33rd European Modeling & Simulation Symposium
18th International Multidisciplinary Modeling & Simulation Multiconference

ISSN 2724-0029 ISBN 978-88-85741-57-7 © 2021 The Authors.
doi: 10.46354/i3m.2021.emss.039

Multidisciplinary design method for product quality based on ResNet50 network

Guodong Yi[1,*], Lifang Yi[1] and Yanwu Feng[1]

[1]State Key Laboratory of Fluid Power & Mechatronic Systems, Zhejiang University, Hangzhou 310027, China

[*]Corresponding author. Email address: ygd@zju.edu.cn

Abstract

The positioning accuracy of the PCB during processing depends on the quality of the MARK point images. The collection of MARK point images is affected by factors such as background, illumination etc., so the classification of images is the key to improve the accuracy of PCB positioning. In this paper, a multidisciplinary design modelling method for product quality is proposed. A classification model through transfer learning based on the ResNet50 network and weights is built. It is verified by a set of customized experiment data that the accuracy of MARK point image classification has reached 98.53%. Compared with traditional classification methods, the accuracy rate of this method is 20% higher, and is more suitable for custom small data sets. It provides a guarantee for the subsequent classification and segmentation of MARK point images with different characteristics.

Keywords: MARK point images; multidisciplinary design model; ResNet50 network; transfer learning

1. Introduction

With the continuous update and iteration of electronic products, the demand for PCBs continues to increase. In the field of PCB production, the maskless double-sided multilayer exposure technology based on DMD is widely promoted, which represents the development direction of this field (Zhang et al., 2018; Peng et al., 2018). The MARK point is the reference used for positioning during the exposure and printing process of the PCB. Affected by uncertain factors in the industrial production environment, such as equipment accuracy, mechanical vibration, unclear background, uncertain brightness, some MARK point images with complex background and low quality are collected. These low-quality images will make it difficult to accurately detect and locate the MARK points, directly reduce the printing alignment accuracy, and ultimately lead to a decline in quality and production efficiency of the PCB production (Peng, 2013).

The existing image detection methods mainly study a specific method to detect a single complex or multiple simple MARK point images and most of them are concentrated on improving the circle detection calculation efficiency. However, due to the complex and changeable industrial environment during the PCB exposure and positioning process, the MARK point images may have various problems such as complex background, mixed size, reflective edge and blurred edge, which will be not conducive to accurate detection. These images are easy to cause false detections, and ultimately lead to the production of waste products due to positioning errors in the double-sided or multi-layer exposure process of the PCB.

Aiming at the problems of low adaptability and accuracy of typical algorithms in the above low-quality MARK point image detection, this paper proposes a novel classification method to achieve four typical low-quality image classifications.

Hence, the contributions are three-fold:

1. A multidisciplinary design modelling method for product quality is proposed.
2. A classification model through transfer learning based on the ResNet50 network and weights is built to complete the task of MARK point image classification.
3. The performance of this classification model is verified by a custom data set through a series of experiments, and the result shows that the model obtained a 98.53% accuracy of MARK point image classification, higher than other models.

The rest of this paper is organized as follows. Section II reviews the related works. Section III introduces the details behind the multidisciplinary design modelling method for product quality based on ResNet50. Experimental results are presented in Section IV. Finally, the conclusion is stated in Section V.

2. Related Work

In recent years, image classification methods based on deep learning have been widely used. Many scholars mainly study how to deepen the network depth and solve the network degradation problem caused by the continuous increase of network depth (Xu and Chen, 2018).

Krizhevsky, Sutskever, and Hinton (2012)deepened the number of layers of Convolutional Neural Network, and used ReLU and dropout to obtain the AlexNet structure. The AlexNet achieved the best classification result applied in ImageNet's LSVRC-12 competition for the first time.

Szegedy et al. (2015) greatly increased the depth of CNN and proposed a CNN structure called GoogLeNet with more than 20 layers. GoogLeNet adopts three convolution operations (1×1, 3×3, 5×5), which reduces the number of parameters while ensuring higher classification accuracy, and improves the utilization of computing resources. It won the first prize in the LSVRC image classification competition.

Simonyan and Zisserman(2015) discussed the importance of "depth" of CNN in their published articles. Experiments show that when the number of VGG network layers reaches 16-19, the performance of the model will be effectively improved.

He, Zhang, Ren, and Sun (2014a) added a spatial pyramid pooling layer (SPP) to the CNN network layer building the SPP-Net model. The application of the SPP layer overcomes the limitation of the previous CNN model requiring fixed-size input, and realizes the effect of different size inputs producing the same size output, and the training speed is improved. He, Zhang, Ren, and Sun (2014b) established ResNet and introduced shortcut connection method to solve the problem of a sudden drop in network accuracy after saturation as the network depth continues to increase. At the same time, they only set the pooling layer after the last convolutional layer and fully trained the bottom layer of the network. ResNet with a depth of 152 layers was used in the LSVRC-15 competition. The network degradation problem was solved very well, and finally won the first place in the image classification competition. Using in a typical classification network, ResNet has greatly improved the accuracy of image classification and training speed as a result of the introduction of the residual block structure.

The deep learning methods widely used in the field of machine vision, but have relatively few applications and researches in MARK point detection (Wu, 2019). Aiming at the problem of MARK point image classification of PCB, this paper proposes using the residual network as the research foundation to study a method with stronger adaptability and higher generalization ability, which will provide reference for related research.

3. Methods

3.1. Feature extraction of marker image

There is a non-linear relationship between the four image features and categories. It is difficult to obtain an accurate mathematical mapping model, and traditional feature detection and extraction algorithms such as Harris, SIFT, ORB, etc. are not suitable for use. The neural network method can imitate the mechanism of human neurons. This method can automatically extract features from a large number of images, thereby establishing a mapping model between images and image categories, and realizing the classification of PCB MARK point images.

The MARK point image classification based on the CNN is essentially to find a specific label assigned to the input image from the set of four known image labels. Problems such as image segmentation and image detection can also be regarded as image classification problems in principle. How to extract a specific feature in a certain marker image and make it a descriptor of the corresponding image is the essential task of image classification based on CNN. The ability of CNN to perform feature extraction comes from the functional cooperation and data transmission between the various functional layers of the network. The main functional layers of CNN can be divided into convolutional layer, pooling layer, ReLU layer and fully connected layer (Krizhevsky et al., 2012). The functions of each layer are as follows.

Convolutional layer: Both shallow convolutional layers and deep neural layers are used. The shallow convolutional layer is used to extract low-level features in the MARK point images, such as the shape of the reflective area; the deep convolutional layer can extract more complex and abstract global features, such as the spatial location of the reflection. The continuous sliding operation of the convolution kernel on the MARK point image is the operation operation of the convolution layer. The deep network finally realizes the feature extraction of different types of images through the convolution layer and the pooling layer. The collected two-dimensional image I of the MARK point is used as the input matrix, the convolution kernel is assumed to be matrix K, and the convolution operation can be expressed as:

$$\mathrm{S(i,j)} = \sum_m \sum_n I(i+m, j+n)K(m,n) \tag{1}$$

Pooling layer: The pooling layer can reduce the parameters and computational complexity, and improve the generalization ability of the model; make the network insensitive to some small changes in the MARK point image, with invariance of translation, rotation and scale (He et al., 2014).

Activation layer: The mathematical principle of the activation operation is ReLU, which can convert all negative values of the input to zero, and output the positive values as they are. The ReLU activation function formula is shown in equation 2, where $\lambda \sim U(l,u), l < u$, and $l, u \in [0,1)$ (Simonyan and Zisserman, 2015).

$$f(x) = \begin{cases} x & if\ x > 0 \\ \lambda x & if\ x \le 0 \end{cases} \tag{2}$$

Fully connected layer: The fully connected layer can be used as an image classifier to combine the local features of the MARK point images extracted from each layer of the network into a complete image to achieve classification.

Combine the convolutional layer, pooling layer and activation layer and increase the depth of the network. A deep neural network is established neural network.

The convolution results are arranged into a column of vectors x_i, which are fully connected to the output four images, and each connection has a weight ω_i. According to the equation: $y = \sum \omega_i * x_i$, the outputs of the two neurons are obtained respectively. Finally, by constructing a specific logic function to scale the result to 0-3, the recognition result of four types of images can be judged by the weight.

The above is the entire process of forward propagation. The obtained set of output is compared with the standard established by the loss function, and then errors are corrected through back propagation during training and learning. Different convolution kernels are used to extract specific features in the MARK point image. Since the position of the original image that the convolution kernel should match is uncertain, the convolution kernel will scan the input image with a specific step size. After each convolution operation, a feature value matrix is obtained, and finally multiple feature value matrices are used to describe the features of this image. In the classification and recognition of MARK point images, input a MARK point image, use several convolution kernels that express different image features, and extract features such as reflective edge and complex background through convolution operations.

The variability of reflective edge images is shown in Figure 1. A convolution kernel is used to detect the reflective edge. In fact, the size and shape of the reflective area of different MARK point images are different. If the convolution kernel is too specific, feature detection will fail, so a backpropagation is needed to determine the value of the convolution kernel.

Backpropagation is the process of comparing the predicted value of a MARK point image with the real image tag value, and then returning to the process of modifying the network parameters. According to a specific method, such as random initialization, or Kaiming initialization, initialize the parameters of the size of the convolution kernel, and then adjust the value of the convolution kernel adaptively with the loss function as the comparison standard, so that minimizes the error between the predicted value and the true value (Simonyan and Zisserman, 2015).

In the image processing process, pixel values are used to express and extract features, and the different quantification of the MARK point image is the difference of the feature value matrix. We use a 4-dimensional vector and one-hot encoding to represent the classification result of the classification model, where each dimension represents the type of the MARK point image. For example, the classification result vector of the images with reflective edge is $[0,0,1,0]^T$.

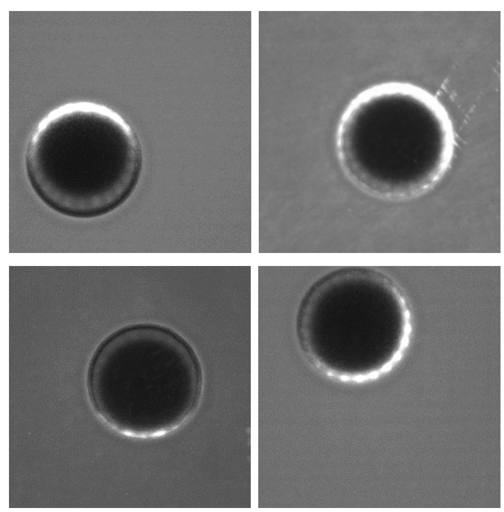

Figure 1. Variability of reflective edge images.

3.2. Transfer learning

Generally, in order to obtain an effective image classification model, the training process of image classification using deep learning requires a large number of images as data sets. The difference between the MARK point image classification and the image semantic segmentation is that the former uses the MARK point image as a unit to classify, while the latter uses the target pixel point as the unit to classify or to mark. Therefore, the data set required for MARK point image classification is much larger than the data set required for MARK point segmentation. However, due to the additional costs of materials, time, and manpower, it is very difficult to construct a large-scale well-labeled MARK point image data set. At the same time, small-scale data sets will lead to long training time and poor results. Transfer learning can adjust and improve models that have been trained in specific tasks for new tasks, reducing the dependence on large-scale data sets (Tan et al., 2018). Inspired by this, we use transfer learning to solve the problem of insufficient training image data in the task of MARK point image classification.

The network based on deep transfer learning can be used as a general image-specific feature extractor. This network structure and pre-training parameters are obtained by training on a large-scale data set of a specific target domain. Then they are selectively transferred to a new target domain task for use, and eventually a sub-network is updated in the fine-tuning strategy. Research shows that deep networks such as AlexNet, VGG, and ResNet are more mobile and can be preferred in network-based transfer learning(Yosinski, Clune, Bengio, and Lipson, 2014).

In order to solve the problem of vanishing gradient, the residual structure is used as the basic unit of the deep residual network. Transform the optimized objective: h(x)=f(x)+x into h(x)-x through a cross-layer summation operation. The training target is equivalently mapped from approximating a function h(x) to zero, thereby reducing the difficulty of training. This method greatly improves the training speed and training accuracy without increasing the number of existing network parameters(He et al.,2014b).

Existing typical ResNet models based on PyTorch are mostly used for training and feature extraction of open-source large data sets (such as ImageNet, etc.). However, if this paper retrains the ResNet network classification model, it will not only lead to a long training time, but also cause the model to not converge due to the limited image of specific marker points, which will not achieve the expected effect. Therefore, we finally choose to pre-train the network model and weights on the ImageNet dataset. On this basis, the classifier is designed. Then a new network model that can effectively perform image classification can be obtain by fine-tuning the network parameters, which is suitable for the custom image classification data set (Tan et al., 2018).

The classification model is built based on the ResNet network structure and the parameters obtained by pre-training on ImageNet are used in the initialization of the labeled point image classification. The ResNet convolutional layer is used to extract the depth features of the labeled image. Freeze the pre-trained layers to ensure that they will not backpropagate during training. Next, modify the number of nodes in the last fully connected layer of the network model to correspond to the four low-quality MARK point images with complex background, mixed size, reflective edge, and blurred edge. During the training iteration, only the weight value of the custom fully connected layer is changed. On this basis, a negative log likelihood loss function is created, and the Adam optimizer is used to achieve adaptive adjustment of the learning rate. The specific step is shown in Figure 2:

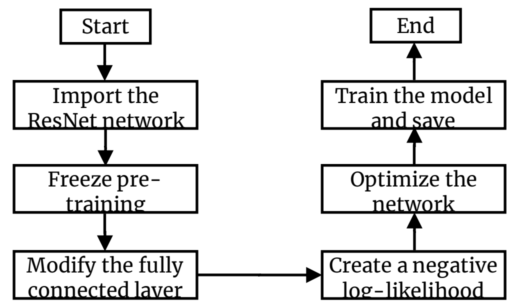

Figure 2. Variability of reflective edge images.

4. Experimental Results

4.1. Dataset

Four representative MARK point images with complex background (Bac), mixed size (Mix), reflective edge (Ref) or blurred edge (Blu) are summarized to construct classification data sets. The four original images are used as the training set: validation set=7:3, (training set and validation set): test set=10:1. The division relationship of the data set is shown in Figure 3.

In different directories, Augmentor-based automated data set enhancement is performed. Then eight data enhancement operations are performed in the custom training set.

The number of each MARK point images before and after data enhancement is shown in Table 1. In order to reduce the classification accuracy fluctuation problem in the training process, it is necessary to ensure that the amount of each type of image data after enhancement is equal, rather than performing proportional enhancement based on the number of original images. Finally, perform training and horizontal comparison on the custom data set, and select the model with the best effect and the best generalization ability.

Table 1. The custom MARK point image classification data set.

Category code	Image category	Number of initial images	Number of enhanced images
Bac	Complex background	33	150
Mix	Mixed size	11	150
Ref	Reflective edge	11	150
Blu	Blurred edge	15	150
Total	Total	70	600

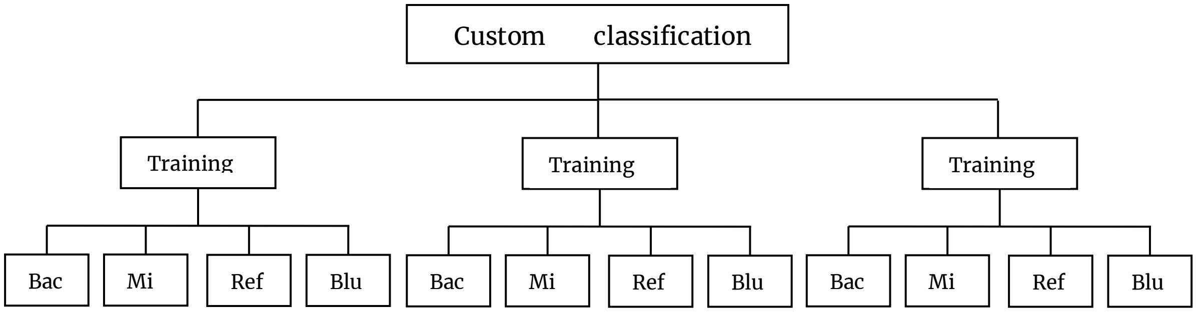

Figure 3. The division and storage relationship of custom classification data sets.

4.2. Comparison of classification models

A typical ResNet network has five types: ResNet18, ResNet34, ResNet50, ResNet101 and ResNet152. In theory, the more network layers and the deeper the network depth, the stronger the ability to extract image features. However, in some cases, it may cause overfitting due to a mismatch with the number of data sets, resulting in performance degradation. Therefore, it is first necessary to set up a horizontal comparison experiment on the ResNet network of each layer on the custom data set, and select the most suitable network layer for the custom classification data set. Given that the pre-training parameters have been used for initialization, the network weights can be fine-tuned by freezing the feature extraction network layer. At the same time, as the Adam optimizer can automatically adjust the learning rate during the training process, the learning rate can be selected as low as possible (Keskar and Socher, 2017; Reddi, Kale, and Kumar, 2019). The training batch size is determined by the number of MARK point images that need to be trained for each iteration, and its size affects the direction of gradient descent in the process of model optimization.

Limited by the fact that the GPU memory cannot be too large, so the following hyperparameters are used for network model performance comparison experiments: epochs = 100, lr = 1e-4, optim = Adam, loss function = NLL, batch_size = 64.

The performance index comparison statistical table is shown in Table 2.

On the custom MARK point classification data set, the horizontal comparison curve of the training accuracy of different residual networks is shown in Figure 4.

From the experimental results, it can be found that taking ResNet50 as the boundary, with the increase of network depth, the training accuracy of image classification network improves as a whole; while ResNet101 and ResNet152 have serious problems in the custom data set as the network depth increases.

ResNet50 with the highest overall accuracy (97.95%) is chosen as the main structure of the MARK point image classification network, and stopping training at the 30th epoch can obtain a better classification model. The feature extraction layers are initialized and frozen by pre-training parameters, and the fully connected layers are fine-tuned. In order to get a better training effect and to obtain higher classification accuracy, the parameters of the feature extraction layers are unfrozen to return back.

Table 2. Image classification network model comparison statistics table

Network model	Number of parameters (k)	Training time per epoch(s)	Accuracy (%)
ResNet152	59195.461	36.56	86.34
ResNet101	43551.813	32.74	96.87
ResNet50	**24559.685**	**23.98**	**97.95**
ResNet34	21549.893	20.24	97.16
ResNet18	11441.733	15.78	95.08

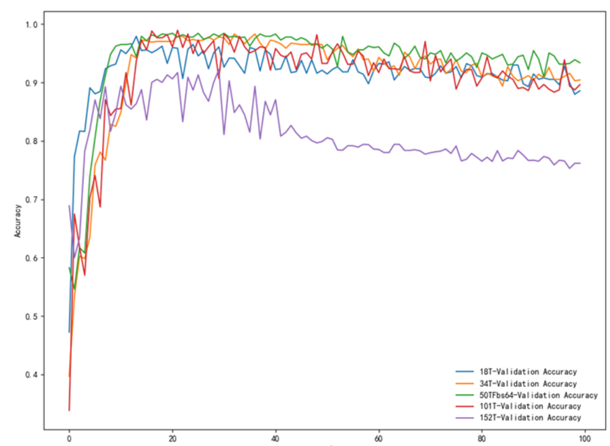

Figure 4. The training accuracy curve of five residual networks.

4.3. Impact of training batches

Training batch refers to the number of MARK point images used for one training in the deep learning network training process, and its size needs to consider the training efficiency and GPU memory (Goyal et al., 2017; Hoffer, Hubara, and Soudry, 2017; Keskar, Nocedal, Tang, Mudigere, and Smelyanskiy, 2017; Smith, Kindermans, Ying, and Le, 2017).

In our experiment, the maximum batch size is set to 64, as the parameter transfer of the abstraction layer requires more memory resources. The smaller batch size is selected according to the interval of 2 times reduction each time to study the influence of the training batch on the classification accuracy of the

MARK point images.

We mainly study the impact of training batches on the training speed and classification accuracy of the MARK point image classification network, and finally the result statistics table shown in TABLE 3 is obtained.

As shown in Figure 5, we adopt ResNet50 model with pre-training parameters for initialization and without freezing the residual network layer parameter transferring, the image classification accuracy increases first and then decreases as the training batch decreases. When the training batch is equal to the cut-off point which is equal to 4, compared with ResNet50 frozen residual network parameter transferring, the classification accuracy is improved overall during the 100 iterations. However, as the training batch continues to decrease, the classification accuracy begins to decline. When the training batch is set to 1, the classification accuracy in the entire iteration period drops significantly.

ResNet50 is finally selected to use the pre-training parameters as the training initialization values, and then pass the parameters forward and backward on the custom training set to train the final MARK point image classification model. According to the experimental analysis, the training batch size is equal to 4 meeting the classification accuracy. At the same time, in order to stop training before the accuracy drops due to over-fitting, the training is carried out for 30 epochs, and the loss curve and accuracy curve are obtained, as shown in Figure 6 and Figure 7 respectively.

It can be seen from Figure 7 that the image classification accuracy of this classification method has reached 98.53%.

Table 3. Statistical table of Comparison of the impact of training batch.

Network model	Training time per epoch(s)	Accuracy (%)
ResNet50TFbs64	23.98	97.95
ResNet50TTbs8	25.67	98.26
ResNet50TTbs4	**30.31**	**98.53**
ResNet50TTbs2	43.17	98.02
ResNet50TTbs1	63.09	51.78

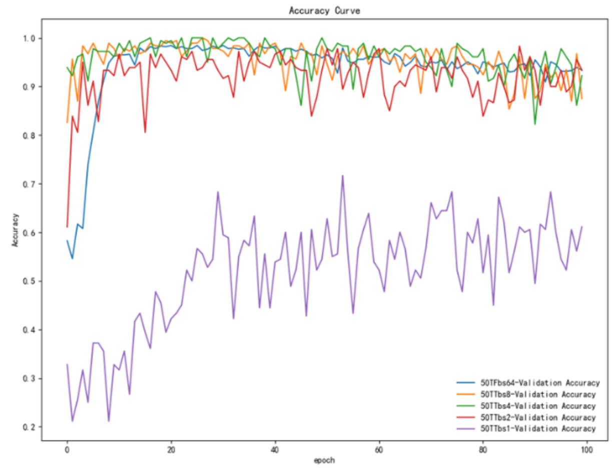

Figure 5. Impact of training batches on the classification accuracy of classifiers.

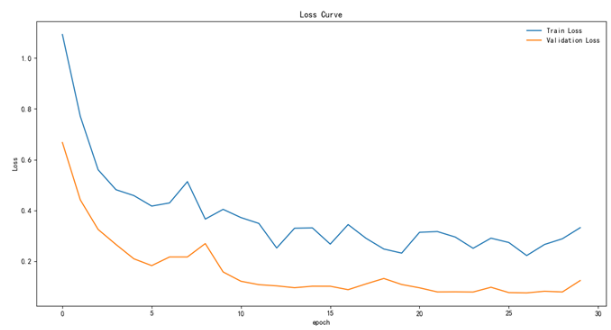

Figure 6. The loss curve of the training network model based on ResNet50.

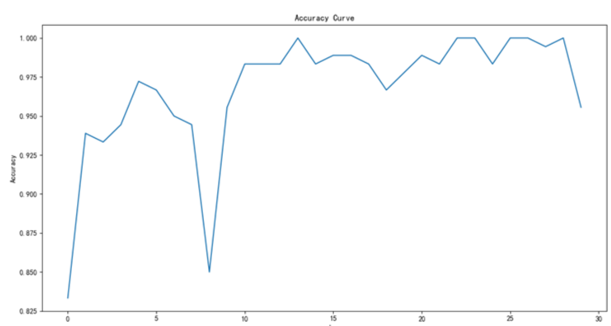

Figure 7. The accuracy curve of the training network model based on ResNet50.

4.4. Comparison of classification accuracy with KNN

In order to further prove the efficiency of the model, this paper constructs a KNN model, and performs a classification test on a custom data set. As shown in the Figure 8, the highest accuracy rate is 0.783333 (K=1). However, when k is greater than 10, as the number of iterations (K) increases, the classification accuracy rate decreases. Finally, at the end of the iteration (K = 34), the accuracy rate is 0.650000. It can be seen from comparison that the classification accuracy rate of the ResNet50-based classifier is 98.53%, which is 20.20% higher than the KNN classifier.

As a result, the classification model through transfer learning based on the ResNet50 network and weights is more suitable for customizing small sample data sets, and can well meet the needs of low-quality MARK point image classification.

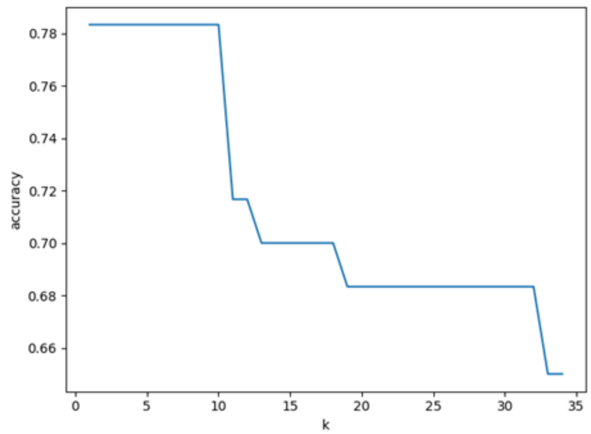

Figure 8. Training process of the KNN classifier.

5. Conclusion

According to the fact that the collection of MARK point images is affected by factors such as background, illumination, resulting in low quality PCB, this paper proposes a multidisciplinary design modelling method for product quality. The method main uses CNN for image feature extraction. The fully connected layer is built on the ResNet50 residual network model, and at the same time, the pre-training parameter layer is frozen to prevent backpropagation to achieve transfer learning based on the network and weights. Iterative training is performed on the constructed custom image classification data set to obtain a classification model. Experiments show that, the method is more suitable for custom small data sets and has obtained a 98.53% accuracy of the MARK point image classification. This work provides a guarantee for subsequent classification and segmentation of MARK point images.

However, the research object in this work is set as widely used circular MARK point images, which means that this method does not cover all shapes of MARK points. Future studies should pay more attention to more typical shapes of MARK points, and further optimize the multidisciplinary design model, which will have practical significance for improving the quality of products produced in complex industrial conditions.

Funding

This research was funded by the National Key Research and Development Program of China, grant number 2018YFB1701600; the National Natural Science Foundation of China, grant number 51875515.

References

Goyal, P., Dollár, P., Girshick, R., Noordhuis, P., Wesolowski, L., Kyrola, A., ... He, K. (2017, June 8). Accurate, large minibatch SGD: Training imagenet in 1 hour. ArXiv. arXiv. Retrieved from http://www.caffe2.ai

He, K., Zhang, X., Ren, S., & Sun, J. (n.d.). Deep Residual Learning for Image Recognition. Retrieved from http://image-net.org/challenges/LSVRC/2015/

He, K., Zhang, X., Ren, S., & Sun, J. (2014). Spatial pyramid pooling in deep convolutional networks for visual recognition. In Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) (Vol. 8691 LNCS, pp. 346–361). Springer Verlag. https://doi.org/10.1007/978-3-319-10578-9_23

Hoffer, E., Hubara, I., & Soudry, D. (2017). Train longer, generalize better: Closing the generalization gap in large batch training of neural networks. Advances in Neural Information Processing Systems, 2017–Decem(June), 1732–1742.

Keskar, N. S., Nocedal, J., Tang, P. T. P., Mudigere, D., & Smelyanskiy, M. (2017). On large-batch training for deep learning: Generalization gap and sharp minima. 5th International Conference on Learning Representations, ICLR 2017 - Conference Track Proceedings, 1–16.

Keskar, N. S., & Socher, R. (2017). Improving generalization performance by switching from ADAM to SGD. ArXiv, (1).

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Advances in Neural Information Processing Systems (Vol. 25). Retrieved from http://code.google.com/p/cuda-convnet/

Reddi, S. J., Kale, S., & Kumar, S. (2019). On the convergence of adam and beyond. ArXiv, 1–23.

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, 1–14.

Smith, S. L., Kindermans, P. J., Ying, C., & Le, Q. V. (2017). Don't decay the learning rate, increase the batch size. ArXiv, (2017), 1–11.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... Rabinovich, A. (2015). Going deeper with convolutions. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 07–12–June, 1–9. https://doi.org/10.1109/CVPR.2015.7298594

Tan, C., Sun, F., Kong, T., Zhang, W., Yang, C., & Liu, C. (2018). A survey on deep transfer learning. Lecture

Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 11141 LNCS, 270–279. https://doi.org/10.1007/978-3-030-01424-7_27

Yosinski, J., Clune, J., Bengio, Y., & Lipson, H. (2014). How transferable are features in deep neural networks? Advances in Neural Information Processing Systems, 4(January), 3320–3328.

Peng, G., Xiong, C., Xia, C., et al. (2018). Yi zhong ji yu MARK dian de dian jiao ji qi ren shi ju ding wei fang fa [A method of vision target localization for dispensing robot based on mark point] [J]. CAAI Transactions on Intelligent Systems, 5, 728-733.

Peng, X. (2013). Fu za chang jing gao wei kuan tu xiang dui bi zeng qiang chu li yan jiu [The Research of Contrast Enhancement Processing on Wide-bits Image in Complex Scenes]. (Doctoral dissertation, University of Chinese Academy of Sciences, 2013).

Wu, Z. (2019). High-precision positioning of circular positioning points on defective printed circuit boards under uneven illumination. (Doctoral dissertation, University of Chinese Academy of Sciences, 2019).

Xu, S., Chen, S. (2018). Ji yu shen du xue xi de tu xiang fen lei fang fa [Image classification method based on deep learning]. Application of Electronic Technique, 480, 122-125.

Zhang, X., Liu, H., Gu, W., et al. (2018). Quan qiu guang ke ji fa zhan gai kuang yi ji guang ke ji zhuang bei guo chan hua. [A survey on the development of global lithography machines and the localization of lithography equipment]. Wireless Internet Technology, 15, 110-111.