Journal of Electronic Testing (2024) 40:669–678
https://doi.org/10.1007/s10836-024-06142-6

JETTA LETTER

# Efficient Selective Image Fusion: A PCB Diagnosis Approach and Implementation

Xueqin Wu[1,2] · Yikai Chen[1,2] · Zekai Wang[1,2] · Chenming Tian[3] · Zhonghua Shen[1,2]

Received: 20 August 2024 / Accepted: 18 September 2024 / Published online: 1 October 2024


## 1 Introduction

With the arrival of the Industry 4.0 era [1], the manufacturing industry has begun to step into the development path of intelligence, high-end and green. Printed circuit boards (PCBs) [2], as the support body of electronic components, play an increasingly critical role in all kinds of electronic devices [3]. High-quality production and repair of printed circuit boards requires not only precise design, but also real-time and accurate PCB image stitching technology to achieve accurate detection of components and ensure product reliability [4].

Traditional image acquisition methods rely on fixed focal lengths and lighting conditions [5], which often fail to provide sufficient detail and clarity when dealing with complex and dense PCB layouts. Scale-Invariant Feature Transform [6] (SIFT) and Speeded-Up Robust Features (SURF) [7] are commonly used local feature detection methods that exhibit good scale and rotation invariance [8, 9]. However, these methods have high computational complexity. Improved SIFT algorithms proposed by [10, 11] effectively reduce inference and computation time, but still perform poorly when handling repetitive patterns and complex backgrounds. Additionally, hybrid methods [12–14] that combine local and global features enhance matching robustness. Deep learning methods [15–17], such as Convolutional Neural Networks [18] (CNNs), demonstrate superior performance in feature extraction but are inefficient when processing large-scale images.

The method proposed in this paper adjusts the focal length based on heatmap [19]-identified component-dense regions to capture high-resolution images, providing sufficient clarity for dense PCB component areas. The Transformer [20–22] model is employed to extract image feature points, using self-attention [23] and cross-attention [24] mechanisms to capture both global and local correlations, thereby improving the accuracy and robustness of feature extraction. The extracted feature points are clustered and filtered using the K-means clustering algorithm [25], reducing redundant data and enhancing matching efficiency and accuracy. This approach has advantages in handling large-scale image data.

The test scenario in this work is focused on the repair and inspection of defective PC boards identified by repair engineers through initial tests, such as voltage and power assessments. In these scenarios, sparse component areas on the board are typically easy for engineers to inspect visually, allowing them to detect issues like missing solder, reversed polarity, or misplaced components. However, in densely populated areas, it becomes challenging to visually inspect small components, and traditional Automated Optical Inspection (AOI) systems struggle to accurately detect

Responsible Editor: V. D. Agrawal

This work was supported in part by the National Natural Science Foundation of China (11804161), the Fundamental Research Funds for the Central Universities, No. 30923010907.

✉ Zhonghua Shen
shenzh@njust.edu.cn

Xueqin Wu
wuxq@njust.edu.cn

Yikai Chen
ykchen@njust.edu.cn

Zekai Wang
zekaiwang@njust.edu.cn

Chenming Tian
tianchenming@njust.edu.cn

1 School of Physics, Nanjing University of Science and Technology, Nanjing 210094, China

2 Engineering Research Center of Semiconductor Device Optoelectronic Hybrid Integration in Jiangsu Province, Nanjing, China

3 School of Cyber Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094, China

and identify component names and polarities due to insufficient imaging resolution. This is where our algorithm plays a crucial role by enhancing detection capabilities in these challenging areas.

The key contributions of our paper are summarized as follows:

1. A novel selective area focus technique that emphasizes component-intensive regions, improving processing precision and efficiency(see Section II. B for details).
2. A unique combination of Transformer models and K-means clustering for effective feature point retention and real-time processing (discussed in Section II. C).
3. Integration of background information to mitigate local component confusion, enhancing robustness and accuracy (detailed in Section III).
4. Empirical validation demonstrating significant improvements in detection accuracy and speed over traditional methods like SIFT and template matching (see Section III).

## 2 Proposed Approach

In this paper, an automated optical path system is built to realize the capturing, fusion and stitching of PCB images as well as the subsequent component detection and identification.

Our system specifically targets these component-dense areas, replacing the original images with high-resolution ones identified through heatmaps. For areas specified by repair engineers, the system can capture real-time high-definition images to inspect for defects, such as missing components or reversed polarity. By utilizing template matching and edge detection techniques, the algorithm accurately identifies components in these densely populated regions, determines their quantity, and assesses chip polarity to ensure proper placement and compliance with design standards.

### 2.1 Experimental Procedure

The experimental flow is shown in Fig. 1: A high-resolution camera (UVA930, UVAER) is mounted on the end of a robotic arm (CR5, DOBOT) and the center of the PCB is aligned with the center point of the camera. The PCB is not powered up during the imaging process to ensure safety and avoid any potential interference with the imaging system.

Adjust the distance between the camera and the PCB to 54 cm above the PCB, ensuring sufficient depth of field and clear imaging. When capturing images of the entire PCB, set the camera magnification to 1.2x, with a focal distance of 5.20 mm, a horizontal field of view (H FOV) of 54.7 degrees, and a vertical field of view (V FOV) of 32.4 degrees.

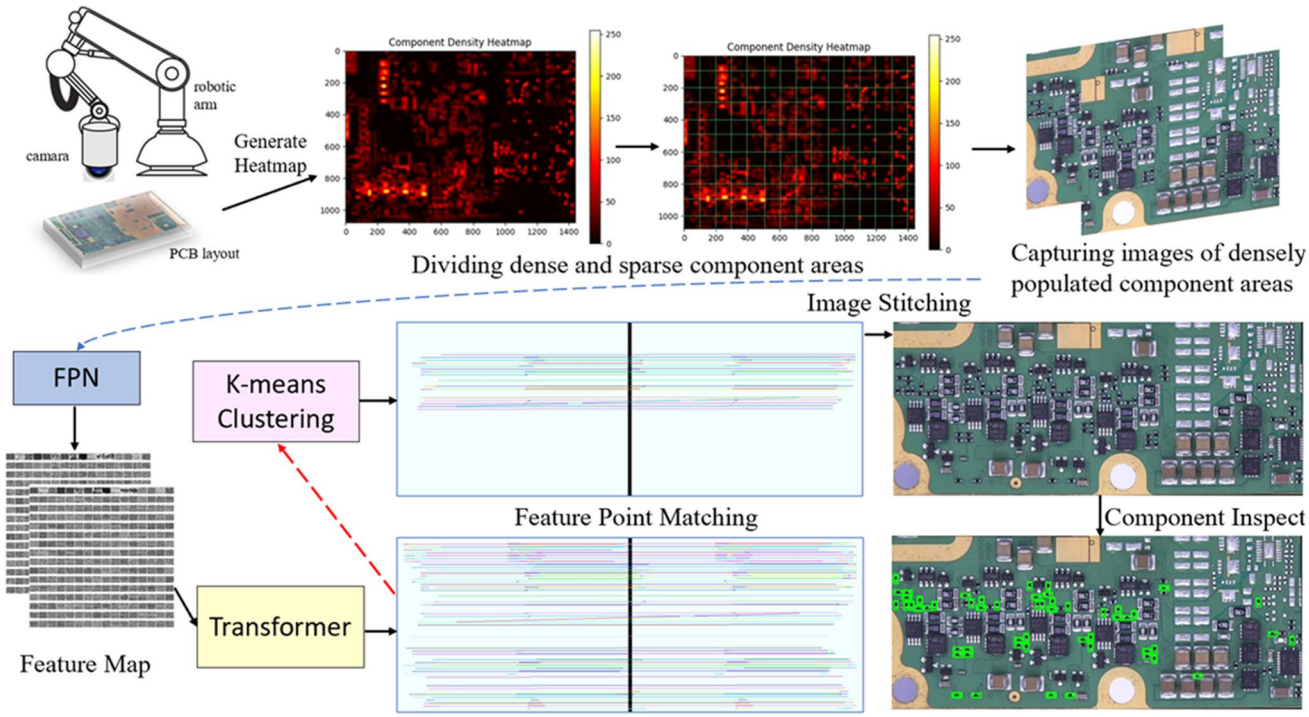


**Fig. 1** Experiment Flowchart

Based on the captured image of the entire PCB, generate a heatmap to identify areas with high and low component density. Adjust the camera magnification to 7.03x, with a focal distance of 30.20 mm, a horizontal field of view (H FOV) of 10.0 degrees, and a vertical field of view (V FOV) of 5.7 degrees.

Maintain the distance between the camera and the PCB at 54 cm. Then take a high-definition image of the component-dense areas. Generate a feature map using the Feature Pyramid Network (FPN), where "features" refer to distinct patterns or characteristics within the image—such as edges, corners, and textures—that are useful for identifying and distinguishing components. Use a Transformer model to extract these effective feature points. Apply k-means clustering to filter these points, then perform feature matching to achieve image stitching, seamlessly combining overlapping images into a single, unified view. Finally, validate the stitching results through component detection.

## 2.2 Heatmap Generation

The acquired image of the entire PCB board is depicted in Fig. 2(a). Due to the non-standard ambient lighting during image capture, conventional fixed global thresholding methods often fail, resulting in the loss of image details. To address this issue, we employ an adaptive thresholding technique as described in [26]. This method dynamically adjusts the threshold according to the local lighting conditions, effectively enhancing the separation between the foreground and background, as illustrated in Fig. 2(b).

Subsequently, all detected contours are traversed and overlaid on an all-zero floating-point array, thereby marking the regions of interest, as shown in Fig. 2(c). Regions with contour areas smaller than $50 \times 50$ pixels are given a weight of 1, while other regions are assigned a weight of 0.5, prioritizing smaller components to better localize dense regions. The heatmap is then obtained by summing these weighted regions across the entire PCB image, effectively highlighting areas of component density. Noise filtering and Gaussian smoothing are subsequently applied to the heatmap, as shown in Fig. 2(d), reducing image noise and enhancing the robustness of subsequent processing steps.

The resulting thermogram visualizes the component distribution within the image. The more saturated the color on the heatmap, the higher the component density in that region. From the heatmap in Fig. 2(d), it is evident that the lower-left corner of the PCB has a more saturated color, indicating a higher density of small components in this area.

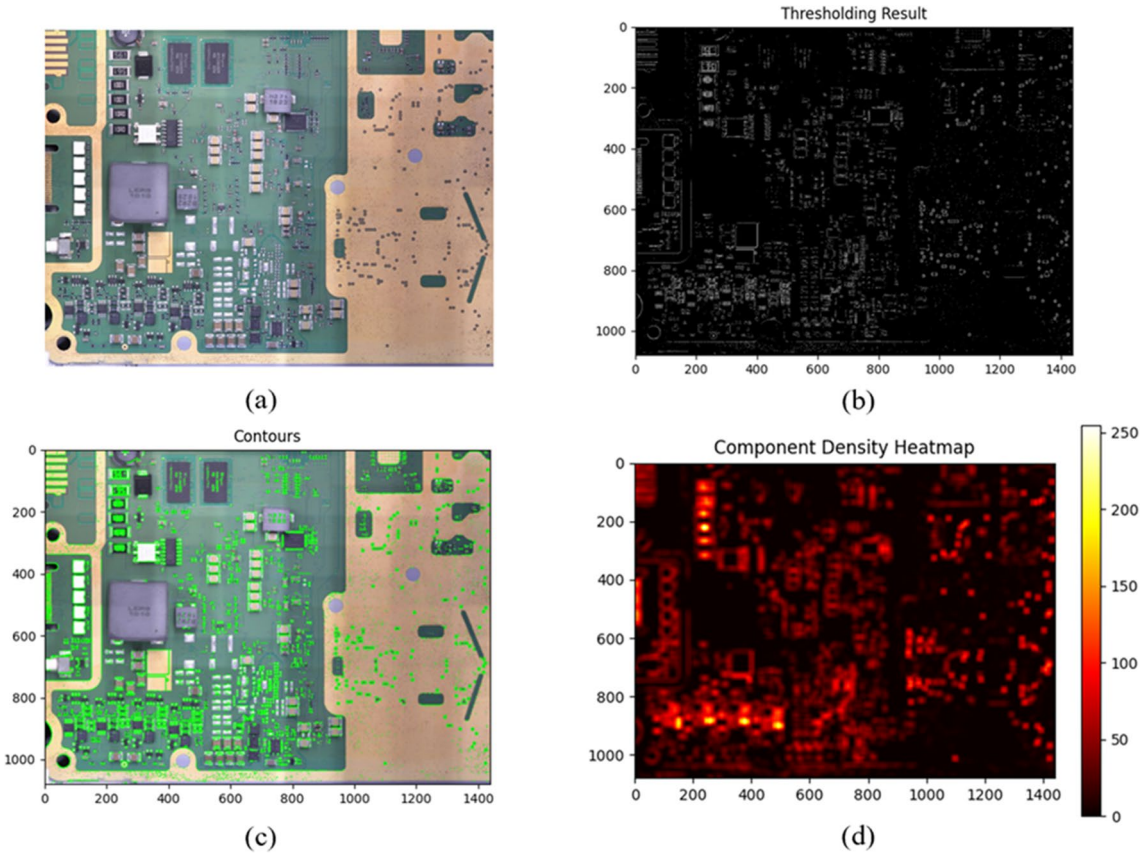


**Fig. 2** Heatmap Generation (**a**) Full PCB Image (**b**) Adaptive Threshold Segmentation Image (**c**) Contour Detection Image (**d**) Heatmap

Based on the generated heat map, this paper traverses this heat map with a grid size of 100*100 and calculates each grid score, setting the threshold as:

$$T_{\mathrm{h}} = mean(h) + 0.5 * \sigma \tag{1}$$

where $T_h$ is the threshold value of the heat map, $mean(h)$ is the mean value of the heatmap and $\sigma$ is the standard deviation of the heat map. If the grid score is higher than, the area is considered to be a component-intensive area. High-definition shooting of the component-intensive area to obtain $I^A and I^B$, in order to realize real-time seamless fusion and splicing of images, where "fusion" refers to the process of integrating multiple images into a single, coherent representation that combines the strengths of each individual image.

## 2.3 Model Architecture

We proposed a model called Selective Feature Fusion and Alignment with Transformer and K-means (SFFA-TK), as illustrated in Fig. 3, which consists of five components:

1) *Feature Pyramid Network (FPN)* [27]*:* The model begins with feature extraction through a convolutional layer followed by batch normalization. It progressively captures deeper features using a hierarchical structure of residual blocks. T Each layer's output is processed by a $1 \times 1$ convolution, and high-level features are fused with low-level features through upsampling. *A* $3 \times 3$ enhances the representation of the fused features. The outputs include coarse-level feature maps $\widetilde{F}^A and \widetilde{F}^B$ from images $I^A and I^B$ at 1/8 resolution, and fine-level feature maps $\widehat{F}^A and \widehat{F}^B$ at 1/2 resolution.
2) *Transformer Model* [28]*:* Coarse-level feature maps $\widetilde{F}^A$ are flattened into vectors with positional encoding and input into the Transformer model—a deep learning architecture designed to capture complex relationships within data by leveraging attention mechanisms. The model employs four layers each of self-attention and cross-attention. Self-attention evaluates the relationships within the same image, enriching each patch's representation with its context and relative positioning. Cross-attention assesses correlations between different images, facilitating feature fusion and information exchange. The final outputs are feature maps $\widetilde{F}_T^A and \widetilde{F}_T^B$, retaining the same size as the coarse-level feature maps.
3) *Matching Module:* The similarity matrix between the two sets of feature vectors $\widetilde{F}_T^A and \widetilde{F}_T^B$ is computed using bidirectional Softmax. A mask is applied to exclude boundary regions and other irrelevant areas, ensuring that only valid feature points are considered for subsequent matching calculations. The masked similarity matrix is then normalized using row and column Softmax to generate a confidence matrix. Based on this confidence matrix, pairs of matching points with confidence scores higher than a preset threshold are selected as candidate matches. By implementing a mutual nearest neighbor filtering condition, it is ensured that each pair of feature points is the best match from each other's perspective. This results in the coarse matching prediction set $\left[\left(\hat{i}_A, \hat{j}_A\right), \left(\hat{i}_B, \hat{j}_B\right), Conf\right] \epsilon M_c$, where $\left(\hat{i}_A, \hat{j}_A\right), \left(\hat{i}_B, \hat{j}_B\right)$

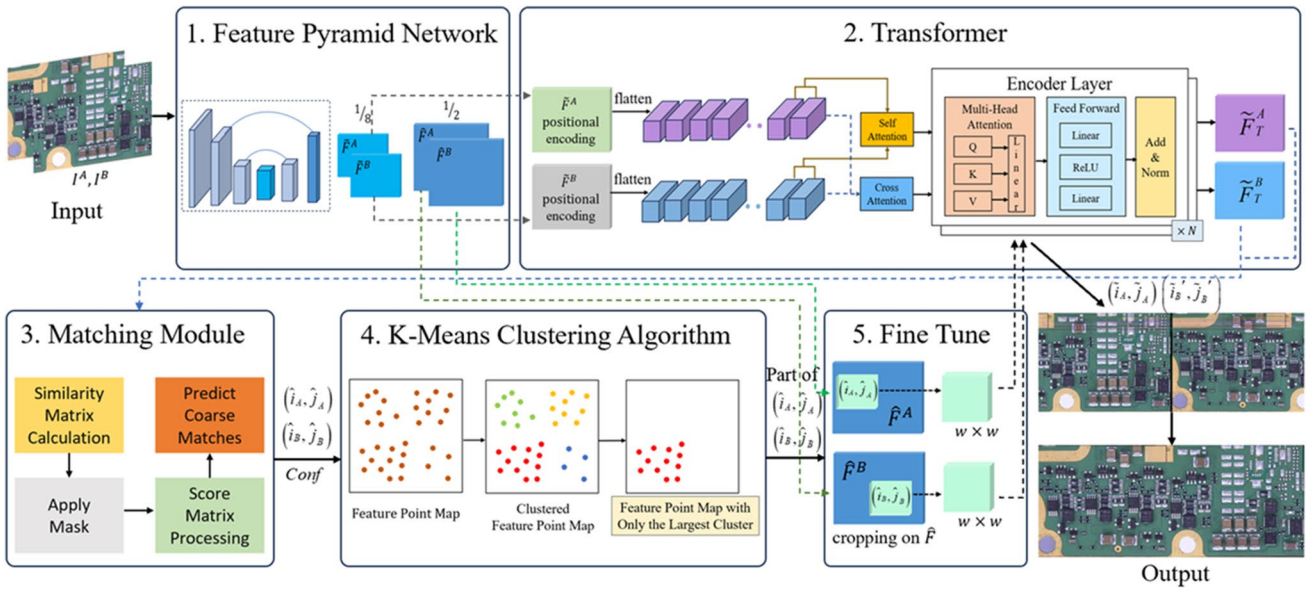


**Fig. 3** SFFA-TK model architecture

includes the coordinates of the matching points between the two images $I^A and I^B$, *Conf* is the confidence score for each pair of matching points.

4) *K-Means Clustering Algorithm:* The top 200 feature point pairs with the highest confidence scores in are retained based on their values. These 200 feature point pairs are then clustered using the K-Means algorithm, a method that partitions the data into distinct groups based on similarity by minimizing the variance within each cluster. The cluster with the highest number of points and the most concentrated distribution is selected for feature matching, resulting in the output of the filtered subset of coarse matches $\left(\hat{i}_A,\hat{j}_A\right),\left(\hat{i}_B,\hat{j}_B\right)\epsilon M_k$.

The number, quality, and distribution of feature points play a critical role in achieving precise matching. It is essential to select highly distinctive and representative feature points while minimizing the interference of redundant points from irrelevant regions in the image. As depicted in Fig. 4(a), K-means clustering analysis effectively categorized feature points into four distinct classes. This study focuses on the class with the highest number of points and the most concentrated distribution for matching, specifically class 1, as shown in Fig. 4(b), where the points predominantly cluster in the upper-left corner of the image.

5) *Fine-tuning Module:* For each rough match prediction $\left(\hat{i}_A,\hat{j}_A\right),\left(\hat{i}_B,\hat{j}_B\right)\epsilon M_k$ filtered by K-means, a $w \times w$ window is cropped on the fine-level feature maps $\hat{F}^A and \hat{F}^B$. The features from the coarse matching stage are linearly dimension-reduced and merged with the fine features. These enhanced features are further integrated through another linear layer to amplify contextual information. A similarity matrix is generated based on the integrated features, using the reference points to fine-tune within this local window, thereby improving feature point matching accuracy at a finer level.

## 3 Results and Discussions

The time required for image stitching, referred to as the "execution time," along with the matching results for different classes obtained via K-means clustering and using all feature points without K-means, were statistically analyzed, as shown in Table 1. The execution time reflects the duration needed to complete the stitching process for each method.

The corresponding stitching results are illustrated in Figs. 5 and 6. Class 0 through 3 represent different categories of feature points, classified based on their distinctiveness and spatial distribution. Specifically, Class 0 through 3 are derived from the K-means clustering results illustrated in Fig. 4, where some of the classes correspond to feature points that are either more spatially scattered (Class 0) or less distinctive (Class 2).

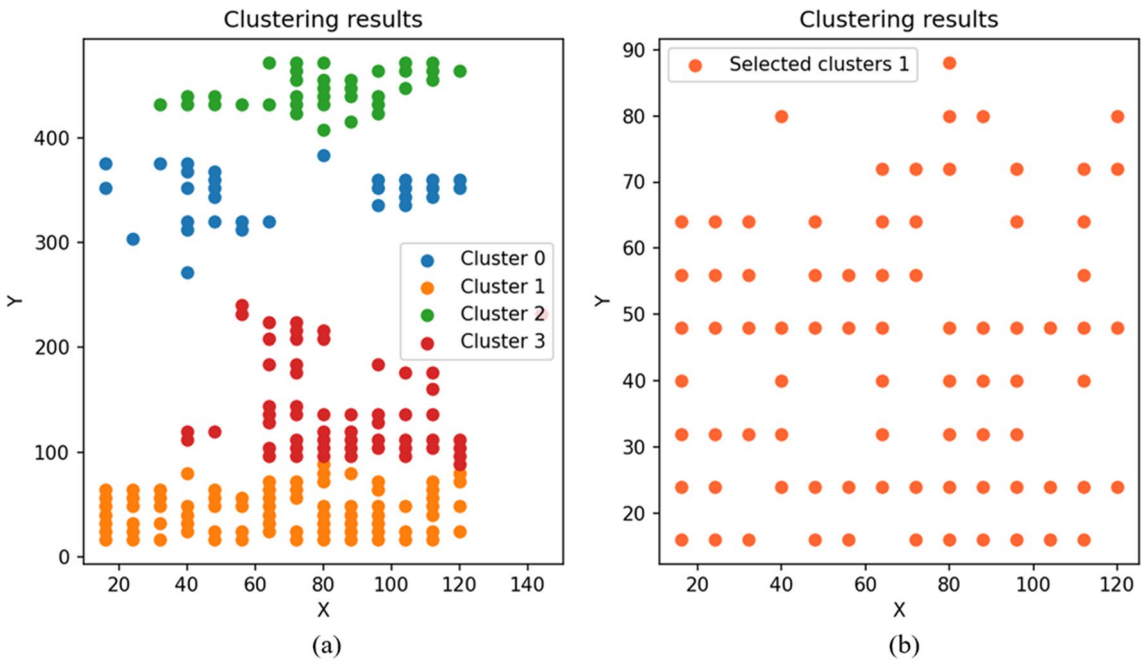


**Fig. 4** Cluster analysis (**a**) Map of clustering results (**b**) Location of class 1 with the highest number of points

**Table 1** Comparison of matching results by using different classes of feature points for stitching

| Adopted Feature Point Categories | Number of Feature Points | Execution Time (s) | Stitching Results |
|---|---|---|---|
| all feature points | 200 | 6.34 | √ |
| class 0 | 31 | 4.66 | × |
| class 1 | 79 | 3.46 | √ |
| class 2 | 35 | 4.58 | × |
| class 3 | 55 | 4.21 | √ |

As shown in Fig. 5(b), (c), these features lead to more mismatches and incorrect splices despite the slight reduction in processing time. As shown in Fig. 5(a), without K-mean clustering, image splicing takes 6.34 s. Using the third class of feature points for splicing, the splicing result is shown in Fig. 5(d), the spliced image does not show significant distortion, but its execution time still takes 4.21 s. We determine that the matches in Fig. 5(a) and (d) are correct, while the other matches are incorrect.

In contrast, selecting feature points from Class 1, which contains the most abundant and concentrated features, significantly improved the stitching process. This selection also

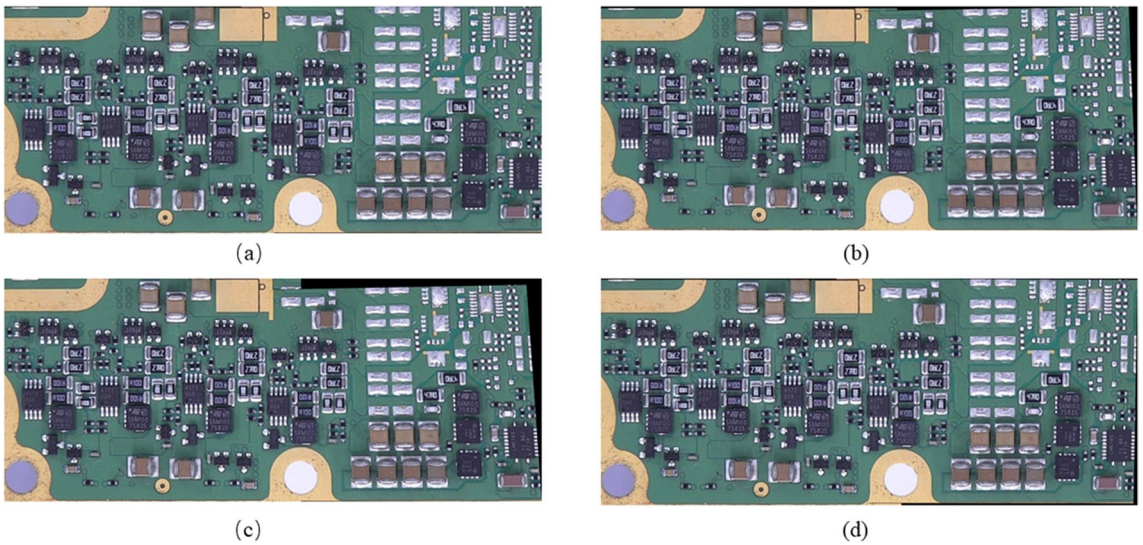


**Fig. 5** Matching results (**a**) using all 200 feature points (**b**) using class 0 feature points (**c**) using class 2 feature points (**d**) using class 3 feature points

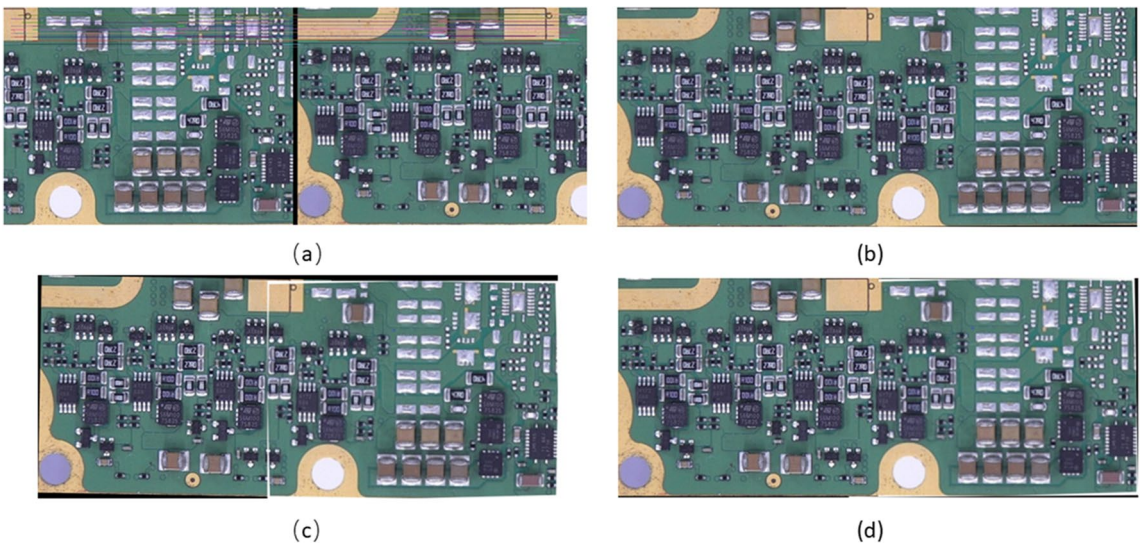


**Fig. 6** Comparison of stitching using SFFA-TK and SIFT

based on the K-means clustering results shown in Fig. 4, reduced the stitching time to 3.46 s—a 45.4% decrease. The effectiveness of this approach is demonstrated in Fig. 6(a) and (b), which illustrate the matching results using the clustered feature points from Class 1.

As illustrated in Table 1, this strategy not only significantly enhances feature matching accuracy but also optimizes the computational process, considerably reducing complexity and memory usage, thus achieving efficient resource utilization. It provides a robust foundation for subsequent image processing and analysis tasks. Additionally, this method underscores the importance of selecting appropriate feature point sets in image stitching technology to improve efficiency and accuracy, laying a theoretical and practical foundation for further advancements in image processing technology.

The matching using clustered feature points from class 1 is illustrated in Fig. 6(a). The SFFA-TK model leverages background features of PCB images for matching, extracting and comparing features on a global scale, yielding superior results. PCB backgrounds, typically rich in contextual information and containing few repetitive structures, compensate for potential confusion from highly similar local components. Therefore, the SFFA-TK model achieves higher accuracy in matching features of PCB images with numerous repetitive or similar components, as shown in Fig. 6(b), resulting in seamless stitching without component misalignment.

To further substantiate the robustness of the SFFA-TK model, we rotated image $I^B$ counterclockwise by 1° and applied a blurring effect. When using the SIFT method for stitching, as shown in Fig. 6(c), noticeable seams and component misalignment are observed. In contrast, the proposed SFFA-TK method, as depicted in Fig. 6(d), shows no significant seams.

To evaluate the effectiveness of different stitching methods in component detection, a template image of a small component, a 0201 resistor (0.6 mm × 0.3 mm), was used, as shown in Fig. 7(a). Component detection was performed on the original image, an image stitched using traditional template matching, and an image stitched using the SFFA-TK method, with results shown in Table 2.

In the original image, only 27 0201 components were detected, as depicted in Fig. 7(b). Using traditional template matching, 59 detection boxes were identified within 3.52 s, including 53 0201 components and 6 false positives, as shown in Fig. 7(c). In contrast, the SFFA-TK method, without requiring reference conditions, completed the image

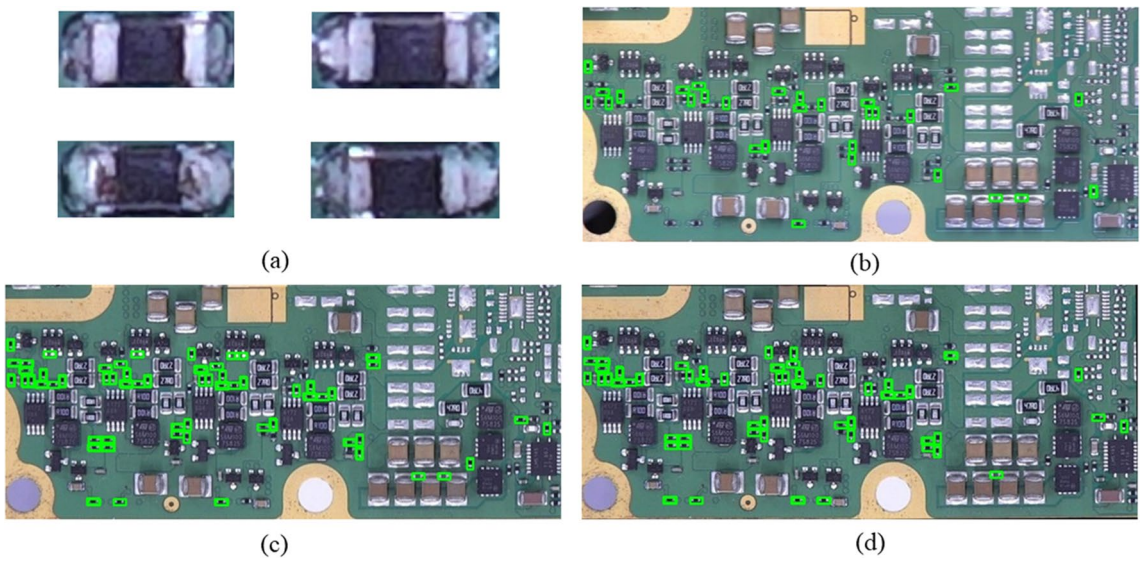


**Fig. 7** Component detection (**a**) component template image (**b**) detection results from the original image (**c**) template matching detection results (**d**) SFFA-TK detection resulos

**Table 2** Comparison of component detection results using different stitching methods

| Stitching Method | Number of detection boxes | Number of correctly detected components | Number of false positives | Precision |
|---|---|---|---|---|
| original figure | 200 | 27 | 0 | 100% |
| Template matching | 79 | 53 | 6 | 89.83% |
| SFFA-TK | 35 | 53 | 1 | 98.15% |

stitching in just 3.46 s, successfully detecting 54 component boxes, with 53 correct detections, significantly reducing false positive rates compared to traditional template matching, as illustrated in Fig. 7(d).

Precision, as depicted in Fig. 2, a key metric in evaluating model performance, measures the proportion of true positive instances among those predicted as positive. It is defined as the number of true positives (TP) over the total number of instances predicted as positive (including both TP and false positives (FP)). TP refer to the instances where the model correctly identifies a positive case (e.g., correctly detecting a resistor as resistor). FP, on the other hand, are instances where the model incorrectly identifies a negative case as positive (e.g., mistakenly identifying an inductor as resistor). Higher precision indicates fewer false positives, meaning the model is more accurate in its positive predictions.

In this study, the proposed SFFA-TK method achieved image stitching in 3.46 s, a substantial improvement in detection speed over traditional methods using all feature points. Moreover, the image quality post-stitching with the SFFA-TK method surpassed that of traditional matching algorithms. The SFFA-TK method achieved an 8.32% increase in accuracy, particularly excelling in small component detection.

To further demonstrate the effectiveness of the proposed stitching method in component identification and polarity determination, we employed adaptive thresholding and edge detection techniques to analyze both the original image and the stitched images obtained using our method. As shown in Fig. 8, Fig. 8(a) and (b) display components extracted from the original image and the detection results, while Fig. 8(c) and (d) present components extracted from the stitched image along with their detection results. It is evident from Fig. 8(c) and (d) that the stitched image allows for accurate reading of chip model numbers, such as the S6M100 chip, and reliable identification of polarity regions, with the red areas indicating the polarity positions of the chips or components.

These results validate the SFFA-TK method's effectiveness in enhancing PCB image processing efficiency and accuracy, especially in handling complex images with numerous repetitive or similar components. This study demonstrates the significant potential of integrating traditional image processing techniques with advanced deep learning algorithms, providing a reference for future high-precision component detection research. This work breaks away from the traditional framework of relying on large-scale sample construction for component detection, achieving efficient, real-time, and accurate detection of circuit board components with limited resources, which is crucial for quality control and automation in the electronics manufacturing industry.

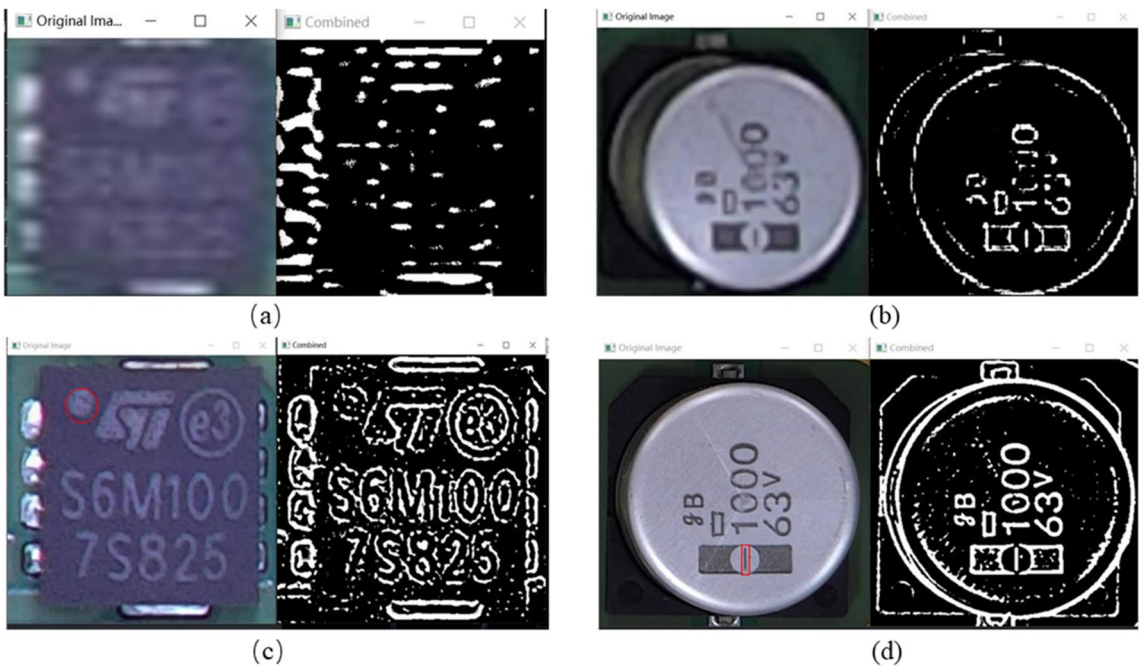


**Fig. 8** Comparison of component detection in original and stitched images: (**a**) and (**b**) Components extracted from the original image; (**c**) and (**d**) Components extracted from the stitched image

## 4 Conclusion

This study presents efficient selective PCB image capturing, fusion, and stitching process, addressing the challenge of accurately locating small components on PC boards. By using a heatmap to distinguish between densely and sparsely populated areas, the method focuses on dense regions and assigns higher weights to smaller components. Incorporating background information effectively reduces confusion caused by the high similarity of local components. The innovative combination of the Transformer model and K-means clustering further optimizes feature extraction, resulting in more accurate and efficient image fusion and stitching. Our experimental results demonstrate a significant improvement in stitching speed compared to traditional methods SIFT and template matching, and further detection tests confirm the accuracy, robustness and effectiveness of the stitching, particularly in identifying components like the S6M100 chip and determining its polarity. This research not only advances PCB maintenance technology but also provides a solid framework for future developments in component detection and image processing technologies.

Future work will focus on exploring and optimizing the scalability and generalizability of the algorithm to accommodate a wider variety of component types and more complex PCB designs. Given the complexity of real-world applications, we also aim to further reduce the computational cost of the algorithm and explore more efficient data processing and analysis methods to handle increasing data volumes and real-time processing demands. Additionally, strengthening collaborations with industry to translate research findings into practical applications is an important direction for future research.

**Code, Data, and Materials Availability** The data supporting this article's findings are not publicly available due to intellectual property rights. They can be requested from the author at wuxq@njust.edu.cn.

## Declarations

**Conflicts of Interest** There are no conflicts of interest.

## References


1. Ghobakhloo M (2020) Industry 4.0, digitization, and opportunities for sustainability, in journal of cleaner production, 252nd ed., New York, NY, USA: McGraw-Hill, pp 119869
2. Ladou J (2006) Printed circuit board industry. Int J Hyg Environ Health 209(3):211–219
3. Perdigones F, Quero JM (2022) Printed circuit boards: the layers' functions for electronic and biomedical engineering. Micromachines 13(3):460
4. Yang W, Liu X (2021) RTStitch: Real-time stitching of high-resolution PCB images. In: 2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Melbourne, Australia. IEEE, pp 541–546. https://doi.org/10.1109/SMC52423.2021.9658780
5. Abd Al Rahman M, Mousavi A (2020) A review and analysis of automatic optical inspection and quality monitoring methods in electronics industry, IEEE Access, vol 8, pp 183192–183271
6. Lowe DG (1999) Object recognition from local scale-invariant features. In: Proceedings of the Seventh IEEE International Conference on Computer Vision, Kerkyra, Greece, vol 2. pp 1150–1157. https://doi.org/10.1109/ICCV.1999.790410
7. Bay H, Ess A, Tuytelaars T, Van Gool L (2008) Speeded-up robust features (SURF). Comput Vis Image Underst 110(3):346–359
8. Panchal P, Panchal S, Shah S (2013) A comparison of SIFT and SURF. Int J Innov Res Comput Commun Eng 1(2):323–327
9. Loncomilla P, Ruiz-Del-Solar J, Martínez L (2016) Object recognition using local invariant features for robotic applications: a survey. Pattern Recogn 60:499–514
10. Liu Y, He M, Wang Y et al (2022) Farmland aerial images fast-stitching method and application based on improved SIFT algorithm. IEEE Access 10:95411–95424
11. Etezadifar P, Farsi H (2020) A new sample consensus based on sparse coding for improved matching of SIFT features on remote sensing images. IEEE Trans Geosci Remote Sens 58(8):5254–5263
12. Lin B, Zhang S, Yu X (2021) Gait recognition via effective global-local feature representation and local temporal aggregation. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp 14648–14656
13. Cao B, Araujo A, Sim J (2020) Unifying deep local and global features for image search," in computer vision – ECCV 2020: 16th European conference, Glasgow, UK
14. Peng Z, Huang W, Gu S, Xie L, Wang Y, Jiao J, Ye Q (2021) Conformer: Local features coupling global representations for visual recognition. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV). pp 367–376
15. Sharmin S, Ahammad T, Talukder MA, Ghose P (2023) A hybrid dependable deep feature extraction and ensemble-based machine learning approach for breast cancer detection. IEEE Access 11:123
16. Li Y, Luo J-H, Dai Q-Y et al (2023) A deep learning approach to cardiovascular disease classification using empirical mode decomposition for ECG feature extraction. Biomed Signal Process Control 79:104188
17. Apostolopoulos ID, Tzani MA (2023) Industrial object and defect recognition utilizing multilevel feature extraction from industrial scenes with Deep Learning approach. J Ambient Intell Humaniz Comput 14(8):10263–10276
18. Li Z, Liu F, Yang W et al (2021) A survey of convolutional neural networks: Analysis, applications, and prospects. IEEE Trans Neural Networks Learn Syst 33(12):6999–7019
19. Metsalu T, Vilo J (2015) ClustVis: a web tool for visualizing clustering of multivariate data using principal component analysis and heatmap. Nucleic Acids Res 43(W1):W566–W570
20. He K, Gan C, Li Z et al (2023) Transformers in medical image analysis. Intell Med 3(1):59–78
21. Xiao H, Li L, Liu Q et al (2023) Transformers in medical image segmentation: a review. Biomed Signal Process Control 84:104791
22. Xu P, Zhu X, Clifton DA (2023) Multimodal learning with transformers: a survey. IEEE Trans Pattern Anal Mach Intell 45(4):1234–1245
23. Niu Z, Zhong G, Yu H (2021) A review on the attention mechanism of deep learning. Neurocomputing 452:48–62

24. Wen Z, Lin W, Wang T, Xu G (2023) Distract your attention: Multi-head cross attention network for facial expression recognition. Biomimetics 8(2):199
25. Ikotun AM, Ezugwu AE, Abualigah L et al (2023) K-means clustering algorithms: a comprehensive review, variants analysis, and advances in the era of big data. Inf Sci 622:178–210
26. Bobin J, Starck J-L, Fadili JM et al (2007) Morphological component analysis: an adaptive thresholding strategy. IEEE Trans Image Process 16(11):2675–2681
27. Lin TY, Dollar P, Girshick R, He K, Hariharan B, Belongie S (2017) Feature pyramid networks for object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp 2117–2125
28. Sun J, Shen Z, Wang Y, Bao H, Zhou X (2021) LoFTR: Detector-free local feature matching with transformers. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). pp 8922–8931



**Xueqin Wu** is currently a graduate student at the School of Physics, Nanjing University of Science and Technology. She received her BS degree in electrical information engineering from Fuzhou University in 2022. Her research focuses on image processing and artificial intelligence.

**Yikai Chen** received his Ph.D. degree in 2015 from Dept. of Optics & Optical Engineering, USTC. Prior to joining the Department of Information Physics and Engineering at Nanjing University of Science and Technology as an Associate Professor, he was a Post-doctoral Researcher in the Department of Chemistry at University of Washington. Yikai's current research is microscopic spectral imaging, sensing with metasurfaces, and R&D of photonic chip devices.

**Zekai Wang** is currently a graduate student at the School of Physics, Nanjing University of Science and Technology. He received his bachelor′s degree in applied physics from Nanjing University of Information Science & Technology in 2022. His research focuses on nanophotonics and plasmonics.

**Chenming Tian** is currently a graduate student at the School of Cyber Science and Engineering, Nanjing University of Science and Technology. He received his BS degree in computer science from Jilin University in 2021. His research focuses on image processing and artificial intelligence.

**Zhonghua Shen** is a Professor and Ph.D. supervisor at Nanjing University of Science and Technology, earned her Ph.D. in Optical Engineering there in 1999. Her research directions are laser ultrasonic non-destructive testing technology and application, laser-substance interaction mechanism and testing. From 2004–2009, she has visited the University of Heidelberg in Germany frequently and has established collaborations with many internationally renowned photoacoustic research groups.