Deep Learning Defect Detection Methods for Industrial Products

Subjects: Automation & Control Systems | Engineering, Electrical & Electronic

Contributor: Alireza Saberironaghi , Jing Ren , Moustafa El-Gindy

Deep learning models based on convolutional neural networks (CNN) have had a lot of success in various computer vision fields, such as recognizing faces, identifying pedestrians, detecting text in images, and tracking targets. Additionally, these models are used in a wide range of industrial settings for defect detection.

defect detection | surface defect detection | defect detection for X-ray images | defect recognition | deep learning

1. Deep Learning Surface Defect Detection Methods for Industrial Products

Deep learning has become increasingly popular in the field of defect detection due to its rapid development. This section summarizes the state of research on inspection of industrial products for detecting surface defects. Learning-based approaches are classified as supervised, semi-supervised, and unsupervised. The performance of learning-based methods is best optimized when large datasets are provided. In particular, supervised techniques perform well when there are sufficient examples of each class in the dataset.

1.1. Supervised

Supervised detection requires large datasets of defect-free and defective samples labeled in a training set. Since all the training data is labeled, detection rates can be very high. It must be noted, however, that supervised detection may not always the most effective approach due to the imbalance of classes in the datasets. There are a number of datasets that supervised learning methods use, including the fabric dataset [1], rail defect dataset [2], and railroad dataset [3].

Deep neural networks and feature extraction and classification methods used in supervised methods differ in their structures. For example, detecting cross-category defects without retraining was proposed using a two-layer neural network in the literature [4]. Based on structural similarities between image pairs, the method learns differential features, which may result in some structural similarities among different classification objects. This method has been shown to be able to detect defects in different types of factories based on experiments in real factory datasets. Literature [5] suggests that the composition of kernels is more important than the number of layers when it comes to detection results. To detect small defects and textures in surface images, it is necessary to use a sample image that is large enough for computational accuracy and reducing the cost of the network. ShuffleNet uses convolution of pointwise groups and channel shuffle as two new techniques to achieve this goal. Literature [6] proposes a novel in-line inspection system for plastic containers based on ShuffleNet V2. The system can be used to inspect images on complex backgrounds as well. In [7], they proposed ShuffleDefectNet, a deep-learning-based defect detection system that achieved 99.75% accuracy on the NEU dataset.

Reference [8] suggested that shallow CNN networks can be used to identify anomalies. To train the model, only negative images are used and the research employs full-size images. The argument is that it is not necessary to have full-size examples of both defective and defect-free samples, as the negative samples already have pixels that correspond to the defect-free regions. Based on the Fast R-CNN model, Faster R-CNN introduces a region proposal network (RPN), which enables an end-to-end learning algorithm. This leads to a near-costless regional recommendation algorithm that significantly improves the speed of target detection. Faster R-CNN was used in [9] to detect PCB surface defects, a new network was proposed combining ResNet50, GRAPN residual units, and ShuffleNetV2. Using a cascaded RCNN structure, as described in literature [10], the defect detection problem of power line insulators can be changed into a two-level target detection problem.

In limited hardware configurations, MobileNet-SSD [11] improves real-time object detection performance. There is no need to sacrifice accuracy for the reduction of parameters in this network. An SSD network classifies regression and boundary box regression using various convolution layers. Translation invariance and variability are resolved in this model, resulting in good detection precision and speed. Object detection is effective when defects have regular or predictable shapes [12]. Additional preprocessing steps can be applied to more complex defect types. Fully Convolutional Networks (FCNs) use all convolutional layers as network layers; label maps can be directly derived using pixel-level prediction. To achieve accurate results, a deconvolution layer with larger data sizes is used. In literature [13], FCN and Faster R-CNN were combined to develop a deep learning model that could detect stains, leaks, and pipeline blockages in tunnels. A method for segmenting defects in solar cell electroluminescence pictures was presented in [14]. A defect segmentation map was obtained in one step by combining FCN with a specific U-net architecture.

1.2. Unsupervised

Research has begun to explore unsupervised methods to overcome the disadvantages of supervised methods. By learning the inherent characteristics of the input training data, the machine can learn some of its own characteristics and connections when there is no label information and automatically classifies the input training data based on the pattern of these unlabeled data [15]. It automatically classifies these unlabeled data based on inherent characteristics and connections between the data. Methods based on reconstruction and embedding similarity are the most commonly used to detect surface defects among unsupervised learning methods. Reconstruction-based methods such as autoencoders (AEs) and Generative Adversarial Networks (GANs) are most commonly used. Popular algorithms include PaDIM [16], SPADE [17] PatchCore [18], etc. In [19], an algorithm based on DBN was proposed for detecting defects in solar cells. Both training and reconstructed images were used as supervision data by the fine-tuning network of the BP algorithm. Literature [20] proposed a multi-scale convolutional denoising autoencoder with high accuracy and robustness that synthesizes the results of multiple pyramid levels.

A SOM-based detection method was proposed in [21] for determining the difference between normal and defective wood. The first stage involves detecting suspected defect areas, and the second stage involves separately inspecting each defect area. A detection method that uses GANs was proposed in reference [22]. The method is divided into two stages: first, a generative network and a learning mechanism based on statistical representation are used to detect new areas. In the second stage, defects and normal samples are directly distinguished using the Frechet distance. The solar panel dataset was used to test the method, and it achieved 93.75% accuracy.

A multiscale AE with fully convolutional neural networks has been proposed [23], in which each FCAE sub-network directly obtains the original feature image from the input image and performs feature clustering. Utilizing a fully convolutional neural network, the residual images were combined to create the defect image. PatchCore, introduced in literature [18], is a technique for identifying and isolating abnormal data in scenarios where only normal examples are available. It balances the need to retain normal context through memory banks of patch-level features extracted from pre-trained ImageNet networks and minimize computational time via coreset subsampling to create a leading system for cold-start image anomaly detection and localization that is efficient on industrial benchmarks. On MVTec, the algorithm demonstrated an AUROC of over 99%, while also being highly efficient in small training set scenarios. Literature [24] presented a GAN-based surface vision detection framework that uses OTSU to segment fusion feature response maps and fuses the responses of the three layers of the GAN discriminator. The framework has been proven effective on datasets of wood cracks and road cracks.

1.3. Semi-Supervised

As a result of combining the properties of supervised and unsupervised methods, semi-supervised methods are developed. Only normal samples are used as training data for semi-supervised defect detection and a defect-free boundary is learned and set, and any samples outside the boundary are considered anomalous. Since there are few defective samples to be obtained, the method is extremely useful. Nevertheless, this method has lower accuracy in defect detection compared to supervised methods. Unlabeled sample data can be automatically generated by semi-supervised methods without manual intervention.

A framework for identifying defects in PCB solder joints was proposed in literature [25], which utilizes a combination of active learning and self-training through a sample query suggestion algorithm for classification. The framework has been demonstrated to improve classification accuracy while reducing the need for manual annotations. A semi-supervised model of convolutional autoencoder (CAE) and generative adversarial network is proposed in [26]. After training with unlabeled data, the stacked CAE's encoder network is retained and input into the SoftMax layer as a GAN discriminator. Using GAN, false images of steel surface defects were generated to train the discriminator. For the detection of steel surface defects, literature [27] developed a WSL framework combining localization networks (LNets) and decision networks (DNets), with LNets trained by image level labels and outputs a heat map of potential defects as input to DNets. Through the use of the RSAM algorithm to weight the regions identified by

LNet, the proposed framework has been demonstrated to be effective on real industrial datasets. The application prospects for weakly supervised methods are also wide because the methods simultaneously combine advantages of both supervised and unsupervised methods. There are few weakly supervised methods for detecting surface defects in industrial products. The literature [28] proposed a deep learning algorithm to learn defects from a variety of defect types with an unbalanced training sample pool for PCBA manufacturing products. In this method, an overall defect recognition accuracy of 98% is achieved in PBCA images using a novel batch sampling method and the sample weighted cost function.

A semi-supervised learning system that generates samples to detect surface defects was proposed according to the literature [29]. As part of the semi-supervised learning part, two CDCGAN and ResNet18 classifiers were used, and the NEU-CLS dataset was used to compare the two classifiers. In this way, supervised learning and transfer learning are both shown to be inferior to the method. A convolutional neural network structure based on residual network structures was proposed in [30] by stacking two layers of residual building modules together, resulting in a 43-layer convolutional neural network, while at the same time by appropriately increasing the network width; a more balanced network depth and network width can be obtained and accuracy can be improved. The network structure shows good performance on the DAGM, NEU steel, and copper clad plate datasets.

2. Deep Learning Defect Detection Methods for X-ray Images for Industrial Products

Non-destructive testing (NDT) is a method that uses radiography or ultrasound technologies to discover faults without causing damage to the detected objects. It is widely used in engineering industries to detect and evaluate defects in materials of all types.

An important technique in non-destructive testing is radiographic testing, which uses X-rays to identify and evaluate flaws or defects, such as cracks or porosities. Defects can appear in X-ray images in many shapes and sizes, making detection difficult. The images are often low contrast and noisy, making identification of defects difficult.

The traditional approach for identifying defects in industrial products is for human operators or experts to visually inspect radiographs. However, this method can be subjective and prone to errors. Additionally, the process of examining a large number of images can be time-consuming and may lead to misinterpretations. However, there have been significant advancements in the field of defect detection in recent years, thanks to the emergence of deep learning techniques. As a result, a number of methods for detecting defects have been proposed, which are more efficient and reliable than the conventional approach. This section aims to provide a summary of current research on industrial product defect detection methods using X-ray images. Specifically, it covers the use of deep learning techniques such as convolutional neural networks and generative adversarial networks to analyze radiographic images and identify defects with a high degree of accuracy. These methods have the potential to reduce the subjectivity and human errors associated with the traditional approach, as well as the time required for inspection. Additionally, they can be trained to improve over time with more data, making them more robust and reliable.

A proposed system in literature [31] aimed to automate the process of inspecting and monitoring the condition of machines in the hard metal industry by analyzing defects in real production samples. Three models were created to analyze different types of data, a method called stacked generalization ensemble was applied and a random forest classifier was utilized to combine and analyze the results of the microprofilometer and ultrasound models. The fusion model was found to have improved performance and higher classification accuracy (88.24%) as compared to the individual models. Additionally, the shop floor model was able to effectively identify breakdowns during the manufacturing process and the ultrasound model was found to have better classification scores compared to the VGG-19 model. According to literature [32], a three-stage deep learning algorithm was proposed for detecting bubble patterns in engines. The algorithm consisted of training an autoencoder using normal images, fixing the coefficients of the encoder, and training a fully-connected network using both normal and defective images. To improve the performance of the network, the entire system was fine-tuned. According to [33], a CNN model was designed with ten layers that belong to six grades for detecting defects in X-ray welding images. It was possible to achieve 98.8% classification accuracy using CNN if the ReLU activation function was used for X-ray welding image recognition. A real-time X-ray image analysis method using Support Vector Machines (SVMs) was presented in [34]. Using a background subtraction algorithm, all potential defects were segmented, and three features were extracted, including the defect area, the grayscale average difference, and the grayscale standard deviation. In order to distinguish non-defects from defects, the extracted features were input into an SVM classifier. A real-time X-ray image defect detection method based on the proposed method reduced undetected defects and false alarms. Another SVM-based method for detecting weld defects was described in [35]. The training SVM is trained by extracting three feature vectors from potential weld defects using grey-level profile analysis. In the last step, the SVM is trained to differentiate between defects that are real and those that are potential. A high percentage of correct detections could be achieved using the proposed method. For detecting insert molding in automotive electronics, ref. [36] proposed a Yolov5-based DR image defect detection algorithm. Width and a window level are adjusted in the preprocessing stage of the acquired data, and fast guided filtering is used for edge retention. Using the overlap, tiny anomalies are detected, and a multi-task dataset is constructed. Using Ghost, which replaces the standard convolutional network with the backbone network with enhanced features, the number of parameters can be further reduced. Moreover, CSP-modules are embedded in the neck and backbone of the network to enhance feature extraction. As a result of adding the transformer attention module after spatial pyramid pooling, over-fitting can be avoided while computational effort can be reduced. DR data-based Yolo series target detection algorithms are used as a final step to conduct consistent experiments. For detecting bead toe errors, ref. [37] proposed a lightweight semantic segmentation network. An encoder extracts the texture features of different regions of the tire in the network first. Then, to fuse the encoder's output feature, a decoder is introduced. A reduction in the dimension of the feature maps has allowed the positions of the bead toe to be recorded in the X-ray image. An index of local mIoU (L-mIoU) is proposed to evaluate the final segmentation effect. YOLOv3_EfficientNet is used as the backbone of the methodology instead of YOLOv3_darknet53. It results in a substantial improvement in YOLOv3 mean average precision, as well as a substantial reduction in inference time and storage space. DR image features are then used to enhance the data, thereby increasing the diversity of the clarity and shape of defects. With depth separable convolution, models can be deployed on embedded devices with acceptable accuracy loss ranges. A method was presented in [38] that utilizes deep learning with X-ray images to detect defects in aluminum casting parts used in automobiles, with the goal of improving the accuracy of both the algorithm and data augmentation. The study found that using Feature Pyramid Networks (FPNs) resulted in a 40.9% increase in Mean of Average Precision (mAP) value, making it the most effective modification. Additionally, using RoIAlign instead of RoI pooling in Faster R-CNN improved the accuracy of bounding box location. The study also proposed various data augmentation methods to compensate for the limited availability of X-ray image datasets for defect detection. The results showed that the mAP values for each data augmentation method reached an optimal value and did not continue to increase as the number of datasets increased. Overall, the proposed improvements to the Faster R-CNN algorithm resulted in better performance for X-ray image defect detection of automobile aluminum casting parts. Using the Faster R-CNN detection model with X-ray preprocessing was applied to the detection of tire defects in [39] to improve curve fitting performance. Faster R-CNN precision and recall of defects were improved by adjusting its feature extractor, proposal generator, and box classifier. According to literature [40], triplet deep neural networks can be used to detect weld defects. X-ray images are first preprocessed into relief images to make defects easier to identify. Following that, a deep network is constructed based on triplets, and a feature vector is obtained by mapping the triplets. The distance between similar defect feature vectors and the distance between different types of defect feature vectors must be closer. The SVM is also used for automatic detection and classification of weld defects. Based on the results of two experiments, the proposed method is capable of effectively detecting multiple defects.

References

1. Silvestre-Blanes, J.; Albero-Albero, T.; Miralles, I.; Pérez-Llorens, R.; Moreno, J. A Public Fabric Database for Defect Detection Methods and Results. Autex Res. J. 2019, 19, 363–374.
2. Faghih-Roohi, S.; Hajizadeh, S.; Núñez, A.; Babuska, R.; De Schutter, B. Deep convolutional neural networks for detection of rail surface defects. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 2584–2589.
3. Gan, J.; Li, Q.; Wang, J.; Yu, H. A Hierarchical Extractor-Based Visual Rail Surface Inspection System. IEEE Sens. J. 2017, 17, 7935–7944.
4. Luan, C.; Cui, R.; Sun, L.; Lin, Z. A Siamese Network Utilizing Image Structural Differences For Cross-Category Defect Detection. In Proceedings of the 2020 IEEE International Conference on Image Processing (ICIP), Abu Dhabi, United Arab Emirates, 25–28 October 2020; pp. 778–782.
5. Park, J.-K.; Kwon, B.-K.; Park, J.-H.; Kang, D.-J. Machine learning-based imaging system for surface defect inspection. Int. J. Precis. Eng. Manuf.-Green Technol. 2006, 3, 303–310.
6. Liang, Q.; Zhu, W.; Sun, W.; Yu, Z.; Wang, Y.; Zhang, D. In-line inspection solution for codes on complex backgrounds for the plastic container industry. Measurement 2019, 148, 106965.

7. Anvar, A.; Cho, Y.I. Automatic Metallic Surface Defect Detection using ShuffleDefectNet. J. Korea Soc. Comput. Inf. 2020, 25, 19–26.
8. Racki, D.; Tomazevic, D.; Skocaj, D. Towards surface anomaly detection with deep learning. Procedia CIRP 2019, 79, 484–489.
9. Hu, B.; Wang, J. Detection of PCB Surface Defects With Improved Faster-RCNN and Feature Pyramid Network. IEEE Access 2020, 8, 108335–108345.
10. Tao, X.; Zhang, D.; Wang, Z.; Liu, X.; Zhang, H.; Xu, D. Detection of Power Line Insulator Defects Using Aerial Images Analyzed With Convolutional Neural Networks. IEEE Trans. Syst. Man Cybern. Syst. 2020, 50, 1486–1498.
11. Chang, C.-F.; Wu, J.-L.; Chen, K.-J.; Hsu, M.-C. A hybrid defect detection method for compact camera lens. Adv. Mech. Eng. 2017, 9, 1687814017722949.
12. Song, L.; Li, X.; Yang, Y.; Zhu, X.; Guo, Q.; Yang, H. Detection of Micro-Defects on Metal Screw Surfaces Based on Deep Convolutional Neural Networks. Sensors 2018, 18, 3709.
13. Gao, X.; Jian, M.; Hu, M.; Tanniru, M.; Li, S. Faster multi-defect detection system in shield tunnel using combination of FCN and faster RCNN. Adv. Struct. Eng. 2019, 22, 2907–2921.
14. Balzategui, J.; Eciolaza, L.; Arana-Arexolaleiba, N. Defect detection on Polycrystalline solar cells using Electroluminescence and Fully Convolutional Neural Networks. In Proceedings of the 2020 IEEE/SICE International Symposium on System Integration (SII), Honolulu, HI, USA, 12–15 January 2020; pp. 949–953.
15. Dike, H.U.; Zhou, Y.; Deveerasetty, K.K.; Wu, Q. Unsupervised Learning Based On Artificial Neural Network: A Review. In Proceedings of the 2018 IEEE International Conference on Cyborg and Bionic Systems (CBS), Shenzhen, China, 25–27 October 2018; pp. 322–327.
16. Defard, T.; Setkov, A.; Loesch, A.; Audigier, R. PaDiM: A Patch Distribution Modeling Framework for Anomaly Detection and Localization. arXiv 2020, arXiv:2011.08785.
17. Cohen, N.; Hoshen, Y. Sub-Image Anomaly Detection with Deep Pyramid Correspondences. arXiv 2020, arXiv:2005.02357.
18. Roth, K.; Pemula, L.; Zepeda, J.; Schölkopf, B.; Brox, T.; Gehler, P. Towards Total Recall in Industrial Anomaly Detection. arXiv 2021, arXiv:2106.08265.
19. Wang, X.; Li, J.; Yao, M.; He, W.; Qian, Y. Solar Cells Surface Defects Detection Based on Deep Learning. Pattern Recognit. Artif. Intell. 2014, 27, 517–523.
20. Mei, S.; Yang, H.; Yin, Z. An Unsupervised-Learning-Based Approach for Automated Defect Inspection on Textured Surfaces. IEEE Trans. Instrum. Meas. 2018, 67, 1266–1277.

21. Silvén, O.; Niskanen, M.; Kauppine, H. Wood inspection with non-supervised clustering. Mach. Vis. Appl. 2003, 13, 275–285.

22. Lai, Y.T.K.; Hu, J.S. A Texture Generation Approach for Detection of Novel Surface Defects. In Proceedings of the 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 7–10 October 2018; pp. 4357–4362.

23. Yang, H.; Chen, Y.; Song, K.; Yin, Z. Multiscale Feature-Clustering-Based Fully Convolutional Autoencoder for Fast Accurate Visual Inspection of Texture Surface Defects. IEEE Trans. Autom. Sci. Eng. 2019, 16, 1450–1467.

24. Zhai, W.; Zhu, J.; Cao, Y.; Wang, Z. A Generative Adversarial Network Based Framework for Unsupervised Visual Surface Inspection. In Proceedings of the 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Calgary, AB, Canada, 15–20 April 2018; pp. 1283–1287.

25. Dai, W.; Mujeeb, A.; Erdt, M.; Sourin, A. Soldering defect detection in automatic optical inspection. Adv. Eng. Inform. 2020, 43, 101004.

26. Di, H.; Ke, X.; Peng, Z.; Dongdong, Z. Surface defect classification of steels with a new semi-supervised learning method. Opt. Lasers Eng. 2019, 117, 40–48.

27. Xu, L.; Lv, S.; Deng, Y.; Li, X. A Weakly Supervised Surface Defect Detection Based on Convolutional Neural Network. IEEE Access 2020, 8, 42285–42296.

28. Ren, J.; Gabbar, H.A.; Huang, X.; Saberironaghi, A. Defect Detection for Printed Circuit Board Assembly Using Deep Learning. In Proceedings of the International Conference Control Science and System Engineering (ICCSSE), Guangzhou, China, 14–16 July 2022.

29. He, Y.; Song, K.; Dong, H.; Yan, Y. Semi-supervised defect classification of steel surface based on multi-training and generative adversarial network. Opt. Lasers Eng. 2019, 122, 294–302.

30. Zheng, X.; Wang, H.; Chen, J.; Kong, Y.; Zheng, S. A Generic Semi-Supervised Deep Learning-Based Approach for Automated Surface Inspection. IEEE Access 2020, 8, 114088–114099.

31. Kotsiopoulos, T.; Leontaris, L.; Dimitriou, N.; Ioannidis, D.; Oliveira, F.; Sacramento, J.; Amanatiadis, S.; Karagiannis, G.; Votis, K.; Tzovaras, D.; et al. Deep multi-sensorial data analysis for production monitoring in hard metal industry. Int. J. Adv. Manuf. Technol. 2021, 115, 823–836.

32. Ren, J.; Ren, R.; Green, M.; Huang, X. Defect Detection from X-Ray Images Using A Three-Stage Deep Learning Algorithm. In Proceedings of the 2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE), Edmonton, AB, Canada, 5–8 May 2019; pp. 1–4.

33. Li, Y.; Gao, W. Research on X-ray welding image defect detection based on convolution neural network. J. Phys. Conf. Ser. 2019, 1237, 032005.

34. Shao, J.; Shi, H.; Du, D.; Wang, L.; Cao, H. Automatic weld defect detection in real-time X-ray images based on support vector machine. In Proceedings of the 2011 4th International Congress on Image and Signal Processing, Shanghai, China, 15–17 October 2011; pp. 1842–1846.

35. Wang, Y.; Guo, H. Weld Defect Detection of X-ray Images Based on Support Vector Machine. IETE Tech. Rev. 2014, 31, 137–142.

36. Wang, B.; Huang, F. A Lightweight Deep Network for Defect Detection of Insert Molding Based on X-ray Imaging. Sensors 2021, 21, 5612.

37. Yi, X.; Peng, C.; Zhang, Z.; Xiao, L. The defect detection for X-ray images based on a new lightweight semantic segmentation network. Math. Biosci. Eng. 2022, 19, 4178–4195.

38. Du, W.; Shen, H.; Fu, J.; Zhang, G.; He, Q. Approaches for improvement of the X-ray image defect detection of automobile casting aluminum parts based on deep learning. NDT E Int. 2019, 107, 102144, ISSN 0963-8695.

39. Chen, J.; Li, Y.; Zhao, J. X-ray of Tire Defects Detection via Modified Faster R-CNN. In Proceedings of the 2019 2nd International Conference on Safety Produce Informatization (IICSPI), Chongqing, China, 28–30 November 2019; pp. 257–260.

40. Liu, X.; Liu, J.; Qu, F.; Zhu, H.; Lu, D. A Weld Defect Detection Method Based on Triplet Deep Neural Network. In Proceedings of the 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 22–24 August 2020; pp. 649–653.