K, Li FF (2009) ImageNet: a large-scale hierarchical image database. Paper presented at 2009 conference on ComputermVision and Pattern Recognition. https://doi.org/10.1007/s00170-022-10470-2/. Accessed 01 Jan 2023

16. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J, Houlsby N (2020) An image is worth 16x16 words: transformers for image recognition at scale. Preprint at https://arxiv.org/abs/2010.11929

17. Tu Z, Talebi H, Zhang H, Yang F, Milanfar P, Bovik A, Li Y (2022) MaxViT: Multi-axis Vision Transformer. In: Avidan S, Brostow G, Cissé M, Farinella GM, Hassner T (eds) Computer Vision – ECCV 2022. ECCV. Lecture Notes in Computer Science 13684. Springer Cham. https://doi.org/10.48550/arXiv.2204.01697

18. Tan M, Le Q (2021) Efficientnet: Rethinking model scaling for convolutional neural networks. In: International conference on machine learning. PMLR, 6105–6114. Efficientnetv2: Smaller models and faster training. In: International conference on machine learning. PMLR, 10096–10106. https://doi.org/10.48550/arXiv.1905.11946

19. Howard A, Sandler M, Chu G, Chen L Chen B, Tan M, Wang W, Zhu Y, Pang R, Vasudevan V, Le Q, Adam H (2019) Searching for mobilenetv3. In: Proceedings of the IEEE/CVF international conference on computer vision (ICCV) 1314–1324. https://doi.org/10.48550/arXiv.1905.02244

20. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. Paper presented at Proceedings of the IEEE conference on computer vision and pattern recognition 2016. https://doi.org/10.48550/arXiv.1512.03385

21. Zagoeuyko S, Komodakis N (2016) Wide residual networks. preprint at https://arxiv.org/abs/1605.07146. Accessed 01 Jan 2023

22. Huang G, Liu Z, Van der Maaten L, Weinberger KQ (n.d) Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 4700–4708. Accessed 01 Jan 2023

23. Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C (2018) A survey on deep transfer learning. Paper presented at 27th International Conference on Artificial Neural Networks, Rhodes, Greece, 2018 4–7. https://doi.org/10.48550/arXiv.1808.01974

24. Vryniotis V (2021) How to train state-of-the-art models using TorchVision's latest primitives. URL: https://pytorch.org/blog/how-to-train-state-of-theart-models-using-torchvision-latest-primitives/ (visited on 09/30/2023). Accessed 01 Jan 2023