Extreme Learning Machine Based Defect Detection for Solder Joints

Liyong Ma[1], Wei Xie[1] Yong Zhang[2], Xijia Feng[1]

[1] School of Information Science and Engineering, Harbin Institute of Technology, China
[2] Department of Civil Engineering, Harbin Institute of Technology, China
maly@hitwh.edu.cn, xw1248@163.com, yz@hit.edu.cn, fengxijiaofficial@outlook.com

Abstract

The quality of solder joints is essential for electronic products, and the detection of defects in solder joints is critical to the quality control of electronic products. A vision inspection is developed to detect defects of solder joints in automatic line. Extreme learning machine is applied to identify defective solder joints from qualified ones. Five low level features and three advanced features are employed as input features. The low-level features include roundness, roughness, entropy, contrast and histogram of oriented gradient. The advanced features include grey-level co-occurrence matrix, local binary pattern, and segmentation-based fractal texture analysis. To solve unbalanced samples problem, Gaussian mixture model based dense estimation scheme is proposed to adjust the classification super plane for extreme learning machine. The experimental results demonstrate that the proposed defect detection method is more efficient than neural network, support vector machines, common extreme learning machine and convolutional neural network-based methods, and it has real-time performance to meet the equirement of the actual production line.

Keywords: Defect detection, Solder joints, Extreme learning machine, Gaussian mixture model

1 Introduction

Electronic components have been widely used in electronic equipment, including mobile phones, digital cameras, sensors, robots, automobiles and so on. Welding is one of the most important tasks in the production process of electronic components connected to electrical equipment. The quality of solder joints is critical to the reliability of electronic products and will determine the quality of electronic products.

During the process of welding, there will be a variety of unqualified solder joints, including air welding, excessive solder, welding bridge, solder tip, split solder and so on. The quality of electronic products is becoming increasingly important due to the high demands for consistency, stability and reliability. The defects of solder joints will lead to potential quality problems of electronic products, so it is of great significant to realize the automatic detection of solder joints defects. In order to ensure reliable connection and electrical functions, the defects of solder joints need to be detected through quality checks during the production process.

The quality inspection of solder joints was mainly accomplished by manual work at the early developing stage of electronic devices. With the development of technology, automatic detection methods gradually appeared. A variety of checking methods were reported in the past several decades, including radiographic testing, thermography testing, ultrasound testing and so on [1-5]. Some of these inspections can detect defects inside the solder joint. Electrical detection used a variety of electrical measuring instruments to detect poor conduction and thermal deterioration in components, and it was off-line detection.

Because visual inspection has the advantages of real-time and low cost, it has been widely used in solder joint or other quality detection [6-8]. Image processing methods were used for the detection of solder joint defects [8-13]. These studies focused on positioning, background processing, and feature extraction of defects. Recently, machine learning based methods have been employed in visual inspection of solder joints. Defect detection can be regarded as a classification problem of images. The images are divided into two categories according to whether the solder joints are normal. Of course, the defects can be further classified in detail. Machine learning methods have made great progress in image classification and they can be used to solve defect detection problem of solder joints. Neural network was used to distinguish shapes of solder joints [14-15]. Support vector machines (SVM) was introduced to solve the problem of weld defects classification [16-17]. With the performance prior to neural network and other methods, SVM has the potential to solve the nonlinear classification problem with a small number of samples without suffering the curse of dimension. Convolution

*Corresponding Author: Wei Xie; E-mail: xw1248@163.com
DOI: 10.3966/160792642020092105025

neural network was used for welding defect classification and solder joint inspection [18-19]. Hybrid learning methods have also been reported, for example, ensemble learning was used for defect classification [20].

With the development in recent years, extreme learning machine (ELM) provides a new method to solve the problem of image classification. ELM is a machine learning method developed on the basis of neural networks. Compared with SVM, ELM not only optimizes fewer parameters, but also has faster calculation speed and better generalization ability, so as to avoid falling into local optimal solution [21-23]. ELM has been widely used in image recognition and vision-based defect detection applications [24-27]. For example, it was employed to detect panel position or fabric defect [28-29]. Besides, it also has some applications in Communication and Internet, like choosing threshold for time estimation of millimeter wave [30], as well as enhancing web caching abilities [31]. Apart from that, many scholars have proposed some novel architectures of ELM recently, and some of them have achieved great performances in classification accuracy [32].

In this paper, ELM is applied to defect detection of solder joints. The contributions of this paper mainly include two aspects. First, an ELM-based real-time defect detection method is applied to solder joint inspection of communication components. Secondly, an ELM method based on Gaussian mixture model density estimation is proposed, which is superior to other learning-based methods for solder joint defect detection. The rest of this paper is organized as follows. In section 2 a product background of solder joint defect detection will be introduced. The details of GMM (Gaussian Mixture Model) density estimation-based ELM for quality evaluation of solder joint will be described in section 3. The performance evaluation and discussions will be provided in section 4. Finally, this paper is concluded in section 5.

2 Solder Joints Defect Detection

Recently, an automatic line for solder joints detection was established in Tong Sheng Electronics Technology Corporation, Weihai, China. The solder joints of an electronic product are checked with machine vision approach, and the product is one component in a communication device which transmits data over long distances with optical fiber in synchronous digital hierarchy mode. The images of solder joints in print circuit boards are captured from the automatic line. These components to be detected are placed in a specially designed assembled installation to ensure that they are at the same level and that the height error was not greater than 0.1 millimeter as shown in Figure 1.

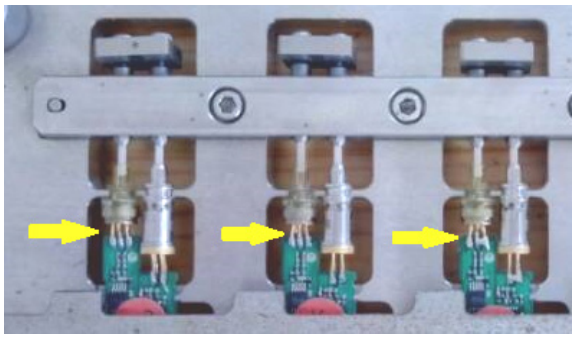

Figure 1. Components to be detected are placed in an assembled installation

Visual inspection of solder joints is performed on the printed circuit board, including pin and pin header solder joint inspection. Typical requirements of solder joint appearance include: (1) the shape is convex, glossy, plump and smooth; (2) no cracks, pinholes and slag inclusion; (3) soldering tin is located in a pad and has good contact with the pad.

Missing component, open solder, excess solder, solder bridge, solder tip, and split solder are the most common types of flaws. Various detection methods have been developed to deal with different flaws. With filtering and threshold segmentation in a solder joint image, the debris can be clearly detected and positioned. The problem of solder bridge, solder joint shift, open solder and excess solder can be accurately detected by means of threshold-based segmentation and the judgement of center of gravity position. However, there is no perfect solution to detect solder tip, solder scanty and split solder of solder joints. The sample images of some defects are shown in Figure 2. This paper will focus on the inspection of these defects.

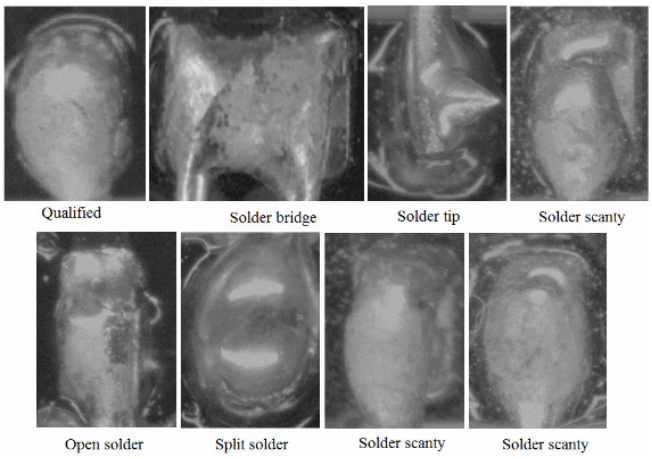

Figure 2. Samples of qualified and defected solder joints

3 ELM Based Defect Detection

The defect detection problem is to judge whether the product has defect in its solder joints, so it can be regarded as a binary classification problem. ELM method has a widespread application in the field of

binary classification. ELM is used to solve the defect detection problem in this work.

3.1 Extreme Learning Machine Algorithm

ELM algorithm has two steps using neural network architecture. Firstly, input data are mapped into the hidden layer. Secondly, the middle results are multiplied with their corresponding weights to get the final output. Different from other traditional three-layer neural networks using parameters adjusting method iteratively, the hidden layer of ELM is designed non-parametric. This obtains the smallest training error and the smallest norm of weights. It has been found that ELM has better generation performance than single hidden layer neural network. The ELM model is illustrated in Figure 3.

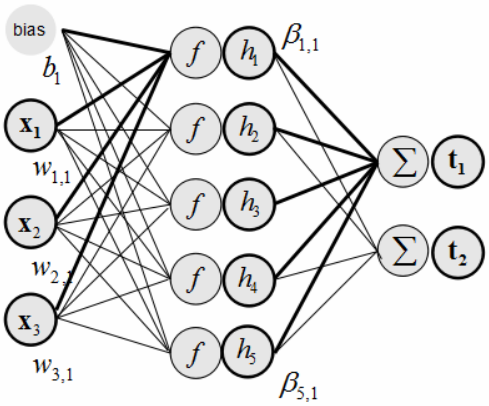

Figure 3. ELM Model

The input feature vectors are denoted as $\mathbf{x}=[\mathbf{x}_1,\mathbf{x}_2,...,\mathbf{x}_N]^T\in\mathbb{R}^{D\times N}$, where N is the sample number, D is the dimensional size of the input feature vectors. The corresponding labels of the input vectors are denoted as $\mathbf{t}_i=[t_{i1},t_{i2},...,t_{iM}]^T\in\mathbb{R}^M$, where $i=1,...,N,M$ is the number of nodes in network output layer. Then the training sample is $(\mathbf{x}_i,\mathbf{t}_i)$.

The output of i-th hidden node is

$$g(\mathbf{x};\mathbf{w}_i,b_i)=g(\mathbf{x}\cdot\mathbf{w}_i+b_i). \quad (1)$$

where $\mathbf{w}_i$ is the input weight vector between the i-th hidden nodes and all input nodes, g is the activation function, and b_i is the bias of this node. In ELM, the input weights $\mathbf{w}_i$ and biases b_i are randomly decided. Let L be the number of hidden layer nodes. A feature mapping function is used to connect the input layer and the hidden layer, which is defined as

$$\mathbf{h}(\mathbf{x})=\left[g(\mathbf{x};\mathbf{w}_1,b_1),g(\mathbf{x};\mathbf{w}_2,b_2),...,g(\mathbf{x};\mathbf{w}_L,b_L)\right]. \quad (2)$$

Let β_{ij} be the output weight between the i-th hidden node and the j-th output node, $i=1,...,L$, $j=1,...,M$. The output of j-th node can be obtained as

$$f_j(\mathbf{x})=\sum_{i=1}^{L}\beta_{ij}\times g(\mathbf{x};\mathbf{w}_i,b_i). \quad (3)$$

The actual output vector is denoted as $\mathbf{Y}$, the input vector is denoted as $\mathbf{X}$. The output vector is represented as

$$\mathbf{Y}=\mathbf{H}\boldsymbol{\beta}, \quad (4)$$

where

$$\mathbf{H}=\begin{bmatrix}\mathbf{h}(\mathbf{x}_1)\\ \vdots\\ \mathbf{h}(\mathbf{x}_N)\end{bmatrix}=\begin{bmatrix}g(\mathbf{x}_1;\mathbf{w}_1,b_1) & \cdots & g(\mathbf{x}_1;\mathbf{w}_L,b_L)\\ \vdots & \ddots & \vdots\\ g(\mathbf{x}_N;\mathbf{w}_1,b_1) & \cdots & g(\mathbf{x}_N;\mathbf{w}_L,b_L)\end{bmatrix}, \quad (5)$$

and

$$\boldsymbol{\beta}=\begin{bmatrix}\beta_1\\ \vdots\\ \beta_L\end{bmatrix}=\begin{bmatrix}\beta_{11} & \cdots & \beta_{1M}\\ \vdots & \ddots & \vdots\\ \beta_{L1} & \cdots & \beta_{LM}\end{bmatrix}. \quad (6)$$

The optimization problem of ELM is to minimize the training error $\|\mathbf{T}-\mathbf{H}\boldsymbol{\beta}\|^2$ and the norm of output weight $\|\boldsymbol{\beta}\|$.

When the matrix $(\mathbf{I}/C)+\mathbf{H}^T\mathbf{H}$ is not singular, the solution of $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta}=\left(\frac{\mathbf{I}}{C}+\mathbf{H}^T\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{T}. \quad (7)$$

Otherwise, the solution of $\boldsymbol{\beta}$ is

$$\boldsymbol{\beta}=\mathbf{H}^T\left(\frac{\mathbf{I}}{C}+\mathbf{H}\mathbf{H}^T\right)^{-1}\mathbf{T}, \quad (8)$$

where $\mathbf{I}$ is an identity matrix. The tradeoff between the closeness to the training data and the smoothness of the decision function can be controlled by the regularization factor C. More details of ELM theory and practical calculation can be found in [21-23].

3.2 Feature Selection for ELM

For the solder joint defect detection application on the production line, the solder joint image is obtained through a camera. And the fixed size solder joint image can be obtained by simple ROI extraction because the solder joint and the component position are fixed. Then mean filter is applied to decrease the surface reflective disturbances of joint solders.

Feature selection is important for ELM applications. After many experiments of different features, 5 low

level features and 3 advanced features are selected for defect detection of solder joints. The low-level features are roundness, roughness, entropy, contrast and histogram of oriented gradients (HOG) [33]. When they are used together as ELM input of samples, better results can be achieved. Because solder joint defect can be observed from the texture, geometrical, shape and other features in the image. These features are calculated as follows and are normalized into the range of $[-1,1]$.

Roundness is defined as

$$f_{roundness}=\frac{P^2}{A}, \tag{9}$$

where P, A is edge girth and area of the region, respectively.

Roughness is calculated as

$$f_{roughness}=\frac{1}{T}\sum_{i=1}^{T}|d(i)-d(i+1)|, \tag{10}$$

where T is the total point number of edge, $d(i)$ is the standardized radius of the i-th edge point.

Entropy is obtained with

$$f_{entropy}=\sum p(i,j)\times\left[-\ln p(i,j)\right], \tag{11}$$

$p(i,j)$ is the element of position (i,j) in grey level co-occurrence matrix.

Contrast is obtained as

$$f_{contrast}=(i-j)^2\times p(i,j). \tag{12}$$

HOG is obtained as described in [33]. The image is divided into non-overlapping blocks with the size of 5×5. Two histograms are calculated in each block, one histogram is cumulative in the direction with the angle of $0^\circ-180^\circ$, and another one with the direction angle of $0^\circ-360^\circ$.

The above 5 low-level features can reflect some of the features of the image, but they are not enough to distinguish the defects of solder joints. Therefore, the following 3 advanced features are also applied to the detection of solder joint defects in this paper.

Grey-level co-occurrence matrix (GLCM) is a matrix describing the gray-scale relationship between a pixel in a local or global area of an image and adjacent pixels or within a certain distance. This is the number of times a pair of pixels of a shape in a gray image appears in the whole image, and it is a texture analysis method.

Local binary pattern (LBP) has obvious advantages such as gray scale invariance, rotation invariance and fast computation. It is used to measure and extract local texture information of images.

$$LBP=\sum_{P=0}^{P-1}s(g_p-g_c)2^P, \tag{13}$$

$$s(x)=\begin{cases}1 & x\geq0\\ 0 & x<0\end{cases}, \tag{14}$$

where g_c is the gray value of center pixel, g_p is the values of center neighbors, and P is the number of neighbors.

Segmentation-based fractal texture analysis (SFTA) automatically divides several binary thresholds according to the characteristics of the image, carries out texture analysis on the binary thresholds, obtains the fractal dimension of the texture.

All these low-level features and advanced features are employed for defect recognition of solder joints in this paper.

3.3 GMM Density Estimation for ELM

As the defect of solder joints has various shapes in different positions, it is impossible to provide all kinds of defect samples. Defect detection is an unbalanced classification problem, and it is often solved by using weighted ELM [34] or normal density estimation-based ELM [35]. However, in this defect detection application, the defect samples are not enough for binary classification with weighted ELM. Meanwhile, in the density estimation method, Gaussian distribution requires a very strict unimodal and convex data model of the samples. However, these assumptions are violated in our application due to the small defect samples. The small samples can be accurately modeled as Gaussian mixture model instead of normal model. A Gaussian mixture model density estimation based unbalanced ELM algorithm is developed for defect detection of solder joints in this paper. Because the defect samples without enough samples make the classification boundary close to the defect class, so some new defect test samples are incorrectly recognized as normal class. The solution in this paper is to estimate the classification boundary of ELM method employing GMM, then the classification boundary is pushed close to the normal class to make the new defect test samples are correctly recognized.

Let F be a family probability density functions on dataset R^D, the function in F is the mixture of Gaussian functions f as an isotropic Gaussian function which is called a component

$$f(\mathbf{x};\theta)=\sum_{l=1}^{K}p_l g(\mathbf{x};m_l,\sigma_l), \tag{15}$$

$$g(\mathbf{x};m_l,\sigma_l)=\frac{1}{(\sqrt{2\pi}\sigma_l)^D}\exp\left(-\frac{1}{2}\left(\frac{\|\mathbf{x}-m_i\|^2}{\sigma_l}\right)^2\right), \tag{16}$$

where $p_l \geq 0$, $\sum_{l=1}^{K} p_l = 1$, and $\theta = ((p_1, m_1, \sigma_1), ..., (p_K, m_K, \sigma_K))$ is a vector containing the mixing probabilities p_l as well as the means m_l, and standard deviation σ_l of the k Gaussian functions in the mixture.

This estimation is solved with expectation-maximization (EM) algorithm, as no parameter is set to influence the optimization algorithm. EM algorithm finds a local maximum for the log-likelihood function $\log f(\mathbf{x};\theta)$ by alternatingly maximizing an auxiliary function over the parameters θ. Denote n as the iteration number, this maximization is called the expectation step as the assignments

$$\alpha_l^{(n)} = \frac{p_l^{(n)} g(\mathbf{x}; m_l, \sigma_l)}{\sum_{k=1}^{K} p_k^{(n)} g(\mathbf{x}; m_k \sigma_k)}. \quad (17)$$

Maximization step results in the parameter estimation with the following update rules

$$m_l^{(n+1)} = \frac{1}{N} \sum_{k=1}^{N} m_k^{(n)}, \quad (18)$$

$$\sigma_l^{(n+1)} = \frac{\sum_{k=1}^{N} \sigma_k^{(n)} \mathbf{x}_k}{\sum_{k=1}^{N} \sigma_k^{(n)}}. \quad (19)$$

More details about EM for GMM parameters estimation can be found in [36]. The component number K is determined as the one which gives the best performance on the dataset. The initial parameters are decided by performing k-means on the data set. The output of ELM can approximate the posterior probability of Naive Bayesian classifier [35]. Consequently, the probability density function can be estimated from the predictive output of qualified class and flaw class after ELM training. The GMM a and b those corresponding to qualified class with 2 components of Gaussian model and flaw class with 3 components respectively are showed in Figure 4. Then the intersection position s can be determined with the estimated GMM of the two classes. When the number of intersections is more than one, we merely choose the one as the selected intersection s which obtains the greatest sum of the area under curve a and b after the intersection position breaks the two curves. We deem s the optimal breakpoint for the binary classification. In our defect detection application, which is lack of defect samples, the sample size of these two classes is imbalanced. So, the intersection position needs to be pushed to the qualified class as showed in Figure 4.

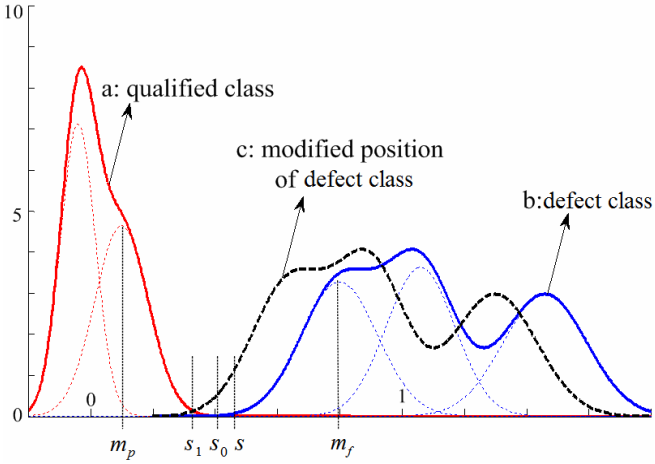

Figure 4. GMM of qualified class and defect class

Mark the mean of the component which is closest to s in qualified class GMM as m_p, and the one in defect class as m_f. Without loss of generality, assume that $m_p < s < m_f$. Let $L = m_f - m_p$, $s_0 = (m_f + m_p)/2$, and $\lambda = (m_f - s)/L$. When the samples of the two classes are enough to provide entire distribution information, s_0 is a good solution for binary classification. However, when defect samples are lacking to supply distribution information, s_0 is required being pushed to a new position as

$$s_1 = m_p + (s_0 - m_p)(1 - \exp^{-\lambda \log(n)}), \quad (20)$$

$$\Delta s = s_1 - s_0, \quad (21)$$

where n is the total number of defect samples during training. When λ is const and n is greater, the new classification position s_1 is farther from m_p and is closer to s_0. When n is constant and s is smaller, which means that samples are enough to provide entire information of the defect class, λ is greater and s_1 is closer to s_0.

3.4 ELM for Defect Detection

In the proposed ELM network, the input feature vector is used as the input for ELM. This input feature vector includes two parts, those are feature vector and pixel vector. The low-level features and advanced features are arranged in a row to get a feature vector. The pixels in regions like defect with the size of 48×48 are also arranged in a row to get a pixel vector. The feature vector and the pixel vector are joined together to form a row input feature vector. This feature vector can be feed into ELM network as input for training or test. The output layer has one node, the value of the output is -1 when there is no defect in solder joint, and otherwise, the value is 1.

During the sample training, the input weights w_i and biases b_i in ELM are randomly selected. After training an unweighted ELM, we obtain the predictive output for each training sample. The two GMMs of qualified class and flaw class are calculated and the only intersection s is determined. We update the predictive output $f_k(x)$ of qualified class and defect class with the following rules

$$f_k(x) \leftarrow f_k(x) - \Delta s, \tag{22}$$

where $1 \le k \le K$. After employing ELM algorithm again to the updated training dataset, we can obtain the result classifier.

4 Results and Discussions

In order to test ELM based defect detection method, a dataset is set up. The images in the data set are collected from the production line. The camera used is Basler acA2500-14gm GigE. The captured image is preprocessed by median filter. Because the standard solder joint is a semi-ellipsoid with smooth surface and the defective solder joint is an irregular shape, the captured image is segmented by watershed algorithm, and the solder joint contour can be obtained. Finally, the segmented contour image is transformed into a uniform size image as sample image. There are 4200 images of qualified solder joints and 360 images of defective solder joints collected from the industry field. The proposed ELM method is cross-validated by dividing the dataset stochastically into two groups. Half of the data set is used for training and the other half for testing. That is to say, there are 2100 images of qualified solder joints and 180 images of unqualified solder joints in the training and testing data sets.

Some other machine learning based methods are employed to compare the performance of different algorithms. There algorithms include neural network (NN) [14], SVM [16], an ELM method applying HOG feature (aELM) [25], weighted ELM (wELM) [34], probability density estimation-based ELM (pELM) [35], convolutional neural network (CNN) [18], and cascaded convolutional neural network (cCNN) [19]. The parameters of NN, ELM and CNN are the same as in the references. The parameters of SVM are optimized for the best classification performance using grid search. Our experiments show that after the number of hidden layers is greater than 400, the results obtained are similar for our proposed method. The number of hidden layer nodes is set to 400 for the proposed method in this paper.

4.1 Performance

According to whether the classification results are correct, TP, TN, FP, and FN can be determined. TP means that the classification result is true and positive. Similarly, TN means true negative, and so on.

Recall, precision, accuracy, specificity and F1-score are employed as classification performance indicators to evaluate different methods. They are defined as follows [37].

$$\text{Recall} = \frac{TP}{TP + FN}. \tag{23}$$

$$\text{Precision} = \frac{TP}{TP + FP}. \tag{24}$$

$$\text{Specificity} = \frac{TN}{FP + TN}. \tag{25}$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}. \tag{26}$$

$$\text{F1-score} = \frac{2TP}{2TP + FP + FN}. \tag{27}$$

Recall measures the proportion of actual positives that are correctly identified as such. Specificity represents the proportion of actual negatives that are correctly identified as such. Accuracy is defined as the proportion of all samples that have been successfully classified. Precision is the ratio of samples correctly classified as positive to all the samples which are classified. F1-score is the harmonic mean of precision and sensitivity. When the above performance index is greater, the classification performance is better.

The above mentioned five indicators of different defect detection methods are listed in Table 1. The proposed method in this paper has the greatest value in all the performance indicators. It means that the proposed GMM density estimation-based ELM method is superior to other classification methods for defect detection of solder joints. The performance comparison is also shown in Figure 5. It indicates that the proposed method has the best classification performance. This also reveals that GMM is efficient for the estimation of unbalanced classification boundary. When GMM is employed to ELM method, the classification boundary is pushed close to the normal class to make the defect samples are correctly recognized.

Table 1. Classification performance of different methods for defect detection of solder joint

Method	NN	SVM	aELM	wELM	pELM	CNN	cCNN	Proposed
Recall	0.888	0.902	0.932	0.949	0.958	0.948	0.950	0.988
Precision	0.950	0.954	0.967	0.975	0.977	0.973	0.972	0.989
Accuracy	0.853	0.870	0.909	0.931	0.941	0.928	0.928	0.979
Specificity	0.453	0.497	0.631	0.714	0.733	0.689	0.678	0.872
F1-score	0.918	0.928	0.949	0.962	0.967	0.960	0.961	0.989

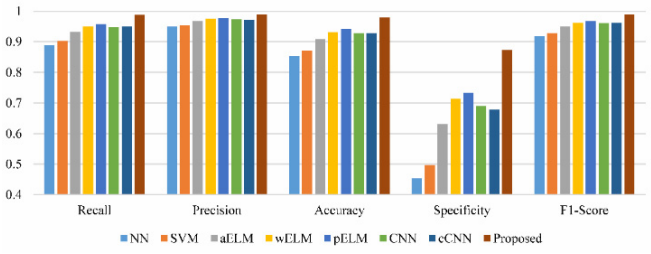

Figure 5. Classification performance comparison of different methods

4.2 Confusion Matrix

Confusion matrix is often used to visualize the performance of supervised learning-based classification. The matrix row represents samples in a predicted class while matrix column indicates the samples in an actual class [37-38]. Confusion matrices of different methods are illustrated in Figure 6.

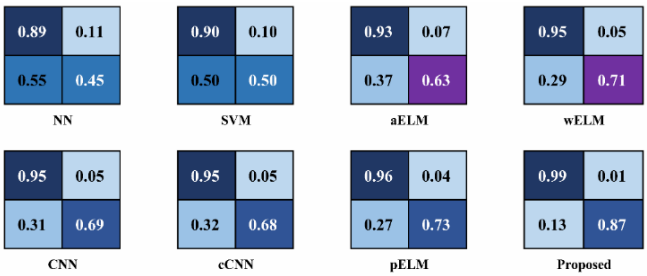

Figure 6. Confusion matrices of different methods

The confusion matrix of the proposed method obtains the maximum value on the main diagonal and the minimum value in the secondary diagonal. It shows that the proposed method in this paper has the best classification performance. This is consistent with the analysis results of the performance data in 4.1. This also indicates that our proposed method is the most efficient one for defect detection of solder joints.

4.3 Realtime

All experiments are performed on an industrial field computer, its memory is 8G, and CPU is 2.66GHz. The test time of each sample is about 0.029-0.049s. Considering the image acquisition time, the detection time of each component is less than 1 second. The average time with manual detection is about 5 seconds. In the actual production line, the component studied in this paper requires an average of 1 second to complete the solder joint detection. The method proposed in this paper can meet the real-time requirement of the actual production line.

5 Conclusions

In this paper, a visual inspection application for defect detection of solder joints is developed. The following conclusions can be drawn from this paper.

(1) Gaussian mixture model density estimation is efficient for the unbalanced classification boundary.

(2) ELM method is efficient for defect detection of solder joints based on Gaussian mixture model density estimation, and it has real-time performance to meet the requirement of the actual production line.

(3) Compared with neural network, support vector machine, and common ELM based methods, the proposed ELM method has superior classification performance.

Acknowledgments

This work was supported by National Key R&D Program of China (2018YFC0114802), Shandong Province Natural Science Foundation (ZR2018MF026), Shandong Province Key R&D Program (2019GGX10 1054, 2019GSF111062, 2018CXGC0925) and University Co-construction Project at Weihai (ITDAZMZ001708).

References

[1] P. Scott, E. Krastev, 2D X-ray Inspection with Materials and Thickness Identification, *SMT Surface Mount Technology Magazine*, Vol. 32, No.6, pp. 70-79, June, 2017.

[2] X. Zhou, Y. Xue, G. Tian, Z. Liu, Thermal Analysis of Solder Joint Based on Eddy Current Pulsed Thermography, *IEEE Transactions on Components, Packaging and Manufacturing Technology*, Vol. 7, No. 7, pp. 1111-1118, July, 2017.

[3] X. Zhou, Y. Chen, X. Lu, Detection of Small-size Solder Ball Defects through Heat Conduction Analysis, *Review of Scientific Instruments*, Vol. 89, No. 2, Article ID 024905, February, 2018.

[4] L. Su, G. Liao, T. Shi, Y. Zhang, Intelligent Diagnosis of Flip Chip Solder Bumps Using High-frequency Ultrasound and a Naive Bayes Classifier, *Insight: Non- Destructive Testing and Condition Monitoring*, Vol. 60, No. 5, pp. 264-269, May, 2018.

[5] R. Murta, F. Vieira, V. Santos, E. Moura, Welding Defect Classification from Simulated Ultrasonic Signals, *Journal of Nondestructive Evaluation*, Vol. 37, No. 3, Article ID 40, September, 2018.

[6] G. Kumar, U. Natarajan, S. Ananthan, Vision Inspection System for the Identification and Classification of Defects in MIG Welding Joints, *International Journal of Advanced Manufacturing Technology*, Vol. 61, No. 9-12, pp. 923-933, August, 2012.

[7] G. Acciani, G. Brunetti, G. Fornarelli, Application of Neural Networks in Optical Inspection and Classification of Solder Joints in Surface Mount Technology, *IEEE Transactions on Industrial Informatics*, Vol. 2, No. 3, pp. 200-209, August, 2006.

[8] L. Ma, W. Xie, Y. Zhang, Blister Defect Detection Based on Convolutional Neural Network for Polymer Lithium-ion Battery, *Applied Sciences*, Vol. 9, No. 6, Paper ID. 1086, March, 2019.

[9] J. Song, Y. Kim, T. Park, Defect Classification Method of

PCB Solder Joint by Color Features and Region Segmentation, *Journal of Institute of Control, Robotics and Systems*, Vol. 23, No. 12, pp. 1086-1091, December, 2017.

[10] L. Bai, X. Yang, H. Gao, A Novel Coarse-fine Method for Ball Grid Array Component Positioning and Defect Inspection, *IEEE Transactions on Industrial Electronics*, Vol. 65, No. 6, pp. 5023-5031, June, 2018.

[11] N. Cai, J. Lin, Q. Ye, H. Wang, S. Weng, B. Ling, A New IC Solder Joint Inspection Method for an Automatic Optical Inspection System Based on an Improved Visual Background Extraction Algorithm, *IEEE Transactions on Components, Packagingand Manufacturing Technology*, Vol. 6, No. 1, pp. 161-172, January, 2016.

[12] J. Ieamsaard, P. Muneesawang, F. Sandnes, Automatic Optical Inspection of Solder Ball Burn Defects on Head Gimbal Assembly, *Journal of Failure Analysis and Prevention*, Vol. 18, No. 2, pp. 435-444, April, 2018.

[13] L. Ma, W. Xie, Y. Zhang, Blister Defect Detection Based on Convolutional Neural Network for Polymer Lithium-ion Battery, *Applied Sciences*, Vol. 9, No. 6, Article ID. 1086, March, 2019.

[14] J. Kim, H. Cho, Neural Network-based Inspection of Solder Joints Using a Circular Illumination, *Image and Vision Computing*, Vol. 13, No. 6, pp. 479-490, August, 1995.

[15] T. Ong, Z. Samad, M. Ratnam, Solder Joint Inspection with Multi-angle Imaging and an Artificial Neural Network, *International Journal of Advanced Manufacturing Technology*, Vol. 38, No. 5-6, pp. 455-462, August, 2008.

[16] W. Mu, J. Gao, H. Jiang, Z. Wang, F. Chen, C. Dang, Automatic Classification Approach to Weld Defects Based on PCA and SVM, *Insight: Non-Destructive Testing and Condition Monitoring*, Vol. 55, No. 10, pp. 535-539, October, 2013.

[17] Y. Chen, H. Ma, M. Dong, Automatic Classification of Welding Defects from Ultrasonic Signals Using an SVM-Based RBF Neural Network Approach, *Insight: Non-Destructive Testing and Condition Monitoring*, Vol. 60, No. 4, pp. 194-199, April, 2018.

[18] A. Khumaidi, E. Yuniarno, M. Purnomo, Welding Defect Classification Based on Convolution Neural Network (CNN) and Gaussian Kernel, *2017 18th International Seminar on Intelligent Technology and Its Application (ISITIA)*, Surabaya, Indonesia, 2017, pp. 261-265.

[19] N. Cai, G. Cen, J. Wu, F. Li, H. Wang, X. Chen, SMT Solder Joint Inspection via a Novel Cascaded Convolutional Neural Network, *IEEE Transactions on Components, Packaging and Manufacturing Technology*, Vol. 8, No. 4, pp. 670-677, April. 2018.

[20] H. Wu, Solder Joint Defect Classification Based on Ensemble Learning, *Soldering and Surface Mount Technology*, Vol. 29, No. 3, pp. 164-170, March, 2017.

[21] G. Huang, Q. Zhu and C. Siew, Extreme Learning Machine: Theory and Applications, *Neurocomputing*, Vol. 70, No. 1-3, pp. 489-501, December, 2006.

[22] S. Ding, H. Zhao, Y. Zhang, X. Xu, R. Nie, Extreme Learning Machine: Algorithm, Theory And Applications, *Artificial Intelligence Review*, Vol. 44, No. 1, pp. 103-115, June, 2015.

[23] G. Huang, G. Huang, S. Song, K. You, Trends in Extreme Learning Machines: A Review, *Neural Networks*, Vol. 61, No. 1, pp. 32-48, January, 2015.

[24] S. Wang, C. Deng, W. Lin, G. Huang, B. Zhao, NMF-based Image Quality Assessment Using Extreme Learning Machine, *IEEE Transactions on Cybernetics*, Vol. 47, No. 1, pp. 232-243, January, 2017.

[25] Z. Huang, Y. Yu, J. Gu, H. Liu, An Efficient Method for Traffic Sign Recognition Based on Extreme Learning Machine, *IEEE Transactions on Cybernetics*, Vol. 47, No. 4, pp. 920-933, April, 2017.

[26] L. Ma, L. Tang, W. Xie, S. Cai, Ship Detection in SAR Using Extreme Learning Machine, *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering*, Vol. 227, No 1, pp. 558-568, January, 2017.

[27] Y. Park, H. Yang, Convolutional Neural Network Based on an Extreme Learning Machine for Image Classification, *Neurocomputing*, Vol. 339, No. 66-76, April, 2019.

[28] H. Liu, C. Zhang, D. Huang, Extreme Learning Machine and Moving Least Square Regression Based Solar Panel Vision Inspection, *Journal of Electrical and Computer Engineering*, Vol. 2017, No.5, Article ID. 7406568, May, 2017.

[29] Z. Zhou, C. Wang, X. Gao, Z. Zhu, X. Hu, X. Zheng, L. Jiang, Fabric Defect Detection and Classifier via Multi-scale Dictionary Learning and an Adaptive Differential Evolution Optimized Regularization Extreme Learning Machine, *Fibres and Textiles in Eastern Europe*, Vol. 27, No. 1, pp. 67-77, January, 2019.

[30] X. Liang, H. Zhang; T. Lu, T. A. Gulliver, Extreme Learning Machine for 60 GHz Millimetre Wave Positioning, *IET Communications*, Vol. 11, No. 4, pp. 483-489, March, 2017.

[31] P. Imtongkhum, C. So-In, S. Sanguanpong, S. Phoemphon, A Two-Level Intelligent Web Caching Scheme with a Hybrid Extreme Learning Machine and Least Frequently Used, *Journal of Internet Technology*, Vol. 19, No. 3, pp. 725-740, May, 2018.

[32] R. Li, X. Wang, L. Lei, C. Wu, Representation Learning by Hierarchical ELM Auto-encoder with Double Random Hidden Layers, *IET Computer Vision*, Vol. 13, No. 4, pp. 411-419, June, 2019.

[33] N. Dalal, B. Triggs, Histograms of Oriented Gradients for Human Detection, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, San Diego, United States, 2005, pp. 886-893.

[34] W. Zong, G. Huang, Y. Chen, Weighted Extreme Learning Machine for Imbalance Learning, *Neurocomputing*, Vol. 101, pp. 229-242, February, 2013.

[35] J. Yang, H. Yu, X. Yang, X Zuo, Imbalanced Extreme Learning Machine Based on Probability Density Estimation, *Lecture Notes on Computer Science*, Vol. 9426, pp. 160-167, November, 2015.

[36] R. Redner, H. Walker, Mixture Densities, Maximum Likelihood and the EM Algorithm, *SIAM Review*, Vol. 26, No. 2, pp. 195-239, April, 1984.

[37] L. Ma, C. Ma, Y. Liu, X. Wang. Thyroid Diagnosis from

SPECT Images Using Convolutional Neural Network with Optimization, *Computational Intelligence and Neuroscience*, Vol. 2019, No. 1, Paper ID. 6212759, January, 2019.

[38] D. Power, Evaluation: from Precision, Recall and F-factor to ROC, Informedness, Markedness & Correlation, *Journal of Machine Learning Technology*, Vol. 2, No. 1, pp. 37-63, January, 2011.

Biographies

Liyong Ma is currently an Associate Professor at Harbin Institute of Technology, Weihai, China. His main research areas include machine learning, intelligent information processing, biomedical imaging and image processing.

Wei Xie is currently a Lecture at Harbin Institute of Technology, Weihai, China. His main research areas include robot, machine learning, intelligent information processing.

Yong Zhang is currently an engineer at Harbin Institute of Technology, Weihai, China. His main research areas include machine learning, intelligent information processing, and civil engineering.

Xijia Feng is a senior undergraduate from School of Information Science and Engineering, Harbin Institute of Technology, Weihai. She is currently interested in areas of computer vision, machine learning, robotics and control.