

Article

Measurement of Metal Surface Temperature Based on Visible Light Images: A Strategy for On-Site Image Acquisition

Xingwang Li ^{1,2}, Wenhua Wu ¹, Chengxiang Lei ² , Yang Chen ², Zheng Tian ¹ and Qizheng Ye ^{1,*} 

¹ School of Electrical and Electronic Engineering, Huazhong University of Science and Technology, Wuhan 430074, China; lxw0214@foxmail.com (X.L.); m202372356@hust.edu.cn (W.W.); c18065893009@gmail.com (Z.T.)

² Electric Power Research Institute, Guangdong Power Grid Co., Ltd., Guangzhou 510062, China; leicx1917@163.com (C.L.); ccycc118@163.com (Y.C.)

* Correspondence: yqzq@hust.edu.cn

Abstract

Based on the mechanism of thermally modulated reflected light, visible light images combined with machine learning methods can be used to estimate the surface temperature of metal equipment at ambient temperature under sunlight conditions. However, the surface conditions of on-site equipment and camera imaging parameters vary greatly across different scenarios, leading to poor generalization of models trained solely on laboratory image databases. To address this, it is necessary to update the original laboratory database by incorporating on-site images and retrain the model accordingly; on the other hand, since most of the on-site equipment is working normally, there are few images capturing fault-induced high temperatures. Even if the method of updating and retraining on-site images is used, the data imbalance in the image database can still cause significant measurement errors in these high-temperature images. This study studies image database update schemes to address both multi-scenario and data imbalance problems and demonstrates that retraining with as little as 5% scenario-specific images or 1% high-temperature images significantly improves temperature prediction accuracy, which was validated through on-site experiments at a substation. By comparing four machine learning algorithms (random forest regression, gradient boosted regression trees, decision trees, and k-nearest neighbors), this study reveals that RFR yields the best performance. These findings enhance the practical applicability of visible light image-based temperature measurement models in engineering contexts.

Keywords: temperature measurement; visible light image; field measurement; machine learning



Academic Editor: Atsushi Mase

Received: 29 January 2026

Revised: 3 March 2026

Accepted: 4 March 2026

Published: 6 March 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and

conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Temperature measurement plays a very important role in various aspects such as industrial production, scientific experiments, and detecting equipment failures. In the power system, some minor faults on electrical equipment may threaten the stable operation of the system, develop to a serious stage, and even lead to system paralysis. These minor faults are usually accompanied by temperature changes, so temperature detection can timely detect abnormalities, eliminate hidden dangers, and ensure stable and reliable power supply.

The temperature measurement methods are generally based on the mechanisms of thermal radiation and light reflection.

Thermal radiation temperature measurement is based on the Planck blackbody radiation principle, and the radiation energy is strong in the infrared band, so infrared temperature measurement is the most widely used [1–3]. Due to the fact that the surface of the equipment is not an ideal blackbody and contains some reflections of incident sunlight, the blackbody radiation law must be scaled by a material-related factor called emissivity. The emissivity depends on the surface properties, geometric structure, wavelength, and temperature of the object. It is necessary to know the emissivity of each surface in order to obtain accurate thermal distribution. Therefore, how to adjust this parameter during on-site testing is an important task. Another issue with infrared temperature measurement technology is its relatively low spatial resolution: the expensive infrared cameras typically offer 0.3 megapixels, while the cheap visible light cameras typically exceed 30 megapixels. Within the same field of view, the spatial resolution of images captured by visible light cameras is much higher. Although visible light cameras have high pixels, the thermal radiation of objects at ambient temperature in the visible light band is much weaker than in the infrared band, and the temperature rise in equipment failures is generally within the ambient temperature range. Therefore, it is difficult to identify or measure the temperature rise in faults through visible light cameras. In addition, the infrared temperature measurement technology is also combined with machine learning methods [4], such as increasing temperature measurement accuracy in train braking components where thermal images are subject to severe temperature variations and the target component position varies continuously [5], and using Mask R-convolutional neural network for infrared insulator images for temperature analysis [6].

Light reflection thermometry, commonly referred to as thermo-reflectance microscopy, utilizes the temperature information of the reflected object contained in the reflected light signal to measure temperature. Since it does not rely on the heat carried by the incident light, it is more suitable to call it thermal modulation reflection thermometry method. In addition, the reflectivity of objects also varies significantly in the visible light band, so this technology can be used in the visible light band. The relative change in reflectance R relates to temperature via [7,8]:

$$\frac{\Delta R}{R} = \left(\frac{1}{R} \frac{\partial R}{\partial T} \right) \Delta T = \kappa(\lambda) \Delta T \quad (1)$$

Here, λ is the incident light wavelength, and κ is a material-specific calibration coefficient influenced by wavelength, incidence angle, surface roughness, and composition. An increase in temperature can lead to an increase (if $\kappa > 0$) or decrease (if $\kappa < 0$) in the reflectivity of the material. Originally used to study semiconductor band structures [9], this technique employs laser scanning or multi-point detection [10–12] and is mainly applied to microelectronics (e.g., semiconductors, magnetic films) [13,14]. The biggest advantage of the thermally modulated reflected light method is that it can utilize visible light, but in the past, it required a specific light source with a single frequency for illumination [15–18], which limits its application under sunlight, as shown in Figure 1a. In addition, there is relatively little research on the temperature measurement of metal materials by light reflection, but the reflectivity of metal surfaces is determined by temperature and has a negatively correlation [19,20]. For most metals and semiconductors, the reflection coefficient is approximately in the range of 10^{-5} – 10^{-2} K^{-1} . Therefore, in practical applications, high-sensitivity measurements are required to obtain the temperature of the equipment under test.

To apply the thermal modulation reflected light method to measure the surface temperature rise in metal equipment, the first step is to solve the problem of light source. Thermo-reflectance microscopy measurement uses a fixed light source, and it is impossible to use a microscope for live detection of power equipment. Fixed light sources are also

easily affected by environmental light sources when illuminated. Therefore, whether it is possible to directly use sunlight and environmental light sources without a fixed light source needs to be studied. The second issue arising from this is that the thermo-reflectance microscopy uses a normal incidence light source, and the relationship between temperature and reflectivity is single, while sunlight strikes surfaces at varying angles, as shown in Figure 1b, invalidating the fixed-angle assumptions used in classical thermo-reflectance models, so there is no single temperature-reflectivity relationship. It is necessary to combine machine learning technology to establish a data-driven model to complete the task of temperature measurement or recognition.

There are numerous precedents for the application of machine learning methods in the field of temperature measurement. For example, artificial neural networks have been applied to predict rectal temperature of dairy cows [21] and temperature measurement of high-voltage electrical equipment [22], but there is relatively little research on their combination with visible light images. We have previously used a combination of machine learning and visible light imaging methods to detect the surface temperature of metals at ambient temperature under sunlight, as shown in Figure 1b, which takes the red, green, and blue gray level histograms (RGB-GLH) of the visible image as the normalized comprehensive state quantity, and uses the regression prediction machine learning method to establish a temperature prediction model. The research also compared the accuracy of different algorithms and determined the optimal features. The temperature measurement error in the laboratory was below 1.0 °C [23,24]. Some researchers have utilized visible light imaging and deep learning techniques to measure common temperature (30–150 °C) [25] and flame high temperature (1400 °C) [26]. Some researchers have also used visible light images for temperature detection of materials with specific shapes [27], or studied algorithms [28,29]. However, the surface conditions of on-site equipment and camera imaging parameters vary greatly, and there is a problem that the model established based on the laboratory image database is poor generalization. It is necessary to add some on-site images to the original laboratory image database for updating and then retrain the data-driven model. On the other hand, most of the on-site equipment is working normally, and there are few high-temperature (such as 50–100 °C) images caused by faults. Even if the method of updating and retraining on-site images is used, the data imbalance in the image database can still cause significant recognition errors in the high-temperature images.

At present, in response to the problem of significant differences in feature distribution between different libraries but their application in similar fields, there are currently four methods, one of which is domain adaptation, often using transfer learning adaptive methods to solve the issue [30]. However, transfer learning is often used for classification problems, while temperature prediction belongs to regression problems. Unlike classification tasks that only require aligning decision boundaries, regression problems require precise matching of conditional distributions in an infinite dimensional continuous space. The fundamental challenge faced in regression tasks stems from the essential characteristics of the continuous output space, which has extremely high complexity in cross-domain scenarios. Therefore, targeted design and validation are also needed when using transfer learning for temperature prediction problems [31]. The second method is to fill in missing data through physical simulation, which is relatively rare in practice due to the difficulty of multi-physics field simulation. The third method is to improve the robustness of the model, such as by using multimodal fusion when there are various interference factors in measuring body temperature on site, considering the factors of distance to IR sensors, subject's facial skin color, and forehead pose, as well as ambient temperature and light intensity [32]. The fourth method is online adaptation, or online update, which is a similar approach used in this article.

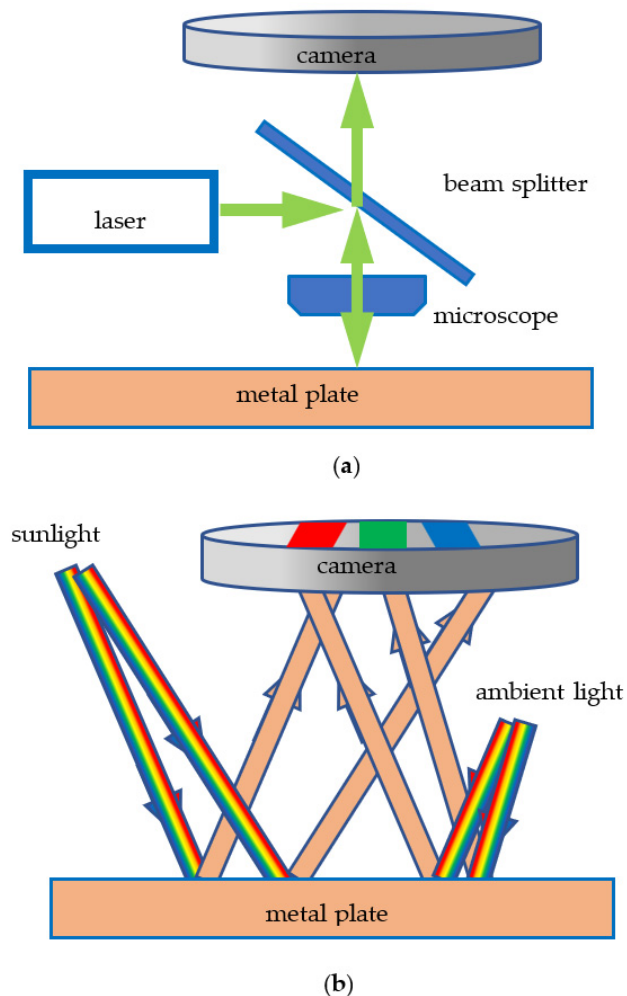


Figure 1. Schematic diagram of metal plate temperature measurement based on the thermally modulated reflected light. (a) Traditional thermo-reflectance microscopy measurement. (b) Temperature measurement under sunlight proposed by us in reference [23].

There are currently some methods to address the issue of imbalanced or missing data filling. One is from the perspective of data sources, such as data filling or data extrapolation [33]; one is algorithm improvement. Scholars have proposed solutions such as resampling techniques and improved algorithms [34,35], but there is a lack of research on the impact of different levels of data imbalance on model performance. Another approach is data evaluation, which quantifies uncertainty and conducts physical validation to avoid overconfidence of the model in missing areas. This leads to the proposal of a database establishment plan and recommendations for on-site measurements, which is similar to the approach used in this article.

Compared to the previous papers [22,23], the focus of this article is to propose a strategy for on-site image acquisition, which directly adds different scene image databases to the original image database in a certain proportion or fills in a certain proportion of scarce samples and then retrain. By comparing four machine learning algorithms and evaluating the measurement effect, the suggestions are made for on-site measurement. The second part of the article introduces the camera, detailed parameters, database construction and machine learning algorithms. The third part introduces the updated training method and calculation result analysis for the multi-scenario problems. The fourth part introduces the updated training method and calculation result analysis for the imbalanced image data

problems. The fifth part introduces the predicted temperature results using this database construction strategy on site. Finally, the conclusions of this article are summarized.

2. Image Database Construction and Machine Learning Algorithms

This section details the experimental apparatus, database construction workflow, image features, and machine learning algorithms.

2.1. Image Database Construction

The experimental setup and image acquisition method are the same as those used in the previous paper [23]. The daylight intensity during this time period was maintained at around 1000 lux. The exposure time was 1/50 s. The temperature measuring device adopts a NR-81531AK (Taiwan, China) thermocouple, with a temperature measurement range of $-50\text{ }^{\circ}\text{C}$ to $800\text{ }^{\circ}\text{C}$ and an accuracy of $(\pm 0.05\%)$. In order to utilize the existing video surveillance system in the transformer substation and improve the efficiency of image acquisition, this article replaces the Nikon D800 (Sendai, Japan) digital camera used in the original paper with a video equipment Hikvision DS-2CD3T86WDV2-I5S (Hangzhou, China) to capture images and collects a multi-scenario image database for this study.

The electrical equipment that needs to monitor temperature on site is metal, and this article chooses iron. There are multiple states on the surface of the equipment, and this article temporarily only selects two different states of corrosion: clean and rusted. Due to the fact that the surface of the on-site equipment may not always face the camera, or the position of the camera may not necessarily face the surface of the equipment to be monitored, the lens is set to face the equipment and offset at a certain angle. In addition, due to the different distances between the equipment and the camera, the pixel size of the images we actually collect may also vary. Therefore, we have set two types of pixel images. Four image databases were established in the laboratory, with temperature ranges from $30\text{ }^{\circ}\text{C}$ to $100\text{ }^{\circ}\text{C}$, for the clean and smooth iron plates as well as the rusted iron plates, as shown in Figure 2 and Table 1, namely, two iron image databases with pixel sizes of 65×80 for front-view (with the lens facing the iron plate) and left-angle view (with the lens slightly left offset), and two image databases with pixel sizes of 120×120 for front-view of iron and rusted iron. Each image database contains 4800 images. The above four image databases simulate multi-scenario and imbalanced data problems in substation sites by using different combinations and selecting some of their contents.

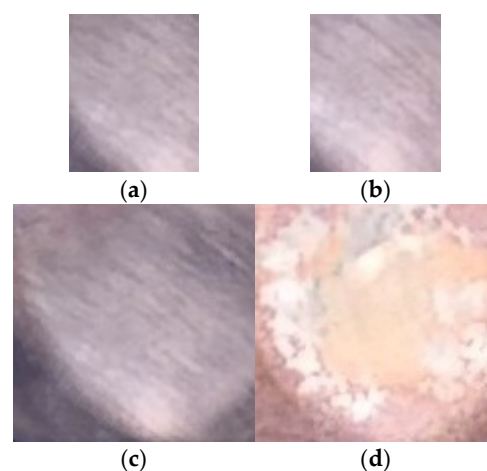


Figure 2. Sample images from distinct image databases. (a), Clean iron, front-view (65×80 pixels); (b), clean iron, left-angle view (65×80 pixels); (c), clean iron, front-view (120×120 pixels); (d), rusted iron, front-view (120×120 pixels).

Table 1. Image databases for different scenarios.

Database Number	Image Database Type (30–100 °C)
<i>a</i>	Clean iron (front-view) 65 × 80 pixels, 4800 images
<i>b</i>	Clean iron (left-angle view) 65 × 80 pixels, 4800 images
<i>c</i>	Clean iron (front-view) 120 × 120 pixels, 4800 images
<i>d</i>	Rusted iron (front-view) 120 × 120 pixels, 4800 images

2.2. Image Features and Machine Learning Algorithms

Each pixel's color is defined by RGB grayscale levels (0–255), where higher grayscale values correspond to greater brightness. The ratio of the number of pixels at a certain gray level to the total number of pixels in the image is the probability of that gray level. For all grayscale levels, a grayscale level probability density function (GLPDF) is formed. For the R, G, and B channels of a color image, a three primary color grayscale probability density function (RGB-GLPDF) is formed:

$$f_L(i) = \frac{Num_L(i)}{\sum_{i=0}^{255} Num_L(i)} \quad (i = 0, 1, \dots, 255; L = r, g, b) \quad (2)$$

where $Num_L(i)$ is the number of pixels at grayscale level i in channel L .

Temperature variations significantly alter grayscale probability distributions. For each channel, nine features were extracted to characterize these changes: mean, standard deviation, kurtosis, skewness, median, peak intensity, and upper/lower α -quantiles ($\alpha = 0.05$). Combining features from all three channels (R, G, B) yields a 27-dimensional color moment-based statistical feature vector [18,19].

The machine learning algorithms used in this article are k-nearest neighbor (KNN) based on Euclidean distance, decision tree (DT), gradient boosting regression tree (GBRT) based on boosting algorithm, and random forest regression (RFR) based on Bagging algorithm.

To ensure the reliability and repeatability of model performance, a random search method was used to optimize the hyperparameters of four machine learning algorithms. Targeting clean iron (front view (65 × 80 pixels) as the object, using its minimum MAE as the criterion, the optimal parameter combination is randomly selected by searching in the preset hyperparameter space. Set the search count to 200. To avoid the influence of randomness on the results, a fixed random seed (random_state = 42) is used for each search to ensure reproducibility of the experiment.

The hyperparameter search range of each algorithm is determined based on algorithm characteristics and previous experimental experience, as shown in Table 2. The final selected hyperparameters (i.e., the values used in all experiments in this article) were obtained from the optimal combination obtained through random search and are listed in the table. For example, the number of trees in RFR is determined to be 200 after searching between 100 and 500, and the minimum number of leaf nodes is determined to be 1 after searching between 1 and 10. After searching between 1 and 20, the size of n Neighbors in KNN is determined to be 2, and after searching between 5 and 50, the size of leaf nodes is determined to be 5. The minimum number of sample splits for DT is determined to be 32 after searching between 2 and 50, and the minimum number of leaf nodes is determined to be 1 after searching between 1 and 20. The number of trees in GBRT is determined to be 300 after searching between 50 and 500, the maximum depth is determined to be 10 after searching between 3 and 15, and the learning rate is determined to be 0.01 after searching in the logarithmic uniform distribution [0.001, 0.1].

Table 2. Hyperparameters of four machine learning algorithms.

Algorithms	Hyperparameter	Search Scope
KNN	n_neighbors	[1, 20]
	weights	['uniform', 'distance']
	leaf_size	[5, 50]
	p	[1, 2]
DT	min_samples_split	[2, 50]
	min_samples_leaf	[1, 20]
	max_depth	[None, 10, 20, 30, 40, 50]
GBRT	n_estimators	[50, 500]
	max_depth	[3, 15]
	learning_rate	loguniform [0.001, 0.1]

To evaluate the accuracy and error of machine learning models in temperature measurement, this paper primarily employs the Mean Absolute Error (MAE) and goodness of fit R^2 as the evaluation metric:

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{y}_i - y_i| \quad (3)$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (4)$$

where $\hat{y}_i$ is the predicted temperature and y_i is the actual temperature. MAE reflects the average absolute error between the predicted temperature and the actual labeled temperature, while R^2 reflects the quality of the regression equation fitting. That is to say, the closer MAE is to 0 and R^2 is to 1, the higher the accuracy of the model.

3. Update Training Methods for Multi-Scenario Issues and Computational Results Analysis

This section investigates how to improve temperature measurement accuracy for new scenarios by retraining models with updated image databases across three scenarios (Cameras in different imaging directions, different corrosion states of metal plates, and different pixel sizes of images). The minimum update image addition ratio is determined.

3.1. Multi-Scenario Issues

Firstly, let us take a look at the problems caused by two different scenarios. At 30 °C, the GLPDF of the gray channel of black and white images for the clean iron and its rusted state are compared in Figure 3.

The results reveal significant feature differences in the GLPDF between clean and rusted iron. Consequently, a model trained on one image database exhibits larger temperature measurement errors when applied to another database with distinct feature characteristics (e.g., rusted surfaces). This underscores the necessity of image database updates to enhance model generalization across varying field conditions.

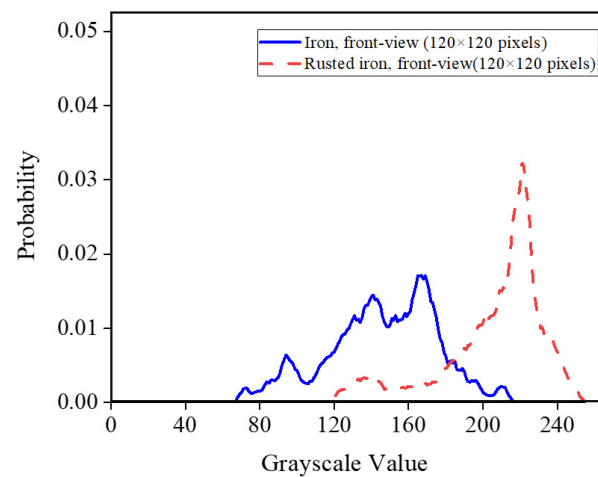


Figure 3. Comparison of the GLPDF for iron and rusted iron at 30 °C.

3.2. Update Training Method for Multi-Scenario Issues

The database structure for addressing multi-scenario issues is shown in Figure 4. Scenario #1 Database (L) represents the Laboratory images (30–100 °C, 4800 images). Scenario #2 Database (S) represents the on-site images (30–100 °C, 4800 images). Table 1 databases are combined to simulate three kinds of on-site scenarios, as shown in Table 3. Randomly select 4000 images from the image S database representing scenario #2 (4800 images) as the update database S_u , and the remaining 800 images as the model's test database S_{test} . Randomly select x images from the update database S_u and add them to the L database, and merge them into a retraining database ($4800 + x$ images). By observing the changes in x , we can observe the impact of η on recognition error. The temperature measurement effect of the retraining model is tested using the S_{test} database in scenario #2, and the test results can reflect the temperature measurement effect on different scenarios. Our ultimate goal is to find the minimum value of x that satisfies the temperature measurement error (e.g., below 3 °C) in engineering by gradually increasing the number of x .

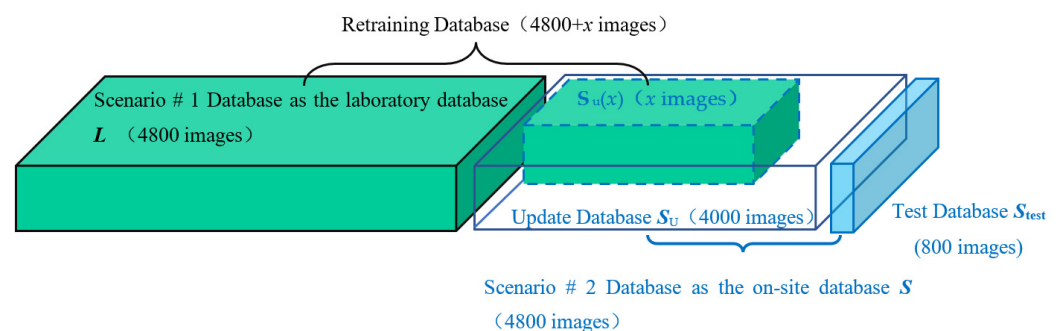


Figure 4. Image database structure for solving multi-scenario problem (0–100 °C).

Table 3. Three simulation research schemes for solving multi-scenario problems.

On-Site Scenarios Problem Type	Scenario #1 Database *	Scenario #2 Database *
Different corrosion states of metal plates	c (clean)	d (rusted)
Cameras in different imaging directions	b (left-angle view)	a (front-view)
Different pixel sizes of images	c (120 × 120 pixels)	a (65 × 80 pixels)

* Databases a , b , c and d are shown in Table 1. Scenario #1 Database (L) represents the Laboratory images, Scenario #2 Database (S) represents the on-site images.

The calculation formula for the proportion η of x images in the update database S_u to the number of images in the L database is as follows:

$$\eta = \frac{[S_u(x)]}{[L]} \quad (5)$$

where $[]$ is the number of total images in the database.

3.3. Impact of Update Proportion on Measurement Errors

Figures 5–7 show that the MAE and R^2 vary with the updated image proportion η curves for the three scenarios across four ML algorithms (RFR, GBRT, DT, KNN).

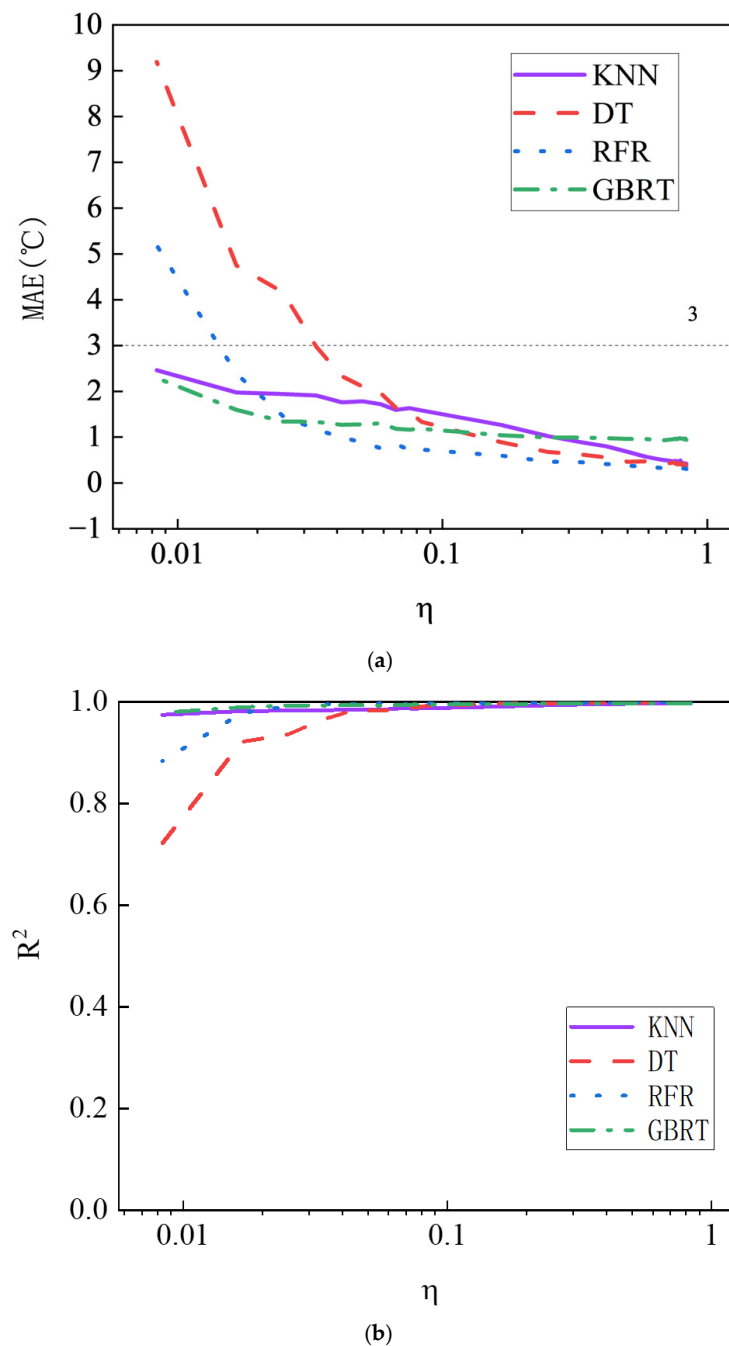


Figure 5. The measurement error MAE, (a), and goodness of fit R^2 , (b), of the rusted iron images vary with the updated image proportion η of the rusted iron database. The horizontal axis is a logarithmic coordinate.

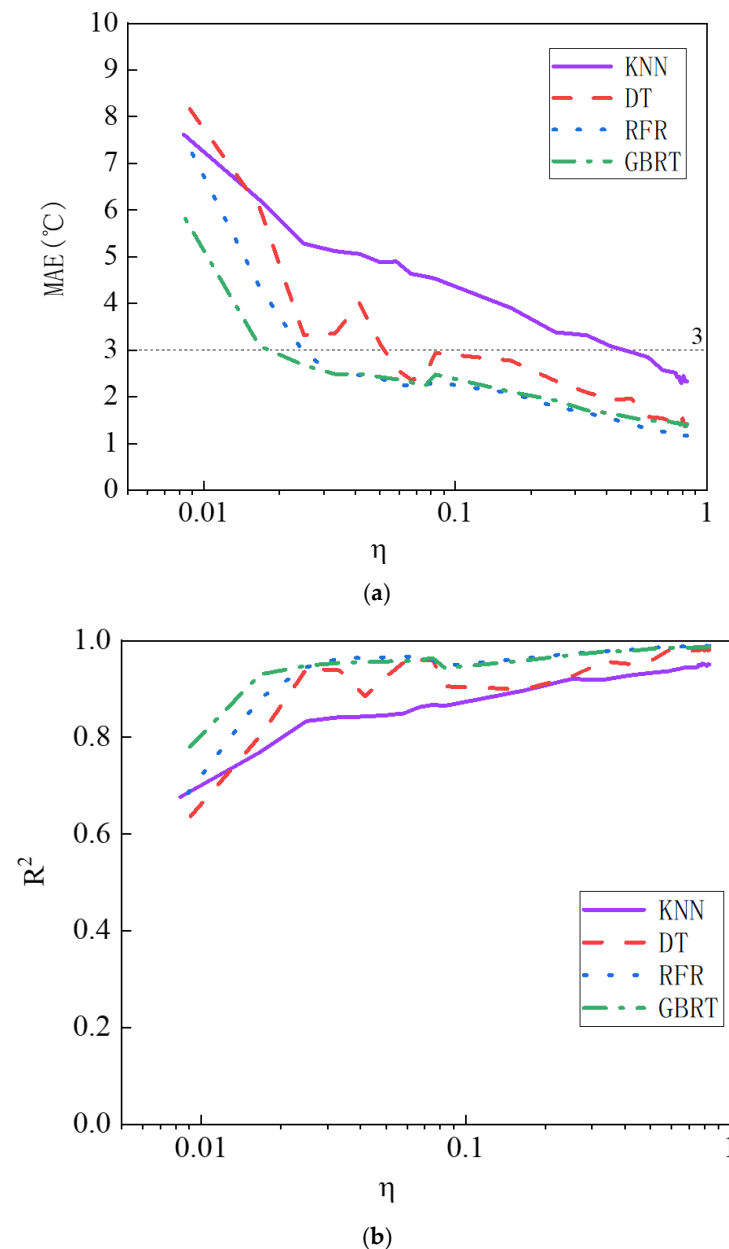


Figure 6. The measurement error MAE, (a), and goodness of fit R^2 , (b), of images with a different imaging direction vary with the updated image proportion η of a different imaging direction. The horizontal axis is a logarithmic coordinate.

It can be seen that before adding new scene images to the existing image database ($\eta = 0$), the temperature errors of all four algorithms are relatively large (MAE close to 10 $^{\circ}\text{C}$), indicating that relying solely on models trained in the laboratory results in significant errors when used in the field. However, with increase in η , MAE shows a rapid downward trend, and R^2 also approaches 1 quickly. As long as η reaches more than 5%, the error can be basically reduced to below 3 $^{\circ}\text{C}$, which meets the needs of general temperature rise fault identification. This indicates that adding a certain proportion of on-site images to the laboratory image database can reduce measurement errors.

When comparing the strengths and weaknesses of the four algorithms at different updated image proportions, it can be found that in the three training sessions, the KNN algorithm exhibits less regularity in its performance compared to the other three algorithms at different stages. This phenomenon occurs because KNN is an instance-based learning method that relies on distances between samples, while DT, GBRT, and RFR are tree-based

algorithms, which have significant differences in principle, leading to less comparability in their performance trends.

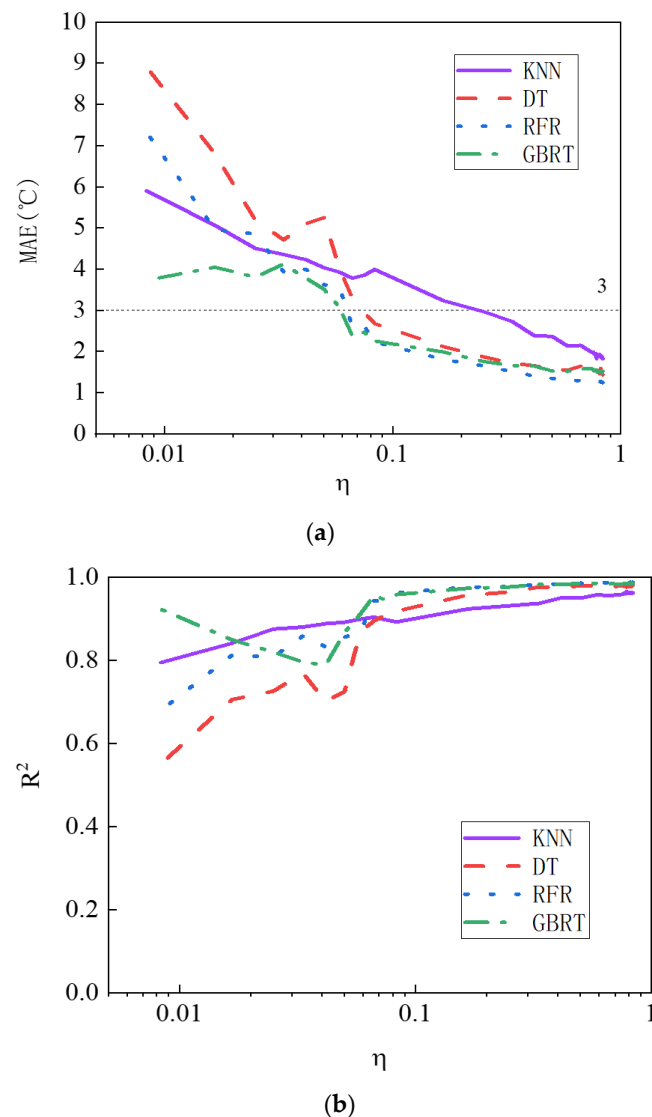


Figure 7. The measurement error MAE, (a), and goodness of fit R^2 , (b), of images with a different pixel size varies with the updated image proportion η of a different pixel size. The horizontal axis is a logarithmic coordinate.

When excluding the KNN algorithm and comparing only the curves of DT, GBRT, and RFR, it can be observed that when η is less than 5%, the GBRT algorithm has the smallest error, while the DT algorithm has the largest error. When η reaches above 5%, among the three algorithms, the training performance of the RFR algorithm becomes the best, while the performance of the GBRT algorithm gradually declines, eventually resulting in the largest error for GBRT and the smallest error for RFR.

RFR has unique advantages in handling high-dimensional features. When training, RFR randomly selects a subset of features for each decision tree, which reduces the correlation between trees, extracts useful information from numerous features, and improves high-dimensional data performance. GBRT and DT are more sensitive to irrelevant or noisy features, which can easily lead to overfitting and a decrease in prediction accuracy. For the DT, GBRT, and RFR algorithms, when the updated image proportion is low, the DT algorithm tends to overfit the features of the base database images, leading to the largest recognition error for the update database. The GBRT algorithm focuses on reducing bias,

enabling the model to better adapt to data complexity, especially when the data volume is small, which results in the smallest recognition error for the updated database. When the updated image proportion reaches above 5%, the RFR algorithm, which emphasizes reducing variance, enhances the stability and generalization ability of the model. As the data volume increases, its performance gradually surpasses that of DT and GBRT.

3.4. Comparative Analysis of Measurement Errors Across Scenarios

Previously, we analyzed the error of updating the image database model to predict the S_{test} database for three different scenarios (different corrosion states, different imaging directions of images, and different pixel sizes). Below, we compare the differences in the measurement error of updating the image database model for S_{test} database for three different scenarios under the same standard. The machine learning algorithm selected is the RFR algorithm. There are two evaluation criteria: When 40 images are randomly selected from the S_u database and added to the L database, updating the model to measure the temperature error of the S_{test} database is used as the first criterion. When the temperature measurement error of S_{test} database drops below 2 °C, the updated image proportion of S_u database is the second criterion. It can be considered that the smaller the measurement error of the S_{test} database and the updated image proportion of S_u database, the better the performance of the updated model and the more excellent the effect. The evaluation results are shown in Table 4. The table also includes test results without retraining.

Table 4. Performance evaluation values of models across scenarios.

On-Site Scenarios Problem Type	MAE Without Retraining /°C	MAE of S_{test} Database (40 Images Updated) /°C	Updated Image Proportion of S_u
Different corrosion states of metal plates	34.25	5.187	0.024
Cameras in different imaging directions	21.44	7.57	0.167
Different pixel sizes of images	29.44 °C	8.142	0.077

It can be seen that the test results are poor without retraining. After updating the model and retraining, the evaluation results obtained by the two evaluation criteria are consistent. It can be concluded that the ranking of the three scenarios from worst to best in terms of measurement error in the S_{test} database is as follows: different imaging directions, different pixel sizes, and different corrosion states. Among these, imaging directions and pixel sizes are differences in image acquisition conditions, while corrosion states are differences in material conditions. Therefore, compared to material condition differences, image acquisition condition differences have a greater impact, and this conclusion is beneficial for on-site temperature measurement because the image acquisition conditions are more controllable.

In summary, during the image acquisition process, priority should be given to standardizing acquisition conditions to reduce temperature measurement errors in practical applications. Additionally, increasing samples with different materials and corrosion states can enhance the model's generalization ability.

4. Update Training Methods for Imbalanced Image Data and Computational Results Analysis

Due to the fact that the electrical equipment on site is operating normally, it is difficult or rare to obtain high-temperature fault images (rare samples). Therefore, even if a portion of the updated image database on site are added to the retrained model, they

are basically ambient temperature images. This section mainly studies how to retrain the model by adding rare sample images, thereby improving the accuracy of temperature measurement for on-site high-temperature samples (rare samples), and determining the minimum updated image proportion for on-site high-temperature images.

4.1. Imbalanced Image Data Issues

Firstly, let us take a look at the problems caused by imbalanced image data. Using the RFR algorithm, the training set consists of the low-temperature images (30–50 °C) from a front-view iron database. The trained model remains unchanged and is directly used to predict the temperature of the high-temperature images (51–80 °C). The results are shown in Figure 8.

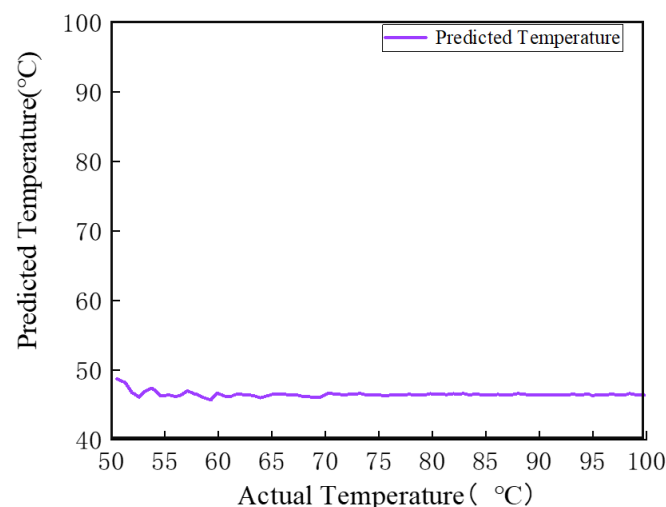


Figure 8. The model trained in the ambient temperature range (30–50 °C) directly predicts the temperature of images in the high temperature range (51–80 °C).

It can be seen that due to the fact that the training set only contains images in the ambient temperature range (30–50 °C), the model has not been able to access the features of the high temperature range, which limits its predictive ability for high temperature images. Therefore, when updating the laboratory image database with the on-site images, increasing the number of samples in the high temperature range is the key to improving the model's predictive ability when the data is imbalanced. It is worth noting that although the predicted results are inaccurate, they tend to be biased towards the upper limit of the ambient temperature range (50 °C), indicating that the temperature information contained in the ambient temperature image also has a certain impact on predicting the high temperature image.

4.2. Update Training Method for Imbalanced Image Data

The image database structure for solving the problem of imbalanced data is shown in Figure 9, with the following details: Use image database *a* (front-view, 30–100 °C, 4800 images) in Table 1 as the laboratory's full temperature range image database *L*; divide the image database *b* (left-angle view, 30–100 °C, 4800 images) in Table 1 into three sub-databases: the ambient temperature database S_n (1370 images at 30–50 °C), the high temperature update database S_{h-u} (3087 images at 51–100 °C), and the high temperature test database S_{h-test} (343 images at 51–100 °C). When retraining, merge the *L* database and S_n database, and then merge them with *x* images extracted from the S_{h-u} database to form the retraining database (4800 + 1370 + *x*). By observing the changes in *x*, we can observe the impact of η on recognition error. The error of the model trained in the

retraining database is tested using the high-temperature database S_{h-test} . The test results can reflect the effectiveness of temperature measurement on high-temperature images and thus find the minimum number of high-temperature images x that satisfies the temperature measurement error in engineering.

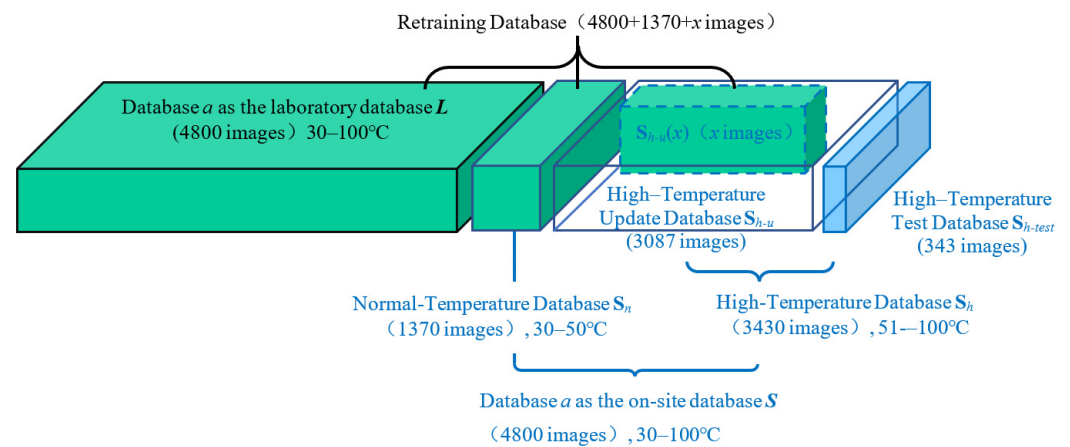


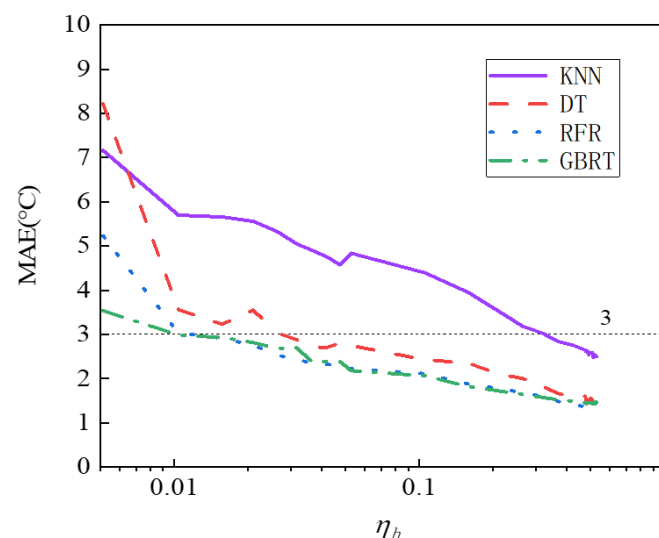
Figure 9. Image database structure for solving imbalanced data problem (30–100 °C).

The calculation formula for the proportion of x images in the high-temperature update database S_{h-u} to the number of images in the $L + S_n$ database is as follows:

$$\eta_h = \frac{[S_{h-u}(x)]}{[L] + [S_n]} = \frac{x}{4800 + 1370} \quad (6)$$

4.3. Impact of High-Temperature Updated Image Proportion on Errors

The machine learning model is obtained by updating the image database, and then the temperature recognition is performed on the S_{test} database. Figure 10 shows that the MAE and R^2 vary with η curves, which displays the results of four machine learning algorithms.



(a)

Figure 10. Cont.

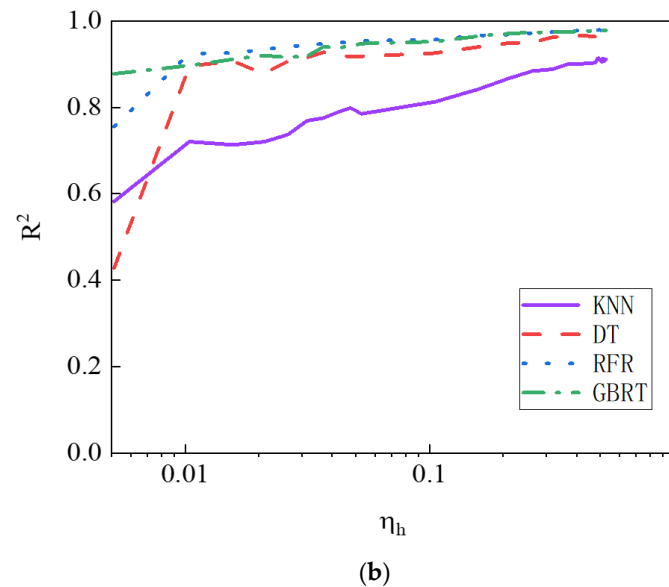


Figure 10. The measurement error MAE, (a), and goodness of fit R^2 , (b), of the high-temperature test database varies with the updated image proportion η_h of $S_{h-u}(x)$. The horizontal axis is a logarithmic coordinate.

As η_h increases, the temperature error MAE for the four algorithms shows a downward trend. When η_h reaches above 1% (61 images), the error generally drops below 3°C , meeting the basic requirements for temperature-rise fault identification in engineering.

In terms of algorithm performance, the ranking from highest to lowest is RFR, GBRT, DT, and KNN. RFR and GBRT, as ensemble learning algorithms effectively reduce overfitting and enhance robustness to data imbalance. Thus, they excel in high-temperature conditions with lower errors. DT is prone to overfitting and sensitive to noise, leading to larger errors. KNN, being sensitive to sample distribution and noise, has the highest errors.

4.4. Impact of Different High-Temperature Setting Ranges on Measurement Errors

Due to the different fault temperature ranges of different equipment on site, as well as the varying difficulty of collecting data under high-temperature conditions ($81\text{--}100^\circ\text{C}$) or ($51\text{--}100^\circ\text{C}$), it is necessary to conduct research on the impact of different fault temperature settings on the model.

Reduce the temperature range of the high-temperature update library S_{h-u} to $81\text{--}100^\circ\text{C}$ (1370 images), expand the temperature range of the ambient temperature library S_n to $30\text{--}80^\circ\text{C}$ (3430 images), and at the same time, reduce the temperature range of the test library S_{h-test} to $81\text{--}100^\circ\text{C}$ (152 images) and name this fault temperature setting range as the high temperature condition. Name the fault temperature range ($51\text{--}100^\circ\text{C}$) in Section 4.3 as the medium-high temperature condition.

When using the RFR algorithm, the error curve for the high-temperature database S_h varying with η_h is shown in Figure 11. In both operating conditions, the temperature error MAE shows a decreasing trend with the increase in η_h . However, the decrease is faster in high-temperature operating conditions at high proportions, indicating that a smaller high-temperature setting range is beneficial for reducing recognition errors. The reason may be that when the high temperature range of the on-site (b database) images is reduced (from $51\text{--}100^\circ\text{C}$ to $81\text{--}100^\circ\text{C}$), under this scheme, the number of on-site (b database) images in the ambient temperature range increases accordingly (from 1370 images to 3430 images), resulting in an increase in the number of on-site (b database) images participating in retraining (although only at ambient temperature). The retrained model contains more image features of the on-site (b database), so adding fewer on-site

(b database) high-temperature images can improve the recognition accuracy of on-site high temperature. However, overall, the impact of high-temperature setting range on prediction error is almost the same, and this conclusion is beneficial for on-site fault temperature identification, because the actual range of high temperature for on-site faults is not known; only the normal operating temperature range is known.

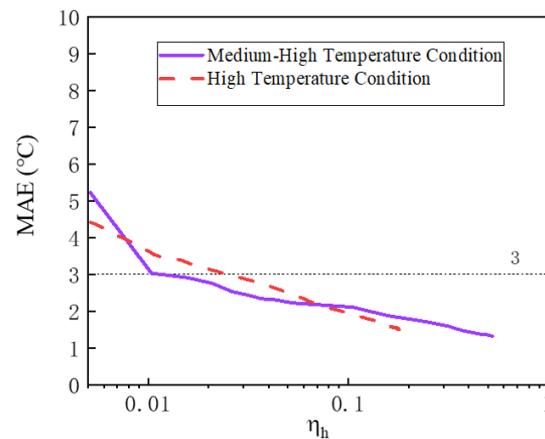


Figure 11. The recognition error MAE of different high-temperature setting ranges varies with the updated image proportion η_h of $S_{h-u}(x)$. The horizontal axis is a logarithmic coordinate.

5. Temperature Monitoring at the Substation Site

We will apply the methods discussed above to temperature monitoring at the substation site. Figure 12 shows the layout of the on-site transformer, with the video equipment positioned 9m away from the transformer, the daylight intensity during this time period was maintained at around 3800–8000 lux.

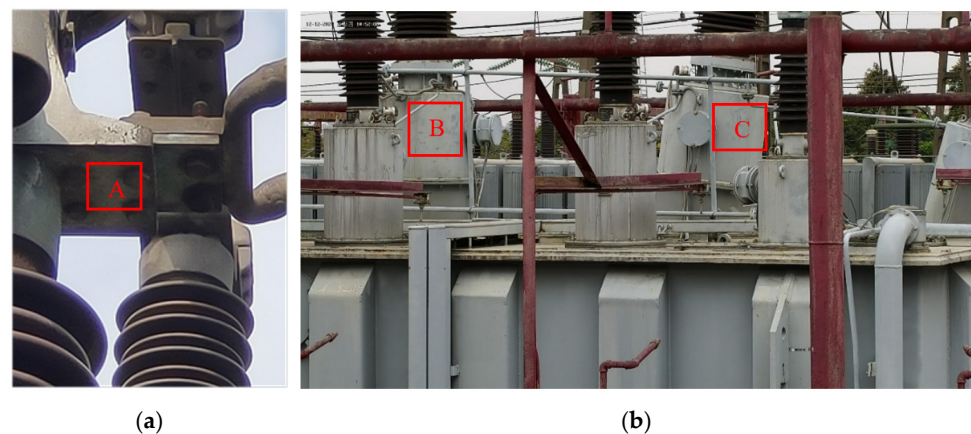


Figure 12. Image of temperature monitoring location for three-phase transformer. (a) A-phase metal wire clamp image, (b) B and C phase transformer bushings. The red boxes display the cropped image of the temperature monitoring location respectively.

5.1. Image Database Construction

The established transformer bushing image library has four types.

- A-phase metal wire clamp image (1300 images), with cropped image pixels of 150×90 . The temperature range of infrared monitoring is 30–45 °C.
- Take 3000 front-view images of phase B transformer bushings, with cropped image pixels of 65×80 . The temperature range of infrared monitoring is 41.5–49 °C.
- Take 3000 front-view images of phase B transformer bushings, with cropped image pixels of 120×1200 . The temperature range of infrared monitoring is 41.5–49 °C.

- (d) Take 3000 images left-angle view (at an angle of about 10 degrees from the front direction) in phase B transformer bushings, and the cropped image has a pixel size of 65×80 . The temperature range of infrared monitoring is 42–45 °C.
- (e) Take 3000 front-view images of phase C transformer bushings, and the cropped image has a pixel size of 65×80 . The temperature range of infrared monitoring is 37–47 °C.

Then select different databases to form multiple scenes with different corrosion states, imaging directions, and pixels, as shown in Table 5.

Table 5. Three scenarios.

Multi-Scenario	L Database	S _u Database
Different corrosion states	C-phase	B-phase
Different imaging directions	B-phase (left-angle view)	B-phase (front-view)
Different pixel sizes	B-phase (120 × 120)	B-phase (65 × 80)

5.2. Prediction Results of On-Site Image Databases Modeling at the Same Location

Compare the differences in predicting S_{test} databases errors using three different scene update image databases models. The machine learning algorithm selected the RFR algorithm, and the proportion of S_u databases (referred to as S_u database proportion) when the predicted temperature error of the S_{test} database decreases to below 0.5 °C and below 2 °C, respectively, is shown in Table 6. Obviously, the higher the proportion, the smaller the error, and as long as the proportion is greater than 1%, the error is less than 2 °C. This result is actually better than the results in Table 4, because the temperature of the equipment tested on site works in a smaller temperature range (about 10 °C), while the temperature range of the laboratory equipment is larger (70 °C). The difficulty of predicting the temperature of the on-site image is lower than that of the laboratory image.

Table 6. Performance evaluation values of models in different scenarios *.

Multi-Scenario	Updated Image Proportion of S_u (MAE of S_{test} Below 0.5 °C)	Updated Image Proportion of S_u (MAE of S_{test} Below 2 °C)
Different corrosion states	0.058	>0.01
Different imaging directions	0.032	>0.01
Different pixel sizes	0.032	>0.01

* Databases are shown in Table 5.

5.3. Prediction Results of On-Site Image Databases Modeling at the Different Location

The previous issue was about the image library update strategy for the same location and device. Now, we will directly test the image library update strategy for different locations and devices (of the same material) using the on-site clamp image library model. Below are two on-site prediction schemes:

(1) Wire clamp model: The model trained using the on-site A-phase wire clamp image library (1300 images) in Figure 12a predicts the temperature of the on-site knife switch wire clamp and reactor wire clamp (10 images each), both made of aluminum alloy. Figure 13 shows the original and cropped images of the on-site knife switch wire clamp and reactor wire clamp.

(2) Update model: Using 1300 on-site A-phase clamp images, as well as 90 on-site images of knife switch clamp and reactor clamp images, establish a merged image library

of updated images, train the updated model, and then predict the results of the knife switch clamp and reactor clamp test library images (10 images each).

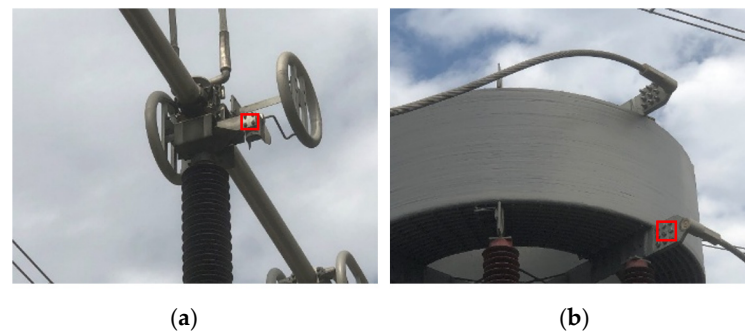


Figure 13. Image of temperature monitoring location for (a) knife switch clamp, (b) reactor clamp. The red boxes display the cropped image of the temperature monitoring location respectively.

Table 7 shows the predicted results of two on-site models. It can be seen that the mean prediction errors of wire clamp model are relatively large (MAE close to -8.9 and -23.6 °C), indicating that relying solely on models trained in the wire clamp library results in significant errors when used in the knife switch and reactor clamp image. However, if an updated model is adopted, even if the proportion of the knife switch and reactor clamp image to the wire clamp image is only $90/1300 = 6.9\%$, a good prediction error (0 or -1.4 °C) can still be obtained. This conclusion is beneficial for building a database for multi-scenario problems in substations.

Table 7. Prediction results of two on-site models.

Actual Temperature /°C		Wire Clamp Model		Updated Image	
		Predicted Temperature/°C	Error/°C	Predicted Temperature/°C	Error/°C
Knife switch clamp	52.2	43.3	-8.9	52.2	0.0
Reactor clamp	55.0	31.4	-23.6	53.6	-1.4

5.4. Prediction Results of Mixed Modeling of Laboratory Image Databases and On-Site Image Databases

A more interesting question is can we use models trained from laboratory image databases to directly predict the temperature of on-site transformer images? We made a simple attempt here by coating the plate in the laboratory with the same paint as the transformer bushing on site. We captured images at a distance of 9 m and established an image database as the L database. Then, we used the on-site images as the S_u database and made predictions on the on-site images. The results are shown in Table 8.

Table 8. Performance evaluation values of models in a hybrid database of on-site and laboratory images.

65×80	Updated Image Proportion of S_u (MAE of S_{test} Below 0.5 °C)	Updated Image Proportion of S_u (MAE of S_{test} Below 2 °C)
η	0.33	0.05

Obviously, the higher the proportion of updated databases, the smaller the prediction error, and as long as the proportion is greater than 5% and the error is less than 2 °C, it still meets the monitoring requirements for engineering temperature rise faults. If we want to

meet higher requirements, such as an error below 0.5 °C, the proportion needs to reach 33%. Such a high proportion is equivalent to reacquiring the image databases.

6. Discussion

It should be pointed out that the focus of this article is on studying on-site strategies, so more advanced machine learning algorithms are not our main focus. Here, we only briefly mention some algorithm comparison results.

Table 9 presents the comprehensive comparison results with other algorithms.

Table 9. Comparison with other algorithms.

Algorithms	Features and Model Architecture	Model Parameter Count	Hardware Requirements	Total Training Time	Single Graph Inference Time	MAE
Proposed Method	RGB-GLH + RFR (27-dimension)	Extremely small (hundreds of decision trees)	Ordinary CPU	3.77 s	0.24 ms	1.60 °C
Traditional Texture (GLCM)	GLCM + RFR (20-dimension)	Extremely small (hundreds of decision trees)	Ordinary CPU	4.14 s	0.26 ms	5.96 °C
Deep Learning (ResNet18)	End to end fine-tuning (training for 10 epochs)	about 11.18 M	Strong dependence on GPU	103.60 s	1.65 ms	4.07 °C

The comprehensive comparative experimental results show that the RGB-GLH method proposed in this paper has significant advantages in prediction accuracy, physical interpretability, and engineering deployment potential. In terms of prediction accuracy, the average error of RGB-GLH is only 1.601 °C, far lower than deep learning ResNet18 (4.079 °C) and traditional spatial texture GLCM (5.963 °C). GLCM, as a classic image feature, performed at the bottom of this task. This result indirectly proves that temperature changes mainly cause global intensity attenuation of diffuse reflectance light (ODRL) in all directions, rather than changes in local spatial texture. Due to the lack of direct statistics on global light intensity, GLCM missed key temperature mapping information; the statistical features of this article accurately capture this core rule.

On the other hand, although the ResNet18 model with end-to-end fine-tuning achieved an average accuracy of 4.079 °C and reached a compromise of 1.528 °C in some data, demonstrating the powerful high-dimensional fitting ability of deep features, its engineering implementation faces enormous challenges. ResNet18 contains massive parameters of up to 11.18 million (11.18 million), which not only heavily relies on GPU computing power, but also takes up to 1.657 ms for single image inference. In contrast, the RGB-GLH method breaks free from GPU limitations and only requires 0.245 ms for single image inference (a nearly 7-fold increase in speed).

Table 10 shows the results of comparing the image database update strategy proposed in this paper with two classic imbalanced data processing methods (weighted loss function and SMOTE oversampling) using the RFR algorithm under identical experimental conditions, taking different shooting directions in a multi-scene problem as an example.

Table 10. MAE and R² of RFR algorithm under three strategies.

η (%)	Proposed Method MAE/°C	Weighted Loss MAE/°C	SMOTE MAE/°C	Proposed Method R ²	Weighted Loss R ²	SMOTE R ²
1.6	4.28	5.77	3.55	0.884	0.785	0.922
3.2	2.76	2.75	3.29	0.956	0.955	0.928
4.7	2.59	2.32	2.95	0.957	0.967	0.936
6.2	2.18	2.25	2.71	0.967	0.967	0.944
7.7	2.34	2.07	2.23	0.951	0.970	0.963
14.3	2.06	1.76	1.92	0.962	0.979	0.973
20.0	1.79	1.62	1.72	0.974	0.982	0.979
25.0	1.63	1.51	1.53	0.979	0.983	0.984
29.4	1.49	1.43	1.50	0.983	0.985	0.984
33.3	1.38	1.36	1.42	0.986	0.986	0.985
36.8	1.30	1.29	1.33	0.987	0.987	0.987
40.0	1.27	1.27	1.31	0.988	0.988	0.987
42.9	1.22	1.22	1.20	0.988	0.988	0.989
44.2	1.21	1.20	1.22	0.988	0.989	0.989

- Under the proportion proposed in the article ($\eta = 5\%$):

When $\eta = 4.7\%$, the image database update strategy proposed in this paper has MAE = 2.59 °C, weighted loss MAE = 2.32 °C, and SMOTE MAE = 2.95 °C, all of which meet the error requirements of 3 °C in engineering. At the same time, the error difference between the three schemes is not significant, and only adding about 5% of real on-site images can achieve an accuracy level comparable to complex methods.

- In the case of a small sample size ($\eta < 5\%$):

The SMOTE algorithm outperforms the image database update strategy proposed in this paper (4.28 °C) in terms of MAE (3.55 °C) when there are very few samples, but its subsequent performance is not satisfactory; The weighted loss performs the worst at $\eta = 1.6\%$ (5.77 °C), indicating its sensitivity to weight settings. The original method has approached 2.8 °C at $\eta = 3.2\%$, reaching an engineering usable level, and the curve is steadily rising.

- In the case of a large sample size ($\eta \geq 5\%$):

The MAE of the three methods is stable in the range of 1.2–2.6 °C. The image database update strategy proposed in this article is comparable to weighted and SMOTE algorithms in most proportions, and even better in some proportions. Weighted loss and SMOTE algorithms have a slight advantage in some proportions, which has limited engineering significance, but it will increase training costs and time.

Therefore, in terms of temperature measurement accuracy, the image database update strategy proposed in this article has no substantial disadvantage compared to weighted loss and SMOTE. At the critical 5% addition ratio, all three can reliably control the error within 3 °C, meeting the actual monitoring requirements. Meanwhile, the training time of the image database update strategy proposed in this article is relatively short, so adopting this scheme is more in line with engineering practice.

7. Conclusions

This article addresses the issues of poor generalization of the laboratory image database model in the on-site environment and the problem of imbalanced data in the on-site image database. It proposes an image database update scheme for multi-scenario problems and imbalanced data problems. Then, a retraining strategy is implemented in the on-site environment, which improves the engineering applicability of the visible light image temperature measurement model.

In response to the multi-scenario problem of different corrosion states, imaging directions, and pixel sizes on site, it has been found that when the updated image proportion on site reaches above 5% or more, the temperature measurement error of the updated model for different scene test databases can be reduced to below 3 °C. It has also been found that using standardized imaging parameters and image cropping methods on site can effectively improve measurement accuracy.

In response to the problem of limited and imbalanced on-site high-temperature image data, it has been found that when the on-site high-temperature updated images proportion reaches 1% or more, the temperature measurement error predicted by the updated model in the high-temperature testing database decreases to below 3 °C. It has also been found that the range of high-temperature settings has little effect on improving measurement accuracy and is convenient for on-site use.

On the basis of comparing the training effects of four machine learning methods, RFR, GBRT, DT, and KNN, it was found that the RFR algorithm is the optimal algorithm for solving the above problems due to its unique advantages in processing high-dimensional features. Although these methods belong to the traditional machine learning methods, these strategies can still provide some guidance and methods for on-site use in dealing with multi-scenario problems and data imbalance issues.

The database construction strategy for multi scenario problems has been validated on site in substations, proving the feasibility of the above conclusions.

Of course, when using machine learning to identify multi-scene problems on site, there are still problems with low recognition accuracy caused by the diversity of on-site scenes and changes in sunlight intensity. It is necessary to establish a larger database and search for amplification of weak physical signals.

Author Contributions: Investigation, X.L., W.W., C.L., Z.T. and Q.Y.; Resources, Y.C.; Writing—review & editing, Q.Y.; Supervision, Q.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Scientific Research Project of China Southern Power Grid (Grant Nos. GDKJXM20231287 and 030000KC23110062).

Institutional Review Board Statement: Not applicable for studies not involving humans or animals.

Informed Consent Statement: Not applicable for studies not involving humans.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: Thank you for the administrative and technical support from the School of Electrical and Electronic Engineering, Huazhong University of Science and Technology.

Conflicts of Interest: Authors X.L., C.L., and Y.C. were employed by the Electric Power Research Institute, Guangdong Power Grid Co., Ltd. The remaining authors declare that the re-search was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Ni, N.; Zhang, K.; Hu, J.; Li, L.; Mi, S.; Zhang, Y.; Zhang, Y. Combined use of blackbody and infrared radiation for accurate measurement of temperature field of aluminum alloys. *Optik* **2022**, *268*, 169763. [\[CrossRef\]](#)
2. Hong, X.; Chen, G.; Chen, Y.; Cai, R. Research on abnormal ship brightness temperature detection based on infrared image edge-enhanced segmentation network. *Appl. Sci.* **2025**, *15*, 3551. [\[CrossRef\]](#)
3. Mao, X.; Xie, Q. Neural-network-based multi-spectral thermometry and emissivity reconstruction in cavity high-temperature environments. *Appl. Opt.* **2025**, *64*, 7304–7314. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Giulietti, N.; Chiariotti, P.; Zanelli, F.; Debattisti, N.; Cigada, A. Combined use of infrared imaging and deep-learning techniques for real-time temperature measurement of train braking components. *IEEE Trans. Instrum. Meas.* **2025**, *74*, 5009908. [\[CrossRef\]](#)
5. Wang, B.; Dong, M.; Ren, M.; Wu, Z.; Guo, C.; Zhuang, T.; Pischler, O.; Xie, J. Automatic fault diagnosis of infrared insulator images based on image instance segmentation and temperature analysis. *IEEE Trans. Instrum. Meas.* **2020**, *69*, 5345–5355. [\[CrossRef\]](#)
6. Farzaneh, M.; Maize, K.; Lüerßen, D.; Summers, J.A.; Mayer, P.M.; Raad, P.E.; Pipe, K.P.; Shakouri, A.; Ram, R.J.; Hudgings, J.A. CCD-based thermoreflectance microscopy: Principles and applications. *J. Phys. D Appl. Phys.* **2009**, *42*, 143001. [\[CrossRef\]](#)
7. Zhang, H.; Wen, S.-B.; Bhaskar, A. Two-wavelength thermoreflectance in steady-state thermal imaging. *Appl. Phys. Lett.* **2019**, *114*, 151902. [\[CrossRef\]](#)
8. Matatagui, E.; Thompson, A.G.; Cardona, M. Thermoreflectance in Semiconductors. *Phys. Rev.* **1968**, *176*, 950–960. [\[CrossRef\]](#)
9. Guidotti, D.; Van Driel, H.M. Spatially resolved defect mapping in semiconductors using laser-modulated thermoreflectance. *Appl. Phys. Lett.* **1985**, *47*, 1336–1338. [\[CrossRef\]](#)
10. Christofferson, J.; Shakouri, A. Thermoreflectance based thermal microscope. *Rev. Sci. Instrum.* **2005**, *76*, 024903. [\[CrossRef\]](#)
11. Grauby, S.; Forget, B.C.; Holé, S.; Fournier, D. High resolution photothermal imaging of high frequency phenomena using a visible charge coupled equipment camera associated with a multichannel lock-in scheme. *Rev. Sci. Instrum.* **1999**, *70*, 3603–3608. [\[CrossRef\]](#)
12. Martin-Horcajo, S.; Pomeroy, J.W.; Lambert, B.; Jung, H.; Blanck, H.; Kuball, M. Transient thermoreflectance for gate temperature assessment in pulse operated GaN-based HEMTs. *IEEE Electr. Device Lett.* **2016**, *37*, 1197–1200. [\[CrossRef\]](#)
13. Yang, H.Z.; Leong, S.H.; An, C.W.; Chen, Y.J.; Ye, K.D.; Ng, L.T.; Zhang, J.; Ng, K.K.; Hu, J.F. Thermoreflection measurement of magnetic thin films. *IEEE Trans. Magn.* **2013**, *49*, 2827–2830. [\[CrossRef\]](#)
14. Kurbanova, B.; Chakraborty, D.; Abdullaev, A.; Shamatova, A.; Makukha, O.; Belarouci, A.; Lysenko, V.; Azarov, A.; Kuznetsov, A.; Wang, Y.; et al. Multiscale phonon thermal transport in nano-porous silicon. *Appl. Phys. Lett.* **2024**, *124*, 252202. [\[CrossRef\]](#)
15. Lundh, J.S.; Pavlidis, G.; Sasaki, K.; Centrone, A.; Spencer, J.A.; Masten, H.N.; Currie, M.; Jacobs, A.G.; Konishi, K.; Kuramata, A.; et al. Assessment of channel temperature in β -(Al_xGa_{1-x})₂O₃/Ga₂O₃ heterostructure field-effect transistors using visible wavelength thermoreflectance thermal imaging. *Appl. Phys. Lett.* **2024**, *124*, 054103. [\[CrossRef\]](#)
16. Zheng, X.; Pomeroy, J.W.; Jindal, G.; Kuball, M. Temperature-dependent thermal impedance measurement of GaN-based HEMTs using transient thermoreflectance. *IEEE Trans. Electron Devices* **2024**, *71*, 2367–2372. [\[CrossRef\]](#)
17. Kumar, N.; Vaca, D.; Joishi, C.; Xia, Z.; Rajan, S.; Kumar, S. Ultrafast thermoreflectance imaging and electrothermal modeling of β -Ga₂O₃ MESFETs. *IEEE Electron Device Lett.* **2020**, *41*, 641–644. [\[CrossRef\]](#)
18. Ujihara, K. Reflectivity of metals at high temperatures. *J. Appl. Phys.* **1972**, *43*, 2376–2383. [\[CrossRef\]](#)
19. Yilbas, B.S.; Danisman, K.; Yilbas, Z. Measurement of temperature-dependent reflectivity of Cu and Al in the range 30–1000 degrees C. *Meas. Sci. Technol.* **1991**, *2*, 668. [\[CrossRef\]](#)
20. Brezov, D.; Hristov, H.; Dimov, D.; Alexiev, K. Predicting the rectal temperature of dairy cows using infrared thermography and multimodal machine learning. *Appl. Sci.* **2023**, *13*, 11416. [\[CrossRef\]](#)
21. Ding, Y.; Dai, X.; Zhang, T. Low-cost fiber-optic temperature measurement system for high-voltage electrical power equipment. *IEEE Trans. Instrum. Meas.* **2010**, *59*, 923–933. [\[CrossRef\]](#)
22. Yuan, Z.; Ye, Q.; Wang, Y.; Shi, H. Temperature measurement of metal surface at ambient temperatures by visible images and machine learning. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 2514516. [\[CrossRef\]](#)
23. Yuan, Z.; Ye, Q.; Li, W.; Nie, X. Temperature difference measurement of metal surfaces at ambient temperatures under sunlight by differential chromatic features of visible images. *IEEE Sens. J.* **2021**, *21*, 21221–21238. [\[CrossRef\]](#)
24. Zhu, J.; He, Z.; Huang, C.; Zheng, J.; Lin, H.-C.; Chen, F.C.; Yu, C.Q.; Li, Y.; Zhang, Y.T.; Chen, H.T.; et al. Deep learning-assisted common temperature measurement based on visible light imaging. *Chin. Phys. B* **2024**, *33*, 080701. [\[CrossRef\]](#)
25. Shan, L.; Tang, C.; Hong, B.; Kong, M. Reconstruction of a three-dimensional temperature field in flames based on ES-ResNet18. *Appl. Opt.* **2024**, *63*, 1982–1990. [\[CrossRef\]](#)
26. Schöttle, M.; Tran, T.; Oberhofer, H.; Retsch, M. Machine learning enabled image analysis of time-temperature sensing colloidal arrays. *Adv. Sci.* **2023**, *10*, 2205512. [\[CrossRef\]](#)
27. Xu, X.; Jia, W.; Xu, S.; Ning, F.; Lei, J. Relationship between high-level color features and temperature mapping of magnesium alloy surface images based on the K-nearest neighbor algorithm. *Appl. Therm. Eng.* **2025**, *259*, 124940. [\[CrossRef\]](#)

28. Zhu, J.; He, Z.; Huang, C.; Zeng, J.; Lin, H.; Chen, F.; Yu, C.; Li, Y.; Zhang, Y.; Chen, H.; et al. Realization of normal temperature detection through visible light images by Retinex-CNN. *J. Opt.* **2025**. [[CrossRef](#)]
29. Pan, S.J.; Tsang, I.W.; Kwok, J.T.; Yang, Q. Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* **2011**, *22*, 199–210. [[CrossRef](#)]
30. Pan, S.J.; Yang, Q. A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
31. Rinanto, N.; Kuo, C.-H. PCA-ANN contactless multimodality sensors for body temperature estimation. *IEEE Trans. Instrum. Meas.* **2021**, *70*, 4005816. [[CrossRef](#)]
32. Qu, Z.; Yang, K.; Li, Y.; Jiang, X.; Zhang, Y.; Zhao, Y.; Wu, W.; Gao, Y.; Gu, Z.; Zhao, Z. On grain security by temperature interpolation: A deep learning method for comprehensive data fusion in smart granaries. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 9519920. [[CrossRef](#)]
33. Husain, G.; Nasef, D.; Jose, R.; Mayer, J.; Bekbolatova, M.; Devine, T.; Toma, M. SMOTE vs. SMOTEENN: A study on the performance of resampling algorithms for addressing class imbalance in regression models. *Algorithms* **2025**, *18*, 1. [[CrossRef](#)]
34. Jawa, M.; Meena, S. Software effort estimation using synthetic minority over-sampling technique for regression (SMOTER). In *Proceedings of the 2022 3rd International Conference for Emerging Technology (INCET), Belgaum, India, 27–29 May 2022*; IEEE: Piscataway, NJ, USA, 2022; pp. 1–6. [[CrossRef](#)]
35. Xu, Y.; Chen, J.; Liang, Y.; Zhai, Y.; Ying, Z.; Zhou, W.; Genovese, A.; Piuri, V.; Scotti, F. Flexible and diverse contrastive learning for steel surface defect recognition with few labeled samples. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 2516014. [[CrossRef](#)]

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