Advances in Computer, Signals and Systems (2023)
Clausius Scientific Press, Canada

DOI: 10.23977/acss.2023.070105
ISSN 2371-8838 Vol. 7 Num. 1

Transmission Defects Localization Network: Towards Wrong Assembly in the Transmission Assembly Process

Wenwen Zhu[1,a], Bin Zhang[2,b], Tao Song[2,c], Shuai Zhang[3,d,*]

[1]Chongqing University of Science and Technology, Qinjiagang Street, Chongqing, China
[2]Luzhou Rongda Intelligent Transmission Co. LTD, Tai'an Street, Luzhou, China
[3]China National Tobacco Corporation Chongqing Company Youyang Branch, Chongqing, China
[a]zhuwenwen_work@163.com, [b]bin.zhang78811@seres.cn,
[c]375380950@qq.com,[d]zhangshuai_2007@163.com
**Corresponding author*

Keywords: Transmission assembly, defects localization network, features fusion

Abstract: Transmission valve component is an important part of automobile manufacture. The success of the assembly of transmission valve is directly related to driving safety of vehicles. While localizing transmission assembly defects is particularly important in assembly of transmission valve component. As an image processing problem, real-time assembly images of transmission valve component are adopted to determine whether the assembly is correct or wrong. Transmission valve component in these images have small, severe reflection, and sparse properties, which increases the difficulty of detection. Therefore, this paper proposes a transmission defects localization network based on Siamese network for improving the performance of assembly of transmission valve components. In our model, we establish an image similarity evaluation network with designed multi-scale features fusion approach. Furthermore, in order to reduce intra-class spacing by similar transmission valve part samples on evaluation action, an improved binary cross entropy and focal loss function is discovered for feature re-processing. Finally, experimental results on real-world transmission assembly dataset indicate that our proposed approach outperforms other compared methods.

1. Introduction

Transmission, as a typical mechanical product, is an important part of automotive transmission system [1]. In the rapidly developing automobile industry, transmission directly affects driving safety and fuel economy characteristics of cars [2]. Therefore, requirements of users for its quality, safety, and performance are also increasingly high.

Transmission valve component is a key part of automatic transmission shift [3]. Its assembly is an important link in the formation of products in the later stages of transmission manufacturing and is the top priority of transmission product quality control [4]. Transmissions are intricate mechanical products with numerous parts, thus the accuracy of its valve component assembly must meet strict standards. At the same time, environmental constraints and unregulated assembly operations have been in the assembly of transmission valve components, resulting in incorrect treatment. Hence, error-proof detection [5] during assembly of transmission valves is of great significance.

In general, transmission valve components are manually assembled and inspected. The assembly process for transmissions involves many different and intricate elements, making it certain that there would be potential risks including incorrect installation, missing installation, and numerous installations. Using machine vision detection [6] in place of manual detection can significantly increase production efficiency and lower the rate of error since manual detection of transmission valve components assembly is labor-intensive and inefficient. Deep learning-based solutions may detect smaller and more complicated product faults under frequent environmental changes, improve detection efficiency [7]. Therefore, deep learning-based solutions become the primary technique to address this issue in complex quality detection scenarios [8].

Currently, image processing algorithms for target [9] detection, such as binary morphological approaches [10], similarity measurement techniques [11], and segmentation techniques [12], have been thoroughly investigated. These methods, whereas, have shortcomings in poor contrast, uneven brightness, and irregular form when identifying and localizing tiny objects. They also demand that identified images be matched with the matching reference samples.

Principal component analysis and support vector machines are used in a revolutionary real-time monitoring system that is designed to automatically detect faults [13]. However, missing data sensitivity is one of their drawbacks. A brand-new self-reference template-guided image decomposition method is created to find surface flaws in strip steel [14]. A raster ROI, as described by [15] in 2021, is an algorithm designed to identify streak flaws on the customer content area. Nevertheless, drawback is that they focus on the component problems specifically.

Literature [16] unique machine learning-based technique for error detection that just needs the raw output data from a susceptibility test. Its shortcoming is also missing data sensitivity. A novel method is designed to gauge the efficacy of error detection techniques [17]. Instead of evaluating assembly outcomes, it evaluates inspection procedures.

For transmission valve components, images have small, severe reflection, and sparse properties, which limit the implementation of localizing transmission assembly defects with traditional detection techniques. Assembled parts in overall transmission images are small, severe reflection, and highly sparse which requires transmission defect localization models are capable of anti-reflection and detailed features extraction. However, unlike transmission assembly dataset, prominent open-source datasets (e.g., ImageNet [18], COCO [19], and PASCAL VOC2012 [20]) place more emphasis on big objects. This means that these datasets cannot be used to pre-train deep neural networks for transmission valve component assembly defects localization.

Hence, in this letter, we reveal that there is no analogous technique to prevent assembly error detection during manual assembly of transmission valve components. To alleviate this problem, we propose a transmission defects localization network, based on Siamese network to evaluate the similarity of transmission assembly by extracting detailed features to localize transmission assembly defects. In conclusion, our main contributions are as follows:

(1) A transmission defects localization network is presented localize transmission assembly defects with evaluating the similarity of transmission assembly by extracting detailed features without any manually operations. The outcomes of the simulation demonstrate its high viability.

(2) To improve the performance of network for various sizes, multi-scale features fusion approach is designed. In addition, to reduce intra-class spacing by similar transmission valve part samples, an improved binary cross entropy and focal loss function is explored on end-to-end training techniques to teach the network.

(3) Great results are obtained when our network is used in the actual transmission assembly shop.

2. Related Work

2.1. Metric Learning

As opposed to contrastive learning [21], metric learning [22] is an approach of thinking that uses data to determine how far apart two objects are. Its fundamental concept is around the measurement and computation of distances. To facilitate learning, the primary principle is to reduce dimensionality to a lower dimensional space. Finding a suitable space is essentially the same as finding an appropriate distance metric. Thus, metric learning is straightforward attempting to learn an appropriate distance metric.

Due to its limited capacity to handle original data, traditional metric learning need first pre-process the input using the expertise of feature engineering before using the metric learning algorithm to learn. As a result, certain traditional metric learning algorithms can learn just linear features. Despite some kernel approaches for extracting non-linear features have been put forth, however, their learning impact has not greatly increased. Since the ability of activation function to learn non-linear features, deep learning approaches can learn high-quality features from original data. In order to achieve the best outcomes, deep neural network architecture and conventional metric learning methods can be combined.

2.2. Siamese Network

Siamese network [23] is a straightforward and remarkable construction. As shown in Figure 1, it is made up of two identical networks with weights that are shared. It transfers two network inputs to new space and obtains the representation of input in new space. The distance measure in new space is then adopted to evaluate the similarity of two inputs. Siamese network has been successfully adopted in many realms, such as semantic similarity analysis [24], handwriting font recognition [25], and visual tracking algorithm [26]. Therefore, it is of great significance in many mission-critical applications.

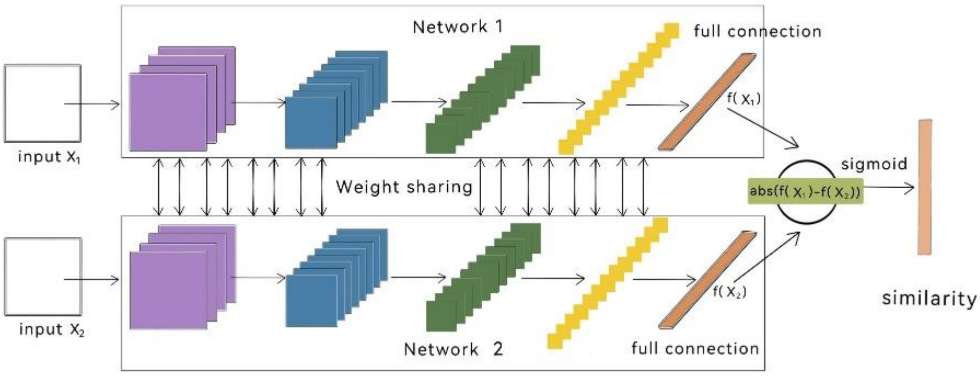

Figure 1: Siamese network structure

Network 1 and Network 2 are identical networks. After feature mapping, $abs(f(X_1) - f(X_2))$ describes the metric distance between two inputs in low-dimensional space. The output of two fully connected layers is represented by $f(X_1)$ and $f(X_2)$.

3. Proposed Methods

Transmission defects localization network involves three parts (Figure 2 illustrates its structure): (a) image pre-processing: key-point localization, image correction, and image de-noising are adopted to improve the quality of images. (b) feature extraction: Based on this understanding, we introduce multi-scale features fusion and improved binary cross entropy and focal loss function, which improve the performance of network for various and reduce intra-class spacing. (c) model application: A model application system is applied in transmission valve component production and assembly process.

3.1. Image Preprocessing

The complicated environment of production workshop has an impact on image quality. To reduce position offset between different images, in this letter, PLC linkage and lifting positioning approach is applied. Meanwhile, because of high resolution of industrial cameras, while areas required to detect transmission valve components take up only a small amount of overall image region.

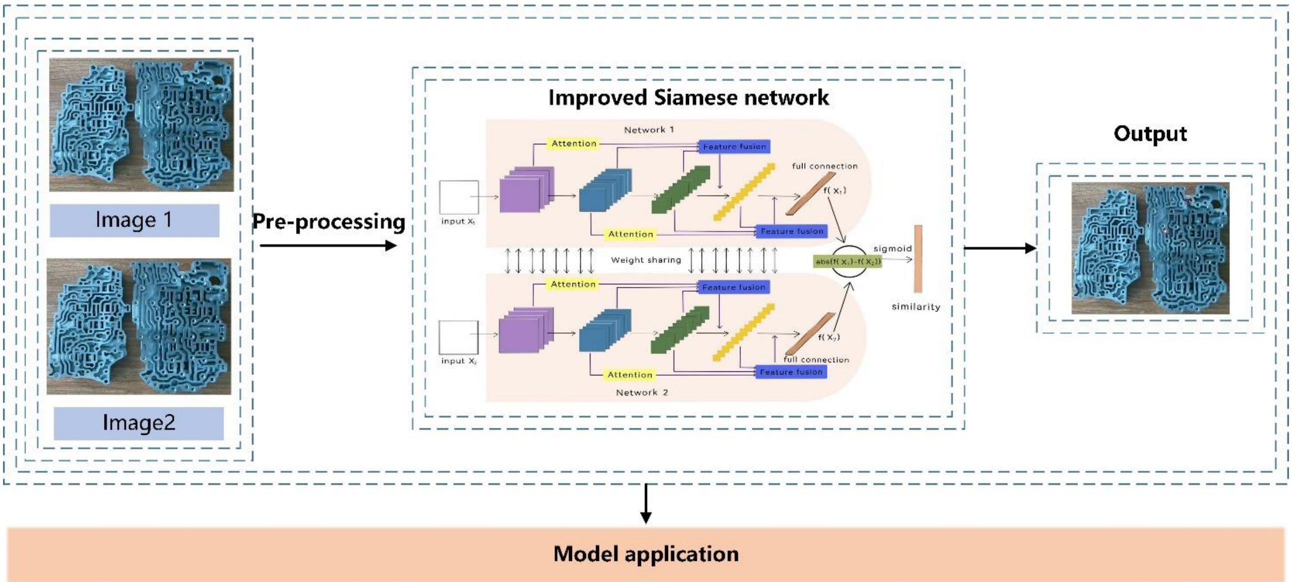

Figure 2: The application architecture of proposed network

Hence, we use overlapping moving crop picture approach and feed cropped small images into network as input. To lessen negative influence of metal components in image, we apply background interference and metal reflection removal methods (as shown in Figure 3).

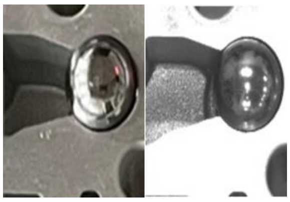

Figure 3: The effect of image reflection and background noise removal

3.2. Feature Extraction

Following images preprocessing, we improved backbone network and loss function of Siamese network. In transmission valve component assembly process, valve components are too tiny. Properties of parts such as marbles and springs are virtually similar except for size, making differentiation difficult. To address those issues, multi-scale features fusion and improved loss function are proposed to further performance of network.

3.2.1. Multi-scale Features Fusion

In order to focus on small objects while reducing redundant shallow feature information in background, we design a multi-attention mechanism in shallow feature map. Its fundamental idea is to gain smaller features utilizing high resolution of shallow feature maps. Visual attention mechanism allows focusing on a portion of image rather of seeing entire region. However, due to small, severe reflection, and sparse properties characteristics of parts, more attention should be paid to characteristics of a specific section of image throughout transmission valve component assembly process. Therefore, multi-attention mechanism is included specifically after shallow features 1 and 2 to increase attention to small features in image (Figure 2 illustrate its structure).

This paper contends that there are two key causes for poor performance of transmission defects localization model: (a) a lack of context information to gain tiny features; (b) characteristics of transmission valve components are taken from shallow features that lack semantic information. Hence, features fusion is specially added to network to improve it. The context features of this network are derived from second and third convolutional layers, however sizes of convolutional layers vary. Accordingly, deconvolution is utilized in feature map to match same size with target feature map. At the same time, batch normalization is used to ensure that feature values of different layers have same scale.

3.2.2. Loss Function

To reduce intra-class spacing by similar transmission valve part samples, an improved binary cross entropy and focal loss function is proposed. Total loss is a weighted sum of binary cross entropy loss and focal loss. It for each image is defined as:

$$L\big(f(X_1)-f(X_2)\big)=a*BCE\big(f(X_1)-f(X_2)\big)+(1-a)*focal(f(X_1)-f(X_2)) \tag{1}$$

where $f(X_1)-f(X_2)$ is input, $BCE()$ and $focal()$ represent binary cross entropy loss and focal loss, separately. a is a hyper-parameter.

Binary cross entropy loss is defined as:

$$BCE(X)=-\sum_{i=1}^{n}q^i\log(\widehat{q^i})+\big(1-q^i\big)log(1-\widehat{q^i}) \tag{2}$$

where X is input of transmission assembly dataset. q^i and $\widehat{q^i}$ are values of true probability distribution and anticipated probability distribution, respectively.

Focal loss is defined as:

$$focal(X)=-b(1-q)^c*Xlog(q)-(1-b)q^c*(1-\mathrm{X})\log(1-\mathrm{q}) \tag{3}$$

where b and c are hyper parameters.

3.2.3. Model Application

We design a model application system in transmission valve component production and assembly process. A visual demonstration platform is created using python-GUI programming and OpenCV interface. It allows for the visualization of anti-error detection during transmission valve component construction process, as well as repair and testing of algorithm. An intelligent transmission valve component assembly industrial computer module is deployed based on a fixed station in actual production workshop of factory. It realizes off-line real-time anti-error detection in the transmission valve component assembly process, allowing workers to check assembly errors and ensure practical application of factory workshop production.

4. Experiments

4.1. Experiment Settings

4.1.1. Transmission Assembly Dataset

Extensive experiments are being conducted on transmission assembly dataset. It is from Luzhou Rongda Intelligent Transmission Co. LTD. Images of transmission assembly dataset have two types: 'Wrong assembly' and 'Correct assembly'. Then, the size of input images is 128x128. As a training set, 20,000 images are chosen at random. Meanwhile, 850 images are chosen as testing set at random.

4.1.2. Training Settings

These experiments use the PyTorch framework and are trained on a GPU GTX-2060. On GPU, each batch contains 2 images. 300 epochs are adopted to experiments. At the same time, batch size is 16. For batches, the learning rate is 0.01 and weight decays to 0.0005.

4.1.3. Evaluation Metrics

In those experiments, in order to explore effective transmission defect localization network in this case. Two indicators are adopted: $accuracy$ and $F1-score$ [27]. The higher $accuracy$ and $F1-score$, the better method.

4.2. Ablation Study

4.2.1. Impact of Hyper Parameters in Loss

We conduct ablation experiment with assessment metrics to determine the hyper parameters in loss function. The result of ablation experiment is depicted in Figure 4. According to result, the best value for hyper-parameter a in loss function is 0.30.

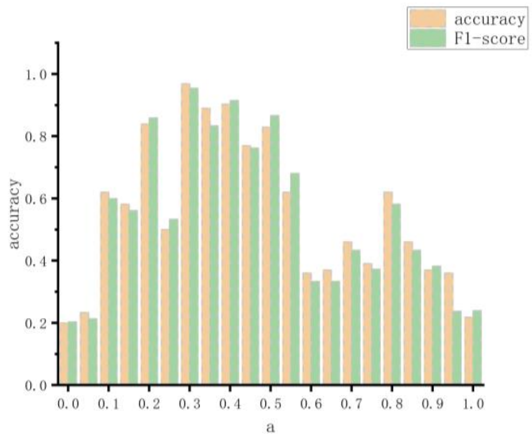

Figure 4: Comparsion on $accuracy$ and $F1-score$ under different values of hyper parameters a

4.2.2. Impact of Training Epoch

We compare training epochs on transmission assembly dataset to examine the effect of training epochs on use of transmission defects localization network. The association between $accuracy$, $F1-score$ and transmission defects localization network training epochs is depicted in Figure 5. The number of our network training epochs is given on the X-axis. $accuracy$ and $F1-score$ are described on the Y-axis.

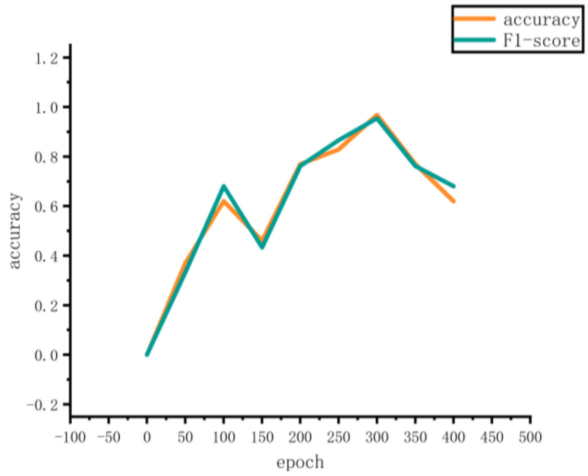

Figure 5: $accuracy$ and $F1-score$ are in connection with respect to training epochs

As seen in Figure 5, experimental outcome of our model has an overall ascending and then diminishing pattern as the number of transmission defects localization network training epochs increased. Furthermore, when training epoch exceeds 300, the performance of transmission defects localization network is excellent. Hence, we set training epochs of transmission defects localization network as 300.

4.3. Comparisons

4.3.1. Comparisons with Popular Approaches

To explore the performance of our network, we also compare proposed model with some state-of-the-arts approaches on transmission assembly dataset. Indicators are used to assess the performance of all methods on transmission assembly dataset. We compare our method to common approaches such as perceptual hashing, hamming distance, cosine distance, KNN [28], VarifocalNet [29], and DETR [30] (as shown in Figure 6). Based on results, our approach outperforms previous comparison methods for both indicators $accuracy$ and $F1-score$.

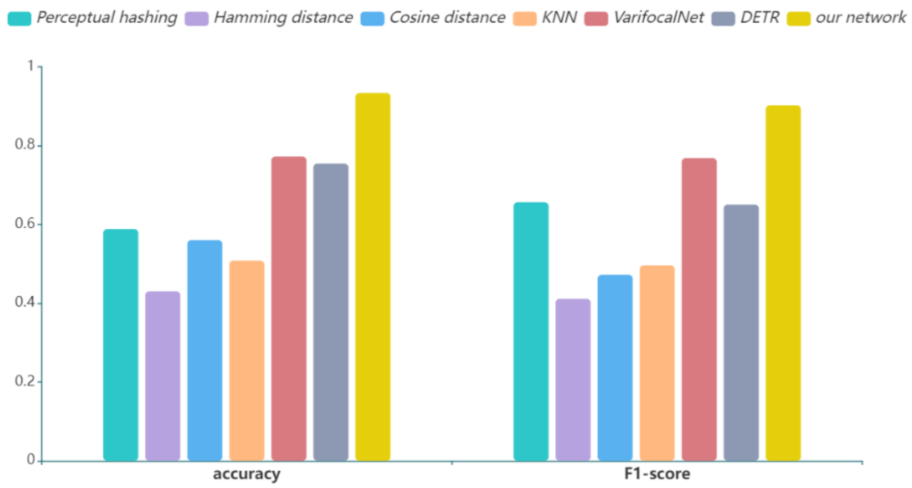

Figure 6: Comparison between our network and state-of-the-arts on transmission assembly dataset

Meanwhile, as shown in Figure 7, our model is integrated into station management system, and an application demonstration is carried out successfully. The experiments demonstrate that our method can significantly improve the performance of assembly error detection during manual assembly of transmission valve components.

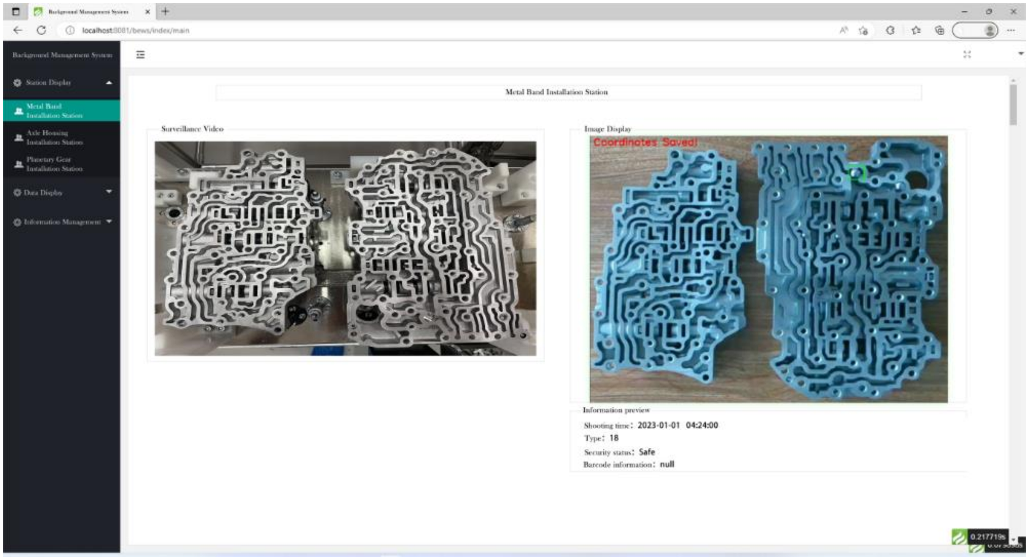

Figure 7: The operation interface of station management system

5. Conclusions

In this paper, we mainly focused on localizing transmission assembly defects. We provided an effective transmission defects localization network which can be evaluated similarity of transmission valve components. Firstly, a depth multi-scale features fusion network was established. Then, we designed improved binary cross entropy and focal loss and this model was trained by this loss function. We more thoroughly examined the impact of the features fusion method in the ablation investigation. We also looked test effects under different models training in order to determine the impact of various models on localization wrong assembly outcomes. We conduct experiments on transmission assembly dataset, which verify and explain the effectiveness of our approach. We hope that our work can provide a new angle of error proofing detection, to facilitate the quality of transmission component assembly in practice.

Acknowledgements

This work was supported by Luzhou Science and Technology Plan Project 2021-WXD-120.

References

[1] X. Liu, D. Xiang, C. Zeng, and F. Huang, "The investigation of coupled vibration of automobile transmission system and its intelligent design," Machines, vol. 10, no. 6, p. 428, 2022.

[2] B. Gao, D. Meng, W. Shi, W. Cai, S. Dong, Y. Zhang, and H. Chen, "Topology optimization and the evolution trends of two-speed transmission of evs," Renewable and Sustainable Energy Reviews, vol. 161, p. 112390, 2022.

[3] H. Zhu, "Application of arma prediction algorithm in gear fault diagnosis of automobile drive axle," in International Conference on Frontier Computing. Springer, 2022, pp. 1369–1374.

[4] C. Kong, K. Ma, and J. Wei, "Information encryption transmission method of automobile communication network based on neural network," in International Conference on Machine Learning for Cyber Security. Springer, 2023, pp. 243–256.

[5] R. Bogue, "The changing face of the automotive robotics industry," Industrial Robot: the international journal of robotics research and application, 2022.

[6] Z. Ren, F. Fang, N. Yan, and Y. Wu, "State of the art in defect detection based on machine vision," International Journal of Precision Engineering and Manufacturing-Green Technology, vol. 9, no. 2, pp. 661–691, 2022.
[7] M. S. Almutairi, "Deep learning-based solutions for 5g network and 5g-enabled internet of vehicles: advances, meta-data analysis, and future direction," Mathematical Problems in Engineering, vol. 2022, 2022.
[8] Z. Wang, Y. Ling, X. Wang, D. Meng, L. Nie, G. An, and X. Wang, "An improved faster r-cnn model for multi-object tomato maturity detection in complex scenarios," Ecological Informatics, vol. 72, p. 101886, 2022.
[9] Z. Cai and N. Vasconcelos, "Cascade r-cnn: Delving into high quality object detection," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 6154–6162.
[10] P. Malge and R. Nadaf, "Pcb defect detection, classification and localization using mathematical morphology and image processing tools," International journal of computer applications, vol. 87, no. 9, 2014.
[11] V. H. Gaidhane, Y. V. Hote, and V. Singh, "An efficient similarity measure approach for pcb surface defect detection," Pattern Analysis and Applications, vol. 21, no. 1, pp. 277–289, 2018.
[12] B. Kaur, G. Kaur, and A. Kaur, "Detection and classification of printed circuit board defects using image subtraction method," in 2014 Recent Advances in Engineering and Computational Sciences (RAECS). IEEE, 2014, pp. 1–5.
[13] T. Liu, L. Huang, and B. Chen, "Real-time defect detection of laser additive manufacturing based on support vector machine," in Journal of Physics: Conference Series, vol. 1213, no. 5. IOP Publishing, 2019, p. 052043.
[14] K. Liu, N. Luo, A. Li, Y. Tian, H. Sajid, and H. Chen, "A new self-reference image decomposition algorithm for strip steel surface defect detection," IEEE Transactions on Instrumentation and Measurement, vol. 69, no. 7, pp. 4732–4741, 2019.
[15] R. Zhang, E. Maggard, Y. Bang, M. Cho, M. Shaw, and J. Allebach, "Streak detection in color printed customer content," Electronic Imaging, vol. 2021, no. 16, pp. 254–1, 2021.
[16] R. Medico, N. Lambrecht, H. Pues, D. V. Ginste, D. Deschrijver, T. Dhaene, and D. Spina, "Machine learning based error detection in transient susceptibility tests," IEEE Transactions on Electromagnetic Compatibility, vol. 61, no. 2, pp. 352–360, 2018.
[17] M. Mahdavi and Z. Abedjan, "Reds: Estimating the performance of error detection strategies based on dirtiness profiles," in Proceedings of the 31st International Conference on Scientific and Statistical Database Management, 2019, pp. 193–196.
[18] L. Yao and J. Miller, "Tiny imagenet classification with convolutional neural networks," CS 231N, vol. 2, no. 5, p. 8, 2015.
[19] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. L. Zitnick, "Microsoft coco: Common objects in ´ context," in European conference on computer vision. Springer, 2014, pp. 740–755.
[20] M. Everingham, S. A. Eslami, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes challenge: A retrospective," International journal of computer vision, vol. 111, no. 1, pp. 98–136, 2015.
[21] C.-Y. Chuang, J. Robinson, Y.-C. Lin, A. Torralba, and S. Jegelka, "Debiased contrastive learning," Advances in neural information processing systems, vol. 33, pp. 8765–8775, 2020.
[22] M. Kaya and H. S¸. Bilge, "Deep metric learning: A survey," Symmetry, vol. 11, no. 9, p. 1066, 2019.
[23] I. Melekhov, J. Kannala, and E. Rahtu, "Siamese network features for image matching," in 2016 23rd international conference on pattern recognition (ICPR). IEEE, 2016, pp. 378–383.
[24] D. Chicco, "Siamese neural networks: An overview," Artificial Neural Networks, pp. 73–94, 2021.
[25] W. Xiao and Y. Ding, "A two-stage siamese network model for offline handwritten signature verification," Symmetry, vol. 14, no. 6, p. 1216, 2022.
[26] Y. Zhang, L. Wang, J. Qi, D. Wang, M. Feng, and H. Lu, "Structured siamese network for real-time visual tracking," in Proceedings of the European conference on computer vision (ECCV), 2018, pp. 351–366.
[27] J. Li, H. He, H. He, L. Li, and Y. Xiang, "An end-to-end framework with multisource monitoring data for bridge health anomaly identification," IEEE Transactions on Instrumentation and Measurement, vol. 70, pp. 1–9, 2020.
[28] S. Sheynin, O. Ashual, A. Polyak, U. Singer, O. Gafni, E. Nachmani, and Y. Taigman, "Knn-diffusion: Image generation via large-scale retrieval," arXiv preprint arXiv:2204.02849, 2022.
[29] H. Zhang, Y. Wang, F. Dayoub, and N. Sunderhauf, "Varifocalnet: An iou-aware dense object detector," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021, pp. 8514–8523.
[30] X. Zhu, W. Su, L. Lu, B. Li, X. Wang, and J. Dai, "Deformable detr: Deformable transformers for end-to-end object detection," arXiv preprint arXiv:2010.04159, 2020.