Computers in Industry 148 (2023) 103911

Contents lists available at ScienceDirect

Computers in Industry

journal homepage: www.sciencedirect.com/journal/computers-in-industry

Deep CNN-based visual defect detection: Survey of current literature

Shashi Bhushan Jha, Radu F. Babiceanu *

Department of Electrical Engineering and Computer Science, Embry-Riddle Aeronautical University, Daytona Beach, FL, USA

A R T I C L E I N F O

Keywords:
Automatic optical inspection
Deep convolutional neural networks
Defect detection

A B S T R A C T

In the past years, the computer vision domain has been profoundly changed by the advent of deep learning algorithms and data science. The defect detection problem is of outmost importance in high-tech industries such as aerospace manufacturing and is extensively employed using automated industrial quality control systems. Defect inspection methods can be mainly grouped into manual inspection, traditional computer vision, and modern computer vision inspection. Initially developed two decades ago, the CNN algorithms recently became popular for solving complex machine vision problems, as big datasets and computationally potent hardware became widely available. Deep learning-based methods form the foundation for modern automatic optical inspection methods and can be grouped based on their network connections into two categories: dense networks and sparse networks. Another method for grouping considers the type of learning: supervised learning used primarily for defect classification and segmentation, and unsupervised learning models, which have the potential to overcome the challenges of supervised models such as labeling images and annotating pixels. In addition, pixel-level based segmentation techniques are considered to cover the state-of-the-art methodologies for the automatic optical inspection. Still, both supervised and unsupervised models pose challenges in regards to model training and attaining the expected detection accuracy. Identified open challenges include algorithmic, application, and data processing challenges. By addressing these challenges, in the future, the demand for automated optical inspection is expected to only grow in both industry practice and academic research.

1. Introduction

The defect detection problem is extensively applied in automated industrial quality control systems. It is broadly used in quality control to monitor the manufactured components to ensure product quality is maintained. Traditionally, quality inspection in the industry is usually performed by human workers. However, the worker-based inspection operation is tedious, time consuming, and requires experienced workforce. In the current times of cyber manufacturing, smart manufacturing, and industry 4.0, new solutions such as automatic optical inspection (AOI) are essential to enable real-time quality assessment and monitoring of production quality assurance by automated inspection. Automatic inspection significantly reduces the workload of human experts, as well as the needed labor cost. In addition, the literature reports it improves the quality of the product (Dong et al., 2021).

In the past decade, computer vision has been transformed by the emergence of deep learning algorithms and big data. Moreover, the advancement in the hardware (GPU, CPU, and TPU) enables powerful and large-scale computations that make possible big data processing and training of complex models (Tulbure et al., 2022). Furthermore, there has been rapid development of programming languages including C, C++, Python, and others, as well as deep learning and machine learning tools ranging from Matlab's image processing capabilities to OpenCV, CNTK, Tensorflow, Keras, Pytorch, and more. In 1998, LeCun (LeCun et al., 1998) developed the Convolutional Neural Network (CNN) algorithm and proposed LeNet CNN architecture. These contributions strengthened the computer vision domain immensely and allowed researchers and practitioners to solve complex machine vision problems. However, until 2012, these solutions were not very popular because of the unavailability of big datasets and inefficient hardware. In 2010, the ImageNet challenge commenced in the machine learning (ML) community with the goal to train the model with big datasets to classify images. Then, in 2012, a CNN architecture named AlexNet (Alex et al., 2012) was developed. Since then, CNN models have been gaining popularity, and year after year, improved CNN architectures were developed such as GoogleNet (Szegedy et al., 2015), VGGNet (Simonyan and Zisserman, 2014), ResNet (He et al., 2016), and others. The emergence of CNN model encouraged several researchers to attain automatic

* Corresponding author.
E-mail address: babicer@erau.edu (R.F. Babiceanu).

https://doi.org/10.1016/j.compind.2023.103911
Received 25 January 2023; Received in revised form 22 March 2023; Accepted 27 March 2023
Available online 3 April 2023

optical inspection in the manufacturing domain with higher accuracy. Before wider acceptance of CNN, in earlier research, simple image processing was applied for the quality inspection process. For example, background subtraction and thresholding were used for this purpose (Dong et al., 2020a). However, these models required stable condition while capturing the picture, which was hard to achieve.

Going forward, CNN models have been widely accepted by the computer vision community in several fields. To achieve a considerable accuracy of the model, a large-labeled dataset is required to train it for the classification task. However, for the defect classification in automatic optical inspection system, acquiring a large-labeled image dataset is difficult to accomplish. Thus, most of the researchers focused on the data augmentation technique to expand the number of images with the help of shifting, rotating, shearing, and other processes. With this technique, a large training dataset can be achieved from the small number of images. However, in the manufacturing domain, it is hard to collect the same number of defective and defect-free samples in the early phase of production, which leads to an imbalanced dataset. To address this issue, researcher was conducted to develop synthesis algorithms in which a small sample of abnormal images synthesize with normal images to make the dataset large and balanced (Dong et al., 2020b). However, even with these solutions, labeling the images is still challenging, time-consuming, expensive, and requires huge human efforts. At the conclusion of the process, if large-labeled images are obtained, the model can classify the image as defective or non-defective, but it is unable to locate the defective regions.

To locate the defective area, research works tackled the defect detection problem by employing object detection and segmentation techniques. In the defect detection literature, defect segmentation approach is well-known for segmenting the defects since it works at pixel level, and thus it is more accurate than object detection approach (Dong et al., 2020a). However, the key challenge of this technique is to annotate the input image at pixel level, which is again expensive and time-consuming, and requires considerable human efforts. To address these limitations, scholars working in the defect detection area developed unsupervised learning techniques to train the model with unlabeled dataset to identify the defects. This model needs a large unlabeled dataset from manufacturing operations, which in many cases can be easily obtained; though, there could be manufacturing domains, where the scarcity of training images may need be addressed. Overall, the unsupervised segmentation technique, namely, once-class classification or outlier detection, with data augmentation can be used to achieve efficient and effective results with a smaller number of unlabeled images (Dong et al., 2021).

The remaining of the paper is outlined next. In the second section, visual defect inspection methods are categorized and discussed, along with traditional and modern automatic optical inspection (AOI) methods. Then, the deep CNN-based defect detection model is described, followed by the open challenges of AOI systems presented in the subsequent section. Finally, the conclusion and future research are discussed in the last section of the paper.

2. Visual defect inspection methods classification

The past literature covering defect inspection methods can be broadly divided into three parts: (i) manual inspection, (ii) traditional computer vision, and (iii) modern computer vision. To be more general, it can be divided into human and machine inspection. For each defect inspection method, the type of product defects identified, and evaluation of visual inspection methods are presented in the Table 1. Furthermore, traditional AOI, modern AOI, and structure of the AOI are also discussed in the following subsections.

Table 1
Evaluation of Defect Inspection Methods.

Visual Defect Inspection Methods	Defects	Benefits	Drawbacks
Manual Inspection	Cracks and wrinkle formation, die casting, scratches, floaters, open circuits, welding defects, steel surface defects, etc.	Diverse set of defects can be identified.	Tedious, time consuming, inconsistent, subjective. Labor intensive, which leads to more mistakes and requires huge human efforts, thus costly.
Traditional AOI	Steel surface defects, fabric defects, material surfaces, architectural glass materials, etc.	Lessens the labor cost and human errors and reduces the time.	Poor feature extraction. Requires skilled labor for the feature selections and extractions. Limited accuracy, lack of flexibility, less reliable and robust.
Modern AOI	Aerospace welding defects, steel, machined, and composite material defects; mobile screen defects, wafer defects, small and large foreign matters, cracks, wrinkle formation, scratches, abrasions, oil stains, dent, chips, light and severe strains, etc.	Offers high detection accuracy, real-time, high speed, more robust. Reduces labor costs significantly and can extract the features by themselves.	Requires large dataset, high-speed processor, GPU, and TPU.

2.1. Traditional automatic optical inspection methods

Traditional AOI technique includes four steps, namely, (a) data acquisition, (b) pre-processing, (c) feature extraction, and (d) defect detection. The first step, data acquisition, is needed to collect the data for the model. It requires lighting and camera for collecting the image dataset from the production line, if available datasets are not enough or useful for the modeling. Secondly, preprocessing step removes the unnecessary parts of the dataset, and this process uses the method such as noise reduction and filtering. For feature extraction and defect detection, there are several methods that can be applied for the automated defect detection. These methods can be classified into four groups: (1) statistical methods, (2) spectral methods, (3) model-based methods, and (4) learning-based methods (Ozseven, 2019). The goal of the statistical methods is to find regions with distinct spatial distribution on the input image. To extract defect features, statistical models based on spatial distribution employ the first order (one pixel), second order (two pixels) statistics and higher order with multiple pixels. First-order statistics are based on a single pixel and include measures such as mean, standard deviation, and variance. Second-order statistics involve two adjacent pixels and can provide information about spatial relationships between pixels. Examples include co-occurrence matrices. Higher-order statistics use multiple pixels and can provide more complex information about spatial relationships between groups of pixels. Some statistical methods used are co-occurrence matrix (Ashour et al., 2019), autocorrelation (Huang and Chan, 2004), histogram properties (Li et al., 2019), and edge detection (Wen and Xia, 1999). Spectral approaches transform the signals from the spatial domain to the frequency domain to identify the defect through wavelet transform (Wen et al., 2014), Fourier transform (Malek et al., 2013), and Gabor filters (Ma et al., 2018). The objective of the model-based approach is to capture the basic characteristic and detect defects by making an image model. Some standard model-based methods for defect identification are autoregressive models (Kulkarni et al., 2019) and the Markov random field (MRF) models (Xu and Huang, 2012). Learning based approaches first train the model to detect defects and then determine the defects using pattern recognition algorithms, namely, support vector machines (SVMs) (Zhang et al., 2018), k-nearest neighbors (kNNs) (Nguyen et al., 2017) and artificial neural networks (ANNs) (Chen et al., 2009).

The above-mentioned approaches were researched extensively in the literature. To consider a method or combination of these methods to apply on AOI problems depends on expert views, success results and the dataset availability. For instance, spectral or learning-based methods can be applied to obtain a defect in a patterned surface. To determine deformation on the steel surfaces, statistical or model-based methods can be used. However, these methods are generally employed in a two-stage manner, explicitly, feature extraction and defect detection, and implemented together in a hybrid style model.

For example, (Celik et al., 2014) considered a fabric defect detection problem and employed AOI methods to extract features using wavelet transform. Then, in the next stage, it classifies the defects using neural network and co-occurrence matrix. In another work, (Wang et al., 2008) discussed a weld defect detection problem and solved again using the hybrid style model. First, the features of X-ray images were extracted using multiple thresholds, and then the SVM algorithm was employed to classify the defective and non-defective features. Further, Hough transform was also applied for removing the noisy pixels in the defective region and later, the defect was isolated. In another research, principal component analysis (PCA) and ANNs were applied to detect and classify the real time arc welding defects (Mirapeix et al., 2007). Using ANN to train the plasma spectrum dataset is reported to have been difficult because of huge number of spectral lines. Thus, PCA was employed first to remove the redundant information and reduce the dimensionality, and then processed data was used to train the ANN model to detect and classify the defects.

To extract features in the traditional AOI methods, a prime role is assigned to human experts who design specific rule and adjust several parameters. Thus, the success behind these methods is highly dependent on experts (Ren et al., 2017). Moreover, these methods can perform well under certain conditions, but are sensitive to changes in real world condition. These drawbacks can be easily overcome by using deep learning. The recent advancements in deep learning can extract the high-level features from given inputs and can classify the defects without any involvement of human expertise to design features sets manually (LeCun et al., 2015). Furthermore, these models are highly robust to variations, adaptable, and can allow detection of several types of defects in various applications.

2.2. Modern automatic inspection methods

Deep learning-based methods for automatic optical inspection problems are widely accepted in the research community. Deep learning networks can be primarily divided into two parts: dense networks, if the model is based on fully connected feed forward network, such as deep neural network (DNN), and (2) sparse networks, if the model is sparsely connected, such as deep convolutional neural network (DCNN). Moreover, dense or sparse networks of deep learning methods are mainly classified into three paradigms: supervised learning, semi-supervised learning, and unsupervised learning. Supervised based learning has been the most widely used model, with convolutional neural networks for defect classification and segmentation being employed on numerous occasions. Given a large training dataset, the supervised based models can attain substantial defect detection accuracy.

For example, (Shu et al., 2021a) addressed a surface quality inspection problem of LED chips using computer vision techniques. This work proposed parallel DCNN model for labelled LED chip defects to classify the defects with considerable detection accuracy. In another work, a commutator surface defect detection problem was considered with several defects such as abrade, dark-spot, scratch, and others (Shu et al., 2021b). To solve the problem, a separable residual CNN-based model has been developed to recognize the defects in a faster way with shallow layers. The solution also reduced the number of parameters of the model due to smaller model size. The proposed model achieved a reasonable accuracy, around 93%. In another research, (Park et al., 2020) collected a large dataset of mobile phone glass cover with 16,800 images with labels of dent, scratch, chips, and other defects. The work employed a multi-DCNN to solve the cover glass defect detection problem and attempted to allow manufacturers to set up a fully automated inspection system operated at a high detection accuracy (99%). Nevertheless, collecting a large training dataset and labeling the data requires huge manpower and makes the model expensive. The scarcity of large-labeled dataset can be mitigated by semi-supervised and unsupervised learning models.

Semi-supervised methods can obtain similar or even better results than supervised methods requiring fewer labeled training dataset. In the following four reported research works employing semi-supervised machine vision techniques, the detection accuracy varies from 92% to 99%. In (Liu et al., 2021), authors considered the automated optical inspection problem with the objective to detect the anomaly. To solve this problem, semi-supervised anomaly detection using dual prototypes autoencoder model was proposed. The model is trained with Aluminum Profile Surface Defect (APSD) dataset and obtained a reasonable accuracy. In addition, the results are compared with state-of-the-art algorithms considering four different publicly available datasets, namely, Magnetic-Tile (MT) defect dataset, Road Surface Defect (RSD) dataset, Carpet Surface Defect (CSD) dataset, and APSD dataset. In another study of automated surface inspection problem (Zheng et al., 2020a), a generic semi-supervised model is developed considering two public datasets (DAGM and NEU) and an industrial dataset (CCL). The proposed model outperforms the several benchmark algorithms with 95% of accuracy.

Two weak-supervision computer vision detection methods were developed with considerable accuracy in (Li et al., 2020a), where a synthesis algorithm was proposed to simulate a large dataset of mobile phone screen defects such as light strains, severe strains, scratches, and floaters, so that the challenges of insufficient amount of training dataset can be overcome. Then the model was trained, fine-tuned, and used for defect recognition. In another research (Xu et al., 2020), automated surface inspection problem was studied, and a model was developed with a new loss function and trained with a small number of defect dataset comprising around 25 defect samples. The approach can identify the anomaly regions at image levels and can address imbalanced data at the pixel level using collaboration learning strategy by utilizing the loss function. The reported detection accuracy of both models with small real-world datasets is 95% and 99%, respectively. Nevertheless, semi-supervised models still need label training samples.

Unsupervised learning is currently one of the most attractive research directions in the machine-learning domain. Unsupervised based models work on unlabeled training samples, and thus do not require manual labeling, which reduces labor cost. The literature suggests that the widely recognized models of unsupervised learning for automatic optical inspection are based on deep autoencoders (AE) and generative adversarial network (GAN). The AE model is a distinctive unsupervised model for high dimensional data comprised of two neural networks namely, encoder and decoder. The encoder extracts the latent features from the input images, while the decoder reconstructs the input image with some loss. GAN is another typical unsupervised learning model consisting of a generative and a discriminative stage. The following paragraphs present a series of research models from the available literature, which are based on AE and GAN approaches.

First, three literature models are explained using both dense and sparse networks, both networks employing autoencoder models. In (Yu et al., 2019), process pattern recognition problem was discussed and a deep autoencoder feature learning approach based on stack de-noising AE models (SDAE) was developed for manufacturing processes to learn important features from the process signals. Moreover, the robustness of the model was checked using a large, simulated dataset and Tennessee Eastman process. To extend the work, a multivariate manufacturing process using DNNs was examined to detect other types of patterns such as cycle, trend, etc. In another work (Yu, 2019), wafer defect detection problem was investigated and an improved SDAE-based feature learning approach to recognize the defects was proposed. The detection accuracy of this improved model was reportedly enhanced to 97%. All previous reported research mentioned so far in this literature review were using normal and abnormal samples to train the model, either with supervised or unsupervised learning. But research work (Mujeeb et al., 2019) studied the automatic optical inspection problem concentrating on solder images of integrated circuits (IC) manufacturing that used only normal samples where one-class based feature learning method was developed to recognize the defects using deep autoencoder. The results of the experiments showed that both sensitivity and specificity were reasonable, around 85% and 7%, respectively. Another one-class-based research was conducted on surface inspection problem focusing on decorated plastic parts to detect the fault with improved area under receiver operator characteristic curve (AUROC) reported around 98% using small datasets (Haselmann et al., 2019).

After the widely gained success of AE models, there were several variants of AE models proposed for defect detection problem. One such example is the simple AE model that resulted in an overfit for complex problem models, but the variational autoencoder (VAE) models turned out to perform well on the same complex problems. For example, machine vision inspection system for anomaly detection was examined aiming to identify the abnormality of Electric Cathode Metal Coating (KTL Coating) (Nejc and Drago, 2019). To solve this problem, VAE model was trained with KTL coating datasets at pixel-level and the results of the model showed more robustness than those obtained using simple methods. This type of VAE model can be also used for generating the datasets to work with segmentation-based models for defect detection.

Besides AE based models, GAN based models, such as pixelGAN and cycleGAN, are also currently popular in unsupervised learning. For example, (Kim et al., 2021) studied the adversarial defect detection problem with the objective of isolating the defects. The research proposed a pixelGAN based model concentrating on semiconductor manufacturing process data to expedite the process. The results outperformed the baseline model such as CenterNet on a real industry dataset. In another research, machine surfaces and medical acne patches inspection problem were examined to detect the surface defects (Tsai et al., 2021). This work developed a two-phase deep learning model to isolate the faults at pixel level without human annotation. In the first phase, it synthesized the defects and annotate fault pixels in the input image using cycleGAN model. Then, the resultant dataset was used to train the model using U-Net semantic network. The results showed that it could be applied to distinct set of surface inspection problems with considerable detection accuracy, in the range of 95%.

The main drawback of unsupervised models is that it is not as accurate or reliable as supervised learning. However, it reduces significantly the effort for labeling the image and manual annotation at pixel-level in the image dataset, since it does not require a large number of defective samples with the semantic network. In some cases, unsupervised anomaly detection models can be trained with only one-class of dataset, thus, only small size of defective dataset can be required during the testing of the model.

2.3. Architecture of automatic optical inspection methods

Fig. 1 describes a generic automated optical visual inspection system for the external surface defect detection problem. The batches of products move on a conveyor belt with suitable lighting system. In addition, the architecture includes industrial camera sets with proper angle to capture the image data of the product and store on the image captured card, so that it can be transferred on the industrial computer.

The choice of camera depends on the specific application and the requirements for resolution, speed, and accuracy. Factors such as lighting conditions, object size, and processing requirements also need

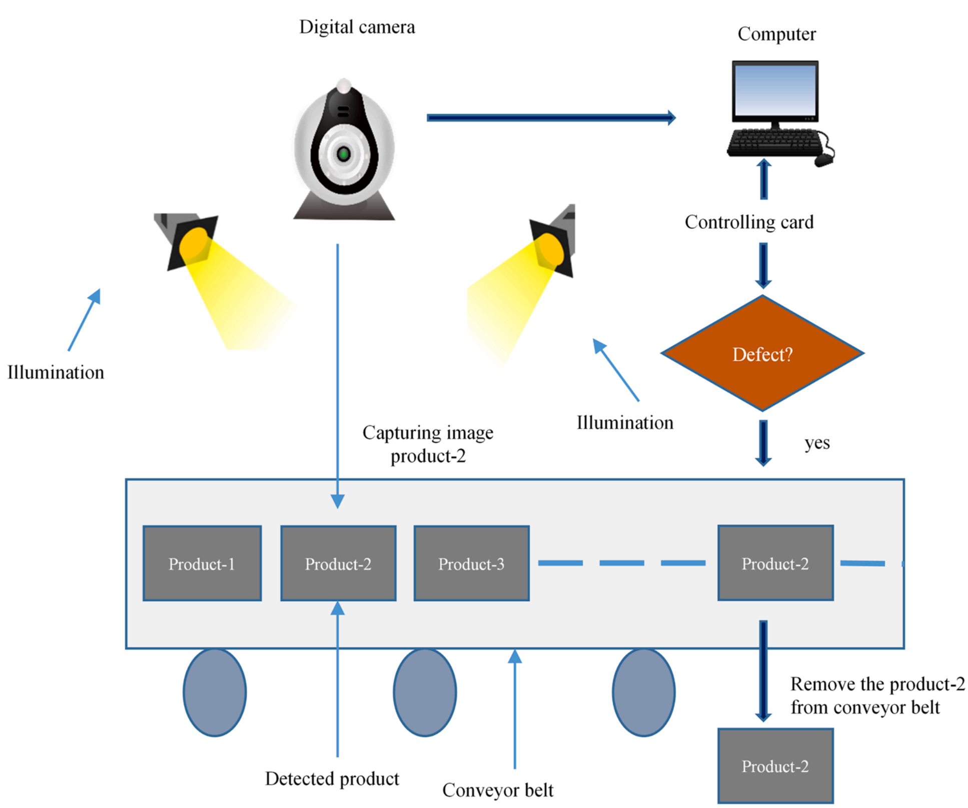

Fig. 1. Generic architecture of an automated optical system.

to be considered when selecting a camera for a particular application. There are several types of cameras that are commonly used in computer vision systems for conveyor applications. These include: (1) area scan camera – it captures an image of a 2D area and is typically used for applications where a relatively high resolution is required; (2) line scan camera – it captures an image of a single line at a time and is commonly used in applications where high-speed inspection is required; (3) 3D camera – it captures depth information, allowing for the creation of 3D images and used in applications where the shape and size of objects are important; (4) smart camera – this camera has built-in processing capabilities and can be used for applications where real-time processing is required. Then, direct communication interfaces between a camera and a computer network interface card (NIC) are commonly used in machine vision applications. These interfaces allow for real-time streaming of image data from the camera to the computer for processing. The most common direct communication interfaces are GigE Vision and USB3 Vision. GigE Vision uses Ethernet technology to transmit image data over a standard network connection, while USB3 Vision uses USB 3.0 technology to transmit data over a USB connection. In addition, an image data buffer, a temporary storage space, is also used to ensure that real-time streaming of image data is reliable and efficient.

Next, the visual system processing continues with the modern machine vision model that is stored on the computer system, and which processes the captured image to make the decision whether the product is defective or defect-free. If the product is defective, then the system sends a signal to the sensor to sideline the product from conveyor belt. Otherwise, it moves the product forward. For internal defect detection, the set up would be different for scanning the product and capturing the X-ray image of the product for decision making.

3. Deep CNN-based defect detection

Since the early 2010 s, solving the computer and machine vision problems using CNN techniques have been gaining momentum. Several computer vision problems are image classification, image segmentation, object detection, feature extraction, and object tracking. However, instead of only classifying the image as normal or defective, the topic of interest here is to localize the defect. Thus, image segmentation and object detection techniques are considered to solve computer vision problems. In this section, the basic structure of CNN is discussed, along with the emergence of CNN models in the past decade. Next, supervised and unsupervised learning of deep CNN-based models are explained with diverse set of computer vision techniques. Then, pixel level segmentation techniques are specifically reviewed since these techniques cover the state-of-the-art methodologies for automatic optical inspection. In addition, computer vision applications using deep CNN-based models are also surveyed.

3.1. Basic structure of CNN

The CNN architecture has three key components, namely, convolutional layer, pooling layer, and full connected layer that can be observed in Fig. 2. Convolutional layers are the essential parts of CNN networks. The function of this layer is to extract high-level features from the input image. Convolutional layers also comprise distinct set of filters that produces a set of feature maps after convolving the kernel over the input image. Convolutional networks stack up the convolutional and pooling layers, with and fully connected layer to complete the construction of the model. Finally, the last layer is output layer that classifies the image or pixel in the classification section.

3.2. The emergence of CNN

Since 1998, LeCun (LeCun et al., 1998) introduced a CNN architecture based on gradient learning that was implemented on hand digit recognition. Since CNN is a sparsely connected network, it has a couple of benefits such as parameter sharing, and less trainable parameters than traditional fully connected feed forward networks. It primarily includes few basic concepts: shared weights, local acceptance, and pooling layers. Particularly, the shared parameters of CNN reduce the degree of freedom parameters without degrading the solutions quality. It also allows CNN to be implemented by normal gradient decent approach. Thus, the CNN-based model emerges as one of the key algorithms to solve any computer vision problems, and it also has become one of the most active research domains in the machine vision community.

LeNet-5, described in (LeCun et al., 1998) was the first CNN architecture with five convolution layers to recognize the handwritten digits. AlexNet, described in (Alex et al., 2012) is another CNN architecture with a few more convolutional layers, which gained higher accuracy in ImageNet Large Scale Visual Recognition Challenge (ILSVRC). Later, GoogleNet (Szegedy et al., 2015), VGGNet (Simonyan and Zisserman, 2014), and ResNet (He et al., 2016) models obtained better accuracy in the ImageNet challenge with thousands of classes and millions of training samples, including millions of parameters. Initially, the researchers were adding more layers and achieving higher accuracy. However, ResNet found that adding more convolutional layers could increase the complexity of the model, while not always achieving higher accuracy. Thus, ResNet introduced the residual concept and found better results. Table 2 shows the Top 1 accuracy, model parameters, and error rate of the CNN model changes in the ImageNet challenge from year 2010 to 2021. Over the years, the percentage of Top 1 accuracy and model parameters in millions of CNN architectures increased tremendously in the ImageNet Challenge, which can be noticed in Fig. 3 and Fig. 4, respectively. On the other hand, the percentage of error rate of the model decreased significantly, result illustrated in Fig. 5. Besides supervised CNN models, convolutional networks also worked with AE

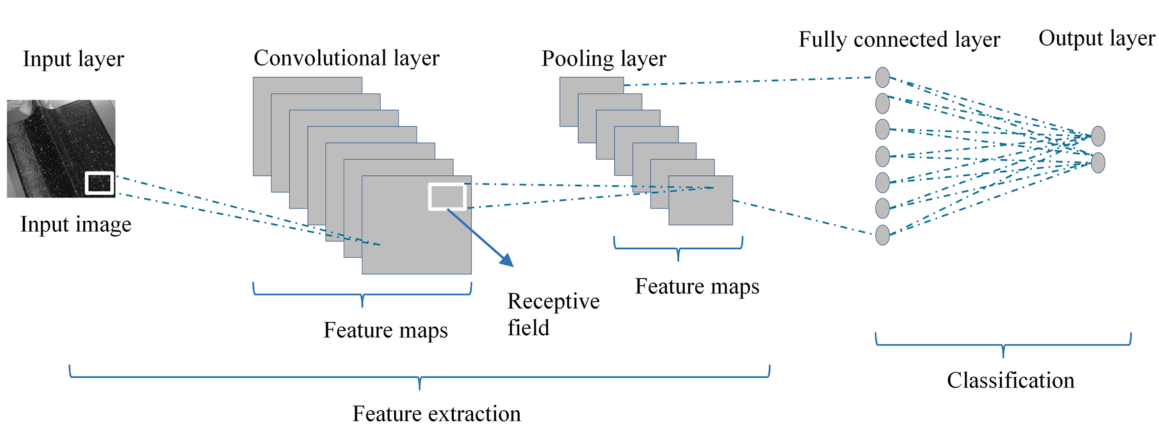

Fig. 2. Basic structure of CNN network.

Table 2
Benefits and Drawbacks of Defect Inspection Methods.

Year	Model	Top 1-Accuracy (%)	Parameters (in Million)	Error Rate (%)	Network
2010	ILSVRC'10 (Lin et al., 2010)	52.9	-	28.2	shallow
2011	ILSVRC'11 (Sánchez and Perronnin, 2011)	54.3	-	25.8	shallow
2012	AlexNet (Alex et al., 2012)	63.3	60	16.4	deep
2013	ZFNet (Zeiler and Fergus, 2013)	64.0	-	11.7	deep
2014	VGG19 (Simonyan and Zisserman, 2014)	74.5	144	7.3	deeper
2014	GoogleNet (Szegedy et al., 2015)	74.8	11.2	6.7	deeper
2015	ResNet (He et al., 2016)	81.2	25	3.57	deeper
2016	GBDNet (Zeng et al., 2016)	66.3	-	2.81	deeper
2017	SENet (Hu et al., 2018)	80.9	1.23	2.25	deeper
2018	MobileNet-V2 (Sandler et al., 2018)	74.7	3.4	-	deeper
2019	FixResNeXt-101 (Sun et al., 2019)	86.4	829	-	deeper
2020	EfficeintNet-L2 (Xie et al., 2020)	88.4	480	-	deeper
2021	ViT-G/14 (Zhai et al., 2021)	90.5	1843	-	deeper

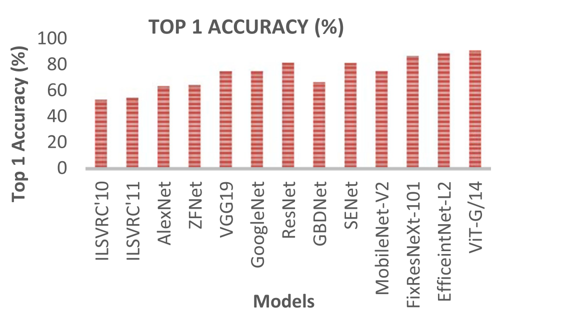

Fig. 3. Top 1 accuracy of various CNN models in ImageNet challenge over the years.

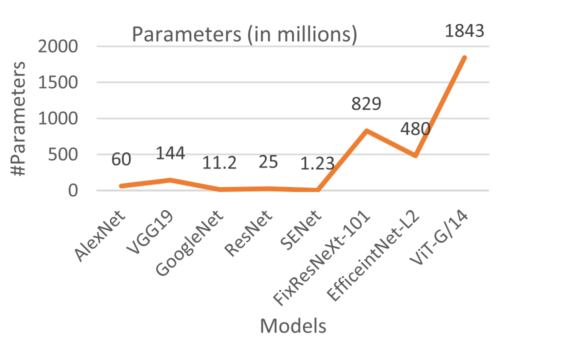

Fig. 4. Number of parameters of various CNN models in ImageNet challenge over the years.

and GAN models, namely, convolutional autoencoder (CAE) and deep convolutional GAN (DCGAN), respectively. These algorithms are widely accepted in solving the unsupervised task in computer vision. CAE was first introduced by (Masci et al., 2011) to extract the high-level features

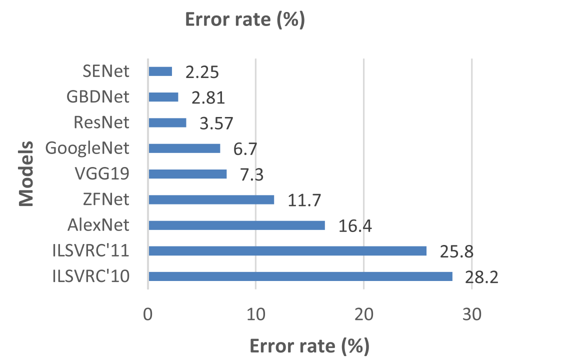

Fig. 5. Error rate of various CNN based model in ImageNet challenge over the years.

and for dimensionality reduction. DCGAN was initially proposed by (Radford et al., 2015) to generate new data with the same distribution of training dataset at image level.

3.3. Supervised CNN

Supervised CNN is one of the most widely used model for defect detection problems. In the past decade, CNN techniques have shown the most promising results for several tasks. In literature, supervised CNN techniques are primarily used for two tasks: defect classification and defect segmentation. Defect classification is a classification technique that works at image level and seeks to recognize the type of object in the image. Defect segmentation is a segmentation technique, which works at pixel level and seeks to find the object type in each pixel of the input image. Both the techniques can be applied on the defect detection problem, and the model can be trained either from the scratch or using pre-trained models. Ref (Chen and Jahanshahi, 2018). proposed a CNN-based model with Naïve Bayes data fusion technique aiming to classify cracks in components' surfaces of nuclear power plants, where regular inspections are required for safety.

For the surface defect detection problem, (Tabernik et al., 2020) developed a deep learning approach with a segmentation network. In the first stage, a segmentation CNN model was trained, and in the next stage, the features extracted from the previous stage were used to train a classification CNN-based model. The final detection task was implemented as a classification task to identify whether an image is normal or abnormal. Another research (Shi et al., 2019) studied the underwater pipeline damage detection problem using pre-trained MobileNet model for defect classification task. Ref (Yang et al., 2018). studied a Mura defect classification problem for a thin film transistor liquid crystal display. To ensure the quality of the displays, the research work developed a new method that blended a CNN feature extractor with a sequential extreme learning classifier. In another study, X-ray images of castings were considered to identify the defects by employing CNN-based spatial attention bilinear network (Tang et al., 2021).

End-to-end deep learning approach can be used for both defect classification and defect segmentation. This approach is exemplified in (He et al., 2020) for a steel surface defect detection problem. To identify the exact class and precise location, a standard CNN model combined with multilevel feature fusion network was proposed to accomplish robust classification ability. Similarly, (Ren et al., 2018) used a pre-trained CNN-based model for both classification and segmentation task to extract the features from the patches for automated surface inspection. Additionally, an image segmentation model is trained in (Dong et al., 2018) by using the segmentation network along with random forest techniques.

In 2015, one popular model for semantic segmentation, fully convolutional network (FCN), was proposed with effective inference and learning (Long et al., 2015). Along these lines, (Balzategui et al., 2020) developed an FCN architecture to generate the defect segmentation map in one step, identifying the localization of the defects precisely. In the same year, (Ronneberger et al., 2015) proposed U-Net, which is a CNN-based network, for biomedical image segmentation with different network and training strategy relying on vigorous data augmentation techniques. Later, several researchers used this technique in their studies for diverse set of applications and obtained efficient results. In (Zambal et al., 2019), an automated fiber placement defect detection was proposed with end-to-end fashion formulated as a defect segmentation problem. Initially, the artificial training data was generated using a probabilistic model and then, a CNN network motivated by U-Net architecture was trained to identify the defects at pixel level. In another research (Zou et al., 2018), a novel end-to-end trainable segmentation-based CNN model was developed for crack detection problem using multi-scale feature learning and results were compared with the benchmark U-Net model. For polycrystalline solar cell defect inspection problem, (Rahman and Chen, 2020) proposed an improved U-Net architecture considering multi-attention networks to classify and segment the complex defects using real photovoltaic images. In another U-Net based architecture, (Huang et al., 2018) developed a model, named as MCuePush U-net, comprising of primarily three components: (1) MCue that creates three channel inputs, (2) U-net that learns the informative regions, and (3) Push network that identifies the specific region with bounding boxes. A hybrid approach of regional proposal networks (RPN) to identify defect regions, and modified U-Net architectures to segment the defects at pixel level, were employed for the silicon wafer defect segmentation problem (Han et al., 2020). In most of the research works surveyed, it has been shown that U-Net performs better than FCN for segmentation networks, which is portrayed in Table 3. Furthermore, Table 3 also presents the advantages and disadvantages of CNN-based classifiers and semantic networks for defect classification and defect segmentation, respectively.

3.4. Unsupervised CNN

The growing demand of unsupervised CNN techniques for defect detection problems is unparalleled. Unsupervised CNN models have the potential to overcome the challenges of supervised models such as labeling images and annotating pixels and are also capable of isolating the defects with reasonable accuracy, without having the limitations of supervised CNN model. The models can work for image labeling as well as pixel level. In addition, they can classify the defective image and can localize the defective area. Unsupervised CNN can be applied in several machine vision applications such as internal defect detection (non-destructive testing, radiography images, etc.) and external defect detection (steel surface, mobile phone screen, etc.). The literature mainly categories unsupervised CNN technique for defect detection in three types: (1) anomaly detection, (2) GAN-based model, and (3) hybrid models.

Table 3
Advantages and Disadvantages of CNN-based Classifiers and Segmentation Networks for Defect Detection.

Category	Method	Advantage	Disadvantage
Defect Classification (CNN-based)	CNN + logistic/ classifier/attention (Chen and Jahanshahi, 2018; Tabernik et al., 2020; Shi et al., 2019; Yang et al., 2018; Tang et al., 2021)	Require only image-based labeling, high speed in training and testing.	Require large training samples (defective and defect-free samples); no defect segmentation.
Defect Segmentation (Semantic Networks)	FCN; U-net He et al. (2020);Ren et al. (2018);Dong et al. (2018);Long et al. (2015); Balzategui et al. (2020); Ronneberger et al. (2015);Zambal et al. (2019);Zou et al. (2018); Rahman and Chen (2020);Huang et al. (2018);Han et al. (2020)	Fast training and evaluation; identify and isolate the defective area.	Requisite of pixel-level human annotation; tiresome work, labor-intensive.

Anomaly detection is one of the most common techniques in unsupervised CNN-based model to detect abnormalities. In the specific domain literature, anomaly detection often uses the CAE model and variants of AE model. As an example, (Yang et al., 2019) studied the machine vision inspection of surface defects problem considering only defect-free samples for model training. To localize the defective area fast and with accuracy, the research proposed a multi-scale fully convolutional autoencoder (FCAE).

Another CAE based model for anomaly detection was developed to identify the abnormality in concrete structures (Chow et al., 2020). The study solved the civil infrastructure inspection problem, requiring no labeled images for training, thus it highlighted processing time savings for data labeling. In (Makhzani and Frey, 2015), authors presented a winner-take-all AE method to learn the shift-invariant sparse representations including lifetime and spatial sparsity in each feature map. For semi-supervised and unsupervised anomaly detection, (Ruff et al., 2019) proposed a generalization of deep support vector description (D-SVDD) model. For several years, AE and variants of AE based model were predominant, yet GANs eventually gained the lead in machine vision domain.

The second category of unsupervised CNN models for defect detection is the GAN-based method. The variants of GAN are further classified as GAN synthesis, GAN scoring, and AnoGAN (anomaly detection with GAN). For GAN synthesis, (Lian et al., 2020) proposed a defect exaggeration model, where GAN is combined with CNN network to generate flawless image and identify tiny surface defects. To improve the defect recognition process, (Niu et al., 2020) developed a surface defect-generation adversarial network (SDGAN) and applied it on defect-free images to enlarge the defective dataset. For GAN scoring, (Zenati et al., 2018) trained a GAN-based model using score function to make the anomaly detection more efficient on high-dimensional dataset. In another research for high-dimensional spaces, (Deecke et al., 2018) proposed a novel method for anomaly detection considering GAN networks to search a good representation of the sample in the generator. Research on AnoGAN is advanced in (Schlegl et al., 2019), where a fast AnoGAN method is developed to identify defective images and segment the anomalous area.

The third category of unsupervised CNN models for defect detection is the hybrid approach. A generic defect detection problem was explored to classify surface defects by extracting features locally and globally using bilinear based model constructed as two symmetric sub-networks based on visual geometry, labeled Double-VGG16 (Zhou et al., 2019). However, the proposed model has some limitations, especially on tasks such as localization of defects in complex textures. Ref (Tsai et al., 2021). developed a novel hybrid approach using CycleGAN and U-Net semantic networks with the objective to detect the pixel-wise defect. More detailed information, including the advantages and disadvantages of anomaly detection, GAN-based, and hybrid approach for defect classification and defect segmentation are presented in Table 4.

3.5. Object detection

In this section, one of the most recent computer vision methods for automatic optical inspection is discussed. In recent years, there has been a significant increase in scholarly research on visual defect detection problems using object detection techniques, namely, YOLO and RCNN based models.

Table 4
Advantages and Disadvantages of Anomaly Detection, GAN-based, and Hybris Approach for Defect Classification and Defect Detection.

Category	Method	Advantage	Disadvantage
Anomaly Detection (CAE-based)	CAE reconstruction; FCAE; D-SVDD Yang et al. (2019); Chow et al. (2020); Makhzani and Frey (2015);Ruff et al. (2019)	Require only defect-free samples to train the model; no need of pixel-level annotation.	Defect localization is hard; detection of tiny defects is difficult.
GAN-based	GAN synthesis; GAN scoring; AnoGAN Lian et al. (2020);Niu et al. (2020);Zenati et al. (2018);Deecke et al. (2018);Schlegl et al. (2019)	Defect synthesis to train the model; no annotation is needed at pixel-level; AnoGAN requires only defect-free images.	Training and validation are slow; require human screening in defect synthesis; imprecise defect segmentation.
`Hybrid Models	Bilinear model; CycleGAN; U-net; (Zhou et al., 2019; Tsai et al., 2021)	CNN-based; require only image labelling; no need of human annotated pixels; real-time defect inspection; defect localization.	Requires a substantial amount of dataset; blurry defective area; slow training.

In (Kou et al., 2021), steel strip production can result in surface defects due to mechanical forces and environmental factors. Identifying these defects is crucial for producing high-quality products, as their presence can cause significant economic losses for the high-tech industry. To address the limitations of current algorithms, researchers developed an end-to-end defect detection model based on YOLO-V3, utilizing an anchor-free feature selection mechanism and specially designed dense convolution blocks to improve feature reuse, feature propagation, and network characterization. Experimental results showed that the proposed model outperformed other comparison models, achieving 71.3% mAP on the GC10-DET dataset and 72.2% mAP on the NEUDET dataset. In (Mushtaq et al., 2023), the aerospace industry involves assembling many fastening elements, such as bolts, washers, and nuts, which are currently identified manually by humans. However, human error can have a significant impact on efficiency and safety. To address this, a deep learning and image processing approach using the YOLO-v5 algorithm was proposed to classify these components based on their shape, along with an image processing method to estimate their spatial dimensions, including thread pitch. Despite the challenges, the proposed system achieved promising results. Concrete is a common building material, but strong wind erosion in Northwest China causes damage to its surface, affecting both the appearance and safety of buildings (Cui et al., 2021). To identify erosion areas in concrete, a deep learning dataset was established through erosion tests and an improved YOLO-v3 algorithm model was proposed. The model demonstrated more accurate recognition of erosion damage to concrete, achieving accuracy, precision, and map of 96.32%, 95.68%, and 75.68%, respectively.

Ref (Li et al., 2023). introduced a deep learning-based automatic defect detection system called YOLO-attention, which was specifically designed for wire and arc additive manufacturing (WAAM) processes. YOLO-attention incorporates improvements in three object detection models and achieves both speed and accuracy in defect detection. The evaluation on the WAAM defect dataset showed that the model achieved a mean average precision of 94.5 and a frame rate of at least 42 frames per second, demonstrating its feasibility in practical industrial applications. A computer vision pipeline was developed to rapidly analyze electroluminescence (EL) images of solar photovoltaic (PV) modules and identify defects using machine learning models such as Random Forest, ResNet models, and YOLO (Chen et al., 2022a). The developed models were tested on a dedicated testing set, resulting in macro F1 scores of 0.83 (ResNet18) and 0.78 (YOLO), and were used to analyze 18,954 EL images of a PV power plant damaged in a vegetation fire, finding increased frequency of certain defects on the edges of the solar module closest to the ground after fire.

In (Cheng et al., 2023), Wheel hub defects have complex types, different location and size, making it difficult to establish an accurate detection model. To address this, a wheel nuclear hub defect detection method based on the DS-Cascade RCNN was proposed that uses spatial attention mechanism, deformable convolution, and pruning algorithm to optimize the model and compress the model space without losing accuracy. Experimental results show that the proposed method can effectively detect six kinds of wheel hub defects, and the mean Average Precision (mAP) is 95.49%. In (Ji et al., 2023), the safety of pipeline transportation relies on Non-Destructive Testing (NDT) to detect weld joint defects. However, traditional manual inspection of X-ray images suffers from accuracy and efficiency issues. To address this, a model integrating Feature Pyramid Network FPN and a new visual attention mechanism SPAM was proposed, along with a data augmentation method based on geometry transformation. Experimental results show that the proposed model outperforms Faster-RCNN in detecting defects, with a 4.0% increase in mAP value. Damage to metro tunnel surfaces caused by environmental changes, train-induced vibration, and human interference can lead to accidents if not adequately and efficiently maintained (Li et al., 2021). The inspection of these surfaces is challenging due to harsh conditions, such as low light and limited inspection time. To address this, an automatic Metro Tunnel Surface Inspection System (MTSIS) has been developed, consisting of hardware and software components, including a high-speed image capture system, image pre-processing methods (contrast enhancement and stitching), and a defect detection method based on a multi-layer feature fusion network. Practical experimental results demonstrate the effectiveness of the proposed MTSIS in detecting defects on metro tunnel surfaces.

3.6. Pixel level segmentation

This section reviews the latest developed computer vision techniques for automatic optical inspection. Scholarly contributions for the pixel level segmentation have been significantly growing in the last few years. The models developed in the articles referenced in the previous sections are mostly evaluated through accuracy metrics. The newly developed pixel level segmentation models are highly encouraged to be compared with the mean intersection over union (MIoU) or Dice coefficient, instead of accuracy metrics. The performance comparisons of recently developed pixel level segmentation models for AOI are presented in Table 5, and include MeanIoU and Dice Coefficient evaluations. The reviewed articles are mainly categorized on supervised and unsupervised deep CNN based segmentation models. In addition, a few papers focus on small or micro defects datasets.

Using RSDD dataset, (Cao et al., 2021) and (Yang et al., 2022a) presented segmentation-based model. Ref (Cao et al., 2021). developed a pixel level segmentation network including deep feature fusion, multi-level feature aggregation module, and multi-branch decoder. Ref (Yang et al., 2022a). proposed an NDD-Net model to create an end-to-end defect segmentation scheme comprising of attention fusion block to obtain discriminative features and improve the performances. The performance achieved by these two models show a MIoU of 0.85 for (Cao et al., 2021) and a Dice Coefficient of 0.835 for the model presented in (Yang et al., 2022a), respectively. Furthermore, (Fioresi et al., 2022) introduced the UCF EL defect dataset and proposed a semantic segmentation model to classify the five defects with MIoU of 0.573 and pixel-level accuracy of 95.4%.

In (Versini et al., 2022), authors presented a U-Net GMP method comprising of SCL_{Dice} using Kolektor dataset with MIoU of 0.56. Ref (He and Liu, 2020). developed a regression-based pixel segmentation model using DAGM dataset to localize the defects with MIoU of 0.845. In another work, (Liu and He, 2022) proposed a TAS2-Net model for small

Table 5
Performance Comparison of State-Of-The-Art Pixel Level Segmentation Model for Automatic Optical Inspection.

Model	Dataset	Accuracy (%)	MeanIoU	Dice Coefficient
Pixel level segmentation on RSDD (Kou et al., 2021)	RSDD	-	0.850	-
NDD-Net (Mushtaq et al., 2023)	RSDD	0.997	-	0.835
Semantic Segmentation of EL Images (Cui et al., 2021)	UCF EL Defect Dataset	95.4	0.573	-
U-Net GMP + SCL_{Dice} (Li et al., 2023)	Kolektor	-	0.56	-
Regression based pixel segmentation (Chen et al., 2022a)	DAGM	-	0.845	-
TAS2-Net (Cheng et al., 2023)	DAGM	-	0.869	-
Automatic deep segmentation (Ji et al., 2023)	GDXray	0.998	-	0.854
Improved super-pixel segmentation model (Li et al., 2021)	Machine surfaces	91.11	-	-
CycleGAN (Tsai et al., 2021)	Machined surfaces	95	0.71	-
Pixel-wise semi-supervised model (Cao et al., 2021)	FID	91.85	0.825	-

surface defects with same dataset but slightly better MIoU of 0.869. To address the class imbalanced or micro defects issue, an automatic deep segmentation model is proposed (Yang et al., 2022b) with an attention-guided segmentation network for pixel level welding defects with decent Dice Coefficient of 0.854. To classify the machined surface defects at pixel level, (Chen et al., 2022b) developed an improved super-pixel segmentation model. In contrast, (Tsai et al., 2021) proposed the unsupervised segmentation CycleGAN model to segment the machined surface defects with MIoU of 0.71. Ref (Shao et al., 2022). authors developed pixel-wise semi-supervised segmentation model with multi-task mean teacher using fabric image dataset. The MIoU performance metric reported is 0.825. Future research in this area should be coordinated to compress models in such a way so that to yield a more lightweight model while ensuring high detection accuracy.

3.7. Applications of deep cnn-based defect detection

Computer vision applications have been widely adopted in quality inspection problems, and mostly solved using deep CNN-based model. They can be mainly categorized into internal surface inspection and external surface inspection, as portrayed in Table 6. Internal defect detection is mostly used in aerospace welding, pipe welding, laser welding, and other similar welding operations. Welding is needed in various manufacturing industries to join two distinct parts into one component, many of these processes being used in aerospace industry. In certain unanticipated cases, defects might occur in aerospace-welded components that can increase the risk of accidents. To ensure the quality and safety for aerospace industry components, several researchers solved this problem using deep CNN-based techniques, such as X-ray images of aerospace welds (Dong et al., 2021, 2020a, 2020b).

In addition, (Gong et al., 2020) considered X-ray images of aerospace composite materials to recognize the defects using transfer learning model. Covering different industries, (Guo et al., 2021) studied the petroleum pipelines welding defect detection problem using conditional GAN and transfer learning with augmenting the X-ray images, while (Xia et al., 2020) presented the Keyhole Tungsten Inert Gas (TIG) welding type inspection using ResNet models to identify the different states of welding. To inspect the laser welding defects of safety vents on power battery, (Yang et al., 2020a) used a pre-trained SqueezeNet model to identify the abnormalities in the images. The SqueezeNet model is a CNN architecture that was reported to attain a high accuracy on ImageNet challenge, even though it uses a small model and low number of parameters.

The external surface defect detection is primarily classified into textured and patterned surfaces. First, textured surface defect detection methods are applicable in several domains. This literature review divides textured surface applications into three parts such as 3 C products, construction, and miscellaneous. The 3 C products include mobile phone-type devices, LCD displays, and printed circuit boards (PCB) components. To improve the quality of 3 C products, the surface defect detection of mobile phone screens is one of the essential tasks. As an example, (Li et al., 2020b) studied machine vision problem of mobile phone screens using a novel deep learning algorithm to extract the features and classify the defects. In addition, a weak-supervised defect detection method was proposed for mobile phone screen defects such as scratches, floaters and strains (Li et al., 2020a). Ref (Park et al., 2020). addressed the smart factory display manufacturing for mobile screens and developed a multi-deep learning neural network to identify the defects.

Table 6
Categorization of Defect Detection Application.

Defect Detection												
Internal Surface			External Surface									
Radiographic Images of			Textured									Patterned
Aerospace welding	Pipe welding	Laser welding	3 C Products			Construction Materials		Miscellaneous				Fabric
Dong et al., (2021, 2020a, 2020b); Gong et al. (2020)	(Guo et al., 2021)	(Yang et al., 2020a)	Mobile phone screen Park et al. (2020);Li et al. (2020a), (2020b)	LCD (Tsai and Jen, 2021; Yang et al., 2020b)	PCB (Li et al., 2021; Hu and Wang, 2020)	Glass panels (Pan et al., 2021)	Concrete structures / building cracks Chow et al. (2020); Zheng et al. (2020b); Choi (2020)	Machined surfaces (Tsai et al., 2021; Wang et al., 2021) die casting (Li et al., 2020b)	Steel (Fu et al., 2019; Zheng et al., 2020c), aluminum profile (Liu et al., 2021), polycrystalline alloy (Alqahtani et al., 2021)	Bottle (Wang et al., 2019), wood (Tsai and Jen, 2021), resin-casting (Nakashima et al., 2021)	Wafer surface (Yu, 2019; Nakazawa and Kulkarni, 2019; Cheon et al., 2019), Voltage dependent resistor (Yang et al., 2020c)	Ouyang et al. (2019); Zhang et al. (2021); Zhu et al. (2020)

With the goal to enhance the quality of the LCD display of 3 C products, (Tsai and Jen, 2021) addressed the homogeneously structured LCD display defect detection problem using autoencoder based anomaly detection technique. While the developed model includes a small size dataset, the experimental results are significant (100%). In another study covering small industrial image dataset of LCD, (Yang et al., 2020b) developed a deep CNN-based method for defect detection. To ensure the quality of 3 C products, one of the key issues is to improve the external surface defect detection of PCB. Within the defect detection domain, (Li et al., 2021) addressed an automatic inspection system for PCB board using effective self-adaption methods to identify the PCB defects with significant detection rate. In another study, also covering PCB defects, (Hu and Wang, 2020) proposed a deep learning method using faster region-based CNN (R-CNN) and feature pyramid network to recognize the PCB surface defects with mean average precision (mAP) of 95%. The mAP is calculated as the average of AP values of the number of defects in a candidate area.

Defect detection of construction materials in the civil infrastructure domain is another vital application of machine vision. Construction materials and their typical defects include glass panels that exhibit scratches and concrete structures that exhibit cracks, among others. Glass pieces are the key components of building materials. For quality assurance of glass products, (Pan et al., 2021) studied the automated scratch detection of transparent glass components. In order to identify the scratches on the surface, this study developed a deep learning approach using mask and region-based CNN (Mask R-CNN) with a significant reported accuracy of 94%. Other important computer vision application in the construction materials domain is the study of concrete structures defects such as building cracks. The data of building surfaces can come from structures such as bridges, houses, roads, and dams. In (Zheng et al., 2020b), the building cracks defect detection using FCN, R-CNN, and richer fully convolutional networks (RFCN) were compared and evaluated for picture performance detection and comprehensive assessment, with RFCN found to exhibit the best outcomes. Ref (Chow et al., 2020). attempted to tackle the anomaly detection problem of concrete structures using CAE with defect-free images. Other machine vision applications for textured surface defect detection are reported for machined surfaces (Tsai et al., 2021; Wang et al., 2021), die casting (Lee et al., 2020), steel surfaces (Fu et al., 2019; Zheng et al., 2020c), aluminum profiles (Liu et al., 2021), polycrystalline alloys (Alqahtani et al., 2021), bottles (Wang et al., 2019), wood (Tsai and Jen, 2021), and wafer surface (Yu, 2019; Nakazawa and Kulkarni, 2019; Cheon et al., 2019).

The last defect detection application category identified in the survey is the patterned surface defect detection. This is another widely researched area within the larger external surface domain, with the main application being fabric defect detection. Fabric inspection system plays a key role for quality assurance in textile manufacturing. There is an ever-growing demand in the textile factory to substitute the human-intensive quality inspection performed with naked eyes by an automated inspection system. This task compelled researchers and practitioners to develop deep CNN-based approaches to isolate fabric defects (He and Liu, 2020; Liu and He, 2022; Yang et al., 2022b). For all machine vision applications in this area, the results of detection accuracy vary from 88% to 99%.

4. Open challenges

For the last few decades, many researchers and practitioners studied the defect detection problem for quality control and assurance. As such, the demand of automated inspection systems for quality control in the manufacturing industry grows by the day. Over the years, researchers proposed a diverse set of deep learning techniques to isolate the defects. However, there are still numerous challenges left to tackle in this domain. This section presents the identified open challenges that are categorized as follows: (1) challenges in algorithms, (2) challenges in applications, and (3) challenges in data processing on high performance computing systems.

4.1. Algorithms challenges

This section covers the many challenges explained at the algorithmic level, such as defect inspection methods, supervised CNN-based methods, unsupervised CNN-based methods, etc. Defect inspection methods include manual inspection, traditional AOI, and modern AOI. Each inspection method brings forth various challenges. Manual inspection has specific well-known drawbacks. First, it requires large number of human experts for the inspection, which significantly increases the labor cost. Being labor-intensive work, manual inspection leads to many mistakes. Also, manual inspection is time-consuming, inconsistent, and subjective. To overcome these disadvantages, traditional AOI methods have been developed where image processing techniques and shallow ML algorithms were used to reduce the labor cost, human errors, and inspection time. However, many challenges still exist when employing the above-named methods. For example, traditional computer vision techniques have poor feature extraction and huge time complexity. Thus, the shallow ML techniques require human experts to find the specific features to feed the model. Though, some of the traditional machine vision methods achieve detection accuracy for one defect pattern as high as 99%, they still do not work as expected with multiple patterns. In contrast, modern AOI methods have been consistently used by researchers and manufacturing industries for addressing the defect detection problems. These techniques are known to attain high detection accuracy, increased real-time model robustness, reduced labor costs, and good performance for high-level feature extractions. Still, these traditional machine vision methods were identified to have challenges such as the need for substantial amount of data to train the model, performance challenges for hyper-parameters tuning, and limitations on the high-speed processor, GPU, and TPU for execution of complex models.

Modern computer vision techniques for defect detection are deep CNN-based models. For supervised deep CNN models, a large, labeled image dataset is required to train the model to attain considerable detection accuracy, since the detection accuracy of the model highly depends on the quality of the dataset. Thus, acquiring, and labeling datasets are essential, but it still carries a series of challenges. First, acquiring the large dataset is one of the major challenges for the researchers and practitioners. Capturing the images in the industry could run into limitations such as non-uniform illumination, motion blur, and camera noise. The quality of the images also depends on the manufacturer's standards. Then, the challenges also depend on the image dataset acquired in different applications. As an example, obtaining images by X-ray can run into issues such as noise and defects in the images are very subtle, so background differentiation makes it difficult to process the images. For steel surfaces, usually the occurring of defects has a very low probability as well as they are visually indistinctive. Therefore, the defective samples are limited in number, which makes it hard to represent the distribution of dataset, which in turn brings even more challenges. Generally, scarce training samples can and many times will lead to a poor generalization ability of the model. One of the major challenges with deep learning models is the problem of generalization. While deep learning models are known to perform well on training data, they may not generalize well to new, unseen data. This is because deep learning models often have a large number of parameters that can be tuned to fit the training data very closely, sometimes resulting in overfitting. Overfitting occurs when the model fits the noise and idiosyncrasies of the training data rather than the underlying patterns and relationships. Another issue related to generalization is the problem of dataset bias. If the training data is not representative of the target population or contains systematic biases, the model may not be able to generalize well to new data. To address these issues, various techniques have been developed, such as regularization, early stopping, data augmentation, and transfer learning. These techniques aim to improve the generalization performance of deep learning models by reducing overfitting and increasing the robustness of the model to new, unseen data.

Secondly, labelling images is another difficult task to complete and as an undesired consequence, it increases the labor cost. In addition, it can only classify images without being able to locate the defective area. To alleviate this issue, several scholars developed the CNN-based semantic networks to account for defect localization, such as FCN and U-Net networks. While addressing the defect localization, these methods also have challenges. One first challenge comes from the labor-intensive process to perform pixel-wise human annotation, which again increases the labor cost. A second labor- and time- intensive task is to annotate the data at pixel-level. To overcome these challenges, computer vision researchers proposed unsupervised deep CNN-based models that do not require any human labeled or annotated data and used them to identify the image and/or localizing the defective area. Unsupervised anomaly detection techniques and the GAN-based models have the advantage not only to require normal samples to train the model, but also there is no need of pixel-level annotation. Even with these advantages, there are challenges still yet to be addressed, such as improvements in the accuracy of defect localization and in the reliability of unsupervised models, which are lower than that of supervised models.

Other challenges related to algorithms are the complicated process of selection of specific algorithm for certain defect detection problem. The size of training samples influences many aspects of the deep CNN models during learning process, such as the detection accuracy to be achieved, training time of the model, number of features, distribution of the dataset, and the computing system required for the model. Besides algorithm selection, tuning the hyper-parameters is another key challenge for deep CNN-based defect detection algorithms. Some of the hyper-parameters' tasks include determining the number of network layers, finding the proper number of filters and setting the filter size for each layer, selecting the stride size and the pooling type, and choosing the number of neurons and activation functions for each neuron. To achieve reliable results for large models, the deep CNN-based models challenges list is completed by another series of tasks that need to be addressed, namely selecting an optimizer for the model from stochastic gradient descent (SGD), performing the adaptive momentum estimation (Adam), setting the adaptive gradient learning algorithm (Adagrad), and training a significantly large number of parameters. These challenges require an efficient hardware infrastructure to execute the complex model.

The challenge with the explainability of CNN network decisions is that they operate as black boxes, meaning that it can be difficult to understand how they arrive at their conclusions. CNNs learn to recognize patterns in data through layers of abstraction, and the final decision is often made based on a combination of these patterns. However, it can be challenging to determine which patterns are being used in the decision-making process and why they were given importance. This lack of transparency is particularly problematic when it comes to safety critical applications such as medical devices, aircraft systems, autonomous vehicles, etc., where it is crucial to understand how a decision was made. Without a clear explanation, it may be challenging to trust the results of a CNN network, and it could lead to serious consequences if incorrect decisions are made. Several techniques have been developed to improve the explainability of CNN network decisions, such as visualization of feature maps and saliency maps, or the use of attention mechanisms to focus on specific regions of an image. However, further research is needed to fully address this challenge and make CNNs more transparent and interpretable.

4.2. Applications challenges

This section discusses the key applications' challenges, such as those encountered in databases for automated optical inspection. The process of acquiring a database is an essential step for all deep CNN-based models. But, as in any other application, databases come with a series of challenges such as data privacy, data cleanness, data labeling and annotation, and data sharing. Data privacy is one of the major challenges for the machine vision-based defect detection methods. Regulations in force in several industries, such as aerospace manufacturing domain, ask that the dataset is kept private and not uploaded in the public domain due to security reasons. Cleaning the dataset is very much a requirement before feeding it into a model. Particularly, cleaning the noises in the dataset improves the quality of the data and makes it easier to explore and understand the model. Also, before training the model, anomalous data needs to be removed, the images need to be resized, and image resolution needs to be fixed. Data labelling and annotation is another essential step in preparing the dataset for the model and it requires significant human efforts and expertise. Lastly, but not the least important, researchers and practitioners do not share data easily, which makes it difficult for other interested parties to collect the dataset and train their models. In certain manufacturing industries, it is relatively easy to gather non-defective samples in the early phase of production but collecting enough defective samples for the robust model is hard, since occurrence of defective samples is rare compared to non-defective ones. Automated machine vision-based non-destructive testing (NDT) methods usually face two challenges: low availability of defective images and lack of precise annotation of defective samples. Insufficient abnormal images create an imbalanced dataset and result in inaccurate representation of the distribution of all defective samples, which makes the training process difficult. Furthermore, the low number of defective samples, with low contrast in many cases, causes ambiguity in both defective samples and normal samples.

4.3. High performance computing challenges

To solve the complex defect detection problems using deep CNN-based models, high performance computing (HPC) is essential. Most of the deep CNN models researchers use high-speed processors, substantial amount of graphic processing unit (GPU), and tensor processing unit (TPU) via cloud computing. Cloud computing facilitates HPC by providing huge computational capabilities to individual researchers and organizations who might have insufficient hardware infrastructure in-house to train complex model. Amazon web services (AWS) provides researchers and practitioners the power to create HPC clusters on demand, train and test ML models, gain valuable insights on complex models, and improve their productivity. But since nothing is free, these services come with external constraints and limitations such as cost, security, data transfer, performance, to name a few. Cost-management to build and use cloud HPC systems is a major concern for most organizations. Data security on cloud is another major concern for deep learning researchers and practitioners. In addition, to run models on the cloud, proprietary data must be moved into the cloud, which is many times a no-solution challenge for many organizations. Lastly, performance in the cloud is another major concern since most of the deep learning scholars expect high performance form HPC systems, but the performance may be reduced due to inter-connect latencies and outside network limited capabilities for data transfer.

Furthermore, federated learning has emerged as a promising approach for defect detection in manufacturing, as it allows multiple parties to collaborate and jointly train a machine learning model without sharing their private data. However, there are several challenges that need to be addressed when applying federated learning to defect detection, including data heterogeneity, communication overhead, privacy concerns, and quality control. While federated learning holds great promise for defect detection in manufacturing, these challenges need to be carefully addressed to ensure the performance, privacy, and reliability of the system (Mehta and Shao, 2022).

5. Conclusions

In the last decade, the research community productivity covering automated quality inspection systems using deep CNN-based models has been increased rapidly. The literature survey confirmed that researchers and practitioners focus not only on supervised CNN-based models but also study the unsupervised CNN-based models since labeling and annotation of datasets are difficult. Both model types pose challenges with regards to training the model and achieving the expected significant detection accuracy. In the future, with the identified key challenges being addressed, the demand for automated optical inspection is expected to only grow continuously in both industry practice and academic research.

CRediT authorship contribution statement

Shashi Bhushan Jha: Conceptualization, Methodology, Formal analysis, Investigation, Resources, Data curation, Writing – original draft, **Radu Babiceanu:** Conceptualization, Resources, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

No data was used for the research described in the article.

References

Alqahtani, H., Bharadwaj, S., Ray, A., 2021. Classification of fatigue crack damage in polycrystalline alloy structures using convolutional neural networks. Eng. Fail. Anal. vol. 119, 104908.
Ashour, M.W., Khalid, F., Halin, A.A., Abdullah, L.N., Darwish, S.H., 2019. Surface defects classification of hot-rolled steel strips using multi-directional shearlet features. Arab. J. Sci. Eng. vol. 44 (4), 2925–2932.
Balzategui, J., Eciolaza, L., Arana-Arexolaleiba, N., 2020. Defect detection on polycrystalline solar cells using electroluminescence and fully convolutional neural networks. Proc. 2020 IEEE/SICE Int. Symp. Syst. Integr. 949–953.
Cao, J., Yang, G., Yang, X., 2021. A pixel-level segmentation convolutional neural network based on deep feature fusion for surface defect detection. IEEE Trans. Instrum. Meas. vol. 70.
Celik, H., Dulger, L.C., Topalbekiroglu, M., 2014. Development of a machine vision system: real-time fabric defect detection and classification with neural networks. J. Text. Inst. vol. 105 (6), 575–585.
Chen, F.C., Jahanshahi, M.R., 2018. NB-CNN: deep learning-based crack detection using convolutional neural network and naïve bayes data fusion. IEEE Trans. Ind. Electron. vol. 65 (5), 4392–4400.
Chen, L.-F., Su, C.-T., Chen, M.-H., 2009. A neural-network approach for defect recognition in TFT-LCD photolithography process. IEEE Trans. Electron. Packag. Manuf. vol. 32 (1), 1–8.
Chen, W., Zou, B., Yang, J., Huang, C., Yao, P., Liu, J., 2022b. The machined surface defect detection of improved superpixel segmentation and two-level region aggregation based on machine vision. J. Manuf. Process. vol. 80, 287–301.
Chen, X., Karin, T., Jain, A., 2022a. Automated defect identification in electroluminescence images of solar modules. Sol. Energy vol. 242 (June), 20–29.
Cheng, S., et al., . 2023. Wheel hub defect detection based on the DS-Cascade RCNN, Meas. J. Int. Meas. Confed. vol. 206.
Cheon, S., Lee, H., Kim, C.O., Lee, S.H., 2019. Convolutional neural network for wafer surface defect classification and the detection of unknown defect class. IEEE Trans. Semicond. Manuf. vol. 32 (2), 163–170.
W. Choi, Deep Learning Implemented Structural Defect Detection on Digital Images, 2020.
Chow, J.K., Su, Z., Wu, J., Tan, P.S., Mao, X., Wang, Y.H., 2020. Anomaly detection of defects on concrete structures with the convolutional autoencoder. Adv. Eng. Inform. vol. 45 (December 2019).
Cui, X., Wang, Q., Dai, J., Zhang, R., Li, S., 2021. Intelligent recognition of erosion damage to concrete based on improved YOLO-v3. Mater. Lett. vol. 302 (June), 1–4.
Deecke, L., Vandermeulen, R., Ruff, L., Mandt, S., Kloft, M., 2018. Image anomaly detection with generative adversarial networks. Jt. Eur. Conf. Mach. Learn. Knowl. Discov. Databases 3–17.
Dong, X., Taylor, C.J., Cootes, T.F., 2018. Small defect detection using convolutional neural network features and random forests. Proc. Eur. Conf. Comput. Vis. (ECCV) Workshops.
Dong, X., Taylor, C.J., Cootes, T.F., 2020a. A random forest-based automatic inspection system for aerospace Welds in X-ray images. IEEE Trans. Autom. Sci. Eng. 1–14.
Dong, X., Taylor, C.J., Cootes, T.F., 2020b. Defect detection and classification by training a generic convolutional neural network encoder. IEEE Trans. Signal Process. vol. 68 (2), 6055–6069.
Dong, X., Taylor, C.J., Cootes, T.F., 2021. Automatic aerospace weld inspection using unsupervised local deep feature learning. Knowl. -Based Syst. vol. 221.
Fioresi, J., et al., 2022. Automated defect detection and localization in photovoltaic cells using semantic segmentation of electroluminescence images. IEEE J. Photovolt. vol. 12 (1), 53–61.
Fu, G., et al., 2019. A deep-learning-based approach for fast and robust steel surface defects classification. Opt. Lasers Eng. vol. 121, 397–405.
Gong, Y., Shao, H., Luo, J., Li, Z., 2020. A deep transfer learning model for inclusion defect detection of aeronautics composite materials. Compos. Struct. vol. 252.
Guo, R., Liu, H., Xie, G., Zhang, Y., 2021. Weld defect detection from imbalanced radiographic images based on contrast enhancement conditional generative adversarial network and transfer learning. IEEE Sens. J. vol. 21 (9), 10844–10853.
Han, H., Gao, C., Zhao, Y., Liao, S., Tang, L., Li, X., 2020. Polycrystalline silicon wafer defect segmentation based on deep convolutional neural networks. Pattern Recognit. Lett. vol. 130, 234–241.
Haselmann, M., Gruber, D.P., Tabatabai, P., 2019. Anomaly detection using deep learning based image completion. Proc. IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2018, 1237–1242.
He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual learning for image recognition. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 770–778.
He, Y., Song, K., Meng, Q., Yan, Y., 2020. An end-to-end steel surface defect detection approach via fusing multiple hierarchical features. IEEE Trans. Instrum. Meas. vol. 69 (4), 1493–1504.
He, Z., Liu, Q., 2020. Deep regression neural network for industrial surface defect detection. IEEE Access vol. 8, 35583–35591.
Hu, B., Wang, J., 2020. Detection of PCB surface defects with improved faster-RCNN and feature pyramid network. IEEE Access vol. 8, 108335–108345.
Hu, J., Shen, L., Sun, G., 2018. Squeeze-and-excitation networks. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 7132–7141.
Huang, Y., Chan, K.L., 2004. Texture decomposition by harmonics extraction from higher order statistics. IEEE Trans. Image Process. vol. 13 (1), 1–14.
Huang, Y., Qiu, C., Guo, Y., Wang, X., Yuan, K., 2018. Surface defect saliency of magnetic tile (-Augus). IEEE Int. Conf. Autom. Sci. Eng. vol. 2018, 612–617.
Ji, C., Wang, H., Li, H., . 2023. Defects detection in weld joints based on visual attention and deep learning. NDT E Int. vol. 133.
Kim, J., et al., 2021. Adversarial defect detection in semiconductor manufacturing process. IEEE Trans. Semicond. Manuf. vol. 34 (3), 365–371.
Kou, X., Liu, S., Cheng, K., Qian, Y., 2021. Development of a YOLO-V3-based model for detecting defects on steel strip surface. Meas. J. Int. Meas. Confed. vol. 182 (May).
Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional networks. Proc. 25th Int. Conf. Neural Inf. Process. Syst. 1097–1105.
Kulkarni, R., Banoth, E., Pal, P., 2019. Automated surface feature detection using fringe projection: an autoregressive modeling-based approach. Opt. Lasers Eng. vol. 121, 506–511.
LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature vol. 521 (7553), 436–444.
LeCun, Y., Bottou, L., Bengio, Y., Haffner, P., 1998. Gradient-based learning applied to document recognition. Proc. IEEE vol. 86 (11), 2278–2324.
Lee, J., Lee, Y.C., Kim, J.T., 2020. Fault detection based on one-class deep learning for manufacturing applications limited to an imbalanced database. J. Manuf. Syst. vol. 57, 357–366.
Li, C., Huang, Y., Li, H., Zhang, X., 2020a. A weak supervision machine vision detection method based on artificial defect simulation. Knowl. -Based Syst. vol. 208.
Li, C., Zhang, X., Huang, Y., Tang, C., Fatikow, S., 2020b. A novel algorithm for defect extraction and classification of mobile phone screen based on machine vision. Comput. Ind. Eng. vol. 146.
Li, D., et al., . 2021. Automatic defect detection of metro tunnel surfaces using a vision-based inspection system. Adv. Eng. Inform. vol. 47.
Li, M., Wan, S., Deng, Z., Wang, Y., 2019. Fabric defect detection based on saliency histogram features. Comput. Intell. vol. 35 (3), 517–534.
Li, W., et al., 2023. Deep learning based online metallic surface defect detection method for wire and arc additive manufacturing. Robot. Comput. Integr. Manuf. vol. 80.
Li, Y.T., Kuo, P., Guo, J.I., 2021. Automatic industry PCB Board DIP process defect detection system based on deep ensemble self-adaption method. IEEE Trans. Compon., Packag. Manuf. Technol. vol. 11 (2), 312–323.
Lian, J., et al., 2020. Deep-learning-based small surface defect detection via an exaggerated local variation-based generative adversarial network. IEEE Trans. Ind. Inform. vol. 16 (2), 1343–1351.
Lin, Y., et al., 2010. Large-scale image classification: fast feature extraction and svm training. CVPR 2011 1689–1696.
Liu, J., Song, K., Feng, M., Yan, Y., Tu, Z., Zhu, L., 2021. Semi-supervised anomaly detection with dual prototypes autoencoder for industrial surface inspection. Opt. Lasers Eng. vol. 136.
Liu, T., He, Z., 2022. TAS2-net: triple-attention semantic segmentation network for small surface defect detection. IEEE Trans. Instrum. Meas. vol. 71.
Long, J., Shelhamer, E., Darrell, T., 2015. Fully convolutional networks for semantic segmentation. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 3431–3440.
Ma, J., Wang, Y., Shi, C., Lu, C., 2018. Fast surface defect detection using improved Gabor filters. 25th IEEE Int. Conf. Image Process. (ICIP) 1508–1512.

Makhzani, A., Frey, B., 2015. Winner-take-all autoencoders. Adv. Neural Inf. Process. Syst. vol. 2015-Jan, 2791–2799.
Malek, A.S., Drean, J.-Y., Bigue, L., Osselin, J.-F., 2013. Optimization of automated online fabric inspection by fast Fourier transform (FFT) and cross-correlation. Text. Res. J. vol. 83 (3), 256–268.
Masci, J., Meier, U., Ciresan, D., Schmidhuber, J., 2011. Stacked convolutional auto-encoders for hierarchical feature extraction. Int. Conf. Artif. Neural Netw. 52–59.
Mehta, M., Shao, C., 2022. Federated learning-based semantic segmentation for pixel-wise defect detection in additive manufacturing. J. Manuf. Syst. vol. 64 (May), 197–210.
Mirapeix, J., Garcıa-Allend, P.B., Cobo, A., Conde, O.M., López-Higuera, J.M., 2007. Real-time arc-welding defect detection and classification with principal component analysis and artificial neural networks. NDT E Int vol. 40 (4), 315–323.
Mujeeb, A., Dai, W., Erdt, M., Sourin, A., 2019. One class-based feature learning approach for defect detection using deep autoencoders. Adv. Eng. Inform. vol. 42.
Mushtaq, F., Ramesh, K., Deshmukh, S., Ray, T., Parimi, C., . 2023. Nuts & bolts: YOLO-v5 and image processing based component identification system. Eng. Appl. Artif. Intell. vol. 118.
Nakashima, K., et al., 2021. Detection of minute defects using transfer learning-based CNN models. Artif. Life Robot vol. 26 (1), 35–41.
Nakazawa, T., Kulkarni, D.V., 2019. Anomaly detection and segmentation for wafer defect patterns using deep Convolutional Encoder-Decoder Neural Network Architectures in Semiconductor Manufacturing. IEEE Trans. Semicond. Manuf. vol. 32 (2), 250–256.
Nejc, K., Drago, B., 2019. Visual inspection system for anomaly detection on KTL coatings using VAE. Procedia CIRP vol. 81, 771–774.
Nguyen, V.H., Pham, V.H., Cui, X., Ma, M., Kim, H., 2017. Design and evaluation of features and classifiers for OLED panel defect recognition in machine vision. J. Inf. Telecom vol. 1 (4), 334–350.
Niu, S., Li, B., Wang, X., Lin, H., 2020. Defect Image Sample Generation with GAN for Improving Defect Recognition. IEEE Trans. Autom. Sci. Eng. vol. 17 (3), 1611–1622.
Ouyang, W., Xu, B., Hou, J., Yuan, X., 2019. Fabric defect detection using activation layer embedded convolutional neural network. IEEE Access vol. 7 (2), 70130–70140.
Ozseven, T., 2019. Surface defect detection and quantification with image processing methods. Theoretical Investigations and Applied Studies in Engineering. Ekin Publishing House, pp. 63–98.
Pan, Z., Yang, J., Wang, X., Wang, F., Azim, I., Wang, C., 2021. Image-based surface scratch detection on architectural glass panels using deep learning approach. Constr. Build. Mater. vol. 282, 1–11.
Park, J., Riaz, H., Kim, H., Kim, J., 2020. Advanced cover glass defect detection and classification based on multi-DNN model. Manuf. Lett. vol. 23, 53–61.
Radford, A., Metz, L., Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv Prepr.* arXiv1511. 06434.
Rahman, M.R.U., Chen, H., 2020. Defects inspection in polycrystalline solar cells electroluminescence images using deep learning. IEEE Access vol. 8, 40547–40558.
Ren, R., Hung, T., Tan, K.C., 2017. A generic deep-learning-based approach for automated surface inspection. IEEE Trans. Cybern. vol. 48 (3), 929–940.
Ren, R., Hung, T., Tan, K.C., 2018. A generic deep-learning-based approach for automated surface inspection. IEEE Trans. Cybern. vol. 48 (3), 929–940.
Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: convolutional networks for biomedical image segmentation. Int. Conf. Med. Image Comput. Comput. -Assist. Interv. 234–241.
Ruff, L., Vandermeulen, R.A., Gornitz, N., Binder, A., Uller, E.M., Kloft, M., 2019. Deep support vector data description for unsupervised and semi-supervised anomaly detection. Proc. ICML 2019 Work. Uncertain. Robustness Deep Learn. Long Beach, CA, USA.
Sánchez, J., Perronnin, F., 2011. High-dimensional signature compression for large-scale image classification. CVPR 2011 1665–1672.
Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.-C., 2018. Mobilenetv2: Inverted residuals and linear bottlenecks. Proc. IEEE Conf. Comput. Vis. Pattern Recognit. 4510–4520.
Schlegl, T., Seeböck, P., Waldstein, S.M., Langs, G., Schmidt-Erfurth, U., 2019. f-AnoGAN: fast unsupervised anomaly detection with generative adversarial networks. Med. Image Anal. vol. 54, 30–44.
Shao, L., Zhang, E., Ma, Q., Li, M., 2022. Pixel-wise semisupervised fabric defect detection method combined with multitask mean teacher. IEEE Trans. Instrum. Meas. vol. 71.
Shi, J., Yin, W., Du, Y., Folkesson, J., 2019. Automated underwater pipeline damage detection using neural nets. ICRA Workshop Underw. Robot. Percept.
Shu, Y., Li, B., Lin, H., 2021a. Quality safety monitoring of LED chips using deep learning-based vision inspection methods. Meas. J. Int. Meas. Confed. vol. 168.
Shu, Y.F., Li, B., Li, X., Xiong, C., Cao, S., Wen, X.Y., 2021b. Deep learning-based fast recognition of commutator surface defects. Measurement vol. 178, 109324.
Simonyan, K., Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv Prepr* arXiv 1409, 1556.
Sun, K., et al., 2019. High-resolution representations for labeling pixels and regions. arXiv1904.04514 arXiv Prepr. arXiv1904.04514.
Szegedy, C., et al., . 2015. Going deeper with convolutions. 2015 *IEEE Conf. Comput. Vis. Pattern Recognit.*
Tabernik, D., Sela, S., Skvarc, J., Skocaj, D., 2020. Segmentation-based deep-learning approach for surface-defect detection. J. Intell. Manuf. vol. 31 (3), 759–776.
Tang, Z., Tian, E., Wang, Y., Wang, L., Yang, T., 2021. Nondestructive defect detection in castings by using spatial attention bilinear convolutional neural network. IEEE Trans. Ind. Inform. vol. 17 (1), 82–89.
Tsai, D.M., Jen, P.H., 2021. Autoencoder-based anomaly detection for surface defect inspection. Adv. Eng. Inform. vol. 48.
Tsai, D.M., Fan, S.K.S., Chou, Y.H., 2021. Auto-annotated deep segmentation for surface defect detection. IEEE Trans. Instrum. Meas. vol. 70.
Tulbure, A., Tulbure, A., Dulf, E., 2022. A review on modern defect detection models using DCNNs – deep convolutional neural networks. J. Adv. Res. no. 35, 33–48.
Versini, E., Snidaro, L., Liani, A., 2022. SCL—segmentation–classification combined Loss for surface defect detection. Expert Syst. Appl. vol. 198.
Wang, J., Fu, P., Gao, R.X., 2019. Machine vision intelligence for product defect inspection based on deep learning and Hough transform. J. Manuf. Syst. vol. 51, 52–60.
Wang, Q., Yang, R., Wu, C., Liu, Y., 2021. An effective defect detection method based on improved Generative Adversarial Networks (iGAN) for machined surfaces. J. Manuf. Process. vol. 65, 373–381.
Wang, Y., Sun, Y., Lv, P., Wang, H., 2008. Detection of line weld defects based on multiple thresholds and support vector machine. NDT E Int vol. 41 (7), 517–524.
Wen, W., Xia, A., 1999. Verifying edges for visual inspection purposes. Pattern Recognit. Lett. vol. 20 (3), 315–328.
Wen, Z., Cao, J., Liu, X., Ying, S., 2014. Fabric defects detection using adaptive wavelets. Int. J. Cloth. Sci. Technol.
Xia, C., Pan, Z., Fei, Z., Zhang, S., Li, H., 2020. Vision based defects detection for Keyhole TIG welding using deep learning with visual explanation. J. Manuf. Process. vol. 56, 845–855.
Xie, Q., Luong, M.-T., Hovy, E., Le, Q.V., 2020. Self-training with noisy student improves imagenet classification. Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. 10687–10698.
Xu, L., Huang, Q., 2012. Modeling the interactions among neighboring nanostructures for local feature characterization and defect detection. IEEE Trans. Autom. Sci. Eng. vol. 9 (4), 745–754.
Xu, L., Lv, S., Deng, Y., Li, X., 2020. A weakly supervised surface defect detection based on convolutional neural network. IEEE Access vol. 8, 42285–42296.
Yang, H., Chen, Y., Song, K., Yin, Z., 2019. Multiscale feature-clustering-based fully convolutional autoencoder for fast accurate visual inspection of texture surface defects. IEEE Trans. Autom. Sci. Eng. vol. 16 (3), 1450–1467.
Yang, H., Mei, S., Song, K., Tao, B., Yin, Z., 2018. Transfer-learning-based online mura defect classification. IEEE Trans. Semicond. Manuf. vol. 31 (1), 116–123.
Yang, L., Fan, J., Huo, B., Li, E., Liu, Y., 2022a. A nondestructive automatic defect detection method with pixelwise segmentation. Knowl. -Based Syst. vol. 242.
Yang, L., Song, S., Fan, J., Huo, B., Li, E., Liu, Y., 2022b. An automatic deep segmentation network for pixel-level welding defect detection. IEEE Trans. Instrum. Meas. vol. 71.
Yang, S., Li, X., Jia, X., Wang, Y., Zhao, H., Lee, J., 2020b. Deep learning-based intelligent defect detection of cutting wheels with industrial images in manufacturing. Procedia Manuf. vol. 48 (2019), 902–907.
Yang, T., Peng, S., Huang, L., 2020c. Surface defect detection of voltage-dependent resistors using convolutional neural networks. Multimed. Tools Appl. vol. 79 (9–10), 6531–6546.
Yang, Y., et al., 2020a. A lightweight deep learning algorithm for inspection of laser welding defects on safety vent of power battery. Comput. Ind. vol. 123.
Yu, J., 2019. Enhanced stacked denoising autoencoder-based feature learning for recognition of wafer map defects. IEEE Trans. Semicond. Manuf. vol. 32 (4), 613–624.
Yu, J., Zheng, X., Wang, S., 2019. A deep autoencoder feature learning method for process pattern recognition. J. Process Control vol. 79, 1–15.
Zambal, S., Heindl, C., Eitzinger, C., Scharinger, J., 2019. End-to-end defect detection in automated fiber placement based on artificially generated data. Fourteenth Int. Conf. Qual. Control Artif. Vis. vol. 11172, 111721G.
Zeiler, M.D., Fergus, R., 2013. Visualizing and understanding convolutional networks. Lect. Notes Comput. Sci. 818–833.
Zenati, H., Foo, C.S., Lecouat, B., Manek, G., Chandrasekhar, V.R., 2018. Efficient GAN-based anomaly detection. arXiv:1802.06222 arXiv Prepr. arXiv:1802.06222.
Zeng, X., et al., 2016. Crafting gbd-net for object detection. IEEE Trans. Pattern Anal. Mach. Intell. vol. 40 (9), 2109–2123.
Zhai, X., Kolesnikov, A., Houlsby, N., Beyer, L., 2021. Scaling vision transformers. arXiv2106.04560 arXiv Prepr. 1–31.
Zhang, C., Feng, S., Wang, X., Wang, Y., 2021. ZJU-leaper: a benchmark dataset for fabric defect detection and a comparative study. IEEE Trans. Artif. Intell. vol. 1 (3), 219–232.
Zhang, Z., Wang, X., Liu, S., Sun, L., Chen, L., Guo, Y., 2018. An automatic recognition method for PCB visual defects. Int. Conf. Sens. Diagn. Progn., Control (SDPC) 138–142.
Zheng, M., Lei, Z., Zhang, K., 2020b. Intelligent detection of building cracks based on deep learning. Image Vis. Comput. vol. 103.
Zheng, X., Wang, H., Chen, J., Kong, Y., Zheng, S., 2020a. A generic semi-supervised deep learning-based approach for automated surface inspection. IEEE Access vol. 8, 114088–114099.
Zheng, X., Wang, H., Chen, J., Kong, Y., Zheng, S., 2020c. A generic semi-supervised deep learning-based approach for automated surface inspection. IEEE Access vol. 8, 114088–114099.
Zhou, F., Liu, G., Xu, F., Deng, H., 2019. A generic automated surface defect detection based on a bilinear model. Appl. Sci. vol. 9 (15).
Zhu, Z., Han, G., Jia, G., Shu, L., 2020. Modified densenet for automatic fabric defect detection with edge computing for minimizing latency. IEEE Internet Things J. vol. 7 (10), 9623–9636.
Zou, Q., Zhang, Z., Li, Q., Qi, X., Wang, Q., Wang, S., 2018. Deepcrack: Learning hierarchical convolutional features for crack detection. IEEE Trans. Image Process. vol. 28 (3), 1498–1512.

Shashi Bhushan Jha received the B.Eng. degree in computer technology from Nagpur University, India in 2013, and the M.S. degree in industrial and systems engineering from the Indian Institute of Technology Kharagpur India in 2018. From 2019, he is a PhD student in Electrical Engineering and Computer Science at Embry Riddle Aeronautical University in Daytona Beach, Florida. His research interests include automated defect detection, computer vision, deep learning, machine learning, and optimization. His previous work was published in renowned journals such as Computers & Industrial Engineering, and Journal of Intelligent Manufacturing.

Radu Babiceanu is the Interim Chair and a Professor of Systems Engineering with the Department of Electrical Engineering and Computer Science at Embry-Riddle Aeronautical University in Daytona Beach, Florida. He received his Ph.D. degree in Industrial and Systems Engineering from Virginia Tech in 2005. At Embry-Riddle, he teaches systems engineering courses, such as Systems Engineering and Analysis, System Quality Assurance, and System Safety and Certification. His research interests are in the aviation and aerospace systems domain, with emphasize in cybersecurity and safety-critical systems assurance and certification, aviation datalink communications, formal system modeling and verification, and AI/ML approaches to enhanced operations. He has authored or co-authored more than 90-refereed publications.