A Feature Engineering-based Method for PCB Solder Paste Position Offset Prediction

1st Binkun Liu
Department of Automation, University of Science and Technology of China
Hefei,China
Anhui Engineering Research Center for Intelligent Applications and Security of Industrial Internet,
Anhui University of Technology
Ma'anshan, China
liubink@mail.ustc.edu.cn

2nd Yunbo Zhao
Department of Automation, University of Science and Technology of China
Institute of Artificial Intelligence, Hefei Comprehensive National Science Center
Institute of Advanced Technology, University of Science and Technology of China
Hefei, China
ybzhao@ustc.edu.cn

3rd Yu Kang
Department of Automation, University of Science and Technology of China
Institute of Artificial Intelligence, Hefei Comprehensive National Science Center
Institute of Advanced Technology, University of Science and Technology of China
Hefei, China
kangduyu@ustc.edu.cn

4th Yang Cao
Department of Automation, University of Science and Technology of China
Institute of Artificial Intelligence, Hefei Comprehensive National Science Center
Hefei, China
forrest@ustc.edu.cn

5th Peng Bai
Department of Automation, University of Science and Technology of China
Hefei, China
baipeng@lenovo.com

6th Zhenyi Xu
Institute of Artificial Intelligence, Hefei Comprehensive National Science Center
Hefei, China
Anhui Engineering Research Center for Intelligent Applications and Security of Industrial Internet,
Anhui University of Technology
Ma'anshan, China
xuzhenyi@mail.ustc.edu.cn

2023 6th International Symposium on Autonomous Systems (ISAS) | 979-8-3503-1615-5/23/$31.00 ©2023 IEEE | DOI: 10.1109/ISAS59543.2023.10164303

***Abstract*—Solder paste printing position offset is a common type of defective printed circuit boards (PCBs) printing, and accurate position offset prediction helps to avoid the production of defects, thus improving efficiency. The existing methods mainly use the powerful nonlinear fitting ability of deep learning to learn the variation pattern of solder paste printing quality to achieve a good prediction. However, factories also focus on the interpretability of the model, and existing methods are difficult to give the basis for decisions, so there are still limitations in the practical application. To solve this problem, we propose a Support vector machine (SVM) approach, in which we manually design 14 statistical features based on the original data, then the resampling reduces the effect of data imbalance and achieves PCB pad-level offset prediction. Finally, we verified about six-day of real solder paste printing production data and achieved good experimental results.**

***Index Terms*—PCB, Feature Engineering, Time series prediction**

I. Introduction

With the booming development of the electronic information industry [1], printed circuit boards (PCBs) as one of the key parts of the electronic information industry [2] [3], efficient and stable production is more and more important. The solder paste printing process is a key process step in PCB production [4] [5], which directly determines whether the electrical function of the PCB is normal. Solder paste position offset as a typical PCB printing defect type [6] [7] [8], if the prediction of the solder paste position offset can be achieved, then the position offset defect can be avoided by adjusting the position compensation parameters of the solder paste printing machine, which reduces the cost of defective product repair and improves the production efficiency [9] [10].

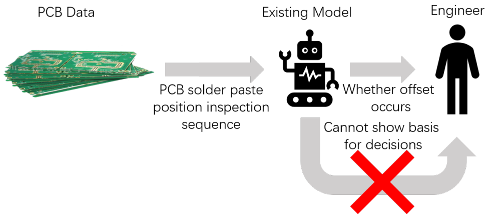
PCB Data
Existing Model
Engineer
PCB solder paste position inspection sequence
Whether offset occurs
Cannot show basis for decisions

Fig. 1: Models are available that can predict more accurately whether a bad paste print offset has occurred, but it is difficult to tell engineers the basis for their judgment.

There are relatively few researches for solder paste position offset prediction [11] [12] [13]. Existing methods utilize deep learning to fit fluctuating position offset time series, and use reweighting to overcome the negative impact caused by far

*This work was supported by the National Natural Science Foundation of China (62173317, 62103124), Major Special Science and Technology Project of Anhui, China (202104a05020064, 202003a07020009), Jingdong Institute Open Fund of Suqian University (2022JDXM14). Corresponding author: Yunbo Zhao (e-mail: ybzhao@ustc.edu.cn).

fewer bad position offsets than good products, to predict whether the offset will be defective in the future period. However, as shown in Fig. 1, these deep learning-based approaches suffer from a shortage of interpretability, which makes it difficult to effectively assist engineers in making predictive judgments, and therefore has major limitations when put into practical industrial production line applications.

To address the above problem, we propose a feature engineering-based PCB solder paste position offset prediction method, aiming to make the method more interpretable by manually designing statistical features that can reflect the characteristics of the offset time series. Specifically, since the number of good products far exceeds the number of defective products, the data is resampled to reduce the percentage of good products. Then, 14 statistical metrics such as mean, variance, kurtosis, skewness, mean of absolute values, and index of maximum absolute values are designed manually. Finally, the manually designed features are nonlinearly transformed by using a support vector machine with a kernel function [14], to achieve the temporal prediction of the offset defect.

In summary, our main contributions are as follows.

1. The 14 hand-designed features are constructed for PCB solder paste position offset prediction in industrial scenarios.

2. Our method is evaluated on about six-day of real solder paste printing production data and the best experimental results are achieved.

II. Related Work

The specific process of solder paste printing is: first set the pressure, speed, X-direction offset compensation, Y-direction offset compensation, cleaning frequency, and other parameters in the solder paste printing machine, and then add the solder paste to the stencil through the squeegee, the mesh of the stencil and the PCB to be printed in the position of one to one correspondence, the solder paste through the mesh to cover the corresponding position. After printing, the finished product is inspected by Solder Paste Inspection (SPI). The inspection items include the absolute value of the solder paste area, the ratio between the area and the standard value, the absolute height value, the ratio between the height and the standard value, the absolute volume value, the ratio between the volume and the standard value, the absolute value of the X-direction offset, the ratio between the X-direction position and the standard value, the absolute value of the Y-direction offset and the ratio between the Y-direction position and the standard value, etc. If any one of them does not meet the requirements, the product will be judged as inferior.

Alelaumi et al. [15] measured the residual amount of solder paste under the stencil and predicted the future residual amount of solder paste on the stencil by LSTM [16] to assist engineers in selecting the appropriate cleaning method. Wang et al. [17] employs a combination of wavelet transform and LSTM to determine the appropriate stencil cleaning method. Alelaumi et al. [18] proposed a new multi-temporal intelligent anomaly prediction (IAP) framework to improve the first pass rate and reduce the rework cost in PCB assembly lines. In the first stage, the highly autocorrelated solder paste printing process is monitored using Random Forest-based Exponentially Weighted Moving Average (RF-Based EWMA), and in the second stage, Adaptive Boosting (AdaBoost) is used to achieve accurate prediction of solder paste printing anomalies before they occur. The above methods suffer from the disadvantage of being difficult to explain the decision.

III. Method

The overall framework of the method is shown in Fig. 2. Initially, we collect the SPI data of PCB to construct the time series of solder paste printing position offset, then build the feature engineering based on the solder paste printing position offset time series, and finally use SVM to classify the feature vector for prediction.

A. Preliminary

PCB Solder Paste Position Offset Sample: Each PCB position offset detection sample $O \in \mathbb{R}^{M\times 2}$, there are M solder paste offset detection results, the ith solder paste detection results $P^i = \left[x^i, y^i\right] \in \mathbb{R}^2, i = 1, 2, \cdots, M$, containing two indicators of X-direction relative offset and Y-direction relative offset.

Solder Paste Position Offset Time Series: Construct the historical offset time series in time order $\mathcal{O} = [O_1, O_2, \cdots, O_t]$. $O_t \in \mathbb{R}^{M\times 2}$ represents the PCB solder paste position offset samples at time t.

Problem: Given a time series $\mathcal{O}$ of solder paste position offsets of time length K, predict whether the offset of each paste will exceed the threshold set by the factory in the next T moments.

B. Data Processing

Using the sliding pane method, the time series of solder paste position offsets $\mathcal{O}$ is divided with a history window size of K, a step size of S for each translation, and a prediction window size of T. This yields a historical window offset sequence $\mathcal{O}^h_{t+K} = [O_{t+1}, O_{t+2}, \cdots, O_{t+K}] \in \mathbb{R}^{K\times M\times 2}$, a predicted window offset sequence $\mathcal{O}^p_{t+K+T} = [O_{t+K+1}, O_{t+K+2}, \cdots, O_{t+K+T}] \in \mathbb{R}^{T\times M\times 2}$. The above historical window offset sequence $\mathcal{O}^h_{t+K}$ and predicted window offset sequence $\mathcal{O}^p_{t+K+T}$ are then decomposed to each solder paste location, i.e., the ith solder paste historical window offset sequence $P^i_{t+K} = \left[P^i_{t+1}, P^i_{t+2}, \cdots, P^i_{t+K}\right] \in \mathbb{R}^{K\times 2}$ and the i-th pad predicted window offset sequence $P^i_{t+K+T} = \left[P^i_{t+K+1}, P^i_{t+K+2}, \cdots, P^i_{t+K+T}\right] \in \mathbb{R}^{K\times 2}$. The history window offset sequence is re-notated as $\mathcal{O}^h_{t+K} = \left[P^1_{t+K}, P^2_{t+K}, \cdots, P^M_{t+K}\right]$. The prediction window offset sequence is re-notated as $\mathcal{O}^p_{t+K+T} = \left[P^1_{t+K+T}, P^2_{t+K+T}, \cdots, P^M_{t+K+T}\right]$. When the i-th solder paste prediction window offset sequence P^i_{t+K+T} has a bad offset in any direction at any moment, then P^i_{t+K} corresponds to a label of 1, otherwise, the label is -1. Considering that there are far more good products than defective products, the proportion of good products in the training set is reduced by resampling and the proportion of defective products is increased.

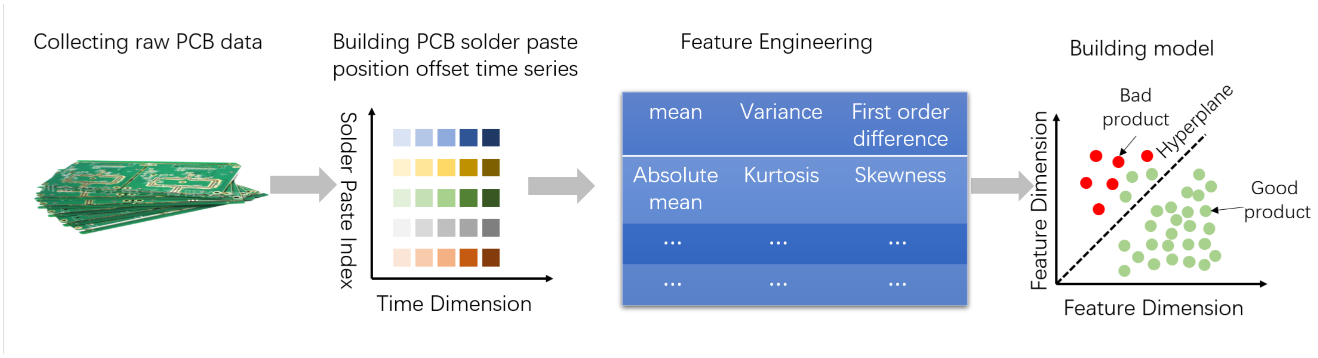

Fig. 2: Overall framework of the methodology.

C. Manual feature construction

TABLE I: FEATURE ENGINEERING DETAILS

Index	Feature	Index	Feature
1	Number of peaks	8	Index of the largest absolute value
2	Skewness	9	root of the mean square
3	Kurtosis	10	Mean of first-order difference
4	Mean of absolute values	11	Mean of first order absolute value difference
5	Mean	12	Mean value of second order difference
6	Standard deviation	13	Timing Data Complexity
7	Maximum absolute value	14	Number of times through the mean

The ith sequence of solder paste history window offsets is denoted as $P^i_{t+K} = \left[X^i_{t+K}, Y^i_{t+K}\right] = \left[x^i_{t+1}, x^i_{t+2}, \cdots, x^i_{t+K} \parallel y^i_{t+1}, y^i_{t+2}, \cdots, y^i_{t+K}\right]$, $\parallel$ means connected. The following 14 statistical features as shown in Tab. 1 are calculated for X^i_{t+K}.

$Peak^{x,i}_{t+K}$ is the number of peaks, i.e. the number of occurrences where $x^i_{t+i} < x^i_{t+i+1}$ and $x^i_{t+i+1} > x^i_{t+i+2}$, or $x^i_{t+i} > x^i_{t+i+1}$ and $x^i_{t+i+1} < x^i_{t+i+2}$. It reflects the degree of time series dithering.

Skewness is a measure of the direction and degree of skewness of a statistical distribution and is a numerical characteristic of the degree of asymmetry of a statistical distribution.

$$Skew^{x,i}_{t+K} = \frac{1}{K} \sum_{j=t+1}^{t+K} \left(\frac{x^i_j - \mu}{\sigma} \right)^3, \tag{1}$$

where μ is the mean of X^i_{t+K} and σ is the variance of X^i_{t+K}.

Kurtosis $Kurt^{x,i}_{t+K}$ is a measure of the height of the peak at the mean of the statistical distribution and reflects the sharpness of the peak.

$$Kurt^{x,i}_{t+K} = \frac{1}{K} \sum_{j=t+1}^{t+K} \left(\frac{x^i_j - \mu}{\sigma} \right)^4, \tag{2}$$

where μ is the mean of X^i_{t+K} and σ is the variance of X^i_{t+K}.

The mean value $Mean^{x,i}_{t+K}$ describes the concentrated trend of the time series of the solder paste position offset.

$$Mean^{x,i}_{t+K} = \frac{1}{K} \sum_{j=t+1}^{t+K} x^i_j. \tag{3}$$

Since the positivity and negativity of the offset represent the direction of the offset, the mean value of the absolute value $Abs_{mean}{}^{x,i}_{t+K}$ is designed to avoid the situation where the mean value is 0, but the offset degree is huge.

$$Abs_{mean}{}^{x,i}_{t+K} = \frac{1}{K} \sum_{j=t+1}^{t+K} \left| x^i_j \right|, \tag{4}$$

where $\left| x^i_j \right|$ denotes the absolute value of x^i_j.

The standard deviation $\mathrm{Std}^{x,i}_{t+K}$ reflects the degree of dispersion in the degree of offset of the solder paste position at different moments.

$$\mathrm{Std}^{x,i}_{t+K} = \sqrt{\frac{1}{K} \sum_{j=t+1}^{t+K} \left(x^i_j - \mu \right)^2}, \tag{5}$$

where μ is the mean value of X^i_{t+K}.

The maximum absolute value $Abs_{max}{}^{x,i}_{t+K}$ characterizes the maximum paste position offset.

$$Abs_{max}{}^{x,i}_{t+K} = Max \left(\left| X^i_{t+K} \right| \right). \tag{6}$$

The index of the maximum absolute value $Arg_{max}{}^{x,i}_{t+K}$ characterizes the impact of the maximum degree of paste position offset on future moments.

$$Arg_{max}{}^{x,i}_{t+K} = arg \left(Max \left(\left| X^i_{t+K} \right| \right) \right). \tag{7}$$

where arg denotes the relative index of the get in the sequence.

$$RMS_{max}{}^{x,i}_{t+K} = \sqrt{\frac{1}{K} \sum_{j=t+1}^{t+K} \left(x^i_j \right)^2}. \tag{8}$$

The difference is used to analyze the smoothness of the time series of solder paste position offsets.

$$Diff_{t+K}^{x,i} = \frac{1}{K-1}\sum_{j=t+1}^{t+K-1}(x_{j+1}^{i} - x_{j}^{i}). \tag{9}$$

$$Abs_{diff}{}_{t+K}^{x,i} = \frac{1}{K-1}\sum_{j=t+1}^{t+K-1}(|x_{j+1}^{i}|-|x_{j}^{i}|). \tag{10}$$

$$Sec_{diff}{}_{t+K}^{x,i} = \frac{1}{K-2}\sum_{j=t+1}^{t+K-2}(x_{j+1}^{i} - 2x_{j+1}^{i} + x_{j}^{i}). \tag{11}$$

The time series data complexity $Cid_{t+K}^{x,i}$ is used to evaluate the complexity of the time series, the more complex the series has more valleys and peaks.

$$Cid_{t+K}^{x,i} = \sqrt{\frac{1}{K-1}\sum_{j=t+1}^{t+K-1}(x_{j+1}^{i} - x_{j}^{i})^{2}}. \tag{12}$$

$Z_{t+K}^{x,i}$ is the number of over-averages, i.e., the number of occurrences where $x_{t+i}^{i} < \mu$ and $x_{t+i+1}^{i} > \mu$, or $x_{t+i}^{i} > \mu$ and $x_{t+i+1}^{i} < \mu$.

The above 14 statistical metrics are repeated for Y_{t+K}^{i}. The final vector F_{t+k}^{i} consisting of 28 statistical metrics is obtained.

D. Model Construction

Construct a support vector machine. Solve the following equations.

$$\begin{aligned} &\underset{w,b,\xi}{min}\frac{1}{2}||w||^{2} + C\sum_{i=1}^{l}s_{n}\xi_{n} \\ &\text{s.t. } y_{i}\left(w\cdot\varphi\left(F_{t+k}^{i}\right)+b\right) \geq 1-\xi_{n}, \end{aligned} \tag{13}$$

where $\xi_n \geq 0$, ξ_n denotes the slack of the nth sample. φ denotes the kernel function, C denotes the penalty term for misclassification, and s_n denotes the classification weight of the nth sample. Here C is set to 1000. Since resampling was performed, the positive and negative samples have the same class weight. The final SVM is obtained.

IV. Experiment

A. Dataset

The data was collected from a laptop production line over six days. Each PCB has 3152 solder paste printing positions, and we use the X-direction relative offset and Y-direction relative offset as the raw data. Set sliding window $K = 20$, sliding step $S = 1$, and prediction window $T = 20$ to obtain a total of 11,352 sequences of solder paste position offsets. Since the number of good products far exceeds the number of defective products, 10% of the good product feature data and 50% of the defective product feature data are randomly selected to form the training set, and 10% of the good product feature data and the remaining defective product feature data are randomly selected to form the test set.

B. Experimental settings

1) Experiment Environment: Our method runs on a high-performance server, the server details are shown in Tab. II.

TABLE II: EXPERIMENTAL ENVIRONMENT CONFIGURATION

Configuration	Parameter Description
OS	Ubuntu 18.04.5 LTS
RAM	32GiB
CPU	i7-10700K

2) Metrics: Considering that the research is a classification problem, we choose accuracy, recall, precision, and $F1$ score as evaluation metrics. Using the defective products as the positive class, the recall rate reflects the proportion of the true positive class samples that are correctly classified. Accuracy reflects the percentage of samples predicted to be defective that is indeed defective. The $F1$ score is a trade-off between recall and precision.

$$F1 = \frac{2\cdot Recall\cdot Precision}{Recall + Precision}. \tag{14}$$

3) Baseline: To show the superiority of the method, we choose the following methods to compare.

Random Forest [19]: The set of multiple decision trees are used, and the output of multiple decision trees is subjected to majority voting thus outputting a most likely prediction. The experiment is set up with 100 decision trees and the impurity is calculated using the Gini coefficient.

Logistic regression [20]: L2 regularization is chosen, with a regularization factor of 1.

Multi-layer perception machine [21]: The activation function is chosen Relu, the optimizer is chosen SGD, the momentum is 0.9, the L2 regularization is chosen, the regularization factor is 0.0001, and the learning rate is 0.001.

C. Experimental Results

Each model is repeated five times for the test, the experimental results are averaged, and the standard deviation of the results of the five experiments is indicated in (.), which reflects the stability of the model, as shown in the following Tab. III.

Since there are far more good products than bad solder paste printing position offsets, the model can achieve an accuracy of 0.999 simply by predicting most of the time series as good products, but this does not reflect the true prediction capability of the model. The recall, precision, and $F1$ score are more indicative of the model's ability to discriminate against defective solder paste printing positions. It can be seen that SVM is substantially ahead of other methods in terms of recall, and $F1$ score, and shows higher stability in five repetition tests. All four methods maintain a high precision rate, proving that the feature engineering we construct can indeed reflect the different characteristics between good and defective products.

Fig. 3 shows the visualization of the SVM prediction results as a confusion matrix. It demonstrates the excellent

TABLE III: COMPARISON OF THE EXPERIMENTAL RESULTS

Method	Accuracy	Precision	Recall	F1
Random Forest	0.999((1.31×10^{-4})	**0.967(0.005)**	0.748(0.072)	0.841(0.042)
Logistic regression	0.999(2.75×10^{-5})	0.660(0.009)	0.515(0.013)	0.578(0.010)
MLP	0.999(1.19×10^{-5})	0.780(0.10)	0.718(0.012)	0.748(0.004)
SVM	**0.999(4.36×10^{-6})**	0.939(0.003)	**0.955(0.002)**	**0.947(0.001)**

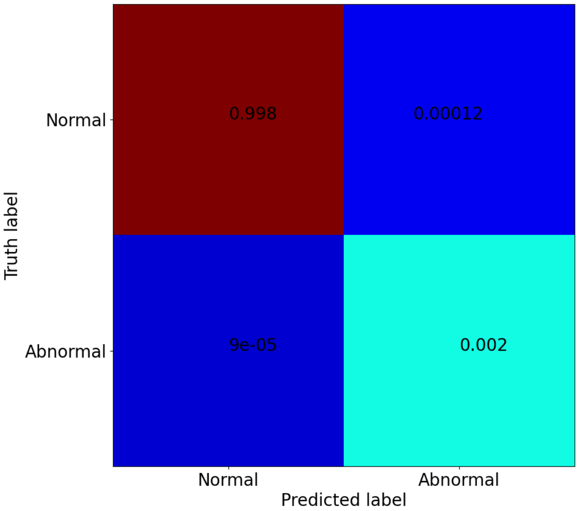

Fig. 3: Confusion matrix of SVM classification percentage results.

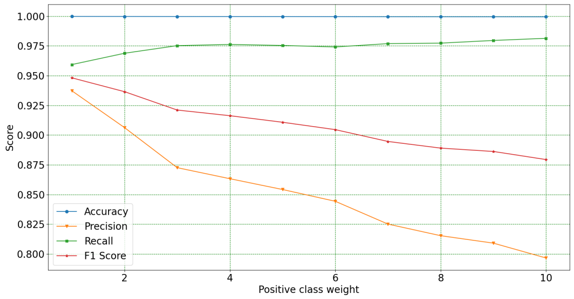

Fig. 4: Correlation chart of model performance with positive class weights.

performance of our method. Fig. 4 shows the comparison of model performance under different weights of defective products. It can be seen that although increasing the defective product weight will slowly increase the recall rate, it will also cause a significant decrease in the precision rate. This suggests that increasing the defect weights causes the model to simply predict more good products as defects, which leads to a gradual decrease in the $F1$ score, which is an indicator of the model's overall performance.

To achieve the interpretability of the model decisions, we analyze the 14 features of the manual design using random forest, as shown in Fig. 5. This can explain the basis on which our model makes its decisions. The overall 14 manual features in the Y-direction can be found to be more important than the 14 manual features in the X-direction because there are more poor offsets in the Y-direction. The maximum absolute value of the paste position offset significantly affects the results, while the mean, standard deviation, root mean square, and mean of the absolute values are also major influencing factors, proving the validity and interpretability of the manual design features.

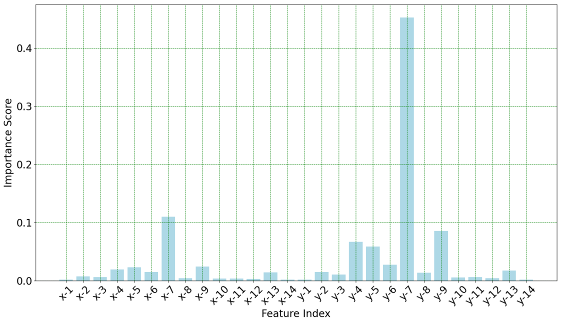

Fig. 5: Visualization of the influence weight of different features on classification results.

Conclusion

In this paper, we propose a feature engineering-based approach for solder paste printing position offset anomaly prediction. By comparing with several methods, the better evaluation metrics prove the superiority of our method. In future work, we will predict position offsets for anomaly detection, as specific offsets can facilitate engineers to adjust machine parameters.

References

[1] Yi-Ming Chang, Chia-Chen Wei, Jeffrey Chen, and Pack Hsieh. An implementation of health prediction in smt solder joint via machine learning. In *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)*, pages 1–4. IEEE.

[2] R. Y. Zhong, X. Xu, E. Klotz, and S. T. Newman. Intelligent manufacturing in the context of industry 4.0: A review. *Engineering*, 3(5):616–630, 2017.

[3] De Jian Zhou and Xiao Yong Chen. Discussion on application of intelligent manufacturing technology in smt product assembly manufacturing system. In *MATEC Web of Conferences*, volume 63, page 02035. EDP Sciences.

[4] Chenlin Zhou, Xiaofei Shen, Peng Wang, Wei Wei, Jia Sun, Yongkang Luo, and Yiming Li. Bv-net: Bin-based vector-predicted network for tubular solder joint detection. *Measurement*, 183:109821, 2021.

[5] M Firat Murat TU, T Martagan Tugce TU, and A Schröder Arjen AME. Using machine learning to reduce the false call problem in electronics manufacturing. 2021.

[6] Gulhan Ustabas Kaya. Development of hybrid optical sensor based on deep learning to detect and classify the micro-size defects in printed circuit board. *Measurement*, 206:112247, 2023.

[7] Sung Yi and Robert Jones. Machine learning framework for predicting reliability of solder joints. *Soldering & Surface Mount Technology*, 32(2):82–92, 2020.

[8] Reinhardt Seidel, Ben Rachinger, Nils Thielen, Konstantin Schmidt, Sven Meier, and Jörg Franke. Development and validation of a digital twin framework for smt manufacturing. *Computers in Industry*, 145:103831, 2023.

[9] Chun-Sheng Chen, Hai Wang, Yung-Chin Kao, Po-Jen Lu, and Wei-Ren Chen. Predictive model of the solder paste stencil printing process by response surface methodology. *Soldering & Surface Mount Technology*, 34(5):292–299, 2022.

[10] Wenjie Chen, Nian Cai, Huiheng Wang, Jianfa Lin, and Han Wang. Automatic optical inspection system for ic solder joint based on local-to-global ensemble learning. *Soldering & Surface Mount Technology*, 33(2):65–74, 2021.

[11] Nourma Khader, Jaehwan Lee, Duk Lee, Sang Won Yoon, and Haeyong Yang. Multi-objective optimization approach to enhance the stencil printing quality. *Procedia Manufacturing*, 38:163–170, 2019. 29th International Conference on Flexible Automation and Intelligent Manufacturing (FAIM 2019), June 24-28, 2019, Limerick, Ireland, Beyond Industry 4.0: Industrial Advances, Engineering Education and Intelligent Manufacturing.

[12] Jingxi He, Yuqiao Cen, Shrouq Alelaumi, and Daehan Won. An artificial intelligence-based pick-and-place process control for quality enhancement in surface mount technology. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 12(10):1702–1711, 2022.

[13] Nourma Khader and Sang Won Yoon. Adaptive optimal control of stencil printing process using reinforcement learning. *Robotics and Computer-Integrated Manufacturing*, 71:102132, 2021.

[14] Corinna Cortes and Vladimir Vapnik. Support-vector networks. *Machine Learning*, 20(3):273–297, 1995.

[15] S. Alelaumi, N. Khader, J. X. He, S. Lam, and S. W. Yoon. Residue buildup predictive modeling for stencil cleaning profile decision-making using recurrent neural network. *Robotics and Computer-Integrated Manufacturing*, 68:102041, 2021.

[16] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9:1735–80, 12 1997.

[17] Haifeng Wang, Hongya Lu, Shrouq M. Alelaumi, and Sang Won Yoon. A wavelet-based multi-dimensional temporal recurrent neural network for stencil printing performance prediction. *Robotics and Computer-Integrated Manufacturing*, 71:102129, 2021.

[18] Shrouq Alelaumi, Haifeng Wang, Hongya Lu, and Sang Won Yoon. A predictive abnormality detection model using ensemble learning in stencil printing process. *IEEE Transactions on Components, Packaging and Manufacturing Technology*, 10(9):1560–1568, 2020.

[19] Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001.

[20] J. S. Cramer. The origins of logistic regression. *Econometrics eJournal*, 2002.

[21] Michael Egmont-Petersen, Jan L. Talmon, Arie Hasman, and Anton W. Ambergen. Assessing the importance of features for multi-layer perceptrons. *Neural Networks*, 11(4):623–635, 1998.