

Unknown defect detection for printed circuit board based on multi-scale deep similarity measure method

Runwei Ding¹, Can Zhang¹ ✉, Qisheng Zhu², Hong Liu¹

¹Key Laboratory of Machine Perception, Peking University, Shenzhen Graduate School, People's Republic of China

²Shenzhen Skyworth-RGB Electronic, CO., LTD., Shenzhen, People's Republic of China

✉ E-mail: can.zhang@pku.edu.cn

eISSN 2051-3305

Received on 14th October 2019

Accepted on 19th November 2019

doi: 10.1049/joe.2019.1188

www.ietdl.org

Abstract: Defect detection with high precision is of great significance for printed circuit board (PCB) fabrication. Due to the lack of priori knowledge of categories and shape features, detection of unknown defects faces greater challenges than that of common defects. Inspired by similarity measurement, this study proposes a multi-layer deep feature fusion method to calculate the similarity between template and defective circuit board. Compared with conventional methods which divide the whole detection into two independent parts of hand-designed features and similarity measurement, the authors end-to-end model is designed to combine these two parts for joint optimisation. First, the Siamese network is utilised as their backbone architecture for feature extraction of pairwise images. And then the spatial pyramid pooling network is incorporated into the feature maps of each convolutional module to fuse the multi-scale feature vectors. Finally, the discriminative feature embedding and similarity metric are obtained by using the contrastive loss during the training process. Experimental results show that the proposed model has better performance in detecting and locating unknown defects in bare PCB images than traditional similarity measurement methods. Moreover, our method is promising for further improvement of defect detection with less training image pairs and more accurate detection results.

1 Introduction

Printed circuit board (PCB) is an important component of most electrical devices, as well as some consumer electronics. In the manufacturing process, a PCB is usually produced by undergoing over 50 fabricating steps [1]. Detection by human eyes is error-prone. Also, manual detection, whose assessments usually cannot be quantified, is rather time-consuming and tedious. A fast and efficient detection system is therefore in great need to replace manual detection method. Due to the good accuracy and efficiency, automated optical inspection (AOI) systems have been widely used in defect detection fields. However, AOI equipment is usually too expensive for small industries to afford [2]. A detection system, both low cost and effective, is the target for current PCB researches.

AOI-based approaches mainly consist of three categories, including referential and non-referential methods, as well as hybrid method, which is the combination of previous two methods [1]. For referential methods, a template image is made comparison with the test image to find defects. Many researches have been proposed in the past few years. Wu *et al.* detect defects using the referential method, and finally classifies them into seven categories [3]. Putera *et al.* classify defects into seven categories by using area characteristics [4]. A referential method is proposed to classify defects into five categories by using edge grey gradient of PCB [5]. Ibrahim *et al.* improve PCB detection by incorporating a geometrical image registration. There were six types of defects are successfully identified by the system in [6]. Kumar *et al.* consider the defect detection and classification equally important. The limitation of this non-referential detection method is that only one defect can be detected on a single image [7]. Inoue *et al.* use bag of Keypoints to form Visual Word dictionary of RootSIFT features from the whole image and then the support vector machine (SVM) is used as the classifier [8]. Although the hybrid methods merge the advantages of both referential and non-referential methods, high computational complexity cannot be ignored. The artificial neural network (ANN) is widely used to do defect detection and classification among hybrid methods. In the literature [9–11], due to the good learning ability, the researchers utilise ANN to detect solder joint defects in two ways, one is the supervised ANNs [10]

and the other one is unsupervised method mentioned in [11]. Hao *et al.* combine multilayer perceptron neural network with a genetic algorithm to do the solder joint defect detection on PCB [12]. Furthermore, some other detection and classification methods have attempted to detect PCB defects, which contain ANN ensembles used in [13], Bayes classifiers in [14] and SVM [15] for inspection of solder joints.

Convolutional neural network (CNN) has been used for many years in AOI systems [16]. In the traditional way, recognition algorithms are usually based on hand-crafted features. Since CNN can automatically learn distinguished features of objects, it has been extensively used in a wide range of areas such as speech recognition, information retrieval, natural language processing, and computer vision [17–19]. To our knowledge, deep learning methods have been already applied to defect detection fields. In [8], an approach is proposed using a neural network with fuzzy rule-based classification method. Experimental results reveal the superiority of the neural network classification method in terms of its classification accuracy. However, this fuzzy rule table needs additional criteria knowledge of a human inspector and the optimisation depends on the initial values of the weight parameters. Method proposed in [20] extract SURF features [21] without reference images to do defect detection and classification of the electronic circuit board. Images are cropped and then used as input of CNN. By using CNN features, SVM is employed for learning and classification. However, this CNN model can only be applicable to the image for manual verification. Caldo *et al.* utilise ANN to visually detect and classify defects on two-layer PCB. The supervised back-propagation learning algorithm is adopted for training PCB images [22]. With great progress in deep learning, deep models evolve from AlexNet [23], VGG [24] to GoogleNet [25] etc. As the deep model used in [26], authors applied the pre-trained CNN model to learn deep discriminative features of bare PCB defects. Compared with the traditional shallow feature-based methods, the proposed deep learning method is much more feasible and effective.

Although referential methods mentioned above all have pretty good performance for PCB defect detection, they are usually computational-cost and time-consuming. The requirement of image

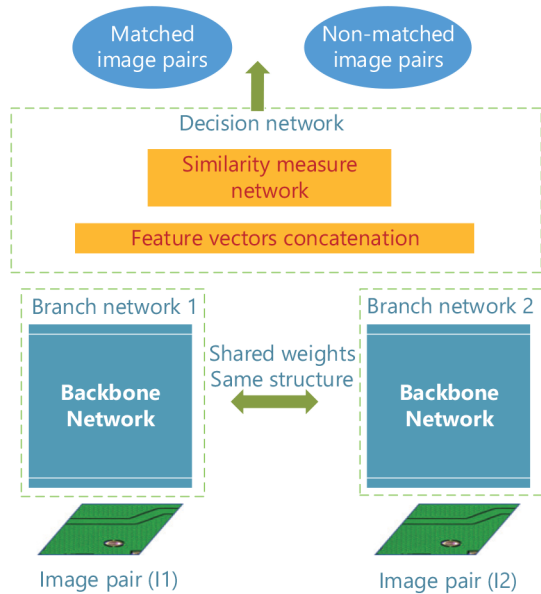


Fig. 1 Framework of two stream Siamese network

registration is relatively high, such as methods in [3–5]. For researches using neural networks, most of them mainly put emphasis on the defect classification. As mentioned in [27], this procedure is very effective when defects are of less categories and common to see. Selecting proper image features and proper thresholds to detect various defects and do classification was not easy [28]. What is more, CNN-based models usually need mass samples for training. When defects are unknown, proposals presented previously are not feasible to make effective detection.

Motivated by the above challenges, we propose to learn a similarity metric with the limited dataset. Then the aim of our task can be generalised as to learn a similarity measure directly from the image pairs without designing handcrafted feature descriptors. When defects that have never been trained previously occur on the PCB, the model can effectively match new defects to previously defined defect categories. Considering these, a Siamese network is applied for defect detection. As Siamese network has been successfully applied to the problems in which the number of categories is very large and samples in each category is not sufficient enough. Based on this method, our small PCB dataset can be efficiently used to train a robust model for the tough unknown defect detection problem.

The main contributions of this paper are threefold:

- (i) We present a similarity measure based on the deep learning method to compare image pairs, which is processed on patches of the whole image. Compared with traditional hand-designed features, our method can better adapt to image block transformation effects, such as illumination and noise. To our best knowledge, this is one of the first attempts to employ the deep neural network to PCB defect detection with the distance metric learning.
- (ii) The untrained data, which has not been exploited previously by the model during training process, is applied to verify the robustness of the deep model. This characteristic is of great importance, as the key task of this paper is to detect the unknown defects on PCB. Experiments verify the good performance of our method.
- (iii) Multi-scale features in deep space are combined to facilitate a more robust feature representation. As features in low layers contain rich texture and spatial information, which could contribute to the localisation of defects. In contrast, features in higher layers are of rich semantic information, which make the model more robust to variations of input image pairs. Considering the complementary feature characteristics, our paper proposes a multi-scale similarity measurement model.

This paper is organised as follows: Section 2 further reviews the previous researches that are relevant to our work. In Section 3, we briefly introduce the Siamese network which is the basis of our model. And then the details of model structure, training process and defect localisation procedure will be introduced. Section 4 presents the experimental results of our proposed model. Finally, conclusions and future works are drawn in Section 5.

2 Multi-scale similarity measurement model

As multiple layers provided by a typical CNN model present different feature characteristics of the target image, the feature hierarchy can be utilised to build a more robust detection system. In order to detect unknown defects with various appearances, our goal is not to train a classification network simply, because the number of unknown defects is large and unknown. There are two main tasks needed to address instead, one is the discriminative feature embedding learning, the other is the architecture construction suitable for a small dataset with large categories. Our method is thus proposed by fusing multi-scale deep features in a Siamese network, which learns a similarity metric between image pairs.

2.1 Two-branch convolutional neural network

Basic structure of the deep network in our proposed method is shown in Fig. 1, which is a Siamese network. The model is composed of two parallel networks, and each contains a basic network architecture including convolutional layers and fully connection layers. Image pairs I_1 and I_2 , similar or dissimilar, are first fed into two networks separately. In the training step, two networks in two streams will be optimised simultaneously under the weight sharing mechanism. Two branches can be regarded as feature extraction modules and outputs are two vectors in low dimensional representation of image pairs. Then the goal is to learn an optimal feature representation of the input image. As a result, images matched in the same categories are pulled together, and on the contrary, unmatched images from different categories are pushed far away. Behind a series of convolutional and activation (ReLU in this paper) layers, a top network is concatenated as a function of the decision network which computes the similarity of image pairs. In our single-scale model, the top network consists of three fully connected layers (Fc1(4096-D), Fc2(512-D), Fc3(2-D)). For the task of computing similarity of two image patches, two feature vectors $Img1_S$ and $Img2_S$ are combined by fully connected layers to compute the loss, which typically employs contrastive loss function (shown in (1)). Then the model is fine-tuned by the back propagation algorithm [29].

Since the structure of AlexNet [24] network is more similar to the human visual cortex [30], we use this deep model as the backbone network of each stream. First five convolutional layers are used to extract features in our paper. Details of the network are shown in Table 1, and top three fully connected layers (Fc1, Fc2, Fc3) are removed. The whole convolutional layers are divided into five slices, denoted as S_1, S_2, S_3, S_4, S_5 . The first slice S_1 contains conv1 and a ReLU layer as that in AlexNet. The second and third slices have the same structure including max-pooling layer, convolutional layer and ReLU layer. For the last two slices, they are identical to the first slice. The feature map dimension of each slice is listed in Table 1 with details.

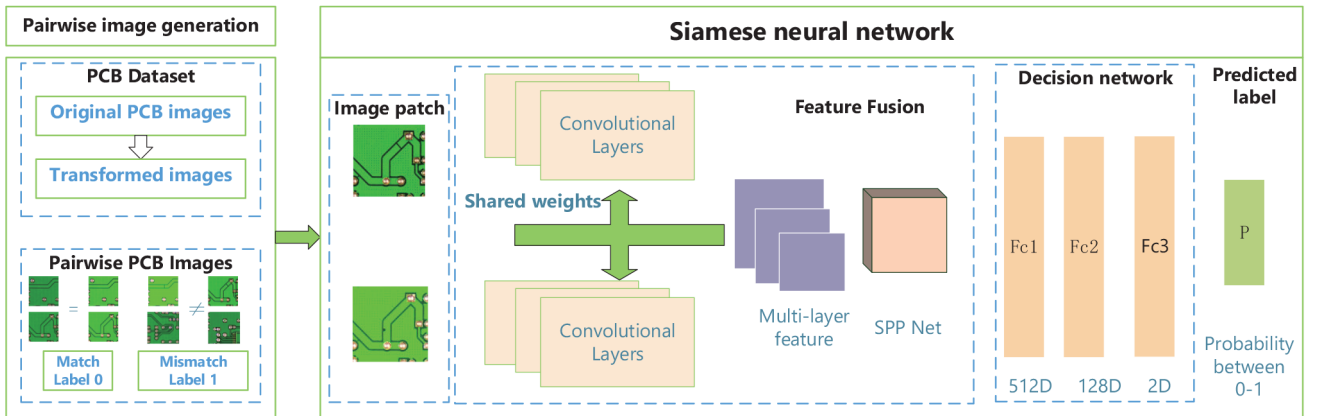
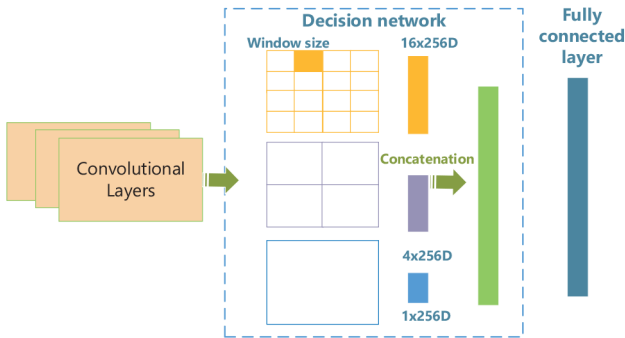
2.2 Multi-layer feature fusion

Considering layers in different levels containing diverse but complementary information, we attempt to fuse multi-scale features in different layers to construct a more discriminative and abundant feature representation of the test images. The strategy, which combines the detailed texture information in lower layers with semantic information in higher layers, has been proven to be effective in many researches [31–33].

Fig. 2 shows an overall framework of the proposed similarity measurement model based on multi-scale feature fusion to detect matched or unmatched image pairs. First, we apply data augmentation technique to raw PCB images to enlarge the limited

Table 1 Details of the backbone network architecture. In the table, channel refers to the dimensions of filters

	Structure	Output feature (Input [512,512])
slice 1	0: Conv2d (in_ch = 3,out_ch = 64, kernel = (11,11),stride = (4,4), padding = (2, 2)) 1: ReLU	channel: 64 map: [127,127]
slice 2	2: MaxPool2d(kernel = 3, stride = 2, padding = 0, dilation = 1) 3: Conv2d (in_ch = 64, out_ch = 192, kernel = (5,5), stride = (1,1),padding = (2, 2)) 4: ReLU	channel: 192 map: [63,63]
slice 3	5: MaxPool2d(kernel = 3, stride = 2, padding = 0, dilation = 1) 6: Conv2d (in_ch = 192, out_ch = 384, kernel = (3,3), stride = (1,1), padding = (1, 1)) 7: ReLU	channel: 384 map: [31, 31]
slice 4	8: Conv2d (in_ch = 384, out_ch = 256, kernel = (3,3), stride = (1,1),padding = (1, 1)) 9: ReLU	channel: 256 map: [31,31]
slice 5	10: Conv2d (in_ch = 256, out_ch = 256, kernel = (3,3), stride = (1,1),padding = (1, 1)) 11: ReLU	channel: 256 map: [31,31]

**Fig. 2** Overall architecture of multi-scale deep similarity measurement model**Fig. 3** SPP architecture composed in multi-scale model

small dataset for training. PCB images with the same layout are divided into two categories, including normal and defective. Small image patches from the template and its transformations are randomly combined in pairs to form a matching image pair. In the unmatched image pairs, one patch comes from the template and the other comes from the defective PCB set. Then, instead of the traditional single-branch network model, we apply Siamese network architecture to learn the feature representations of pairwise PCB images. Two parallel networks which share weights perform a local operation on the input PCB images. After a series of convolutional layers and activation layers as a function of linear and non-linear transformations, multi-resolution feature maps are encoded by a multi-scale feature fusion layer for further similarity computation. Finally, the output of two feature vectors are concatenated as fusion features to be transferred into the decision network. In the training step, with the distance between pairwise image vectors, the contrastive loss is applied to fine-tune the feature extraction network.

2.2.1 Single-feature fusion model: In this model, the fully connected layers are reused which are removed in the feature

extraction stage, including Fc1 and Fc2. After the concatenation, a multi-feature representation of image pairs is built. At this time, only the decision layer is fine-tuned in training the model, and network parameters of the feature extraction part are fixed and not affected by the training process.

2.2.2 Multi-feature fusion model: As shown in Fig. 3, compared with the previous naive one, there is no constraint on the input image size because of the SPP net. As the last convolutional layer is not connected with the fully connected layers of predefined dimension, input images can be of arbitrary sizes. In our paper, the image patches are all of the same size $512 \times 512 \times 3$ and will first be resized to the required spatial dimensions in the first structure. Inserting a spatial pyramid pooling layer, we can input image patches of multi-scale size without deteriorating the resolution. Making the difference of each slice (S_1, S_2, S_3, S_4, S_5) in Table 1 to get the difference feature maps, which can be shown in (1)

$$\text{Diff}_{S_i} = (\text{Img1}_{S_i} - \text{Img2}_{S_i}). \quad (1)$$

After spatial average pooling layer, each 2D difference feature map is followed by SPP net layer. Before the fully connected layer, we fuse feature vectors from different scales (S_1, S_2, S_3, S_4, S_5) to enhance convergence representation of network features. Thus, the feature representations of each slice are concatenated to form multi-level fusion feature vectors, which is denoted as F_{fusion} in (2)

$$F_{\text{fusion}} = \text{Conc}(\text{Diff}_{S_i}) \cdot (S_i \subseteq S, i \in [1, 5]) \quad (2)$$

For the decision network, the architecture consists of three fully connected layers (Fc1 (512-D), Fc2 (128-D), Fc3 (2-D)). The input vector is a combination of two image feature vectors. The output of Fc1 layer is given by:

$$Y_{\text{fusion}}(F) = h(W_{\text{fusion}}F_{\text{fusion}} + b_{\text{fusion}}), \quad (3)$$

where Y_{fusion} represents the fusion feature vectors, $h(\odot)$ denotes the activation function of the decision network. W_{fusion} and b_{fusion} are the weight and offset, respectively.

Our goal is to fine-tune the parameters of the decision network. ReLU is used as the activation layer and softmax is employed as the similarity decider. With the label 0 for match and 1 for the mismatch, the contrastive loss is computed by the distance calculated by the multi-scale feature vectors. After that, backpropagation process can update the decision network parameters for the optimal similarity prediction.

2.3 Loss function

The network is optimised by the loss function to distinguish the difference between standard PCB and defective PCB. To be more specific, similar images should be encouraged as close as possible and dissimilar image pairs as far as possible. In the implementation of this task, we use the contrastive loss proposed in [30], and its definition can be presented with a margin as follows:

$$L = \frac{1}{2}YD^2 + \frac{1}{2}(1 - Y) \max(0, m - D)^2, \quad (4)$$

$$Y = \begin{cases} 1, & \text{if two images are similar} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

here image pairs' label can be denoted as Y , 1 for similar pairs and 0 for dissimilar pairs. The marginal m denotes that there should be at least margin m between dissimilar samples, and the value of $m > 0$. D is the Euclidean distance between two feature vectors of the input image pairs.

Only if the distance is less than the margin m , dissimilar images will contribute to the loss value. Since this loss function is made to pull together similar image pairs closely and put dissimilar images far away from each other. The penalisation of this loss function is the squared Euclidean distance for the similar image pairs. For the dissimilar image pairs, the difference between margin m and Euclidean distance is squared when m is larger than the image distance.

3 Experiments and discussions

3.1 Dataset description

Our experiments are conducted on PCB dataset from [34]. In this data set, there are 693 defective circuit boards with six types of defects, including short circuit, open circuit, mouse bite, spur, missing hole and copper. Details are listed in Table 2.

3.2 Configuration

3.2.1 Hardware environment: In all our experiments, the network is trained using stochastic gradient descent (SGD) with a standard back-propagation method [29] and AdaDelta [35]. We use weights from a deep network pretrained on Imagenet as the initialisation to train our Siamese approach. Specifically, we finetune a pretrained model using similar technique as in [36], e.g. setting the learning rate 10^{-5} for the last fully-connected layer (fc7) and 10^{-6} for other layers. The model was trained using publicly available deep learning framework Keras on one NVIDIA GeForce GTX 1080 GPU. The total number of training epochs is 15 and batch size for each iteration is 12. It took around 1.5 h to finish the whole training when the change of loss value is < 0.0001 .

3.2.2 Dataset augmentation: In the public bare PCB dataset [34] released recently, only ten kinds of layout of single-layer PCBs are

contained. And the number of defective boards is just about 690 for all six kinds of defects. With such limited PCB images, we first conduct data augmentation on the original dataset. Some configurations are set to the augmentation strategies, such as the Gaussian noise mean value being set as 0.2 and random rotation angles being in the range of -10° to 10° . These strategies effectively enlarge the PCB datasets.

3.2.3 Evaluation protocol: There are many ways to measure the accuracy of image similarity methods [37]. In our experiments, we show some precision-recall (PR) curves to better describe the advantages of our proposed method. In the following section we provide results on test dataset in details.

We also adopt the evaluation protocol for image classification task applied in [38]. In order to have a quantitative evaluation of the algorithms, the average precision (AP) is used in the defect classification. Precision is the percentage of positive identifications across all selected samples, while recall is the probability of correctly identified instances in the entire positive samples. The computation formula can be represented by the true positive (TP), false positive (FP) and false negative (FN). The specific expression form can be given by:

$$\text{precision} = \frac{TP}{TP + FP}, \quad (6)$$

$$\text{recall} = \frac{TP}{TP + FN}. \quad (7)$$

The precision-recall curve (PRC) is computed to evaluate the performance of classification problems, which set recall value as the x -axis and precision as the y -axis. Area below PRC is averaged upon recall levels spaced at a fixed interval value between 0 and 1. F_1 score can be obtained by the precision and recall metrics, which is widely used to evaluate the algorithms effectiveness. The F_1 score is given by:

$$F_1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}. \quad (8)$$

3.2.4 Baseline methods: In order to compare the deep model with conventional similarity measure methods, we evaluate the feature extraction ability to better discriminate the distance of pairwise images. In this paper, normalised cross-correlation (NCC) [39], smallest eigenvalue [39] and similarity rank [40] are listed as similarity measures baseline. As mentioned before, original PCB data set experiences a series of transformation to increase data set size, such as light change and noise addition, which brings certain challenges to traditional similarity measure methods. In order to vividly present the difficulty of distinguishing pairwise PCB images' similarity, histograms of matched and non-matched image pairs are illustrated in Fig. 4.

To evaluate the performance of deep features model with Siamese network architecture, we make a comparison on network configurations explored in our paper. Basic deep model is the pretrained AlexNet. Using parameters learned from the ImageNet, our PCB image patches are fed into the network and then the image deep feature representations are obtained by the output of the fc7 layer. The second model merges two feature vectors into one, which is put into three fully connected layers. Parameters thus can be learned using contrastive loss. This fine-tuning technique can well avoid the over-fitting phenomenon when we train the model due to small PCB dataset. The last model is the proposed multi-scale feature network.

Through inserting the spatial pyramid pooling layers into the last convolutional layer and the bottom fully connected layer, the

Table 2 Details of original PCB dataset

Defect category	Open circuit	Short circuit	Mouse bite	Copper	Missing hole	Spur
image number	116	115	115	116	115	115
defects number	482	491	492	503	497	488

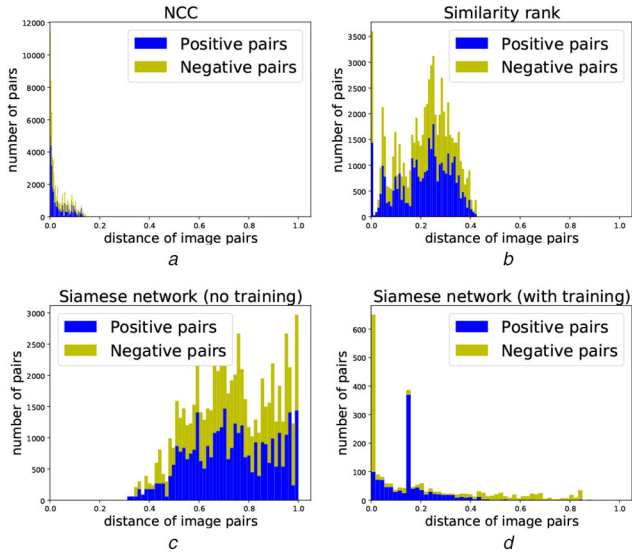


Fig. 4 Distribution of positive and negative pairs under different similarity measure methods on the test dataset
(a) NCC, (b) Similarity rank, (c) Siamese network (no training), (d) Siamese network (with training)

Table 3 Detailed comparison of different deep learning approach on test dataset (area under the ROC curve)

Defect type for testing	AlexNet + SVM	Ours (single-scale)	Ours (multi-scale)
open	0.768	0.927	0.976
Short	0.737	0.893	0.949
mouse bite	0.774	0.876	0.926
Copper	0.728	0.904	0.970
missing hole	0.745	0.873	0.982

overall network can adapt to the input images' sizes greatly. With the help of powerful feature extraction ability of pretrained deep model, we can easily obtain more precise similarity score of image pairs under better feature representations.

In all the experiments, the similarity measurement of image pairs for evaluation is based on the Euclidean distance between feature vectors of pairwise images. Our goal is to obtain better feature descriptors of PCB images.

3.3 Defect localisation

In order to localise defects on the whole big PCB images, a window slides on the PCB to get rather smaller local regions. A large variety of small PCB patches are then fed into the deep model to get feature descriptors. Large sliding window can cause the missing detection problem easily. On the contrary, when the window is too small, only part of the whole defect can be framed out, whose features cannot represent the characteristics of the defect well. Weighing the pros and cons of both sides, we prefer to choose a smaller window, because similarity measurement can still detect the defect area out. However, it further leads to more local area patches from the whole large PCB image and thus reducing the detection speed.

An improvement is thus made to speed up the detection. As the initial model is to apply deep network on the raw PCB image patches directly, the time-consuming feature extraction step will be processed on each patch. Considering decreasing the feature extraction times, we attempt to transfer the sliding window into the feature map. In that case, we only extract the whole PCB image's features once for much quicker detection. The corresponding relationship between the size of the window in the feature graph and that in the original graph is related to the convolutional and pooling layers' parameters of the network.

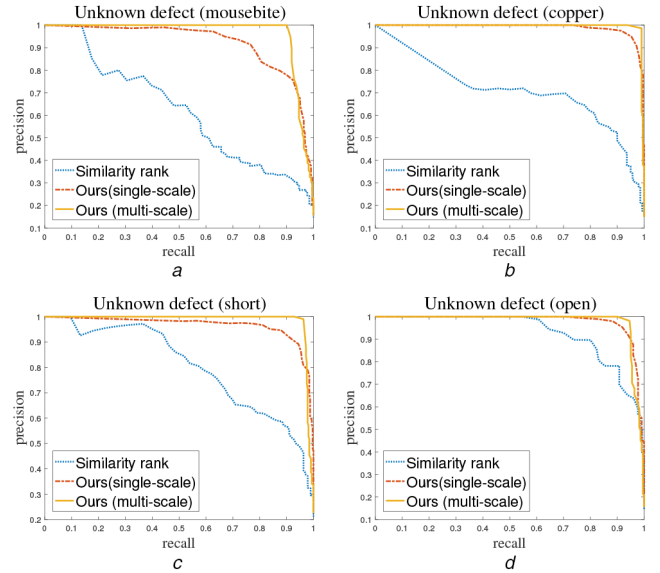


Fig. 5 Set of PR curves for four different test datasets as unknown defect type
(a) Mouse bite, (b) Copper, (c) Short, (d) Open

3.4 Results and discussions

To demonstrate what has been learned by different similarity measure methods on the test dataset, we present the histogram of pairwise Euclidean distances of image representations in Fig. 4. The blue bars and yellow bars represent distances of positive pairs and negative pairs, respectively. The pair distance distribution of Siamese network shows the initial distance distribution of transformed test dataset without learning and the difficulty of discriminating image pairs. When using traditional methods such as NCC and similarity rank metrics, we cannot detect defects out effectively because of various data augmentation techniques imposed on the raw PCB dataset, such as illumination changes. However, it can be clearly seen that training process of Siamese network on image pairs could make the distance between pairwise images of the same kind more concentrated, and separate distributions of two different kinds by a certain distance.

Table 3 summarises classification accuracy of considered approaches on different types of PCB defect test datasets in terms of area under the ROC curve. Experiments show that the proposed method consistently outperforms other algorithms. To be more specific, the Siamese network outperforms traditional NCC methods by % on average. Furthermore, training with a large number of image pairs (with hybrid image sizes), the proposed multi-scale model performs better than the single-scale feature structure.

The detailed evaluation of the methods on the test dataset is shown in Fig. 5. From this set of PR curves, we can make the following observations. Compared with traditional similarity rank metric and our network using a single feature layer, respectively, the proposed multi-scale network has the best performance in all four cases. When the recall value is close to 1, the multi-scale model still has much higher precision. That means positive and negative PCB pairs can be discriminated effectively by our proposed method.

In Table 4, we show the evaluation metrics of different similarity measurement methods, including precision, F_1 score and recall value. It can be seen that the traditional NCC and similarity rank methods could only improve the detection accuracy by a small value compared with random guess. It reflects that our extended data set, after a series of transformations, makes the traditional method lose the ability of effective defect detection. However, our multi-scale method can still achieve a high accuracy rate, and the mAP value reaches 96.3%, which is more than 30% of the method proposed in the paper [39].

Table 4 Evaluations of different similarity measure methods on test dataset

Similarity methods	Precision	Recall	F_1 -score	mAP
NCC [40]	0.55	0.52	0.53	0.529
smallest eigenvalue [40]	0.57	0.55	0.56	0.564
similarity rank [39]	0.63	0.62	0.62	0.627
ours (single-scale)	0.87	0.91	0.89	0.892
ours (multi-scale)	0.96	0.97	0.96	0.963

4 Conclusions

This paper presents a multi-scale feature similarity measurement model based on the Siamese network architecture to detect PCB unknown defects. First, compared with the conventional similarity measurements, the proposed network can learn from data with arbitrary sizes by using the SPP structure. Incorporation with feature fusion strategy, our method can achieve better precision on images with noise addition and illumination changes. Secondly, in the process of defects localisation, detection speed can be effectively improved by transferring the sliding window to feature maps, because feature extraction only needs to be performed once on the original PCB image. Moreover, Siamese network architecture together with contrastive loss is a good approach in learning features for image similarity measure problems. Our proposed method shows great generalisation on unseen PCB defect datasets which provides a good perspective for unknown defect detection problem. However, the selection strategies for different levels of features remain to be further studied, we make it as part of the future work of this paper.

5 Acknowledgments

This work was supported by the Specialised Research Fund for Strategic and Prospective Industrial Development of Shenzhen City (number ZLZBCXLJZI20160729020003).

6 References

- [1] Moganti, M., Ercal, F., Dagli, C.H.: 'Automatic PCB inspection algorithms: a survey', *Comput. Vis. Image Underst.*, 1996, **63**, (2), pp. 287–313
- [2] Guo, F., Guan, S.: 'Research of the machine vision based PCB defect inspection system'. Proc. Int. Conf. on Intelligence Science and Information Engineering, Wuhan, China, August 2011, pp. 472–475
- [3] Wu, W.Y., Wang, M.J.J., Liu, C.M.: 'Automated inspection of printed circuit boards through machine vision', *Comput. Ind.*, 1996, **28**, (2), pp. 103–111
- [4] Putera, S.H.I., Dzafaruddin, S.F., Mohamad, M.: 'MATLAB based defect detection and classification of printed circuit board'. Proc. Int. Conf. on Digital Information and Communication Technology and its Applications, Bangkok, Thailand, May 2012, pp. 115–119
- [5] Ren, S., Lu, L., Zhao, L., et al.: 'Circuit board defect detection based on image processing'. Proc. Int. Congress on Image and Signal Processing (CISP), Shenyang, China, October 2015, pp. 899–903
- [6] Anu, A., Anuncia, M., Rajaji, L.: 'Automatic visual inspection of PCB using CAD information'. Proc. Int. Conf. on Radar, Communication and Computing (ICRCC), Tamil Nadu, India, December 2012, pp. 21–22
- [7] Kumar, S., Iwahori, Y., Bhuyan, M.K.: 'PCB defect classification using logical combination of segmented copper and non-copper part'. Proc. Int. Conf. on Computer Vision and Image Processing, Singapore, 2017, pp. 523–532
- [8] Inoue, H., Iwahori, Y., Kijisirikul, B., et al.: 'SVM based defect classification of electronic board using bag of keypoints'. Proc. Int. Techn. Conf. on Circuits Systems, Computers and Communication, Seoul, Korea, June 2015, pp. 31–34
- [9] Kim, J.H., Cho, H.S.: 'Neural network-based inspection of solder joints using a circular illumination', *Image Vis. Comput.*, 1995, **13**, (6), pp. 479–490
- [10] Acciani, G., Brunetti, G., Fornarelli, G.: 'Application of neural networks in optical inspection and classification of solder joints in surface mount technology', *IEEE Trans. Ind. Inf.*, 2006, **2**, (3), pp. 200–209
- [11] Ong, T.Y., Samad, Z., Ratnam, M.M.: 'Solder joint inspection with multi-angle imaging and an artificial neural network', *Int. J. Adv. Manuf. Technol.*, 2008, **38**, (5–6), pp. 455–462

- [12] Hao, W., Xianmin, Z., Yongcong, K., et al.: 'Solder joint inspection based on neural network combined with genetic algorithm', *Int. J. Light Electr. Opt.*, 2013, **124**, (20), pp. 4110–4116
- [13] Luo, B., Zhang, Y., Yu, G., et al.: 'ANN ensembles based machine vision inspection for solder joints'. Proc. Int. Conf. on Control and Automation, Guangzhou, China, June 2007, pp. 3111–3115
- [14] Kim, T.H., Cho, T.H., Moon, Y.S., et al.: 'Visual inspection system for the classification of solder joints', *Pattern Recognit.*, 1999, **32**, (4), pp. 565–575
- [15] Yun, T.S., Sim, K.J., Kim, H.J.: 'Support vector machine-based inspection of solder joints using circular illumination', *Electr. Lett.*, 2000, **36**, (11), pp. 949–951
- [16] Zeiler, M.D., Fergus, R.: 'Visualizing and understanding convolutional networks'. Proc. Int. Conf. In European Conf. on Computer Vision, Zurich, Switzerland, September 2014, pp. 818–833
- [17] LeCun, Y., Bengio, Y., Hinton, G.: 'Deep learning', *Nature*, 2015, **521**, (7553), p. 436
- [18] Noda, K., Yamaguchi, Y., Nakadai, K.: 'Audio-visual speech recognition using deep learning', *Appl. Intell.*, 2015, **42**, (4), pp. 722–737
- [19] Wu, G., Lu, W., Gao, G., et al.: 'Regional deep learning model for visual tracking', *Neural Comput.*, 2016, **175**, pp. 310–323
- [20] Roh, B., Yoon, C., Ryu, Y., et al.: 'A neural network approach to defect classification on printed circuit boards', *J. Inst. Contr. Robot. Syst.*, 2001, **67**, (10), pp. 1621–1626
- [21] Bay, H., Tuytelaars, T., Van Gool, L., et al.: 'Surf: speeded up robust features'. Proc. Int. Conf. In European Conf. on Computer Vision, Heidelberg, Berlin, May 2006, pp. 404–417
- [22] Caldo, R.B.: 'ANN diagnosis for defect detection and classification in two-layer printed circuit boards using supervised back-propagation algorithm', *Adv. Comput. Commun. Eng. Technol.*, 2016, pp. 577–593
- [23] Krizhevsky, A., Sutskever, I., Hinton, G.E.: 'Imagenet classification with deep convolutional neural networks'. Proc. Int. Conf. in Advances in Neural Information Processing Systems, Lake, Tahoe, USA, December 2012, pp. 1097–1105
- [24] Simonyan, K., Zisserman, A.: 'Very deep convolutional networks for large scale image recognition', arXiv preprint arXiv:1409.1556, 2014
- [25] Szegedy, C., Liu, W., Jia, Y.Q., et al.: 'Going deeper with convolutions'. Proc. Int. Conf. on Computer Vision and Pattern Recognition, Boston, MA, USA, June 2015, pp. 1–9
- [26] Zhang, C., Shi, W., Li, X., et al.: 'Improved bare PCB defect detection approach based on deep feature learning', *J. Eng.*, 2018, **16**, pp. 1415–1420
- [27] Moganti, M., Ercal, F.: 'Automatic PCB inspection systems', *IEEE Potent*, 1995, **14**, (3), pp. 6–10
- [28] Ibrahim, I., Abdul Rahman, S., Abu Bakar, S., et al.: 'A printed circuit board inspection system with defect classification capability', *Int. J. Innov. Manage. Inf. Prod.*, 2012, **3**, (1)
- [29] LeCun, Y., Boser, B., Denker, J. S., et al.: 'Backpropagation applied to handwritten zip code recognition', *Neural Comput.*, 1989, **1**, (4), pp. 541–551
- [30] Chandrasekhar, V., Lin, J., Morere, O., et al.: 'A practical guide to CNNs and Fisher vectors for image instance retrieval', *Int. J. Signal Process*, 2016, **128**, pp. 426–439
- [31] Kong, T., Yao, A., Chen, Y., et al.: 'Hypernet: towards accurate region proposal generation and joint object detection'. Proc. Int. Conf. in Computer Vision Pattern Recognition, Las Vegas, NV, USA, June 2016, pp. 845–853
- [32] Bell, S., Zitnick, C.L., Bala, K., et al.: 'Inside-outside net: detecting objects in context with skip pooling and recurrent neural networks'. Proc. Int. Conf. in Computer Vision Pattern Recognition, Boston, MA, USA, June 2015, pp. 2874–2883
- [33] Hariharan, B., Arbelaez, P., Girshick, R., et al.: 'Hypercolumns for object segmentation and fine-grained localization'. Proc. Int. Conf. in Computer Vision Pattern Recognition, Boston, MA, USA, June 2015, pp. 447–456
- [34] Weibo, H., Peng, W.: 'A PCB dataset for defects detection and classification', arXiv preprint arXiv:1901.08204, 2019
- [35] Zeiler, M. D.: 'ADADELTA: an adaptive learning rate method', arXiv preprint arXiv:1212.5701, 2012
- [36] Lin, T.Y., Cui, Y., Belongie, S., et al.: 'Learning deep representations for ground-to-aerial geolocalization'. Proc. Int. Conf. on Computer Vision and Pattern Recognition, Boston, MA, USA, June 2015, pp. 5007–5015
- [37] Davis, J., Goadrich, M.: 'The relationship between precision-recall and roc curves'. Proc. Int. Conf. on Machine learning (ICML), New York, NY, USA, 2006, pp. 233–240
- [38] Everingham, M., Van Gool, L., Williams, C.K., et al.: 'The Pascal visual object classes (VOC) challenge', *Int. J. Comput. Vis.*, 2010, **88**, (2), pp. 303–338
- [39] Gaidhane, V.H., Hote, Y.V., Singh, V.: 'An efficient similarity measure approach for PCB surface defect detection'. *Pattern Anal. Appl.*, 2018, **21**, (1), pp. 277–289
- [40] Tsai, D.M., Yang, R.H.: 'An eigenvalue-based similarity measure and its application in defect detection'. *Image Vis. Comput.*, 2005, **23**, (12), pp. 1094–1101