Automatic Micro-Crack Detection of Polycrystalline Solar Cells in Industrial Scene

TAO FAN[1], TAO SUN[2], XIANGYING XIE[2], HU LIU[1], AND ZHIXIONG NA[ID][2]

[1]State Grid Corporation of China, Beijing 100031, China
[2]State Grid E-Commerce Company Ltd., Beijing 100053, China

Corresponding author: Zhixiong Na (hellonazx@163.com)

ABSTRACT Solar energy can be a clean and renewable alternative to traditional fuels, which enables its wide application in our life and the industry. However, some defects inevitably occur in the solar cells during production, transportation, and installation, which will reduce the power generation efficiency. In this paper, we propose a ResNet-based micro-crack detection method to detect the micro-cracks on polycrystalline solar cells. Specifically, a novel feature fusion model is introduced to aggregate the low-level features and deep semantically strong features by self-attention mechanism to obtain accurate geometry information. This method boosts the detection accuracy to 99.11%, which significantly surpasses other counterparts, e.g., some state-of-the-art deep neural networks, by a large margin. Since it is difficult for other methods to precisely detect other defect types apart from micro-cracks, we further propose a transfer learning method based on MK-MMD to guide the training process of defect detector with another pre-trained micro-crack detector. With the help of transfer learning, the accuracy of solar cell defect detection increases by 11.6%.

I. INTRODUCTION

In recent years, the practical application of new energy sources, such as solar energy has been attracted much attention. Photovoltaic power generation utilizes solar energy to generate electricity and has been widely used in all aspects of life [1]. The operation of photovoltaic power generation is inseparable from solar cell panels. However, solar cells may get various defects due to external forces or aging during transportation, installation, and utilization. This will weaken the power generation effect of the cells and even make the cells fail to generate electricity, reducing the entire efficiency of the photovoltaic power generation system. Therefore, it is necessary to check the battery board promptly to prevent the defect problems of solar cells.

Electroluminescence (EL) has been widely used in the detection of defects in photovoltaic system components, especially solar cells, due to the high-resolution images it can produce. The acquisition process of EL images is shown in Fig. 1. During this process, the solar cell emits infrared light due to being activated by the voltage. Then cooled
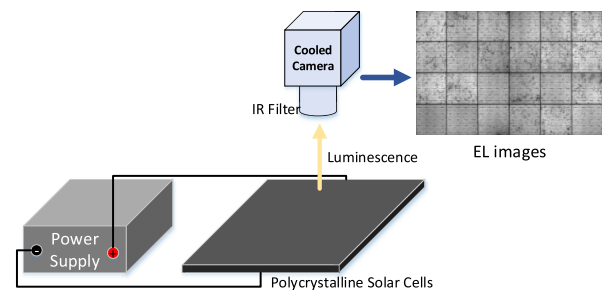
FIGURE 1. The process of obtaining EL images.

Si-CCD cameras capture these infrared lights and output sensed images. The defective parts of the cells cannot be activated normally, so they appear as black in the sensed images. Therefore, it is possible to determine whether the cells contain defects by detecting the EL images.

As shown in Fig. 2(a)(b), Solar cells can mainly be divided into two categories which are polycrystalline and monocrystalline. Among them, because of the presence of impurities in the polycrystalline solar cell, there will be dark spots in
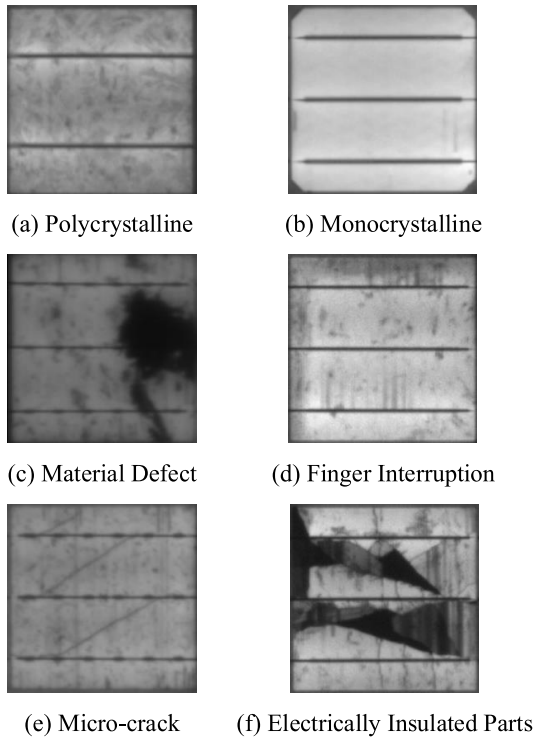
(a) Polycrystalline

(b) Monocrystalline

(c) Material Defect

(d) Finger Interruption

(e) Micro-crack

(f) Electrically Insulated Parts

FIGURE 2. The image examples of polycrystalline cells, monocrystalline cells and the possible defects of the cells.

the obtained EL image, which will affect defect detection. Therefore, the defect detection of the polycrystalline cell is more difficult compared with the monocrystalline cell.

As shown in Fig. 2, the defects of solar cells contain many types, including material defects and micro-cracks [2], [3]. Wherein material defects such as finger interruption generally will not affect the life span and efficiency of the battery. However, the micro-cracks will become more and more serious over time, thereby affecting the power generation effect of solar cells.

The defect detection process is mainly completed by manual identification before the era of machine learning, so that the detection efficiency is inevitably limited. Recently, many researchers begin to adopt machine learning methods for the detection of solar cell defects, such as SVM, AdaBoost and so on [3]. With the continuous development of deep learning and convolutional neural networks [3], more and more studies tend to use deep learning methods for defect detection of solar cells. However, most of these methods use conventional CNN networks to extract features from images for two-class classification. That is, only the feature outputs of networks' last layer are used for the classification of images. Considering that defect features such as micro-cracks are contained in the low-level features of the image, many useful features from the middle layers of the deep networks are ignored.

To solve this problem, we propose a ResNet-based method for detecting the micro-cracks of polycrystalline cells. In our method, we combine feature fusion module with ResNet50 backbone and use the fused features to perform

classification. We aim to make full use of both the low-level and high-level information in the network for classification. This method leads to significant improvements on the polycrystalline micro-crack dataset that we collected from industrial production lines. However, in industrial scenarios, the defects of polycrystalline cells are diverse. The effects of the detection methods we propose will also be affected when targeting other defect datasets. Considering this, we conduct some analyses between different defect datasets of polycrystalline solar cells. The source domain is polycrystalline micro-crack data, and the target domain is various polycrystalline defects. The obtained transfer method adapts the output features of the training models corresponding to the two domains, which makes the micro-crack detection model guide the training process of the polycrystalline defect detection task. To summarize, the main contributions of this work are as follows:

- We propose a micro-crack detection method for polycrystalline solar cells, including image preprocessing, feature extraction, feature fusion module, and image classification network.
- Referring to the mainstream method of transfer learning, we develop an algorithm that can be utilized when transferring from the micro-crack detection task of polycrystalline cells to the polycrystalline defect detection task.

Based on the collected polycrystalline EL images including micro-cracked images, we created a polycrystalline micro-crack dataset. Experiments conducted on this dataset show that the micro-crack detection method based on deep learning and feature fusion can obtain satisfactory detection results. After that, based on the micro-crack detection model obtained in the previous step, we conducted transfer learning experiments on the polycrystalline defect dataset obtained from the network. The results showed that the transfer learning method we propose can effectively improve the effect of the defect detection task.

II. RELATED WORKS
A. IMAGE CLASSIFICATION

The purpose of the image classification task is to assign a corresponding label to an image from a specific label set, that is, to identify the category of the image.

Traditional image classification methods include the process of feature extraction, classification, and so on. Through feature extraction, a large amount of local description information which is robust and not affected by noise such as light can be extracted. The main methods of feature extraction include Scale-Invariant Feature Transform (SIFT) [4], Speeded Up Robust Features (SURF) [5], Histogram of Oriented Gradient (HOG) [6], Local Binary Pattern (LBP) [7] and so on. Besides, some works propose a combination of multiple extraction methods used in the feature extraction process to prevent the loss of useful information. A vector with fixed-dimension can be obtained through feature encoding [8] after feature extraction and fed into the classifier for classification. The classifiers mainly used for image
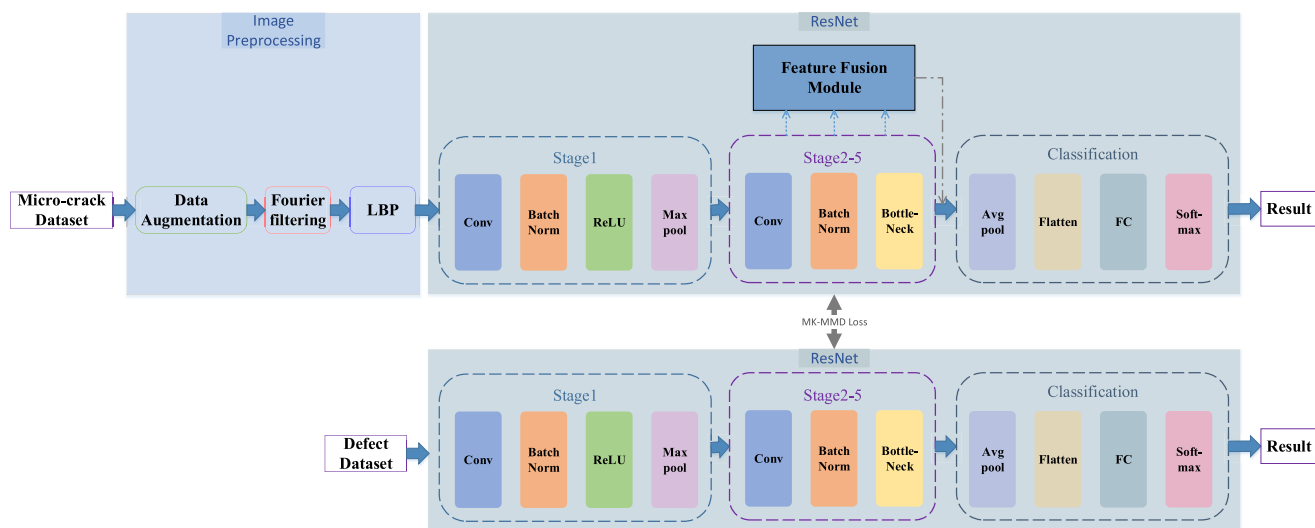
FIGURE 3. The pipeline of the micro-crack detection model for polycrystalline cells, which contains image preprocessing, feature fusion module, and ResNet50 architecture.

classification include K-Nearest Neighbor (KNN) [9], Support Vector Machine (SVM), Random Forest (RF) [10] and various ensemble methods.

With the development of deep learning, convolutional neural networks (CNN) are gradually applied to image classification. The pioneering deep network used for large-scale image classification which is named AlexNet starts the era of CNN. The CNN models generally used for classification include convolutional layers, fully connected layers, softmax multi-class classifiers, and multi-class cross-entropy loss functions, etc., which are continuously trained to achieve image classification effects.

B. DEFECT DETECTION
The methods for defects detection are generally categorized into Photoluminescence (PL) [11], [12] and Electroluminescence (EL). Since some defects in solar cells only display under EL imaging of photovoltaic modules, most current methods [13], [14] use EL for solar cells' defect detection. Table 1 summarizes relevant literature and methods.

In [15], a Fourier image reconstruction method is proposed to detect defects in EL images of solar cells. However, the shape assumption makes defects detection with different shapes more difficult. Therefore, [16] proposes to utilize anisotropic diffusion filtering and shape analysis to detect defects in solar cells. This method produces a better detection effect on the micro-cracks, but the detection performance on other defects is weaker. Besides, [17] proposes a detection method for finger interruption, which mainly uses binary clustering of candidate region features. In addition, methods using CNN for defect detection have gradually emerged in recent years. Reference [3] proposes a CNN-based network for defect classification, and compares it with prior traditional

TABLE 1. The summary of relevant defect detection methods for solar cells.

Method	Image Type	Defect Type
Fourier Image Reconstruction [15]	EL	Line- or Bar-shaped Defects
Anisotropic Diffusion Filtering Image Segmentation [16]		Micro-crack
Binary Clustering [17]		Finger Interruption
Feature Extraction Deep Learning [3]		All Defects

classification methods, proving the superior performance of CNN for defects detection.

C. TRANSFER LEARNING
Transfer learning (TL) refers to the transfer of knowledge learned from one domain (source domain) to another domain (target domain) to enable the model to achieve better results in the target domain. According to the specific domains and tasks in transfer learning, transfer learning can be divided into three types: Inductive, Transductive, and Unsupervised Transfer Learning [18]. Our proposed framework belongs to Inductive Transfer Learning, in which the target domain contains labels, and the ground truth labels can be used to train the target model while transfer learning.

The inductive transfer learning can also be categorized into Instance-based TL, Feature-representation-transfer, Parameter-transfer, and Relational-knowledge-transfer [18]. TrAdaBoost [19] belongs to Instance-based TL, which obtains the target model by adjusting the weight of the source domain label and target domain label, using

instance re-weighting and importance sampling. Feature-representation-transfer is a widely studied transfer learning method that can decrease classification and regression errors by reducing the difference between the source domain and the target domain. Fine-tuning commonly used in model training belongs to parameter-based transfer learning, which improves the detection level of the target model by sharing parameters between the source domain and target domain. In addition, Relational-knowledge-transfer transfers the similarity relations between the source and target domain.

III. METHODS

Our proposed framework, which is illustrated in Fig. 3, can be composed of two important parts. We first perform an effective micro-crack detection method for polycrystalline cells, which can classify the cells as two categories (whether micro-cracks exist) in Section III-A. In Section III-B, a training method of defect detection model of polycrystalline cells based on transfer learning is proposed.

A. MICRO-CRACK DETECTION

As a common type of defect in polycrystalline cells, micro-cracks will affect the power generation efficiency of the cell to a certain extent. Therefore, the detection of micro-cracks is critical in the practical application of polycrystalline cells. According to whether it contains cracks, we can classify polycrystalline cells into normal and micro-cracked types, which can be represented by 0 and 1, respectively. As a result, the micro-crack detection problem of polycrystalline cells can be converted into an image binary classification problem. In this part, we propose a deep learning-based polycrystalline micro-crack detection method, which uses a deep convolutional neural network as the backbone and uses the fused network features to perform binary classification operations.

1) DATA AUGMENTATION

In practical industrial situations, the cells with micro-cracks are only a minority of all cells. In our industrial polycrystal cell micro-crack dataset, the number of non-cracked cells is about 24 times larger than that of micro-cracked batteries. Considering that the micro-crack represents the local features of the image, the commonly used image processing methods such as cropping and affine transformation are not applicable here. Therefore, we choose the flip method for data augmentation. The whole flip process for each image contains 3 steps, which are:

- Flip Horizontal
- Flip Vertical
- Flip Horizontal + Flip Vertical

After these flips, we can get 4 images from the original one.

2) IMAGE PREPROCESSING
a: FOURIER FILTERING

As shown in Fig. 2(a), there are busbars in the polycrystalline cells which are similar to micro-cracks. Therefore, the

(a) Original Image

(b) Spectrogram

(c) Fourier Filtering

(d) Filtered Image
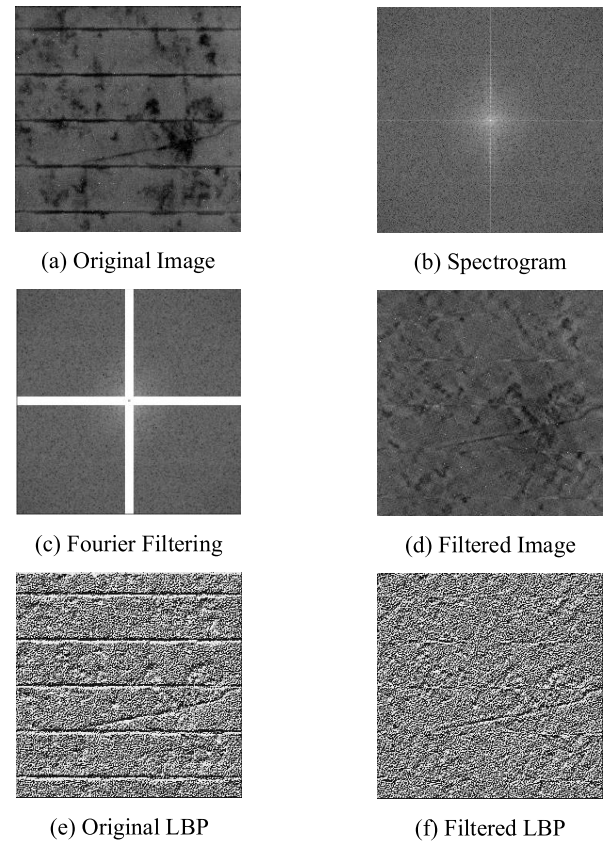

(e) Original LBP

(f) Filtered LBP

FIGURE 4. An example of image preprocessing.

busbars will cause a lot of noise interference to the detection of micro-cracks, which should be removed as much as possible before image classification.

Here we choose Fourier transform [20] to conduct image filtering to remove the busbars from the original images of polycrystalline cells as much as possible. For an $M \times N$ grayscale image $f(x, y)$, the Fourier transform formula for converting it from the time domain to frequency domain is shown in Eq. (1), and the Fourier inversion formula is shown in Eq. (2).

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)} \quad (1)$$

$$f(x, y) = \frac{1}{MN} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)} \quad (2)$$

As shown in Fig. 4, the entire Fourier filtering process can be mainly divided into three steps:

- Transforming the image Fig. 4(a) from the time domain to frequency domain to obtain the spectrogram Fig. 4(b).
- Performing the filtering operation in the frequency domain as shown in Fig. 4(c).
- Converting the filtered spectrogram back to the time domain to get the image Fig. 4(d).
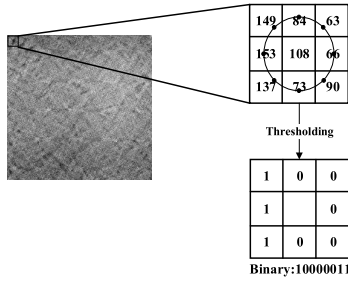
FIGURE 5. Example of LBP feature extraction process.

TABLE 2. The architecture of used ResNet50 (The pictures are resized to 224 × 224 × 3).

Layer Name	Layer Structure		Output Feature Size ($H \times W \times C$)
Stage1	conv, $7 \times 7, 64$, stride=2 maxpool,3×3, stride=2, padding=1		$112 \times 112 \times 64$
Stage2	conv,	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$, stride=1	$56 \times 56 \times 256$
Stage3	conv,	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$, stride=1	$28 \times 28 \times 512$
Stage4	conv,	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$, stride=1	$14 \times 14 \times 1024$
Stage5	conv,	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 6$, stride=1	$7 \times 7 \times 2048$
Classification	average pool 2-d fc		$1 \times 1 \times 2$

After trying various filters, we got the following filter:

$$H(u, v) = \begin{cases} 0 & (u, v) \in S(u, v) \\ 1 & (u, v) \notin S(u, v) \end{cases} \tag{3}$$

$$S(u, v) = \{u \in (m - 120, m - 2) \cup (m + 2, m + 125),$$
$$v \in (n - 4, n + 4)\}$$
$$\cup \{u \in (m - 4, m + 4),$$
$$v \in (n - 120, n - 2) \cup (n + 2, n + 125)\} \tag{4}$$

m and n represent half of the width and height of the image respectively.

This filter can keep other data of the polycrystalline images from being lost, especially information such as irregular lines like micro-cracks, under the premise of achieving the goal of removing the busbars in our observations. And we also find that the use of this Fourier filter can reduce the influence of dark spots inside the polycrystalline cells on the micro-crack detection while removing the busbars.

b: LOCAL BINARY PATTERN

LBP is the operator used to describe the local texture characteristics of an image. In the images after Fourier filtering, the components with obvious line shapes are basically only possible hidden cracks. Therefore, refer to related works [21]–[24], we extract the LBP feature of the image to highlight the texture. This enables the possible line-like micro-cracks to become more obvious, which is beneficial to the detection of micro-cracks.

The neighborhood radius of LBP for calculating is set to 1, and the number of pixels sampled is 8. The final sampling points selected by LBP operator are shown in Fig. 5. The corresponding images obtained after calculating the LBP feature are shown in Fig. 4(e) and Fig. 4(f).

3) BACKBONE FOR CLASSIFICATION

After preprocessing, the images are fed into a ResNet50 [25] network for feature extraction. The structure of the ResNet50 is shown in Fig. 3. Table 2 and Fig. 6 show more details of the network composition. Based on the original network, the ResNet50 used in this work changes the output number
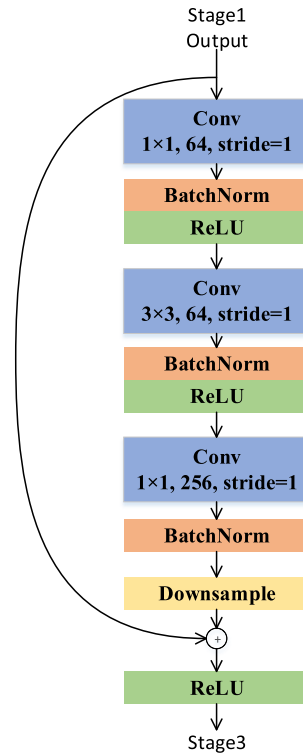

FIGURE 6. The details of Stage2 in ResNet50.

of the fully connected (FC) layer to 2, so for each image, a corresponding array $F = [F_0, F_1]$ can be output, in which there are two elements.
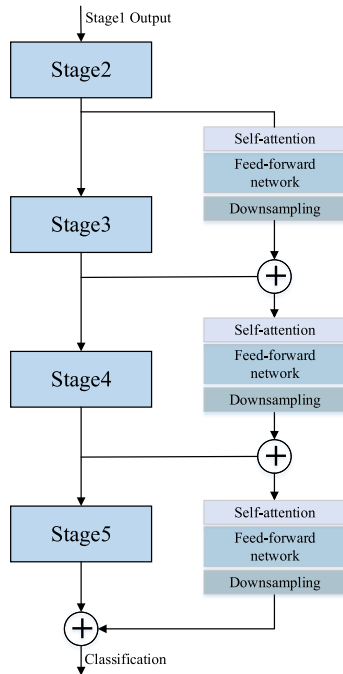
FIGURE 7. The detailed architecture of feature fusion module.

In addition, we add the softmax layer at the end of ResNet50, the softmax calculation formula is

$$S_i = f_{softmax}(F_i) = \frac{e^{F_i}}{e^{F_0} + e^{F_1}}, \qquad (5)$$

where i can be 0 and 1, S_0 and S_1 respectively represent the possibility that this image belongs to category 0 and category 1.

4) FEATURE FUSION MODULE

It is intuitive that the stripes in the polycrystalline solar cell image contain rich discriminative information that is beneficial for defects detection. The features of the deepest layer in the network are semantically strong but lacks precise low-level information. Low-level features are usually deemed rich in containing geometric patterns and stripes, which is contrary to high-level features. Inspired by this, we propose a feature fusion module to fuse as much as possible low-level information to enhance the defects detection ability of the network.

Unlike some previous works integrate multi-scale features by convolutions, we propose to facilitate multi-scale feature fusion by attention mechanism in this paper. The detailed architecture of feature fusion module is illustrated in Fig. 7. We first resize the features of the prior stage by downsampling operation, which is a convolution layer with kernel size of 3 and stride of 2 in our implementation. Then we adopt attention mechanism to fuse the downsampled low-level feature and features at the higher level. We model the relationships among all spatial positions in the features

to get precise and accurate low-level information by a self-attention layer. To be consistent with the transformer architecture, we insert another feed-forward network into the bottom of the self-attention layer to enhance the feature semantics. A residual connection is added in the self-attention layer and feed-forward network to ease optimization. We insert three feature fusion modules before Stage 3, Stage 4 and Stage 5 to construct a bottom-up pathway which is parallel to the backbone.

5) CROSS ENTROPY LOSS FUNCTION

Since this is a relatively simple binary classification task, we choose a typical Cross-Entropy (CE) loss function as the calculation method of the loss value here. The formula of the loss function is

$$L_{CE} = \sum -(y_{true} \log(S_1) + (1 - y_{true}) \log(S_0)), \qquad (6)$$

where y_{true} represents the ground truth label of this image, which is likely to be 1 or 0.

B. DEFECT DETECTION BASED ON TRANSFER LEARNING

In industrial scenarios, the specific conditions and possible defects of different batches of cells are different. In the previous section, we propose the training method of a micro-crack detection model for the collected polycrystalline cell dataset from the industrial scene. In this batch of cells, the defects of the battery are mainly micro-cracks, so the data used in training is relatively regular, only containing two types: without and with micro-cracks. As a result, without using other auxiliary designs, the convergence effect of the model has been well.

However, when the types of defects contained in the dataset increase, or the dataset has less data, the model obtained by only using the method proposed in the above section will have a poor detection effect on defect detection and cannot be truly applied to actual industrial scenarios. Therefore, in this section, we propose a defect detection model training method based on transfer learning. With the help of the micro-crack detection model obtained in the previous part, knowledge transfer is carried out through transfer learning, so that the effect of detection can be improved.

1) FINE-TUNE

As the most basic transfer learning method, fine-tuning has been widely used in the training process of deep learning models. The most commonly used pre-trained model is the one trained on ImageNet [26]. But for the defect dataset, the images contained in ImageNet have a certain gap with those in it. So our intuition is to consider using the model obtained on a more similar dataset as a pre-trained model to get better fine-tune effects.

In the previous part, we have obtained a detection model for polycrystalline cell micro-crack based on the industrial dataset. Therefore, we consider using the model obtained previously as the pre-trained model, only to replace the old
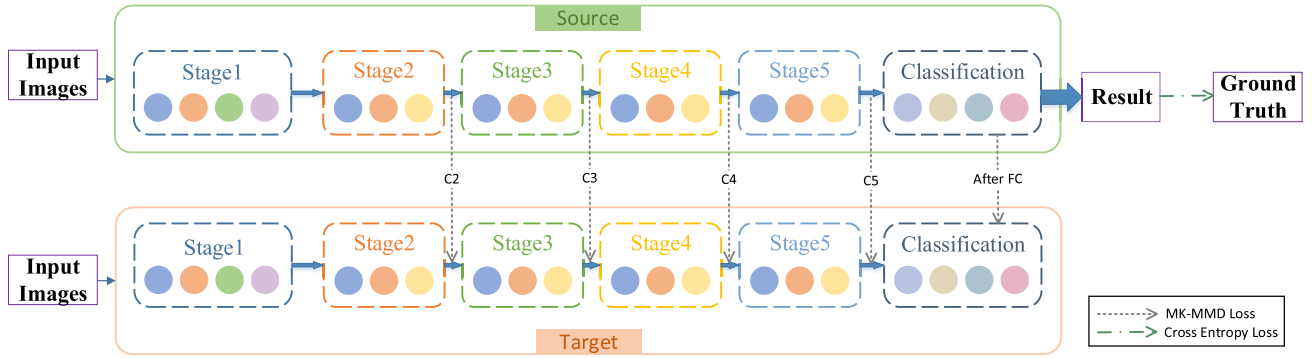
FIGURE 8. The pipeline of the proposed transfer learning method. The picked layers for loss calculation contain the output layers of every stages and the FC layer.

FC layer of the micro-crack detection model with a new one. Experiments have proved that using this type of model as the pre-trained model can get better results than using those trained on ImageNet.

2) MODEL DETAILS

Referring to the previous work related to transfer learning, we use a transfer learning model to obtain a better defect detection model by learning the output of the trained micro-crack detection model. Here, $\mathcal{D}_S = \{(x_i^S, y_i^S)\}_{i=1}^{n_S}$ denotes the source domain, where n_S is the total amount of source domain data got from the micro-crack dataset. Besides, the target domain is obtained from the defect dataset, notated as $\mathcal{D}_T = \{(x_i^T, y_i^T)\}_{i=1}^{n_T}$, where n_T is the number of target domain data. P and Q represent the probability distribution of the source domain and target domain respectively.

a: DAN AND MK-MMD

In the representative work [27] of transfer learning, Deep domain confusion (DDC) is proposed, which uses the results obtained from the 7th layer of AlexNet to calculate the Maximum Mean Discrepancy (MMD) distance, and optimizes this value to reduce the distance between the source domain and the target domain. Deep Adaptation Network (DAN) [28] further improves based on DDC, adapts to the multiple layers of the deep convolutional network. By reducing the multiple kernel variant of MMD (MK-MMD) distance between the source domain and target domain among more layers, DAN can obtain better transfer results than DDC.

MMD is widely used in transfer learning tasks to measure the distance between the source and target domains mapped to the reproducing kernel Hilbert space (RKHS) space. Let F represent a function set in the sample space, MMD can be calculated as

$$MMD[\mathcal{F}, p, q] = [\frac{1}{n_S(n_S-1)} \sum_{i \neq j}^{n_S} k(x_i^S, x_j^S)$$

$$- \frac{1}{n_S n_T} \sum_{i,j}^{n_S, n_T} k(x_i^S, x_j^T)$$

$$+ \frac{1}{n_T(n_T-1)} \sum_{i \neq j}^{n_T} k(x_i^T, x_j^T)]^{\frac{1}{2}}, \quad (7)$$

where k is the kernel function. MK-MMD is the multiple kernel variant of MMD, so when calculating MK-MMD, we only need to change the kernel function in the above MMD formula to a convex combination of multiple kernel functions, which means

$$k(x_1, x_2) = \sum_{m=1}^{n_k} \beta_m k_m(x_1, x_2), \quad s.t. \sum_{m=1}^{n_k} \beta_m = 1, \quad \beta_m \geq 0. \quad (8)$$

In the above equation, $k_m(\cdot, \cdot)$ denotes the kernel function selected for calculation, n_k represents the number of kernel functions, and β_m is the corresponding weight for each kernel function added up.

b: TRANSFER LEARNING TRAINING METHOD

As shown in Fig. 8, the training method of transfer learning includes the optimization of MK-MMD distance between the source domain and the target domain and the distance between prediction results and the ground truth labels in the target domain.

The initial parameters of the target network are obtained by those of source network trained in the source domain. During the training process, MK-MMD distances between the source domain and the target domain are calculated by the outputs of the bottlenecks and the final FC layer in ResNet50. There are 5 layers picked for distance calculation in total. Let $D_S^l = \{x_{l,i}^S\}_{i=1}^{n_S}$ and $D_T^l = \{x_{l,i}^T\}_{i=1}^{n_T}$ respectively represent the target domain and source domain of the lth picked layer. The corresponding distributions are respectively p_l and q_l. Then MK-MMD distance of layer l is

$$d_{mmd}^l(D_S^l, D_T^l) = MMD[\mathcal{F}, p_l, q_l]. \quad (9)$$

TABLE 3. Details of used datasets.

Dataset	w/ defect (1)	w/o defect (0)	sum
Micro-crack Dataset	392	9402	9794
Defect Dataset	402	920	1322

Therefore, according to the MK-MMD distance calculation method, the loss value corresponding to all layers can be obtained as

$$L_{mmd} = \sum_{l=1}^{L} d_{mmd}^l(D_S^l, D_T^l), \quad (10)$$

where L represents the total number of the picked layers, which is 5 in this work. The overall loss of the training process is

$$L_{TF} = L_{CE} + \alpha L_{mmd}, \quad (11)$$

where L_{CE} represents Cross Entropy loss function for binary classification in the target domain, considering the prediction results can be optimized according to the ground truth labels. Besides, α is hyperparameters, which means the weight of L_{mmd} in the fina loss, being set to 0.25 normally.

IV. EXPERIMENTS

In order to prove the effectiveness of the micro-crack detection model and the transfer learning model, we conduct a series of experiments on the collected polycrystalline micro-crack dataset and the polycrystalline defect dataset obtained from the network. The following show the basic configurations of the experiments and the corresponding results.

A. DATASET

The dataset prepared for micro-crack detection contains the collected polycrystalline images from industrial assembly lines. These images are segmented from the EL images of the entire polycrystalline cell panels, which are all grayscale images. The size of each image is 240×240. In these cell images, there may be micro-crack defects. Therefore, according to whether the images contain micro-cracks, we divide them into two types: with micro-cracks (micro-cracked) and without micro-cracks (normal). The classification result of this dataset is shown in Table 3. An overview of the dataset is shown in Fig. 9. In addition, it can be seen that the number of cell images without micro-cracks is much larger than that of cells with micro-cracks, as a result, we should enrich the data of images with micro-cracks before training.

Also, we find a solar cell defect dataset from the Internet [29] and select the polycrystalline cell part of it to perform transfer learning experiments. The quantity of the dataset is shown in Table 3.

B. EXPERIMENTAL SETTING

Except for the model based on transfer learning, the deep networks of other detection models are all pretrained on ImageNet. The main deep convolutional neural network used

TABLE 4. Confusion matrix.

		Actual Class	
		w/ defect (1)	w/o defect (0)
Predicted Class	w/ defect (1)	TP	FP
	w/o defect (0)	FN	TN

is ResNet50, and the number of output channels of the FC layer is changed to 2 to better suit the binary classification. In all experiments in this work, we use mini-batch stochastic gradient descent (SGD) with a learning rate of 0.005 and momentum of 0.9 for training.

C. EVALUATION METRICS

The detection models in this work are essentially classification models for polycrystalline cells, so here we introduce the confusion matrix, which is shown in Table 4, as the calculation basis for the model evaluation metrics. From Table 4, True Positive (TP) represents the predicted result and the actual class are both 1, which refers to cells with defects (micro-cracks as for micro-crack dataset). The concepts of True Negative (TN), False Positive (FP), and False Negative (FN) are similar.

With the help of the confusion matrix, we can get the calculation methods of accuracy (ACC), precision, and recall, which are

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}, \quad (12)$$

$$Precision = \frac{TP}{TP + FP}, \quad (13)$$

$$Recall = \frac{TP}{TP + FN}, \quad (14)$$

The values of precision and recall in the results are hoped to be as large as possible. However, because of the interaction between these two values, we need to make some sacrifices here. To find the balance point, we introduce F1 score [30], which can consider both precision and recall. When it increases, both precision and recall will increase. The formula for calculating the F1 score is

$$F1 = \frac{1}{\frac{1}{Precision} + \frac{1}{Recall}} = \frac{2TP}{2TP + FP + FN}. \quad (15)$$

Since this work is concerned with binary classification, we further introduce the evaluation metric Area Under Curve (AUC), which means the area under the receiver operating characteristic (ROC) curve. In the ROC curve, the abscissa is False Positive Rate (FPR), and the ordinate is True Positive Rate (TPR). FPR and TPR can be calculated as

$$FPR = \frac{FP}{FP + TN}, \quad (16)$$

$$TPR = \frac{TP}{TP + FN}, \quad (17)$$

Considering the characteristics of the binary classification problem and the imbalance problem between two classes,
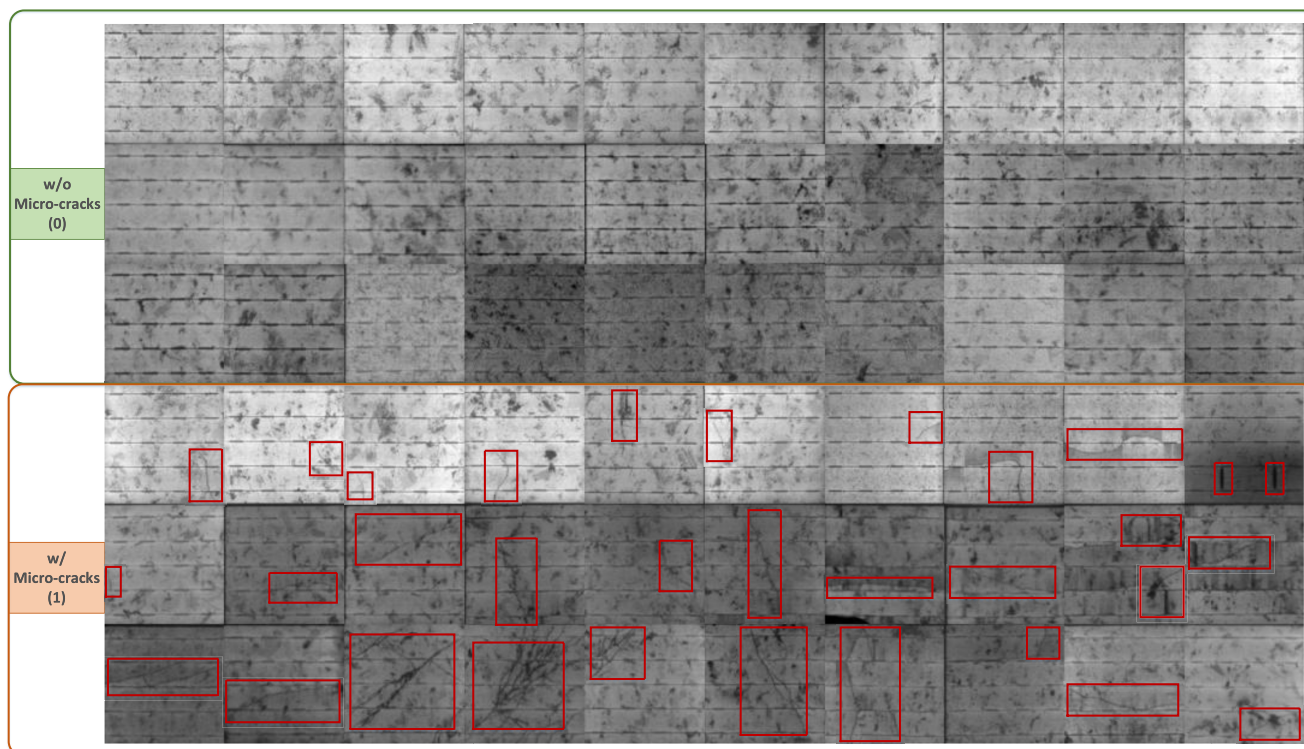
FIGURE 9. The overview of the used dataset. The upper part of the figure shows the polycrystalline cells without defects, while the lower part contains the micro-cracked cells (the micro-cracks are marked with the red frames).

TABLE 5. Splitting strategy for micro-crack dataset.

	train	valid	sum
w/ micro-cracks (1)	1304	264	1568
w/o micro-cracks (0)	8436	966	9402
sum	9740	1230	10970

TABLE 6. Performance of proposed micro-crack detection model with and without image preprocessing (with data augmentation).

	ACC	Precision	Recall	F1	AUC
w/o preprocessing	70.57%	25.00%	18.56%	0.2130	0.5167
w/ Fourier Filtering	78.05%	33.33%	2.27%	0.0426	0.5052
w/ LBP	97.24%	100.00%	87.12%	0.9312	0.9356
w/ Fourier Filtering+LBP	**98.29%**	**100.00%**	**92.05%**	**0.9586**	**0.9602**

we finally select ACC, Precision, Recall, F1 score, and AUC as the evaluation metrics in the experiments of this work.

D. EXPERIMENTAL RESULTS OF MICRO-CRACK DETECTION

Experiments are carried out on the industrial micro-crack dataset to prove the importance of image preprocessing in the proposed method and the effectiveness of this method. All experiments in this part are trained for 10 epochs. This micro-crack dataset is divided into two parts, train and valid, after data augmentation, as shown in Table 5. The model is trained on train and tested on valid.

1) INFLUENCE OF IMAGE PREPROCESSING

In this section, our goal is to prove that image preprocessing helps to improve the final detection results. We conduct a series of experiments on the basic ResNet50 without feature fusion.

Table 6 shows the gap between the detection effects obtained with and without image preprocessing. It can be seen from Table 6 that the addition of image preprocessing can improve the detection of the model under various evaluation metrics, respectively increasing ACC, F1 score, and AUC by 27.72%, 0.7456, and 0.4435. Also, the two main subcomponents of image preprocessing Fourier Filtering and LBP extraction are helpful to improve the effectiveness of the model more or less.

In addition, as shown in Table 7, we also try other filtering methods and edge feature extraction methods to find the best options for image preprocessing. From the evaluation results in Table 7 and Fig. 10 which shows the filtered images, we can find that Fourier filtering has the best effect, outperforming other methods including Mean Blur, Median Blur, and Gaussian Blur. Similarly, after comparing the results of different edge feature extraction methods shown in Table 7 and Fig. 10, we find using the LBP feature for detection is the best choice,

TABLE 7. Performance of proposed micro-crack detection model with different image preprocessing methods.

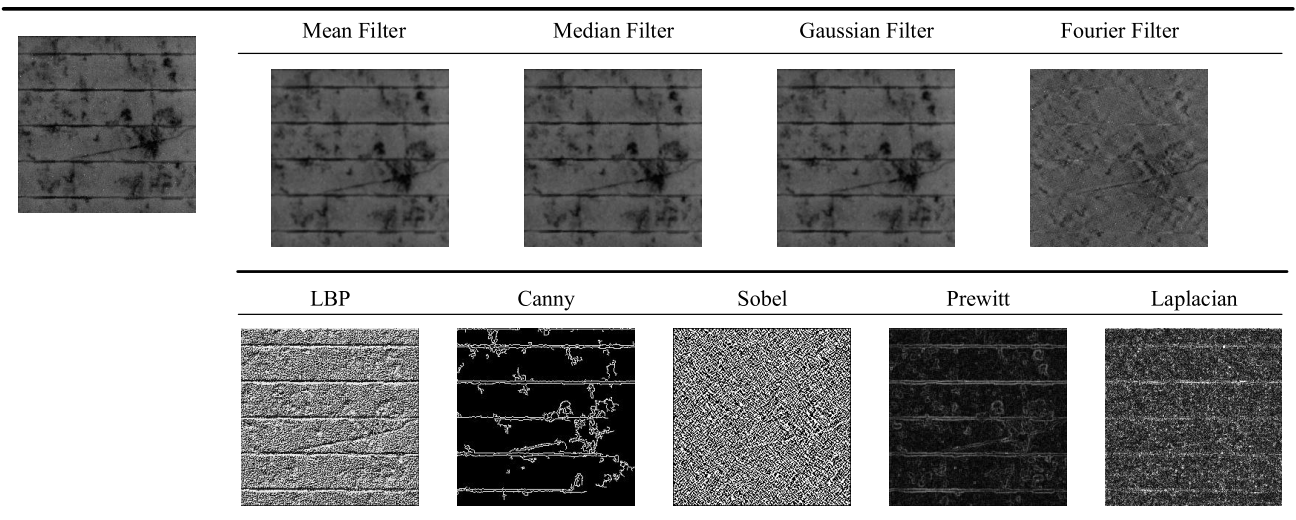|	Filter				Edge Feature Extraction					Metrics				
	Mean	Median	Gaussian	Fourier	LBP	Canny	Sobel	Prewitt	Laplacian	ACC	Precision	Recall	F1	AUC
w/ Filter	√									76.26%	30.56%	8.33%	0.1310	0.5158
		√								71.30%	26.20%	18.56%	0.2173	0.5214
			√							75.37%	34.65%	16.67%	0.2251	0.5404
				√						**78.05%**	**33.33%**	**2.27%**	**0.0511**	**0.5117**
w/ Edge Extraction					√					**97.24%**	**100.00%**	**87.12%**	**0.9312**	**0.9356**
						√				74.31%	21.74%	7.58%	0.1124	0.5006
							√			95.69%	100.00%	79.92%	0.8884	0.8996
								√		78.86%	100.00%	1.52%	0.0299	0.5076
									√	94.96%	94.96%	100.00%	0.8670	0.8826
w/o preprocessing										70.57%	25.00%	18.56%	0.2130	0.5167
w/ preprocessing				√	√					**98.29%**	**100.00%**	**92.05%**	**0.9586**	**0.9602**

FIGURE 10. The image processing results of different filters and edge feature extraction methods.

TABLE 8. Performance of proposed micro-crack detection model with and without data augmentation (with image preprocessing).

	ACC	Precision	Recall	F1	AUC
w/o data augmentation	93.74%	81.48%	91.67%	0.8627	0.9299
w/ data augmentation	**98.29%**	**100.00%**	**92.05%**	**0.9586**	**0.9602**

increasing ACC, F1 score, and AUC by 26.67%, 0.7182, and 0.4189. Compared with other extracted features, it is obvious that LBP is more likely to highlight the line-like micro-cracks in polycrystalline cells.

2) INFLUENCE OF DATA AUGMENTATION
It can be seen from Table 3 that the used dataset is unbalanced, so we propose to use image flipping for data augmentation in Section III-A1. In the previous experiments above, the dataset we use is one after data augmentation. In this section, we conduct related experiments based on image preprocessing, which proves that data augmentation is necessary

and can greatly improve the effectiveness of the proposed micro-crack detection model. The results of the experiments are shown in Table 8, which shows that data augmentation can increase ACC, F1 score, and AUC by 4.55%, 0.0959, and 0.0303 respectively.

3) INFLUENCE OF FEATURE FUSION
The above experiments are all implemented based on the basic ResNet50. In this section, we add feature fusion to the original network to make full use of the shallow features of the network, and conduct related experiments to verify the effectiveness of feature fusion on the detection results.

The results of the experiments about feature fusion are shown in Table 9. The first column of the table represents the features used in the following classification, and the numbers in it represent the output features of the corresponding stages. It can be seen from Table 9 that using the fusion features obtained by fusing the outputs of the four stages for micro-crack detection can get a better detection effect.

TABLE 9. Performance of proposed micro-crack detection model with and without data augmentation (with image preprocessing).

Classification-used Feature	ACC	Precision	Recall	F1	AUC
5(w/o feature fusion)	98.29%	100.00%	92.05%	0.9586	0.9602
2+5	98.53%	97.67%	95.45%	0.9655	0.9742
3+5	98.94%	98.83%	96.21%	0.9750	0.9795
4+5	98.37%	98.80%	93.56%	0.9611	0.9663
2+3+5	98.37%	99.19%	93.18%	0.9609	0.9649
2+4+5	99.02%	98.84%	96.59%	0.9770	0.9814
3+4+5	98.46%	99.60%	93.18%	0.9628	0.9654
2+3+4+5	**99.11%**	**98.84%**	**96.97%**	**0.9790**	**0.9833**

After adding the feature fusion module, ACC, F1 score, and AUC can be increased by 0.82%, 0.0204, and 0.0231 on the basis of using the basic ResNet50 for detection.

4) COMPARISONS WITH OTHER CLASSIFIERS

Keeping the same image preprocessing method, we try other commonly used machine learning methods to detect micro-cracks of polycrystalline cells. For each method, we try to get the configurations that gave the best detection results. The methods used and the corresponding parameters are as follows:

- KNN: We attempt to use one of the simplest classification algorithms, the k-Nearest Neighbors (KNN) algorithm, to classify polycrystalline images, keeping the number of neighbors as 5 and the leaf size as 30.
- GaussianNB: Gaussian Naive Bayes (GaussianNB), which belongs to Naive Bayes algorithms, is also attempted.
- SVM: For the Support Vector Machine (SVM) algorithm, we set the penalty coefficient C to 3, and keep the others as default.
- DT: As a traditional classification method, Decision Tree (DT) is tried here. A classification and regression tree (CART) is built to classify the cells.
- RF: Combining Decision Tree with ensemble learning, we can get Random Forest (RF). Here, we set the number of decision trees in the forest to 100.
- GBDT: We set the number of estimators to 100, learning rate to 0.08, max depth to 9 when using Gradient Boosting Decision Tree (GBDT).
- AdaBoost: As a boosting method, AdaBoost combines the weak classifiers to get a strong classifier. The weak classifier we use is Decision Tree, the number of iterations is set to 50.
- CatBoost: We also try another boosting method proposed in recent years named CatBoost. According to experience, we let the loss function type be Logloss, learning rate be 0.01, iterations be 100000, and depth be 5.
- MLP: To build Multi-Layer Perceptron (MLP), also called Artificial Neural Network, we set two hidden layers which have 100 and 50 neurons respectively.

TABLE 10. Results of detection using different methods.

	ACC	Precision	Recall	F1	AUC
KNN	96.18%	100.00%	82.20%	0.9022	0.9435
GaussianNB	93.41%	100.00%	69.32%	0.8188	0.9371
SVM	93.58%	100.00%	70.08%	0.8241	0.9650
DT	90.33%	93.98%	58.71%	0.7226	0.7884
RF	83.25%	100.00%	21,97%	0.3602	0.9500
GBDT	92.28%	99.42%	64.39%	0.7816	0.9695
AdaBoost	95.37%	97.70%	80.30%	0.8815	0.9676
Catboost	96.10%	100.00%	81.82%	0.9000	0.9796
MLP	97.64%	100.00%	89.02%	0.9419	0.9574
VGG-16	94.07%	94.07%	95.05%	0.8240	0.8631
GoogleNet	95.93%	99.54%	81.44%	0.8953	0.9067
ViT-small	81.06%	100.00%	11,74%	0.2102	0.5587
ViT-base	91.86%	89.05%	70.83%	0.7890	0.8423
MLP-Mixer-base	84.23%	100.00%	26.52%	0.4192	0.6326
Ours	**99.11%**	**98.84%**	**96.97%**	**0.9790**	**0.9833**

When training the network, the learning rate is 0.005, the optimizer is SGD.
- VGG-16/GoogleNet: The VGG-16 [31] and GoogleNet [32] we use for detection are the standard networks from torchvision, while the category number of classifiers is changed to 2. The training optimizer is SGD with a learning rate of 0.005.
- ViT-small/ViT-base/MLP-Mixer-base: The patch size of used ViT [33] and MLP-Mixer [34] is 16 × 16. The depth of these three models is 12. The used optimization methods are consistent with the original article.

Table 10 shows the classification effect of the above methods. It can be seen that the MLP can get the best result among the proposed learning methods, which is a simple feedforward neural network containing two hidden layers. However, compared with our proposed ResNet-based micro-crack classifier, the traditional classification methods like MLP still have certain gaps considering that deep neural network ResNet has stronger feature extraction and recognition ability.

E. EXPERIMENTAL RESULTS OF TRANSFER LEARNING

Further experiments on the effectiveness of the transfer learning method are conducted. The source domain of these experiments is composed of the industrial micro-crack dataset, and the target domain is the defect dataset. Similar to the previous part, we divided the data into train set and valid set after aligning these two datasets. The optimizer in experiments of this part is SGD with a learning rate of 0.005. The results are recorded after the models are trained for 100 epochs.

Table 11 shows the results of transfer learning, including the detection effects of models obtained by fine-tune and MK-MMD-based transfer learning method. It can be found that compared with the ResNet-based classification model directly training on defect dataset, the evaluation metrics of the model obtained using fine-tune have some increase. Besides, since the image type of both the micro-crack dataset

TABLE 11. Defect detection results of methods with and without transfer learning on defect dataset.

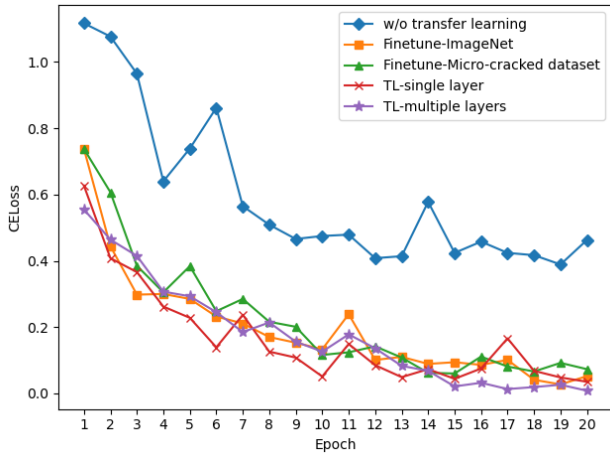|	ACC	Precision	Recall	F1	AUC
w/o transfer learning	72.00%	54.55%	40.00%	0.4615	0.6286
FINE-TUNE					
ImageNet	77.60%	69.39%	48.57%	0.5484	0.6838
micro-crack dataset	79.20%	73.47%	51.43%	0.5806	0.7029
TRANSFER LEARNING WITH MK-MMD					
single layer	81.60%	87.18%	45.33%	0.5965	0.7124
multiple layers	**83.60%**	**94.74%**	**48.00%**	**0.6372**	**0.7343**

FIGURE 11. The loss graph of the transfer learning experiments.

and defect dataset is polycrystalline cells, the parameters of the model based on the micro-crack dataset are more suitable for defect detection. As a result, the effect of fine-tune with the help of the micro-crack dataset is better than that with the aid of ImageNet, with ACC, F1 score, and AUC increase of 7.20%, 0.1191, and 0.0838.

In addition, we conduct single-layer and multi-layer MK-MMD related experiments. The results in Table 11 report that compared to only performing the optimization of MK-MMD distance after the FC layer output of the source domain and the target domain, optimizing the output of multiple layers can better reduce the distance between the source and target domains, to improve the model training effect of the target domain.

Fig. 11 shows the changes of Cross-Entropy losses between predictions and true labels in the first 20 epochs of the experiments mention in Table 11. From the change trends of losses in 20 epochs, we can tell that transfer learning plays a great role in promoting the convergence of the loss.

To find the most suitable hyper-parameter α in (11), we conduct a series of experiments about α based on multi-level transfer learning. As shown in Table 12, when α is set to 0.25, ACC, F1 score, and AUC can achieve the best results.

As shown Fig. 12 and Fig. 13, we further carry out visualization experiments. The heat maps in the figures are calculated from the output features of the last convolutional

TABLE 12. Defect detection results of methods with different hyper-parameter α.

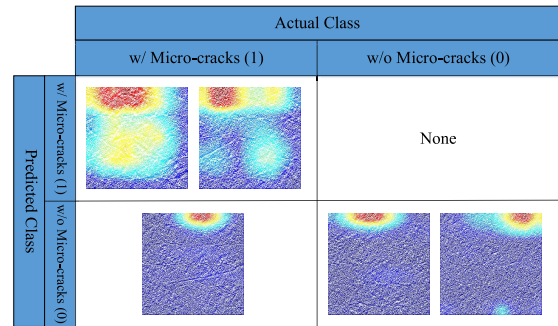| α	ACC	Precision	Recall	F1	AUC
0.1	81.60%	80.85%	50.67%	0.6230	0.7276
0.25	**83.60%**	**94.74%**	**48.00%**	**0.6372**	**0.7343**
0.5	83.20%	92.31%	48.00%	0.6315	0.7314
0.75	82.80%	90.00%	48.00%	0.6000	0.7286
1	82.00%	87.50%	46.67%	0.6097	0.7190

FIGURE 12. The heat maps of the features obtained from the micro-crack detection model. For example, the images in the left upper quadrant of the figure is classified as class 1, whose ground truth label is also 1.
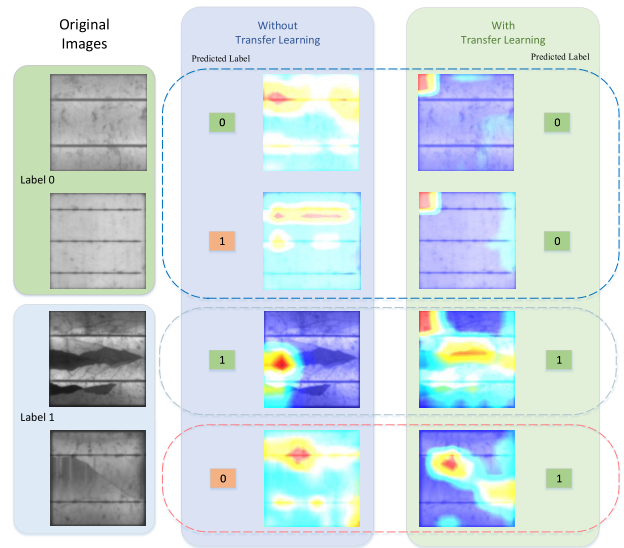

FIGURE 13. The heat maps of the features obtained from the defect detection model trained with and without transfer learning. Their predicted results are shown in the nearby boxes. The green ones mean the corresponding images are correctly classified, and the red ones mean the predicted results are wrong.

layer in the network. The features of images got from ResNet are activated in different regions, and the redder parts in the heat map have the higher activation degree.

It can be seen in Fig. 12 that the pictures with the large activated areas are more likely to be classified as 1. On the contrary, the picture with less activated parts is classified as 0. Besides, there is a trend that the heat maps of class 0 pictures are similar.

Fig. 13 clearly shows that transfer learning improves the feature activation effect of the network, to improve the detection effect. With transfer learning, the activated areas of the heat maps obtained by features of class 0 images become smaller and more similar to each other to a certain extent. And the activated areas of images that should be classified as class 1 become larger, of which the positions are closer to the defects.

V. CONCLUSION

This paper proposes a ResNet-based micro-crack detection method for polycrystalline solar cells. This method contains image preprocessing and a backbone network. The image preprocessing includes data augmentation, Fourier Filtering, and LBP extraction. After image preprocessing, the busbars of the polycrystalline cells can be removed, and the micro-cracked part can be enhanced. The backbone network is ResNet50, which is used for feature extraction and final classification operations. For the industrial micro-crack dataset, this method can achieve a detection accuracy of 98.29%.

In industrial scenarios, in addition to hidden cracks, polycrystalline solar cells have other defect forms, and training defect detection networks from scratch often fails to achieve good results. We propose a transfer learning method based on the micro-crack detection network, using fine-tune and the obtained micro-crack detection model to guide the defect detection. This method can significantly improve the effect of defect detection.

As for future work, we propose to focus on defect detection of polycrystalline cells based on one-shot learning and online learning, which aims to solve the difficulties of initial dataset acquisition. Besides, we observe the performance of current deep networks can be further improved by some tricks, e.g., knowledge distillation.

REFERENCES

[1] S. Eftekharnejad, V. Vittal, G. T. Heydt, B. Keel, and J. Loehr, "Impact of increased penetration of photovoltaic generation on power systems," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 893–901, May 2013.

[2] M. Demant, S. Rein, J. Krisch, S. Schoenfelder, C. Fischer, S. Bartsch, and R. Preu, "Detection and analysis of micro-cracks in multi-crystalline silicon wafers during solar cell production," in *Proc. 37th IEEE Photovoltaic Spec. Conf.*, Jun. 2011, pp. 001641–001646.

[3] S. Deitsch, V. Christlein, S. Berger, C. Buerhop-Lutz, A. Maier, F. Gallwitz, and C. Riess, "Automatic classification of defective photovoltaic module cells in electroluminescence images," *Sol. Energy*, vol. 185, pp. 455–468, Jun. 2019.

[4] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004, doi: 10.1023/B:VISI.0000029664.99615.94.

[5] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Speeded-up robust features (SURF)," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, Jun. 2008, doi: 10.1016/j.cviu.2007.09.014.

[6] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, San Diego, CA, USA, Jun. 2005, pp. 886–893, doi: 10.1109/CVPR.2005.177.

[7] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002, doi: 10.1109/TPAMI.2002.1017623.

[8] H. Jégou, F. Perronnin, M. Douze, J. Sánchez, P. Pérez, and C. Schmid, "Aggregating local image descriptors into compact codes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 9, pp. 1704–1716, Sep. 2012, doi: 10.1109/TPAMI.2011.235.

[9] D. W. Aha, D. F. Kibler, and M. K. Albert, *Instance-Based Learning Algorithms*. Norwell, MA, USA: Kluwer, 1991, pp. 37–66.

[10] L. Breiman, "Random forest," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.

[11] T. Trupke, R. A. Bardos, M. C. Schubert, and W. Warta, "Photoluminescence imaging of silicon wafers," *Appl. Phys. Lett.*, vol. 89, no. 4, p. 1854, 2006.

[12] S. Johnston, H. Guthrey, F. Yan, K. Zaunbrecher, M. Al-Jassim, P. Rakotoniaina, and M. Kaes, "Correlating multicrystalline silicon defect types using photoluminescence, defect-band emission, and lock-in thermography imaging techniques," *IEEE J. Photovolt.*, vol. 4, no. 1, pp. 348–354, Jan. 2014.

[13] J. Balzategui, L. Eciolaza, N. Arana-Arexolaleiba, J. Altube, J.-P. Aguerre, I. Legarda-Ereno, and A. Apraiz, "Semi-automatic quality inspection of solar cell based on convolutional neural networks," in *Proc. 24th IEEE Int. Conf. Emerg. Technol. Factory Autom. (ETFA)*, Zaragoza, Spain, Sep. 2019, pp. 529–535, doi: 10.1109/ETFA.2019.8869359.

[14] S. Spataru, P. Hacke, and D. Sera, "Automatic detection and evaluation of solar cell micro-cracks in electroluminescence images using matched filters," in *Proc. IEEE 43rd Photovoltaic Spec. Conf. (PVSC)*, Jun. 2016, pp. 1–6.

[15] D.-M. Tsai, S.-C. Wu, and W.-C. Li, "Defect detection of solar cells in electroluminescence images using Fourier image reconstruction," *Sol. Energy Mater. Sol. Cells*, vol. 99, pp. 250–262, Apr. 2012.

[16] S. A. Anwar and M. Z. Abdullah, "Micro-crack detection of multicrystalline solar cells featuring an improved anisotropic diffusion filter and image segmentation technique," *EURASIP J. Image Video Process.*, vol. 2014, no. 1, p. 15, 2014, doi: 10.1186/1687-5281-2014-15.

[17] D. C. Tseng, Y. S. Liu, and C. M. Chou, "Automatic finger interruption detection in electroluminescence images of multicrystalline solar cells," *Math. Problems Eng.*, vol. 2015, Oct. 2015, Art. no. 879675.

[18] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[19] W. Dai, Q. Yang, G. Xue, and Y. Yu, "Boosting for transfer learning," in *Proc. 24th Int. Conf. (ICML)* (ACM International Conference Proceeding Series), vol. 227, Z. Ghahramani, Ed. Corvallis, OR, USA: ACM Press, Jun. 2007, pp. 193–200, doi: 10.1145/1273496.1273521.

[20] K. Yildiz, A. Buldu, M. Demetgul, and Z. Yildiz, "A novel thermal-based fabric defect detection technique," *J. Textile Inst.*, vol. 106, no. 3, pp. 275–283, Mar. 2015.

[21] K. Yildiz, "Identification of wool and mohair fibres with texture feature extraction and deep learning," *IET Image Process.*, vol. 14, no. 2, pp. 348–353, Feb. 2020.

[22] K. Yildiz, "Dimensionality reduction-based feature extraction and classification on fleece fabric images," *Signal, Image Video Process.*, vol. 11, no. 2, pp. 317–323, Feb. 2017.

[23] V. G. Bocekci and K. Yildiz, "Classification of textures using filter based local feature extraction," in *Proc. MATEC Web Conf.*, vol. 75, 2016, p. 03001.

[24] M. Garg and G. Dhiman, "A novel content-based image retrieval approach for classification using GLCM features and texture fused LBP variants," *Neural Comput. Appl.*, vol. 33, no. 4, pp. 1311–1328, Feb. 2021, doi: 10.1007/s00521-020-05017-z.

[25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[26] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. S. Bernstein, A. C. Berg, and F. Li, "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, no. 3, pp. 211–252, Dec. 2015, doi: 10.1007/s11263-015-0816-y.

[27] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," 2014, *arXiv:1412.3474.*

[28] M. Long, Y. Cao, J. Wang, and M. I. Jordan, "Learning transferable features with deep adaptation networks," in *Proc. 32nd Int. Conf. Mach. Learn. (ICML)* (JMLR Workshop and Conference Proceedings), vol. 37, F. R. Bach and D. M. Blei, Eds. Lille, France: JMLR, 2015, pp. 97–105. [Online]. Available: http://proceedings.mlr.press/v37/long15.html

[29] C. Buerhop-Lutz, S. Deitsch, A. Maier, F. Gallwitz, S. Berger, B. Doll, J. Hauch, C. Camus, and C. J. Brabec, "A benchmark for visual identification of defective solar cells in electroluminescence imagery," in *Proc. 35th Eur. Photovolt. Sol. Energy Conf. Exhib.*, Brussels, Belgium, Sep. 2018, pp. 1287–1289, doi: 10.4229/35thEUPVSEC20182018-5CV.3.15.

[30] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation," 2020, *arXiv:2010.16061*.

[31] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2015, pp. 1–9, doi: 10.1109/CVPR.2015.7298594.

[32] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. E. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 2014, *arXiv:1409.4842*.

[33] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "An image is worth 16×16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

[34] I. O. Tolstikhin, N. Houlsby, A. Kolesnikov, L. Beyer, X. Zhai, T. Unterthiner, J. Yung, A. Steiner, D. Keysers, J. Uszkoreit, M. Lucic, and A. Dosovitskiy, "MLP-Mixer: An all-MLP architecture for vision," 2021, *arXiv:2105.01601*.

TAO FAN received the M.S. degree from Shandong University. He is currently a Professorate Senior Engineer with the State Grid Corporation of China. He is also the Vice Director of the Internet Department, State Grid Corporation of China, and the Secretary General of central enterprise e-commerce alliance. His research interests include enterprise informatization, power system automation, and energy internet. He has won the First Prize of China Electric Power Science and Technology Award and the First Prize of Science and Technology Progress of State Grid Corporation of China.

TAO SUN received the B.S. degree from the Electrical Engineering Department, Shanghai Jiaotong University, in 1990. He is currently the Vice General Manager of State Grid E-Commerce Company Ltd. His research interests include power system automation, energy internet, new energy, and distributed photovoltaic power station operation monitoring. He received the Senior Engineer Qualification Certificate, in 2001.

XIANGYING XIE received the B.S. degree in computer science and technology from the Wuhan University of Science and Technology, in 2003, and the M.S. degree in software engineering from Sichuan University. He is currently pursuing the Ph.D. degree with Beihang University. He is currently a Deputy General Manager of photovoltaic cloud (e-commerce poverty alleviation) with the Business Department, State Grid E-Commerce Company Ltd., and a General Manager of State Grid New Energy Cloud Technology Company Ltd. His research interest includes the new energy industry internet. He won the Honor of "Outstanding Individual" in SASAC Central Enterprise Innovation Achievement Exhibition.

HU LIU received the B.S. degree in mechanical engineering and automation, computer science and technology and the M.S. degree in mechanical and electrical engineering from North China Electric Power University, in 2003 and 2006, respectively. He is currently the Deputy Director with the Internet Department, State Grid Corporation of China. His research interests include power marketing, power grid production, and energy services. He won the First Prize of Geographic Information Technology Progress of China Geographic Information Industry Association, the Special Award of Scientific Progress of State Grid Corporation of China, and the First Prize of China Electric Power Innovation.

ZHIXIONG NA received the B.S. degree in thermal engineering from Changzhou University, in 1996, the M.S. degree in computer application from the University of Science and Technology of China, in 2001, and the M.B.A. degree from the University of International Business and Economics, in 2009. He is currently a Professor with State Grid E-Commerce Company Ltd. His research interests include industrial automation and new energy industrial internet.

• • •