REVIEW PAPER

State of the Art in Defect Detection Based on Machine Vision

Zhonghe Ren[1] · Fengzhou Fang[1,2] · Ning Yan[1] · You Wu[1]

Abstract

Machine vision significantly improves the efficiency, quality, and reliability of defect detection. In visual inspection, excellent optical illumination platforms and suitable image acquisition hardware are the prerequisites for obtaining high-quality images. Image processing and analysis are key technologies in obtaining defect information, while deep learning is significantly impacting the field of image analysis. In this study, a brief history and the state of the art in optical illumination, image acquisition, image processing, and image analysis in the field of visual inspection are systematically discussed. The latest developments in industrial defect detection based on machine vision are introduced. In the further development of the field of visual inspection, the application of deep learning will play an increasingly important role. Thus, a detailed description of the application of deep learning in defect classification, localization and segmentation follows the discussion of traditional defect detection algorithms. Finally, future prospects for the development of visual inspection technology are explored.

Keywords Machine vision · Defect detection · Image processing · Deep learning

1 Introduction

Advanced industrial systems require increasingly improved product performance along with an increased need for quality control during production [1–3]. However, defects, such as scratches, spots, or holes on the surface of the product, adversely affect not only the aesthetics of the product and the comfort in using it but also its performance [4–7]. Defect detection is an effective method to reduce the adverse impact of product defects [8, 9].

Artificial visual inspection is a traditional method to perform quality control for industrial products [10]. Although in some cases, artificial visual inspection may be superior, it is inefficient and prone to fatigue. Artificial visual inspection is not feasible for some applications that have dangerous consequences in the event of a failure [11]. Because of its shortcomings, such as a low sampling rate, poor real-time

performance, and low detection confidence, artificial visual inspection cannot meet the efficiency and quality requirements of modern industrial production lines [12]. Hence, more efficient and reliable visual inspection technologies need to be developed.

Machine vision is one of the key technologies used to perform intelligent manufacturing, and it has become an effective way to replace artificial visual inspection [13, 14]. Machine vision is a system that automatically receives and processes images of a real object through optical devices and noncontact sensors. Vision is one of the highest levels of human perception. Images play a very important role in human perception [15]. However, human perception is limited to the visible band of the electromagnetic spectrum. Machine vision inspection technology can cover the whole electromagnetic spectrum, ranging from gamma rays to radio waves [16]. Through powerful vision sensors, ingeniously designed optical transmission methods, and image processing algorithms, machine vision can accomplish many tasks that cannot be performed by artificial vision. With the development of computer equipment and artificial intelligence, machine vision, as a measurement and judgement technology, has been used widely in industry. Machine vision detection technology can improve the detection efficiency and degree of automation, enhance the real-time performance and accuracy of detection, and reduce manpower

✉ Fengzhou Fang
fzfang@tju.edu.cn

[1] State Key Laboratory of Precision Measuring Technology and Instruments, Laboratory of Micro/Nano Manufacturing Technology (MNMT), Tianjin University, Tianjin 300072, China

[2] Centre of Micro/Nano Manufacturing Technology (MNMT-Dublin), University College Dublin, Dublin 4, Ireland

Springer KSPE

requirements, especially for some large-scale repetitive industrial production processes. As a non-contact and non-destructive detection method, machine vision can be easily employed to perform information integration, automation, intelligence, and precise control. It has become the basic technology required in computer integrated manufacturing and intelligent manufacturing. Moreover, machine vision has a wider range of spectral responses and a greater ability to work for a long time in harsh environments. Thus, the application of machine vision in manufacturing processes can benefit a large number of industrial activities [17–19].
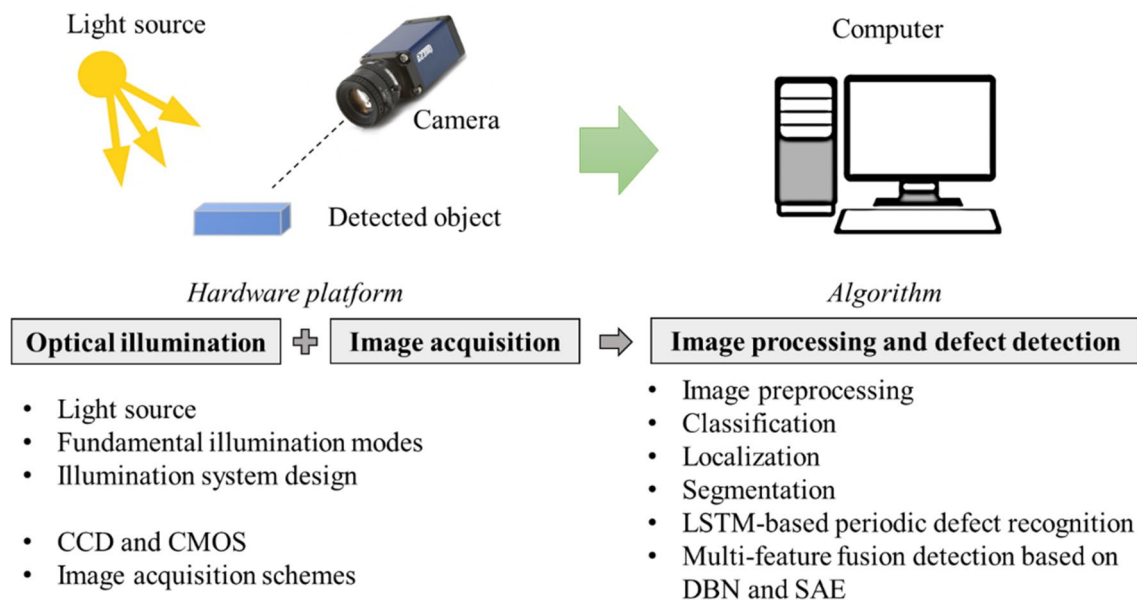
A typical industrial visual inspection system mainly consists of three modules—optical illumination, image acquisition, and image processing and defect detection [11, 20]—as shown in Fig. 1. First, based on the product characteristics and inspection requirements, an optical illumination platform is designed. Next, CCD cameras or other image acquisition hardware are used to convert the target objects placed in the light field into images and transmit them to a computer. As an information carrier, the images that can reflect the features of the objects constitute the core element of visual inspection; hence, their quality is very important. Excellent optical illumination platforms and suitable image acquisition hardware are the prerequisites for obtaining high-quality images. Finally, based on some traditional image processing algorithms or deep learning algorithms, various operations are carried out on the images to extract features and to perform classification, localization, segmentation and other operations. Image processing is a key technology in machine vision. Through image processing and analysis, a computer can automatically understand, analyze, and judge image features, and then control the actuator of the automatic production line for further operation [21].

In industry, the architecture can be used as a step guideline for designing a visual inspection system. For instance, investigating surface characteristics was the first step in designing a strongly reflective metal surface visual inspection system; hence, diffuse bright filed back light illumination was adopted. Light-sensitive components were then used for image acquisition. After image acquisition, wavelet smoothing was used for image preprocessing, and Otsu threshold was employed to segment the image. Finally, support vector machine classifier was designed for defect classification [22].

The main evaluation indexes of a visual inspection system are accuracy, efficiency and robustness. The goals of the system are high precision, high efficiency and strong robustness. In order to achieve these goals, it needs an excellent coordination of optical illumination, image acquisition, and image processing and defect detection.

This study is focused on the current state of development of industrial defect detection utilizing machine vision. Visual inspection modules, including optical illumination, image acquisition, image processing and defect detection are discussed in detail. The light source and illumination system design are discussed in Sect. 2. Section 3 describes the image sensors and image acquisition design for particular scenarios. Then, as the main portion of this study, Sect. 4 focuses on defect detection tasks such as defect classification, localization and segmentation, and it discusses representative traditional image processing methods and intelligent methods based on deep learning. Finally, insights into

Fig. 1 Typical industrial visual inspection system architecture and main components

future research in defect detection based on machine vision are presented in Sect. 5.

2 Optical Illumination

Visual inspection technology is based on an image, and encompasses image acquisition and image processing [23]. The key to the success of the visual inspection system lies in getting high quality images. In general, the image quality is mainly affected by two factors: optical illumination and image acquisition [24, 25]. The main function of an optical illumination platform is to overcome the interference of environmental lighting, ensure the stability of the image, and obtain an image with a high contrast. Thus, the main goal of the optical illumination platform is to make the important features of the objects visible and reduce undesired features of the objects.

The research on optical illumination has a long history. In the 1980s, the commercial white light source for machine vision was not available in the market and some light sources designed for workbenches could not be easily integrated into vision detection systems. With the transition of vision detection systems from laboratory to industry, the necessity of optimizing optical illumination systems has gradually become a research area of focus, and the importance of optical illumination in visual systems has been understood at a preliminary level. In 1987, Mersch [26] systematically discussed the importance of optical illumination in visual systems. Based on the technical conditions at that time, he analyzed the application of polarization and color filters and pointed out the advantages of optical fiber lighting for the illumination of a small area. Furthermore, he discussed the fluorescent marking lighting method and frequency flash lighting technology. Later, Cowan [27] designed the positioning of a camera and a light source by using their models and surface reflectivity to meet the requirements of a vision system. Sieczka et al. [28] presented a detailed exposition and discussion on some important issues related to light sources, such as light source efficiency, light divergence, spectral content, light source size, and packaging. Combined with mathematical programming, Yi et al. [29] discussed the placement design of sensors and light sources. Kopparapu [30] proposed a design method, using multiple light sources to achieve uniform illumination, which regarded the solution of the optimal position of light source as a minimization problem, and used simulations to verify the effectiveness and applicability of the method.

Despite the rapid growth of computer digital image processing and calculations, optical illumination still plays a significant role in visual inspection systems. For an on-line visual inspection system, compared to the long calculation period to process the image by advanced algorithms,

a specially designed optical illumination for field lighting can achieve a higher detection accuracy. Furthermore, a specially designed optical illumination can also meet the real-time requirements of the production line visual inspection in a better way. Therefore, as an important part of the machine vision application, optical illumination deserves further discussion.

2.1 Light Source

Light is a typical energy source for image formation. Common light source devices include LED lamps of various shapes, high frequency fluorescent lamps, optical fiber halogen lamps, etc. Currently, LED lamps have become available for every type of machine vision application [31, 32]. An LED light source can be customized in several array configurations to achieve the desired irradiance [33, 34]. In vision applications, the most popular light source is a circular ring array of LEDs [35]. The circular ring array of LEDs possesses high brightness and can be conveniently installed. It can effectively avoid the shadow phenomenon and highlight the features to be detected. It is often used for IC chip appearance and character detection [36], printed circuit board (PCB) substrate detection [37], microscope illumination [38], etc. In structured lighting, the linear array of LEDs is widely used [35]. Furthermore, it has good heat dissipation and flexibility of usage, and can be used for defect detection of some large structural parts, such as copper strip [39] and steel sheet [40].

Visible light is a common light source. Different wavelengths of light have distinct characteristics and applications. As its wavelength changes, visible light assumes different colors [41, 42]. White light source is a multi-wavelength compound light, which is widely used. High brightness white light source is suitable for color image shooting. The wavelength of blue light is between 430 and 480 nm, and is suitable for sheet metal, machining parts, and other products with a silver colored background, as well as metal printing on film. The wavelength of red light is typically between 600 and 720 nm, which is relatively long and can pass through dark objects. It is used in applications, such as line detection and light transmission film thickness detection. A red light source can significantly improve the contrast of an image. The wavelength of a green light source is typically between 510 and 530 nm and lies between the wavelengths of the red and blue lights. The green light source is mainly used for products with red or silver colored backgrounds.

Invisible light could be infrared light, ultraviolet light, or X-rays. The wavelength of infrared light is generally 780–1400 nm. Infrared light has a strong propagation ability and is generally used in liquid crystal display (LCD) screen detection and video monitoring industries [43]. The wavelength of ultraviolet light is generally 190–400 nm. The

ultraviolet light has a short wavelength and strong penetration and is mainly used in certificate detection, ITO detection of touch screens, scratch detection of metal surfaces [44], etc. X-ray is a type of electromagnetic wave, whose wavelength range is from 0.01 to 10 nm. X-rays have a short wavelength and good perspective effect and are widely used in various perspective tests in industry [45]. These wavelengths of light are invisible to the human eye; however, they can be applied in machine vision. This is also another important advantage of machine vision over artificial vision.

To enhance the visibility of certain features, it is important to consider the interaction between light and objects, including the propagation mode of light, when it reaches the surface of objects, and the relationship between the wavelength of light and the color of objects [22]. The propagation of light is different in different materials. The defective part of an object would also affect the propagation of light. The common defects in surface inspection can be categorized into two categories: (i) geometric defects, such as pits, scratches, cracks, burrs, bulges, scratches, and bumps; (ii) surface strength defects or density defects, such as oxidation, rust, and stains. The geometric defects change the surface reflection, and surface strength defects or density defects change the surface reflection, as well as absorption. In visual inspection, opaque objects are common. The opaque objects have the ability to reflect or absorb color light of different wavelengths. The absorbed color light cannot be seen and only the reflected color light can directly act on the image acquisition devices. Using a black-and-white camera, reliable and stable detection can be achieved by selecting a specific wavelength of light source and highlighting the grayscale difference between the part to be detected on the surface of the object and the other parts. Therefore, the contrast of the image can be enhanced by effectively selecting the wavelength of light or combining multiple wavelengths of light.

2.2 Fundamental Illumination Modes

With the development of optical illumination technology, various types of designs of illumination structure have emerged [46]. In the field of machine vision, based on different positional relationship among the light source, object, and camera, the illumination can be divided into forward and back illuminations. According to the performance characteristics of light source, it can be divided into structured light and stroboscopic light.

2.2.1 Forward and Back Illuminations

In forward lighting, the light source and the camera are located on the same side of the object. Being the most widely used illumination method, forward lighting is suitable for detecting surface defects, scratches, and the important details of objects, especially the surface texture features. The angle between the light beam and the object surface affects the illumination effect. Depending on whether the light is directly reflected onto the camera, the forward lighting is divided into bright field forward lighting and dark field forward lighting, as shown in Fig. 2a, b. For dark field forward lighting, reducing the incident angle of the light forms a low angle dark field forward lighting. Low angle dark field forward lighting can highlight the edge and height of the surface, enhance the topological structure of the image, and provide a strong performance on the surface concavity and convexity. Coaxial forward lighting is a special forward lighting mode. Coaxial light source refers to a high-intensity uniform light passing through the half mirror to form the light coaxial with the lens, as shown in Fig. 2c. Coaxial forward lighting provides more uniform illumination than traditional lighting mode, while avoiding the reflection of the object. Therefore, it improves the accuracy and reproducibility of machine vision. The coaxial forward lighting can be used to detect surface defects, cracks, scratches, etc. For a highly reflective object with a smooth surface, the light is first projected onto the rough cover to produce a nondirectional and soft light, and then projected on the surface of the detected object, which can avoid the strong reflection produced by the direct lighting mode, as shown in Fig. 2d. Scattering forward lighting of a dome structure is commonly used in solder joint detection, chip pin detection, etc. In back lighting, the light source is placed behind the object, as shown in Fig. 2e. A significant feature of back lighting is that it can highlight the shadow of opaque objects or observe the interior of transparent objects. Its advantage is that it can clearly outline the edge of the object to be measured. It is often used in object shape detection and dimension detection. Table 1 compares these typical illumination modes.

2.2.2 Structured Light Illumination

A structured light illumination causes the light to have a certain shape by specific means, so as to facilitate the detection of three-dimensional object information using two-dimensional vision [47], as shown in Fig. 3. Here, firstly, the specific light information is projected on the object surface and the background. Then, a camera is used to collect the image containing the change in the information of the light signal caused by the structure of the object. Finally, the position and depth of the object are calculated by digital image processing technology, and the whole 3D space is restored [48].

Structured light illumination technology is widely used in visual measurement and inspection. Based on laser structured light vision, Li et al. [50] developed an inspection system for weld bead profile monitoring, measuring, and defect detection with scale calibration. Using triangulation with
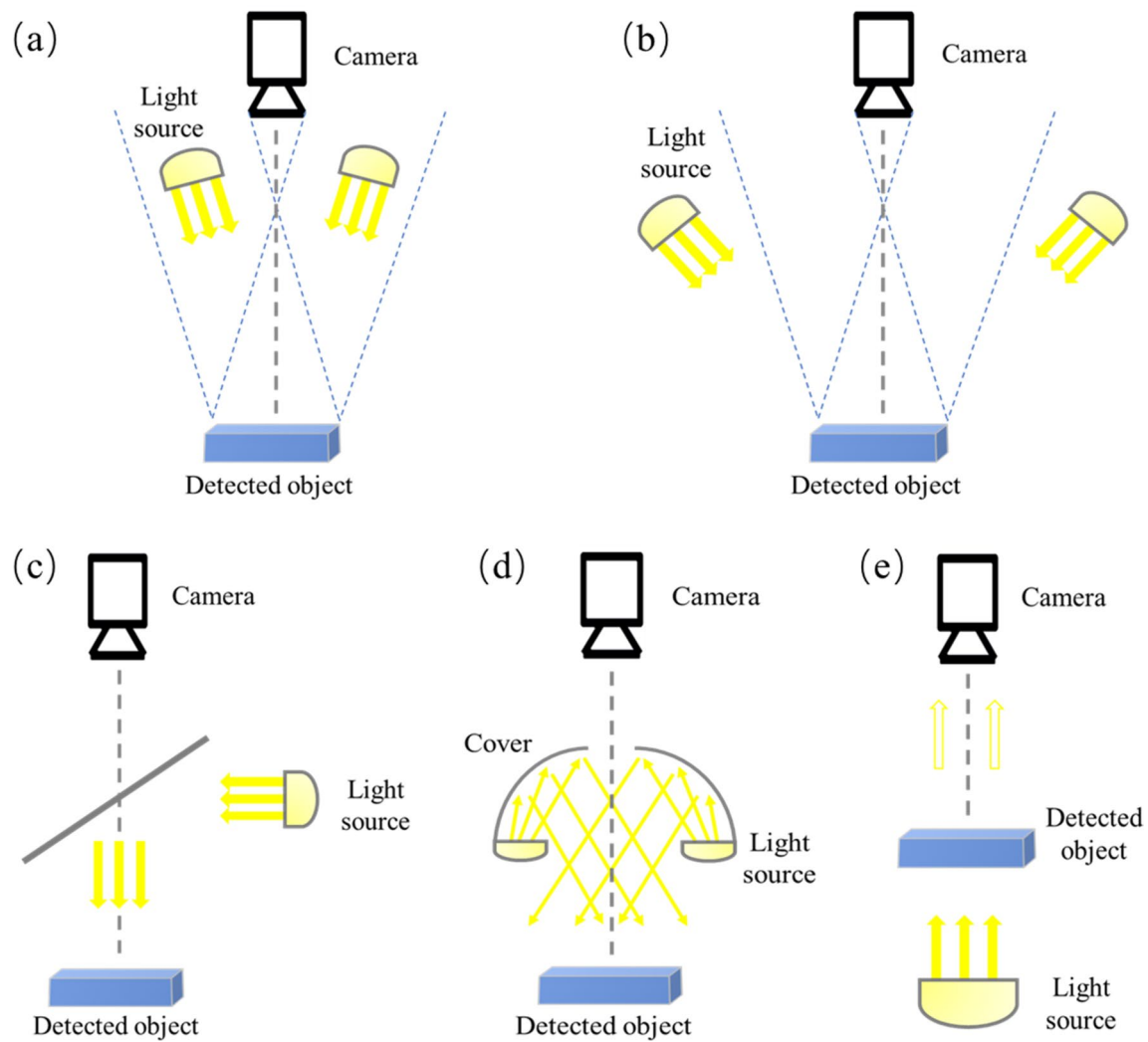
Fig. 2 Schematic diagram of typical illumination modes: **a** bright field forward lighting; **b** dark field forward lighting; **c** coaxial forward lighting; **d** scattering forward lighting of dome structure; e back lighting

Table 1 The performances of typical illumination modes

| Illumination modes	Performances
Bright field forward lighting	It is good for high contrast, but reflective surfaces produce specular reflections
Dark field forward lighting	It can obtain the diffuse light caused by the change of surface texture, and avoid the specular light
Coaxial forward lighting	It can overcome the interference caused by surface reflection, and is mainly used to detect bumps, scratches, cracks and foreign matters on the smooth surface of objects
Scattering forward lighting of dome structure	It can avoid the strong reflection produced by the direct lighting mode
Back lighting	It can highlight the outline of opaque objects or observe the interior of transparent objects

line-scan cameras in a 2D plane, Lilienblum and Al-Hamadi [51] presented a novel technique for optical 3D surface reconstruction by using a combination of line-scan cameras and structured light. It can measure continuously, whereby a single surface scan is sufficient to calculate a high-quality 3D reconstruction.

2.2.3 Stroboscopic Light Illumination

Stroboscopic light is a type of illumination technology applied in optical imaging. It can achieve the effect of freezing the motion of a moving object. An appropriate optical pulse can eliminate the motion blur in the images

Fig. 3 Schematic diagram of structured light illumination [49]. Copyright 2018, Elsevier
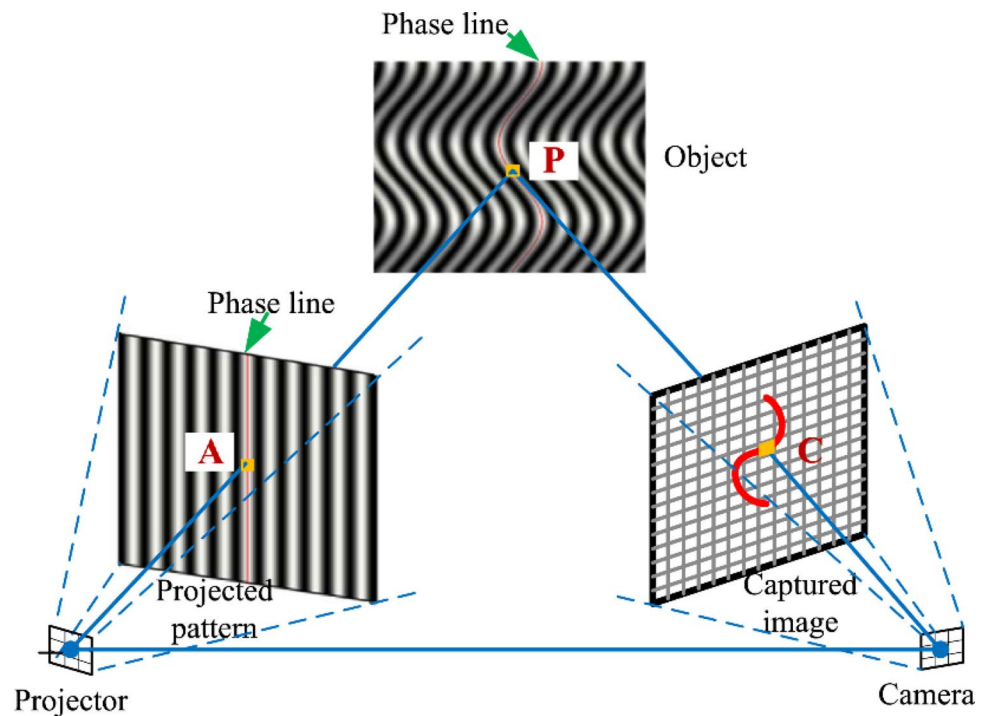
of fast-moving objects, which is very suitable for on-line high-speed detection of machine vision. By improving the brightness of the stroboscopic light, the exposure time can be reduced, and the whole vision detection system can run faster. In a stroboscopic illumination environment, the aperture can be reduced to get a better image depth of the field. To solve the problem of fuzzy images when high-speed moving objects are photographed in a continuous light source, Chen et al. [52] designed a narrow-pulse and high-current strobe light, with a high-illumination LED as the light source. A field-programmable gate array (FPGA) chip generates a pulse signal to control the timing of the stroboscopic light source.

2.2.4 Auxiliary Optical Devices for Illumination

In practical applications, production lines and working environments have different requirements on the brightness, working distance, and irradiation angle of light sources. They are sometimes limited to specific application environments, and it is very difficult to obtain a good visual image directly through the adjustment of light source type or irradiation angle. In this case, some special auxiliary optical devices are needed.

The common auxiliary optical devices include a filter, reflector, spectroscope, prism, polarizer, diffuser, optical fiber, screen, etc. In the image acquisition stage, some noise interferences can be eliminated and the signal-to-noise ratio (SNR) of the image can be improved by using a filter, and consequently, improve the efficiency of the system. A reflector can change the path and angle of the light, change the distance between the observation points, realize simultaneous or time-sharing observations of multiple targets, and provide more choice space for the installation of the light source. In a spectroscope, the ratio of the reflected light to the refracted light can be adjusted by changing the coating parameters. The coaxial illumination is a special case of a spectroscope. A prism can separate multi-colored compound light and get a single frequency light source. A polarizer can eliminate the reflection of light on non-metallic surfaces. A diffuser can make light more uniform and reduce unwanted reflections. An optical fiber can gather the light beam in an optical fiber tube for transmission, which makes the installation of the light source more flexible and convenient. The application of auxiliary optical devices can be of great help in industrial defect detection. For example, metal surface has a high reflection coefficient that makes it difficult to design a proper lighting system for defect enhancement. To suppress this light, Zhang et al. [22] designed a diffuse bright-field back light illumination and mounted a polarizing filter in front of the camera, and oriented it in such a way that the polarized light would be suppressed.
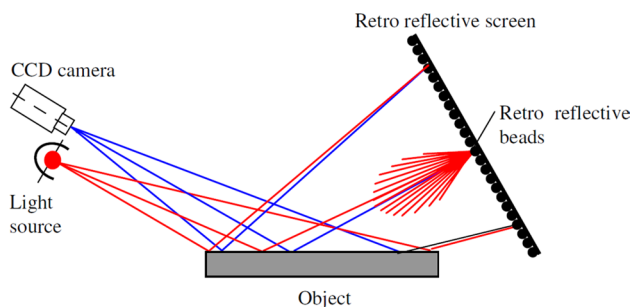
2.3 Illumination System Design

A light source can be designed in various shapes and structures, so that the light emitted has different characteristics. An effective way to achieve a specific lighting function is through an innovative design, which combines various fundamental illumination methods and some auxiliary optical

devices. For some special occasions, there are special-purpose illumination methods available, which include point light source illumination, shadow less illumination, parallel light optical unit illumination, microscope illumination, and customized illumination based on the customer requirements.

For a visual inspection project that aims to obtain high quality images, it is necessary to design a targeted optical illumination system. Firstly, according to the specific needs of the project, the key factors, such as the characteristics and motion state of the objects, surrounding environment and type of camera should be analyzed. Then, the difference between the target and the background is studied to find out the difference in the optical phenomenon between them. According to the characteristics of the materials and the interaction between the light source and the objects, a preliminary determination of the type and color of the light source should be conducted. Finally, experiments should be carried out and from the test results the illumination system should be adjusted until it can meet the requirements of visual inspection. The following is the analysis of several application cases, respectively about highly reflective surfaces, heteromorphic structure, moving objects, and minimally invasive surgery (MIS).

1. Highly reflective surfaces are widely used in automobile, aviation, life science and aerospace industry. These application scenarios have high requirements for surface quality. Optical double-pass retro-reflection surface inspection technique is a typical optical detection technique realized by cleverly designing light reflection path, as shown in Fig. 4. It can inspect very small out-of-plane surface distortions on a specularly reflective surface, such as indentations and protrusions [53]. The advantage of optical double-pass retro-reflection surface inspection technique is that large surface area can be observed in real time, so it can be used for online real-time visual inspection.

Fig. 4 Schematic diagram of double-pass retro-reflection illumination system [53]. Copyright 2007, Elsevier

2. For a belt condition monitoring system, due to the special shape of belt, unique design requirements are put forward for the illumination system. Yang et al. [54] arranged high-brightness linear light sources in a vaulted shape. This lighting design can adapt to the structural characteristics of the upper belt and improve the detection efficiency.

3. To cater to the diverse reflection characteristics of the surface of tin steel strips and different speeds of a tinning line, Peng and He [55] proposed an adaptive illumination light source. This light source was integrated with a time delay integration charge-coupled device to capture the images of the moving objects and facilitate inspection of the surface quality of the tin steel strips.

4. The combination of structured light and white light can take advantage of their advantages to achieve the desired effect. Clancy et al. [56] proposed a MIS stroboscopic illumination system, in which structured light and white light are interleaved during a high-speed camera acquisition. Besides playing its role in the corresponding cycles, the structured light is not perceived and white light can be used solely for navigation and visual assessment during the shielding period of structured light.

Optical illumination plays an important role in visual inspection. To achieve an appropriate illumination effect for a specific scenario, an appropriate light source should be employed based on considering the characteristics of the light source and the interactions between the light and the objects. To realize an innovative design of the optical illumination system, an effective combination of some fundamental illumination models is the preferred approach, and additional auxiliary optical devices will also help significantly.

3 Image Acquisition

In an appropriate optical illumination environment, an object surface can be imaged on a camera sensor by an optical lens. The optical signal is then converted into an electrical signal, and into a digital signal that can be processed by a computer to complete the acquisition process of the product surface image.

Image acquisition technology focuses on the characteristics of sensor devices and the field of view design. The typical photosensitive devices of industrial cameras are mainly based on charged coupled device (CCD) or complementary metal oxide semiconductor (CMOS) chips [57, 58]. The image acquisition technology of many conventional scenarios has become relatively mature, which this study does not elaborate. However, for some special detection requirements, a reasonable fields of view design and an effective photosensitive sensor selection can be very important. Several

representative image acquisition schemes for some particular image acquisition scenarios will be discussed.

3.1 CCD and CMOS

CCD or CMOS image sensor technology is essential for image capturing. They convert optical signals into electrical signals. However, these two types of chips adopt different methods and means in the transmission of this information and their respective designs are totally different.

The CCD, which is a photoelectric converter, originated in the early 1970s and developed to maturity in the 1990s [59, 60]. In 1974, White et al. [61] discussed the image array characteristics of a low illuminance area array CCD. In 1978, Dillon et al. [62] discussed a color imaging system using a single CCD area array. In 1990, Beyer [63] discussed the calibration of CCD for machine vision and robotics. In the CCD chip, the charge of the photosensitive pixel shifts and is converted into a signal. The CCD has a series of advantages, such as small distortion, small volume, low system noise, self-scanning, light weight, small power consumption, long life, wide sensing spectrum range, and high reliability. It can be made into a highly integrated assembly. The CMOS image sensors have been around for almost as long as the CCD; however, it was not until the 1990s that commercial CMOS sensor chips were manufactured [60].

Currently, CCD sensors are widely used in machine vision [64–66]. CMOS image sensors are still in their early stages and yet to mature [67, 68]. The CMOS image sensors can get an image quality similar to that of CCD product and have made great breakthroughs in terms of power consumption and integration.

3.2 Image Acquisition Schemes

This section discusses the state of the art in the image acquisition system design from the aspects of multiple views, omnidirectional vision, micro-domain vision, multispectral.

3.2.1 Multiple Views

In visual inspection, for parts with complex structures, it is difficult to capture all the key information based on a single image. In this case, only a collection of multiple images can show the features to be inspected.

Sun et al. [69] designed a machine vision system to acquire three-view images of one electric contact (EC). For each view, the system incorporated different image pre-processing and feature extraction methods to enhance and detect the surface defects. Chiou and Li [70] proposed a multi-view system for the inspection of PU-packing. Their system consisted of three inspection stations. Station 1 focuses on obtaining image information of the top and bottom surfaces of the package. Station 2 uses another camera to check the interior of the packing incorporated. Station 3 uses two line-scan cameras to simultaneously scan the inner and outer cylindrical surfaces. Through this method, each of the inspection stations would perform its assigned tasks, and multiple view images of the PU-packing could be effectively collected on an efficient work line. For detection in bearings, there are many parts that need to be inspected, such as the inner and outer rings. Shen et al. [71] designed a new image acquisition system for bearing cover inspection. To get the enhanced deformation information, three bearings were captured in one image. The left and right bearings were inspected for deformation defects, while the center bearing was inspected for other defects besides the deformations. This was an efficient and ingenious image acquisition system.

3.2.2 Omnidirectional Vision

Omnidirectional vision is mainly implemented by installing a fisheye lens [72]. Pipes are used to transport gas, liquid, or fluid with solid particles. The detection of their security often involves a visual detection of the inner wall of the pipes. For perspective stereo cameras with limited viewing angle, it is necessary to build a ring of cameras. Hansen et al. [73] introduced a visual odometry-based system, using calibrated fisheye imagery and sparse structured lighting to produce high-resolution 3D textured surface models of the inner pipe wall. The prototype robot with a fisheye lens and a fiberglass pipe network used for testing are shown in Fig. 5. Their research results showed that using a single fisheye camera, high-precision pipe mapping could be achieved. The advantage of this wide-angle fisheye lens system is that it can use a single camera to achieve the full pipe coverage, thus avoid the challenge of multiple camera calibration, and keep the overall size compact. This method is obviously of great significance to improving the efficiency of pipe inspection.

Contact lenses possess the characteristics of contact, lightness, and convenience. The quality of contact lenses has a major influence on the human eye. For contact lens detection, Chen et al. [74] presented an omnidirectional image of a fisheye lens for contact lens inspection system and proved the feasibility of the same. The optical reflection of the object surface depends on the material and microstructure. In the detection of industrial parts, light reflection measurement is an important task. Kogumasaka et al. [75] developed a small reflection measurement system using a fisheye camera, and demonstrated that the fisheye camera system was an effective means for high-precision surface finish inspection.
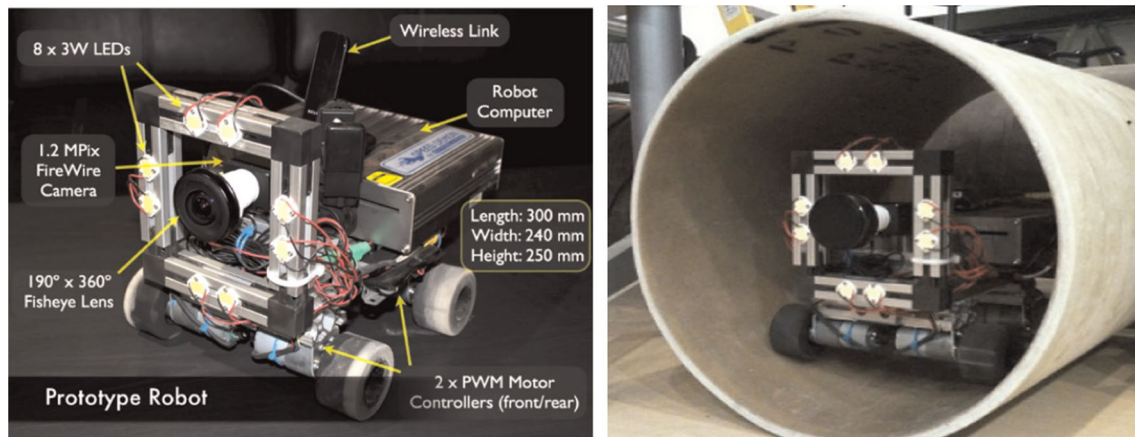
Fig. 5 Prototype robot and fiberglass pipe network used for testing [73]. Copyright 2015, SAGE Publications

3.2.3 Micro-Domain Vision

Quality inspection within mass production of micro-parts is a big challenge [76, 77]. During a micro-manufacturing process, the occurrence of surface imperfections is a critical problem [78]. Nevertheless, some conventional detection platforms are often unable to detect micro-defects on micro-parts [79]. In this regard, some researchers have put forth, micro-domain vision detection technologies to acquire and analyze 2D textures and 3D shape information, which effectively solved this problem.

For metallic micro components, Weimer et al. [80] proposed an image acquisition technology based on plenoptic cameras. The design of plenoptic cameras is relatively compact and can easily realize integrated manufacturing. Effective 2D and 3D information can be obtained in one measurement step by using plenoptic cameras to acquire images of micro-components. This method meets the requirements of quality detection in a micro-domain. To realize on-line surface detection, Scholz-Reiter et al. [76] designed an image acquisition system for micro-part surface imperfections using confocal laser microscopy and realized automatic detection of defects. Li et al. [81] designed a quality inspection system by using micro-vision technology to get graphic information of the micro-accessory. In these methods, the micro-domain vision technology played a significant role in the task of acquiring high-resolution images.

3.2.4 Multispectral

In some industrial detection scenarios, it is necessary to select the multiple photosensitive imaging devices for an effective combination based on the wavelength characteristics of the light, so as to fully represent the characteristics of the objects to be detected in the collected images. A multispectral imaging system, can make up for the shortcomings of traditional CCD photosensitive imaging.

Machine vision has great potential for detecting locomotive and rolling stock condition. Multispectral imaging allows recording of physical and thermal conditions, and their correlations. Combining multispectral imaging with machine vision, Hart et al. [82] proposed a multispectral machine vision technology, in which some visible and infrared (thermal) cameras were placed below the track to capture images. This technology can monitor the physical and thermal state of railway equipment more effectively than the existing methods and technologies.

In addition to the above methods, there are high-dynamic range imaging [83, 84] and multi-vision imaging [85] systems, etc. In each specific visual detection project, we need to consider the characteristics and detection requirements of the objects to be tested, to select the appropriate image acquisition method.

4 Image Processing and Defect Detection

Images are the information carriers of machine vision. Image processing and analysis are the key technologies for automatically obtaining an understanding of the images acquired by hardware in vision detection systems [86].

Image processing has a long history of development. In the 1920s, the first image was successfully transmitted using digital compression technology, from London to New York via submarine cables. This was the origin of digital image processing technology [87]. In the early days, simple defect detection could be achieved through primitive filtering methods. For example, in 1973, in an early attempt to apply visual inspection to industrial production, Ejiri et al. [88] described a method that employed two-dimensional nonlinear logical filtering to detect defects in complicated patterns such

as PCBs. It could detect defects in complicated patterns in real time. Subsequently, Hara et al. [89] proposed an algorithm for comparing the local features of the patterns to be inspected with those of a reference pattern, with intended applications to an automatic PCB inspection system.

Currently, with the development of computer technology and mathematical theory, image processing and analysis methods have become more abundant and advanced. Flexible configurations in modern manufacturing systems can allow them to quickly switch from one product to another [90, 91]. For conventional machine learning, complex feature extractors need to be designed for particular cases so that the desired features can be retrieved. In addition, new products may present complex texture patterns or intensity changes, and surface defects can be of any size, direction, and shape. Therefore, manually designed features may lead to insufficient or unsatisfactory inspection performance in complex surface scenarios or dynamic processes. Compared with traditional machine learning, the main advantage of deep learning is that these rich features are not designed by human engineers but are learned automatically through convolutional neural networks from raw data [92]. Deep learning has been proven to be very adept at discovering complex structures in high-dimensional data [93]. Therefore, for defect detection by machine vision systems based on image processing technology, deep learning can play an important role in inaugurating the era of intelligent detection with machine vision.

In industrial production, there are three kinds of representative defect detection tasks based on machine vision: classification, localization and segmentation. Some primitive image preprocessing methods can help the subsequent image analysis, and sometimes may deal with a few simple defect detection tasks. For most defect detection scenarios, more image processing methods are needed to extract enough features for understanding defect information. For image feature learning, the main types of deep learning network architecture include convolutional neural networks (CNNs) [94], deep belief networks (DBNs) [95], and stacked auto-encoders (SAEs) [96]. Furthermore, long short-term memory (LSTM) [97] plays an important role in images with time-sequenced characteristics. DBNs and SAEs can help multi-feature fusion detection achieve better effect and accuracy.

4.1 Image Preprocessing

The purpose of image preprocessing is to enable the machine to understand the image better and prepare for the next step of image analysis [98]. The principle of image preprocessing is to eliminate irrelevant information and recover useful real information. Some factors may cause image noise, such as the field environment of machine vision, photoelectric conversion of the CCD image, transmission circuit, and electronic components. These noises reduce the image quality, which in turn, adversely affects the image analysis. Therefore, denoising is the main objective of image preprocessing.

Image preprocessing generally comprises spatial domain methods and frequency domain methods [86]. The main preprocessing algorithms include grayscale transformation, histogram equalization, various filtering algorithms based on spatial and frequency domains [99, 100], etc. In addition, mathematical morphology can also be used for image denoising [101].

The basic method for conversion from spatial domain to frequency domain is the Fourier transform and the representative tool for image processing in the frequency domain is the wavelet transform.

4.1.1 Fourier Transform

Fourier transform has helped the industry and academia prosper in an unprecedented manner [102]. Before the Fourier transform, image processing was confined to spatial domain operations. The function of various spatial filtering algorithms is to convolute the image with various templates. For example, the direct grayscale transformation transforms each pixel of the image according to a certain function to get the enhanced image. In generally, a spatial filtering algorithm is easy to operate and has high real-time performance; however, it is not suitable for complex image processing.

The Fourier transform can transform the image from the spatial domain to the frequency domain, and its inverse transform can transform the image from the frequency domain back to the spatial domain [103, 104]. For image processing, the two-dimensional discrete Fourier transform (DFT) is represented as:

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-j2\pi(ux/M + vy/N)}, \tag{1}$$

and the inverse discrete Fourier transform (IDFT) is

$$f(x, y) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(ux/M + vy/N)}, \tag{2}$$

where $f(x, y)$ represents a digital image of size $M \times N$, and then the frequency domain representation $F(u, v)$ can be obtained by using DFT formula (1) [87]. In formulas (1) and (2), u ($u = 0, 1, 2, …, M-1$) and v ($v = 0, 1, 2, …, N-1$) represent the frequency domain variables; x ($x = 0, 1, 2, …, M-1$) and y ($y = 0, 1, 2, …, N-1$) represent the space domain variables. In addition, j is an imaginary number, equal to the square root of -1.

Through the Fourier transform, the image can be converted to frequency domain for transformation and

operation. In the frequency domain, the data reflect the intensity of grayscale changes in the image. The frequency domain filtering modifies the Fourier transform of the image and then, calculates its inverse transform to get the processed result. For example, the moving average window filter and Wiener linear filter use a low-pass filter to denoise, based on the premise that noise energy is concentrated in high frequency, and the image spectrum is distributed in a limited range [87]. For noise removal, Bai and Feng [98] introduced a new class of fractional-order anisotropic diffusion equations by using the DFT. Their experiments showed that the proposed equations yielded good visual effects and better SNR on denoising the real images. However, the frequency domain transformation is complex, and the operation cost is high.

4.1.2 Wavelet Transform

In recent years, the wavelet transform has been demonstrated to be a powerful approach for noise reduction and became a prime field of image processing research [105, 106]. The wavelet transform provides the localization analysis of time or space frequency and gradually refines the signal by scaling and translation [107]. The wavelet transform can subdivide time at high frequency and frequency at low frequency, thus automatically adapting to the requirements of the time–frequency signal analysis.

The wavelet transform plays an important role in image processing. Luisier et al. [108] introduced an inter-scale orthonormal wavelet thresholding algorithm. In this method, the denoising process was parameterized to the sum of the basic nonlinear processes with unknown weights, and the mean square error of the denoised image and the clean image was minimized. Jain and Tyagi [109] presented an edge preserving denoising technique based on wavelet transforms. They decomposed the noisy image and improved the denoising performance by clustering. Yan et al. [110] presented a novel wavelet thresholding procedure to suppress the additive Gaussian noises in images. This method effectively overcame the discontinuity of the hard threshold function. For inspection of strongly reflective metal surface defects, Zhang et al. [22] removed the noise effectively from the image by setting certain coefficients to zero by wavelet smoothing. In addition, the wavelet transform has also been widely used in image fusion [111, 112], image coding [107, 113], image compression [114], image encryption [115], and image enhancement [116, 117].

4.2 Classification

Defect classification is usually used to detect whether a certain defect exists in an image. The traditional way to solve the problem of computer vision is to classify the preprocessed images according to hand-crafted features. Most of the research has focused on the construction of hand-crafted features and classification algorithms, and some outstanding work has emerged.

Feature extraction extracts the information that describes the characteristics of the target from the image pixels and then maps the differences between the different targets to a lower-dimensional feature space to help compress the amount of data and improve the recognition efficiency. The common defect features used in visual inspection include greyscale features, shape and size features, and texture features. The greyscale features are the most intuitive features of the image, such as greyscale statistical characteristics and greyscale difference statistical characteristics. Shape and size features are important information for identifying various defects. Common defects can be detected by shape information, such as lines, curves, ellipses and rectangles, and size information, such as area and perimeter. The texture is an important feature of an image. It does not use color or brightness to reflect the homogeneity of images. It represents important information about the arrangement of the surface structures and their relationships with their surroundings [118, 119].

According to the characteristics of the defects, there are many feature extraction methods that can be used for defect classification.

As simple and effective feature descriptors that are based on statistical characteristics, histograms are widely used in the field of computer vision. For example, Li et al. [120] proposed a defect classification algorithm based on histogram features for automatically detecting defects in both nonpatterned and patterned fabrics. Common statistical features of histograms include the maximum, minimum, mean, median, range, entropy, variance, L1 norm, L2 norm, Bhattacharyya distance, and normalized correlation coefficient. The calculations are simple and are invariant in translation and rotation. However, these features reflect only the probability of the greyscale level of the image and not the spatial distribution of the pixels [121, 122].

The grey-level cooccurrence matrix (GLCM) is a common method of describing a texture by studying the spatial correlation properties of the greyscale. It reflects the comprehensive information from the image grey levels regarding the direction, adjacent interval and change amplitude, which can be used to analyze the image primitives and arrangement structure [123–125]. The Gabor transform is a type of windowed short-time Fourier transform. The window function is the Gaussian function. This transform simulates the biological action of human eyes and can extract relevant features in different scales and directions in the frequency domain [126, 127]. Raheja et al. [128] presented a new scheme for an automated fabric defect detection system using the GLCM and Gabor filter method. The experimental results showed

that, compared with the Gabor filter method, the GLCM has greater accuracy and computational efficiency in the same environment.

The local binary pattern (LBP) expresses the relationship between the local neighborhood point and the center point through binary bits [129]. It has strong robustness to changes in the image greyscale level caused by changes in illumination [127, 130, 131]. For fabric defect classification, Zhang et al. [132] proposed an algorithm that combines the LBP and GLCM. The LBP and GLCM are used to extract the local feature information and overall texture information of the defect images, respectively. However, the LBP algorithm constructs a histogram of the defect images based on spatial neighborhood pixel coding, which may result in losing the discrimination information of the defect images.

The scale-invariant feature transform (SIFT) is an image descriptor for image-based matching and recognition [133, 134]. It can achieve reliable feature matching in different perspectives by extracting unique invariant features from images. The extracted features are invariant with respect to the image zoom, the rotation, 3D affine transformations within a certain range, noise superposition and illumination changes. Dunderdale et al. [135] used the SIFT descriptor combined with a random forest classifier to identify defective photovoltaic modules. The SIFT descriptor showed good performances and could be used to both detect and describe local feature points. However, SIFT has high requirements for image quality, which limits its application.

Histograms of oriented gradient (HOG) features are formed by computing statistical histograms of gradient directions in local regions of the image [136]. It can maintain good invariance to geometric and optical deformations of the image. Halfawy and Hengmeechai [137] presented an efficient pattern recognition algorithm that employed the HOG and support vector machine (SVM) to automate the detection and classification of pipe defects. Compared with the LBP, the HOG can more easily extract the edge information and consider the structural information of the image. However, the HOG algorithm may face the problems of having high dimensionality and neglecting the texture information.

Speeded up robust features (SURF) [138], binary robust independent elementary features (BRIEF) [139], and oriented FAST and rotated BRIEF (ORB) [140] are also used in feature extraction. Furthermore, there are many variations of the classical method; for example, the LBP family includes the completed local binary pattern (CLBP) [141], elliptical local binary pattern (ELBP) [142], adjacent evaluation completed local binary pattern (AECLBP) [5], and robust local binary pattern (RLBP) [143]. Based on these classical algorithms, some novel feature-extraction algorithms have also been proposed in recent years; for instance, Zhao et al. [144] proposed a discriminant manifold regularized local descriptor (DMRLD) algorithm for steel surface defect classification. Compared with hand-crafted histograms, DMRLD achieves better robustness by using the structure of a manifold with a learning mechanism to represent the information contained in the image.

There are many kinds of feature extraction methods with their own advantages and disadvantages. For specific visual inspection items, we should consider whether the feature extraction method makes full use of the global information, whether its calculations are convenient, whether it can meet the real-time needs, etc. For many application requirements, using a combination of multiple feature extraction methods is also a good way to increase efficiency and accuracy.

To identify the defect categories of an image, it is necessary that the selected features not only describe the image properly but also distinguish different categories of images. The primary mission of defect classification is to train the classifier according to the extracted feature set and then make it identify the type of each surface defect correctly based on supervised or unsupervised pattern recognition methods.

The support vector machine (SVM) [145] and K nearest neighbor (KNN) [146] are representative classifiers in supervised pattern recognition.

SVMs are suitable for small and medium-sized data samples and for nonlinear, high-dimensional classification problems, and they have been widely used in the field of industrial vision detection. For example, Jia et al. [147] described a real-time machine vision system that uses an SVM to automatically learn complicated defect patterns. Li and Huang [148] proposed a binary defect pattern classification method that combines a supervised SVM classifier with unsupervised self-organizing map clustering, in which the SVM is used to classify and identify manufacturing defects. The results showed that this method could achieve more than 90% classification accuracy, which was better than that of the back-propagation neural network. However, this study focused only on binary map classification. Valavanis and Kosmopoulos [149] proposed a method of multi-class defect detection and classification based on a multi-class SVM and a neural network classifier for weld radiographs. For real-time analysis of spectrum data, Huang et al. [150] established an improved SVM classification model based on a genetic algorithm to accurately estimate different types of porosity defects in an aluminium alloy welding process. Furthermore, the SVM classifier has played a significant role in the inspection of surface defects in copper strips [151, 152], laser welding process monitoring and defect diagnosis [153], defect detection for wheel bearings [154], etc.

The KNN algorithm has been proven to be simpler and more stable than neural networks [155, 156]. To detect fabric defects, Yıldız et al. [157] preprocessed images with wavelet, threshold, and pathological operations and then used the

GLCM method to extract features. Finally, defect images were classified based on a KNN algorithm with an average accuracy rate of 96%. Cetiner et al. [158] proposed a method of feature extraction based on the wavelet moment and defect image classification based on KNN, which can be used in automatic defect classification systems in the forest industry. Das and Jena [159] presented a method combining image texture feature extraction techniques. First, LBP and the grey level run length matrix (GLRLM) were combined to extract image features, and then KNN and an SVM were used for classification. The experimental results showed that the combination of LBP and GLRLM can improve the performance of feature extraction, and the SVM has better classification performance than the nearest neighbor approach in texture feature classification. Therefore, Lei and Zuo [156] proposed a weighted K nearest neighbor (WKNN) algorithm based on the two-stage feature selection and weighting technique (TFSWT) to improve the performance of the KNN algorithm, and they successfully applied the WKNN method to identify gear cracks.

An unsupervised algorithm can also be used for defect classification. Based on K-means clustering, Mjahed et al. [160] presented an efficient algorithm for solving a multi-objective fault signal diagnosis problem using a genetic algorithm. Hamdi et al. [161] introduced an unsupervised defect detection algorithm for patterned fabrics. An image filtered by non-extensive standard deviation was divided into a series of blocks, and then the squared difference between each block median and the mean of all block medians was input into K-means clustering to classify the blocks as defective or non-defective, with an overall detection success rate that reached 95%.

Table 2 compares some traditional feature extraction and defect classification methods.

In recent years, artificial intelligence technology has greatly benefited industrial production. Neural networks are an important branch in the development of artificial intelligence [162]. With the improvement of computing power and the advent of big data, deep learning, with the core idea that machines can automatically learn from data by increasing the number of network layers, has developed rapidly and has significantly impacted the field of machine vision. Deep learning methods can automatically extract and combine the essential feature information of objects, and they are especially adept at image classification.

The CNN is the most popular architecture for image classification. In 1998, the emergence of LeNet opened the era of CNNs [94]. In 2012, the success of AlexNet [163] in the ImageNet competition promoted the application of deep learning in computer vision. After that, a series of CNN models appeared, such as Network-in-network [164], VGGNet [165], GoogLeNet [166–169], ResNet [170], and DenseNet [171]. There are three main types of neural layers that play different roles in a CNN: convolutional layers, pooling layers, and fully connected layers [172, 173]. The convolutional layers are designed to detect local combinations of features from a previous layer, pooling layers are designed to merge semantically similar features into one, and fully connected layers ultimately convert the feature maps into a feature vector [174], as shown in Fig. 6.

Table 2 The performances of traditional feature extraction and defect classification methods

| Methods	Performances
Histograms	The calculations are simple and are invariant in translation and rotation. However, these features reflect only the probability of the greyscale level of the image and not the spatial distribution of the pixels
GLCM	It reflects the comprehensive information of image gray levels about direction, adjacent interval and change amplitude, which can be used to analyze image primitives and arrangement structure
Gabor transform	It simulates the biological action of human eyes and can extract relevant features in different scales and directions in the frequency domain
LBP	It has strong robustness to changes of image grayscale level caused by changes of illumination. It is often used to extract local texture features
SIFT	It could be used to both detect and describe local feature points. However, it has high requirements for image quality, which limits its application
HOG	It can more easily extract the edge information and consider the structural information of the image. However, it has poor real-time performance and is sensitive to noise
SURF	Compared with SIFT, its computing speed is improved
BRIEF	Great real-time performance, but rotation invariance needs to be improved
ORB	It has fast speed and good anti rotation ability. However, the scale registration error is large and the registration rate is low
SVM	It has fast speed, strong generalization ability, and the results are easy to explain. The disadvantage is that it is not suitable for multi classification, sensitive to missing data, and sensitive to the selection of parameters and kernel function
KNN	It is simple, easy to understand and implement without training. But the computation is too large and the robustness is poor
K-means clustering	It is simple in principle and easy to implement. However, the parameters need to be set manually. It is sensitive to noise and outliers
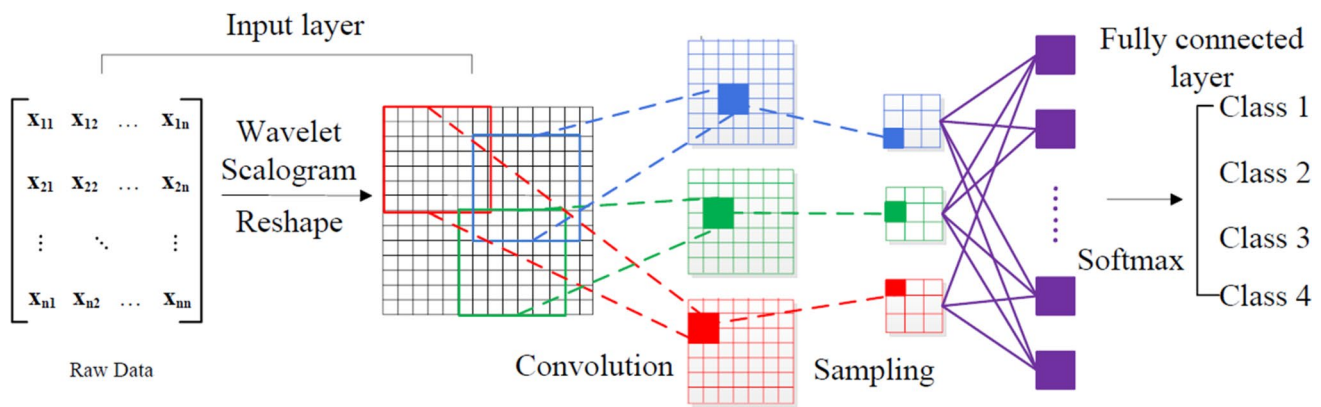
Fig. 6 Architecture of a CNN model [174]. Copyright 2018, Elsevier

The CNN was originally designed for image analysis; therefore, it is a good fit for automated defect classification in visual inspection [175–177]. According to the relevant literature in recent years, the application of deep learning in industrial defect classification involves many fields, such as industrial production and electronic components. For supervised steel defect classification, Masci et al. [178] presented a max-pooling CNN approach. Compared to SVM classifiers, the CNN obtains much better results and can work properly with different types of defects. The surface quality affects not only the appearance of products but also their performance. Park et al. [14] proposed a generic approach based on a CNN for the automatic visual inspection of dirt, scratches, burrs, and wears on part surfaces. Their results showed that a pretrained CNN model works well on small datasets with improved accuracy for a surface quality visual inspection system. To detect casting defects by X-ray inspection, Lin et al. [179] proposed a robust detection method based on a visual attention mechanism and feature-mapping deep learning and established a CNN to extract defect features from potentially defective regions and obtain a deep learning feature vector. Then, the similarity of suspicious defective regions could be calculated by using the feature vector. Their results showed that the method was effective in solving the problem of false and missing inspections. Nguyen et al. [180] proposed an inspection system based on a CNN to achieve defect classification in casting products. However, the CNN deep learning model can only perform well under the condition of having a large number of high-quality datasets. Kim et al. [181] proposed an indicator that can distinguish between defects and the background area for the classification of defect types in thin-film-transistor liquid–crystal display panels. For the process of industrial production, automatic defect classification was performed based on a CNN.

As one of the representative algorithms of machine vision, the CNN has played an important role in defect classification. However, CNNs are becoming increasingly deep, and they require large-scale datasets and massive computing power for training. In addition, collecting labelled datasets requires great human effort. Thus, as a further exploration, unsupervised learning by a CNN may be a meaningful research direction.

Transfer learning is a method of machine learning in which a pre-trained model is reused in another task. Transfer learning can help solve the problem of a lack of labelled data. Imoto et al. [182] proposed a CNN-based transfer learning method for automatic defect classification. The results showed that this method is robust against a lack of labelled data and can achieve more than 80% accuracy with only a few dozen labelled data points.

4.3 Localization

Defect localization needs to accurately determine the location of the defect in a given image and mark the defect category. Generally, defect localization is performed by a series of object detection methods.

The traditional object detection strategies and algorithms include Viola-Jones [183], HOG + SVM, non-maximum suppression (NMS) [184], the deformable part model (DPM) [185], selective search [186, 187], and edge boxes [188]. Ding et al. [189] proposed a detection scheme based on a HOG and SVM. The HOG was used to encode each block-based feature, and the SVM was used to classify the fabric defects. The experimental results showed that this method based on a HOG and SVM is relatively simple and easy to realize in online applications. Dou et al. [190] proposed a fast template matching-based algorithm (FTM) for railway bolt detection and a nearest-neighbor classifier to determine whether a bolt is in the correct position, which achieved a lower false positive rate than previous methods. The DPM is one of the most effective template-based approaches used in object detection. For railway fastener defect detection, He

et al. [191] proposed a Gaussian mixture deformable part model (GMDPM) algorithm based on HOG features. Wei et al. [192] proposed an effective express box defect detection algorithm to identify the shape and size of defects, and this method achieved a 95.83% correct rate.

In recent years, after the successful application of CNN-based image classification methods, object detection technology based on deep learning has also made significant progress. The object detection methods based on deep learning can be divided into two major categories. One generates regions and then classifies each region to obtain different object categories. The other regards object detection as a regression or classification problem and uses a unified framework to obtain the final categories and locations directly [193]. The region proposal-based methods mainly include regions with CNN features (R-CNN) [194], spatial pyramid pooling (SPP-net) [195], Fast R-CNN [196], Faster R-CNN [197], region-based fully convolutional networks (R-FCNs) [198], feature pyramid networks (FPNs) [199], and Mask R-CNN [200]. The regression- and classification-based methods mainly include MultiBox [201], AttentionNet [202], G-CNN [203], You Only Look Once (YOLO) [204], the single-shot MultiBox detector (SSD) [205], YOLOv2 [206], RetinaNet [207], YOLOv3 [208], and YOLOv4 [209]. In terms of performance, the region proposal-based methods are high in accuracy but low in speed; the regression- and classification-based methods are high in speed but low in accuracy.

Based on a cascaded mixed FPN, Wu et al. [210] proposed a two-stage fabric defect detector. The end-to-end defect detection architecture is shown in Fig. 7. The feature extraction backbone model of matching parameters with fitting degrees was proposed to solve the problems caused by a small defect feature space and background noise. Stacked feature pyramid networks were set up to integrate cross-scale defect patterns for feature fusion and enhancement in a neck module. Cascaded guided region proposal networks (RPNs) were proposed for refining the anchor centers and the shapes used for anchor generation. The experimental results showed that this method could improve the recognition performance of included and size-variant fabric defects.

Faster R-CNN is a state-of-the-art method for detecting objects with real-time object detection, which can generate

regions of interest (ROIs) with an RPN instead of selective search [197, 211]. Lei et al. [211] adopted Faster R-CNN to implement the detection of defects in the polarizer and to perform the rapid detection and effective positioning of defects. To further improve the detection accuracy and efficiency, the number of layers of the network could be changed, and some of the network parameters should be adjusted to optimize the test model. Lei and Sui [212] proposed a Faster R-CNN method to perform intelligent fault detection for high voltage lines. To detect defects in an image, Faster R-CNN chooses a random region as the proposal region and then obtains the corresponding category and location of a certain component after training. The experiments showed that the detection method based on the ResNet-101 network model could effectively locate insulator damage and bird nests on a high voltage line. Sun et al. [213] proposed an improved Faster R-CNN method for surface defect recognition in wheel hubs. The last maximum pooling layer was replaced by an ROI pooling layer, as shown in Fig. 8. ROI pooling technology was used in order to employ a single feature map for all the proposals generated by the RPN in a single pass. It enabled object detection networks to use an input feature map with a flexible size and output a fixed-size feature map. The experimental results showed that the improved Faster R-CNN method has a higher detection accuracy. However, the detection speed of the Faster R-CNN method may not meet the real-time requirements of industrial applications.

YOLO is an object recognition and location algorithm based on a deep neural network that performs object detection by using fixed-grid regression [214]. Its primary characteristic is that it runs quickly and can be used in real-time systems. Based on the idea of regression, YOLO takes a whole image as the input of the network and directly regresses the object border and the category of the object in multiple positions of the image. Adibhatla et al. [215] adopted a YOLO/CNN model to detect PCB defects and achieved a defect detection accuracy of 98.79%. However, the defect types that can be detected by the method are limited and need to be optimized. Lv et al. [216] proposed an active learning approach for steel surface defect inspection based on YOLOv2. This model achieves high efficiency but at the expense of precision.

Fig. 7 End-to-end fabric defect detection architecture [210]
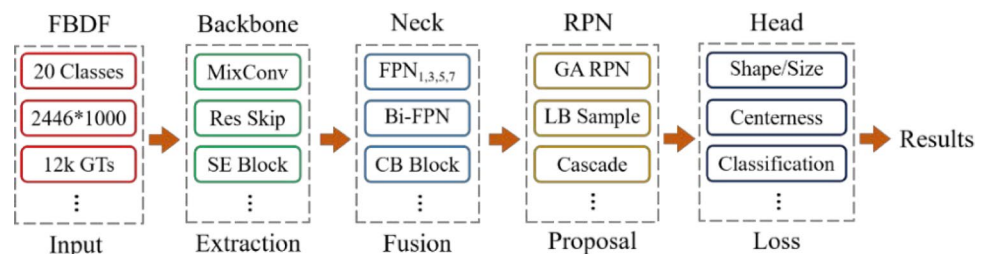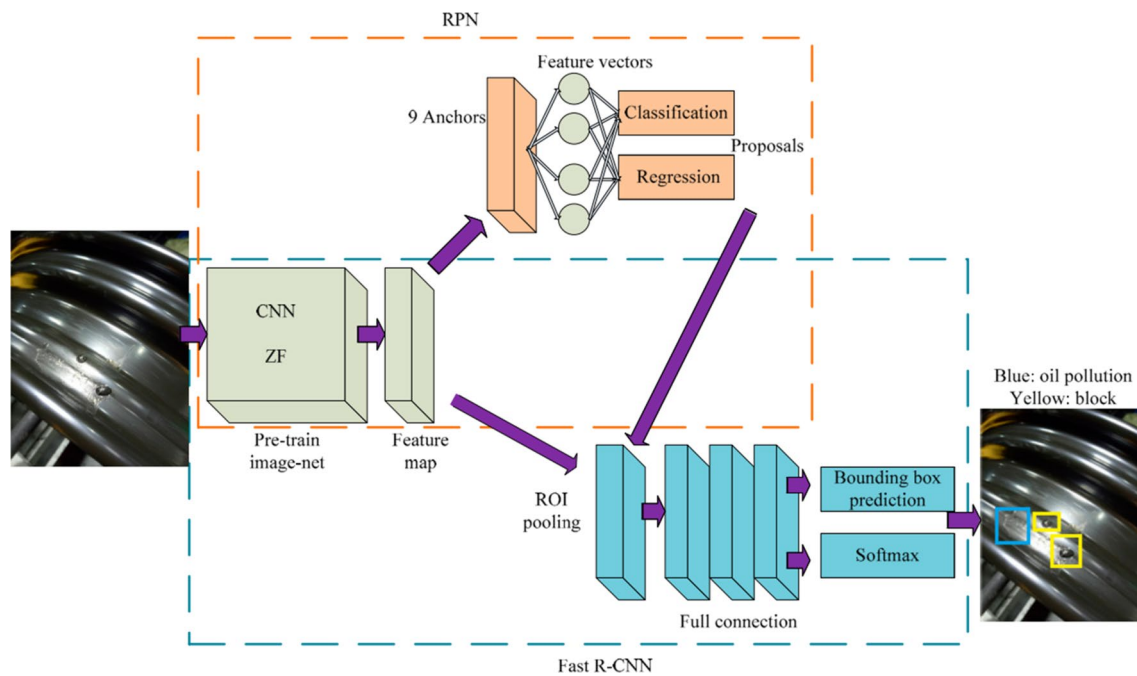
Fig. 8 The structure of the improved Faster R-CNN [213]

Jing et al. [217] proposed an improved YOLOv3 model by using the K-means algorithm to cluster the marker data. The experimental results showed that the improved YOLOv3 model achieves better performance in fabric defect detection. However, the real-time performance needs to be improved. As a regression-based detection method, the YOLOv4 network has an excellent detection speed. However, the detection accuracy for small targets needs to be improved. To detect iron material cracks, Deng et al. [218] proposed a cascaded YOLOv4 (C-YOLOv4) network. The experimental results showed that C-YOLOv4 has better robustness and crack detection accuracy.

SSD combines some strategies of YOLO and Faster R-CNN, and it uses multi-scale regional features for regression, which not only maintains the high speed of the YOLO method but also ensures a certain accuracy of performance. Zhai et al. [219] proposed a DF-SSD object detection method based on DenseNet and feature fusion. The feature extraction network DenseNet-S-32-1 was designed to replace VGG-16 in SSD. To effectively integrate low-level visual features and high-level semantic features, they also designed a fusion mechanism for multi-scale feature layers. The experimental results showed that the proposed DF-SSD method could achieve an advanced performance in the detection of small objects and objects with specific relationships.

4.4 Segmentation

Defect classification and localization can provide information on the defect types and their relative positions in images. Furthermore, in intelligent vision detection, defect segmentation, especially pixel-level segmentation, can provide important references for evaluating the defect severity and performing condition assessment.

Image segmentation is a process that divides an image into several specific and unique regions and proposes objects of interest [220]. The purpose of image segmentation is to predict the category of each pixel in the image. To solve the problem of image segmentation for different features, researchers have proposed numerous segmentation methods. Table 3 lists some traditional image segmentation methods and their characteristics.

These methods are based on different image models, use different characteristics, and have a certain scope of application. Some researchers have also integrated genetic algorithms [233] and wavelet methods [234] into image segmentation and have achieved positive results. Among these methods, the clustering algorithm is widely used for defect segmentation. The clustering algorithm is an unsupervised algorithm that does not require a training set. It is simple and fast. Image segmentation divides the image into several disjoint regions, which is a pixel clustering process [235]. There are many clustering algorithms, such as

Table 3 Traditional image segmentation methods and their characteristics

Classifications	Methods	Illustrations	Characteristics
Region-based	Threshold segmentation [221, 222]	The image pixels are divided into several categories by setting different grayscale thresholds	It is simple; it has less calculations, stable performance, and is widely used. It is critical in selection of a threshold
	Regional growth [223]	Based on certain growth criteria, several similar sub regions are aggregated into larger regions	It is simple and has a good segmentation effect for even connected objects. However, it depends on the selection of initial conditions and the calculation is large; therefore, it is not suitable for real-time detection
	Clustering [224, 225]	Based on clustering, the feature space is segmented and then mapped to the original image space to get the segmentation results	Its advantage is that it has good flexibility and no need for training sets. However, sometimes the results depend on the selection of some empirical parameters
	Splitting and merging [226]	Based on the non-uniformity of a region, it is divided into new sub regions; then, regions with the same content are combined into new larger regions, and finally the segmented image is obtained	The key is the design of split and merge criteria. This method has a good effect on the segmentation of complex images; however, the algorithm is more complex, the calculation is large, and the segmentation may also destroy the boundary of the region
Edge-based	First-order differential operator [227, 228]	Based on the discontinuity of the local characteristics in the image, an algorithm is used to extract the boundary between the object and the background. It Includes Robert, Sobel, Prewitt, and Kirsch operators	The calculation is simple and fast; however, the positioning is not accurate
	Second-order differential operator [227, 228]	The edge is located by finding the zero-crossing point of the second derivative of the image gray-scale, using operators, such as Canny, Log, and Laplacian operators	The method is accurate in edge location; however, it is sensitive to noise
Specific-theory-based	Wavelet transform [229, 230]	The main advantage of multi-scale edge detection method is that it detects the edge accurately at a larger scale and locates the edge points accurately at a smaller scale	Wavelet transform can focus on any detail of the objects and has good localization property. However, there is still a contradiction between the noise resistance and the detection accuracy in edge detection
	Mathematical morphology [231, 232]	To achieve the purpose of image analysis and recognition, the corresponding shape in the image is extracted by structural elements with certain shape	The selection of structural elements is flexible; however, at the same time, it also leads to poor adaptability of the algorithm

fuzzy c-means (FCM) [236], BIRCH [237], CURE [238], CLARANS [239], K-means [240], CLARA [241], CHAMELEON [242], K-medoids [242][242], DBSCAN [244], K-prototypes [245], and MAPK-means [246]. The choice of clustering algorithm depends on the purpose of clustering and the type of data. Xiong et al. [247] proposed a novel 3D laser profiling system for rail surface defect detection. In the process of rail surface defect detection and classification, K-means clustering was used to merge the candidate defect points into candidate defect regions. Jian et al. [248] designed a surface defect detection system for mobile phone screen glass. In this system, improved fuzzy c-means (FCM) clustering was proposed to segment the surface defects more accurately. Melnyk and Tushnytskyy [249] proposed a PCB defect detection and classification system that implemented the K-means clustering algorithm. Li et al. [250] proposed a clustering algorithm that links the regions that are close to each other to detect cluster defects composed of many small point defects. The schematic diagram of the process of connecting domains A and B in the clustering method is shown in Fig. 9.

Deep learning has also brought great progress for image segmentation technology. The fully convolutional network (FCN) is a breakthrough semantic segmentation model that has higher accuracy than traditional approaches [251]. FCNs can efficiently learn to make dense predictions for per-pixel tasks, for example, semantic segmentation, as shown in Fig. 10.

FCN-based segmentation methods also play an important role in industrial applications. Yu et al. [253] presented a novel 2-stage FCN framework for surface defect segmentation. The 2-stage framework improves the generality and reusability of FCNs. Li et al. [254] adopted region-based fully convolutional networks (R-FCNs) to inspect insulator defects. The experimental results showed that the R-FCN algorithm has good robustness and environmental adaptability. In crack inspection, conventional approaches are unable to identify and measure diverse types of cracks concurrently at the pixel level. Yang et al. [255] applied an FCN to study automatic pixel-level crack detection and measurement, and their results showed that the prediction had improved at the pixel level and that the training time was greatly reduced. However, the resolution of the feature maps generated by

the FCN was low, and the prediction results were coarse owing to the large amount of spatial information loss during down-sampling. Qiu et al. [256] presented a 3-stage FCN for pixelwise surface defect segmentation. The FCN is a state-of-the-art algorithm for generic object segmentation. However, for small datasets, its performance cannot meet the requirements. The experimental results showed that the slicing method could improve the efficiency of FCNs in small datasets in industrial environments.

The current common image segmentation algorithms, in addition to FCNs, include U-Net [257], SegNet [258], Mask R-CNN, and PSPNet [259]. These models have an encoder-decoder architecture, where a CNN is used as an encoder to extract features, and a deconvolution network and skip connections are used as decoders to map features to the output image. U-Net was originally proposed to segment the grey-scale of biomedical images. SegNet achieves a good trade-off between efficiency, the memory footprint and precision. The Mask R-CNN can be used for instance segmentation. PSPNet adopts a pyramid pooling module structure, which can extract smaller and more localized features, while the large-size layers can extract global information.

Furthermore, according to the specific visual inspection application scenario, additional defect segmentation methods have been continuously proposed. For example, Yu et al. [260] proposed an adaptive depth and receptive field selection network. In this method, an adaptive depth selection mechanism was designed to extract features of various depths, and an adaptive receptive field block was proposed to select the best acceptance domain. The experimental results for a casting defect segmentation dataset showed that the proposed method achieved better performance than the existing segmentation algorithms. Tabernik et al. [261] proposed a segmentation-based deep-learning architecture for surface defect detection. The network architecture was designed in two stages, as shown in Fig. 11. In the first stage, a segmentation network was used to locate the surface defects accurately at the pixel level. After defect segmentation, each pixel was trained as an independent sample, which increased the effective number of training samples. Then, the second stage was a decision network for binary image classification. The experimental results showed that this method could complete training on a small-scale defect sample dataset,

Fig. 9 The process of connecting domains A and B. Two separate domains A and B are shown in (**a**). The result of the Boolean union of domains A′ and B′ is shown in (**b**). The result of linking domains A and B is shown in (**c**) [250]. Copyright 2020, Elsevier
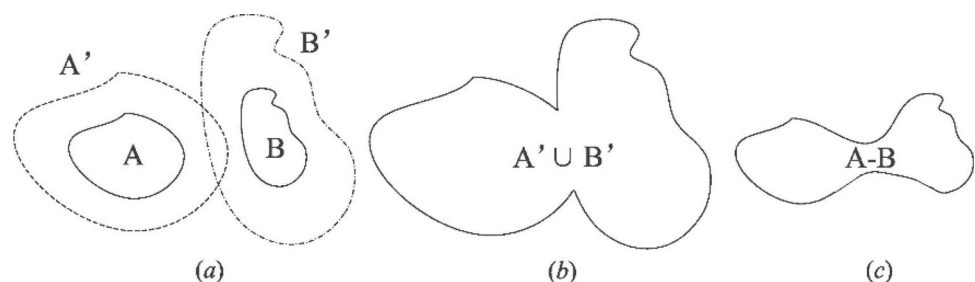
Fig. 10 Fully convolutional
network [252]
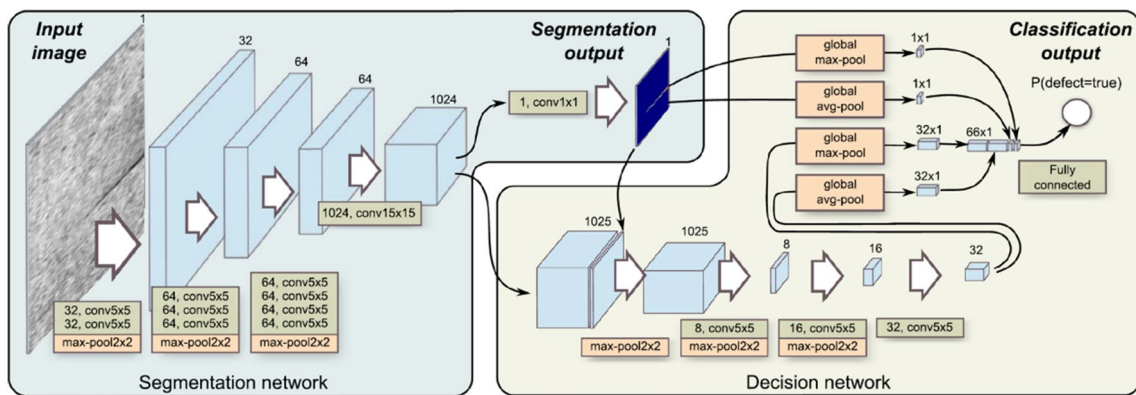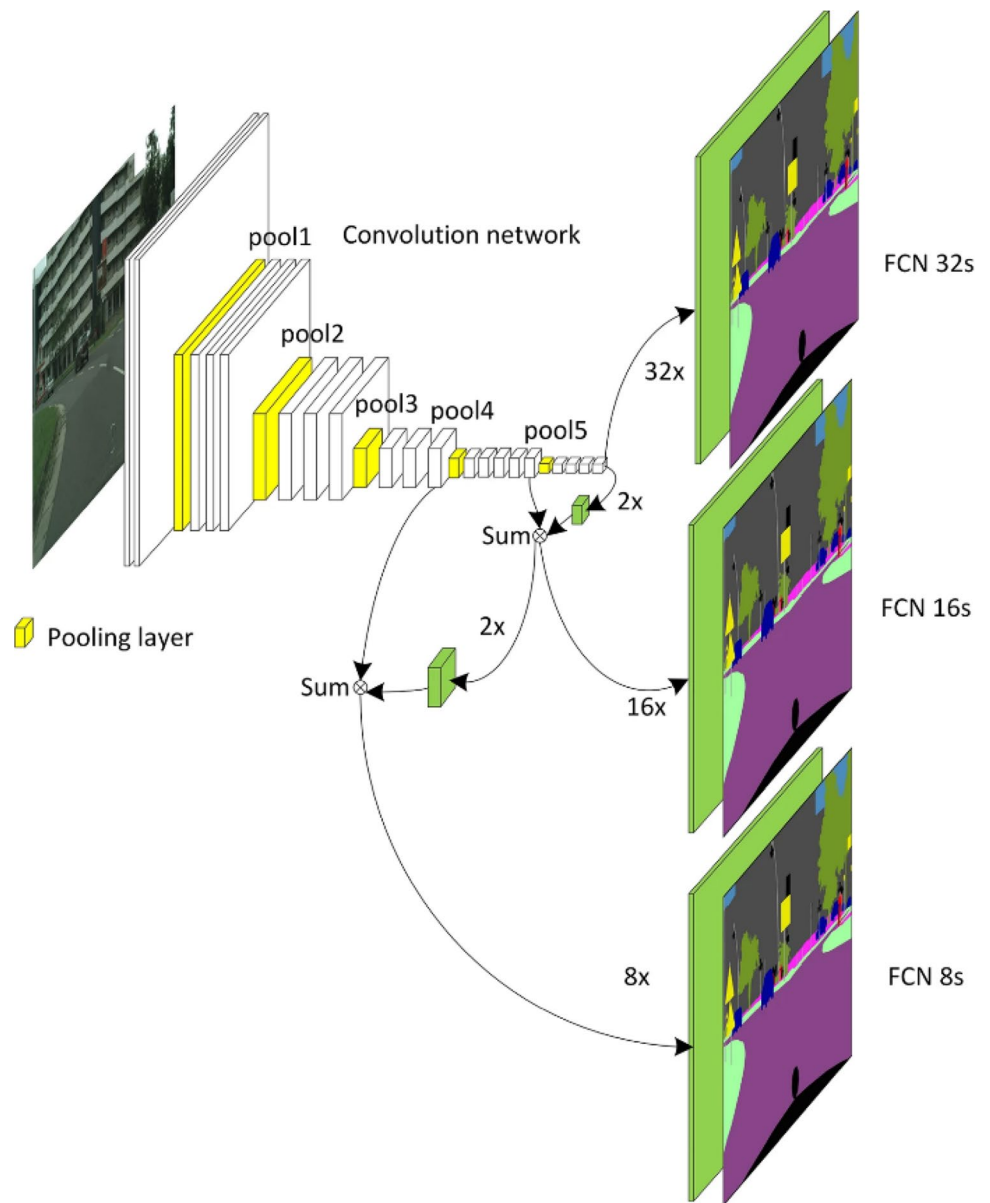
Fig. 11 Two-stage architecture with segmentation and decision networks [261]. Copyright 2020, Springer Nature

which needed only 25–30 training samples. This has great significance for some industrial application scenarios with limited training samples, and this method effectively improves the practicability of deep learning methods.

4.5 LSTM-Based Periodic Defect Recognition

As a deep learning architecture specifically designed for time-series forecasting, the RNN shares parameters among all time steps to learn the information that has been repeated in the past [262]. LSTM is one of the representative architectures in RNNs [97]. In industrial visual inspection, LSTM is an effective method for defects with strong time-sequenced characteristics.

Hu et al. [262] proposed an LSTM recurrent neural network (LSTM-RNN) model to classify common defects in an infrared thermography-based nondestructive testing task for honeycomb materials. Similarly, Wang et al. [263] adopted the LSTM-RNN method to determine the defect depth inside carbon fiber reinforced polymer structures, achieving better performance than a CNN.

A fusion algorithm of a CNN and LSTM is also a widely used defect detection method. For a molten pool online monitoring task, Liu et al. [264] proposed a CNN-LSTM algorithm combining the advantages of a CNN and LSTM. First, feature vectors were extracted from molten pool images through the CNN, and then LSTM was used for welding defect recognition. The experimental results showed that the accuracy of the CNN-LSTM algorithm could reach 94% in

the defect detection task for the CO_2 welding molten pool described in the literature and that it had high efficiency (the time consumption of each image was 0.067 ms), which fully met the industrial requirement of real-time monitoring.

According to the features of periodic roll mark defects in plates, Liu et al. [265] proposed a defect detection method based on a hybrid CNN and LSTM. To improve the detection performance, an attention mechanism algorithm was also integrated into the detection method. The complete network architecture is shown in Fig. 12. As the final output, O represents whether there is a periodic defect in the image sequence. The experimental results showed that the detection method had good performance in identifying periodic defects and that it had an 86.2% detection rate under the experimental conditions described in the literature. However, the integration of the attention mechanism increases the complexity of the algorithm and requires higher computer performance.

4.6 Multi-Feature Fusion Detection Based on a DBN and SAE

The SAE is an unsupervised pre-training method that encodes the input data from a high-dimensional space into a low-dimensional space and then decodes the low-dimensional space data into a high-dimensional space stack by stack [266, 267]. Seker and Yuksek [268] performed fabric defect detection based on the SAE method. After fine-tuning the hyper-parameters of the deep learning
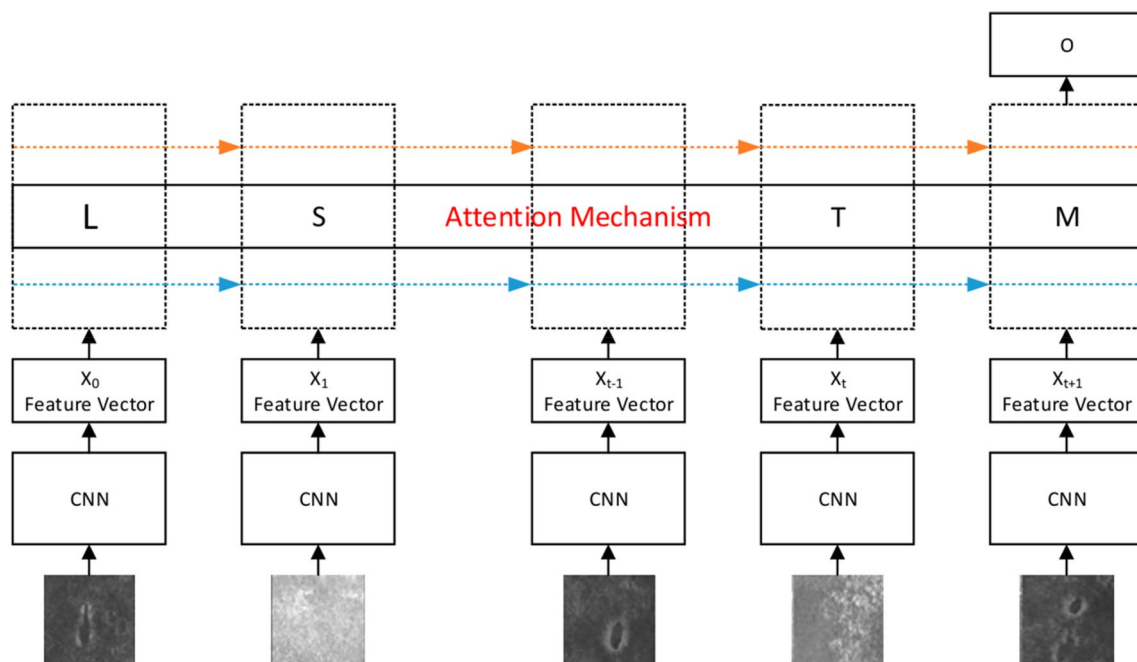

Fig. 12 Network architecture of CNN + LSTM + attention algorithm [265]

model, they achieved a detection rate of 96% on their own datasets. Yang and Jiang [267] proposed a unified deep neural network with multi-level features for weld defect classification. To detect weld defects from radiographic images, they investigated SAEs for pre-training and fine-tuning strategies. As a kind of unsupervised pre-training algorithm, SAEs can improve the generalization performance and reduce the possibility of overfitting, as shown in Fig. 13. The results show that a unified deep neural network can take full advantage of the multi-level features extracted from each hidden layer.

The DBN utilizes the restricted Boltzmann machine (RBM) as a learning module [269]. In a DBN, the top two layers form an undirected graph, and the remaining layers form a belief network with directed, top-down connections [173]. A graphic depiction of a DBN is shown in Fig. 14.

Chen et al. [270] constructed a DBN-based fault softmax classifier for bearing fault classification. A DBN can be used to automatically classify raw data into corresponding classes. Furthermore, a new multi-sensor feature fusion method for bearing fault diagnosis based on a DBN and SAE was proposed [271]. In this study, the SAE extracted features from multiple sensors and merged them into one stream. Then, the features fused by the SAE were used to train a DBN for fault diagnosis and classification. The experimental results showed that this SAE-DBN method could effectively identify the machine running conditions.

With the advent of Manufacturing III [272], defect detection will develop from image-based detection to comprehensive detection combined with multiple sensors. Deep learning architectures such as DBNs and SAEs can help multi-feature fusion detection achieve better effect and accuracy, which is worthy of further research.
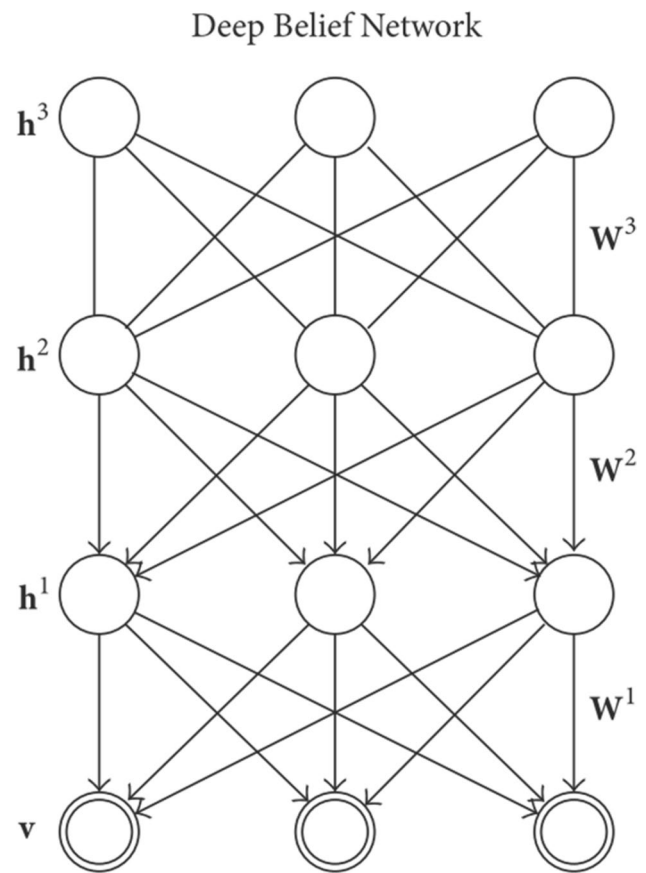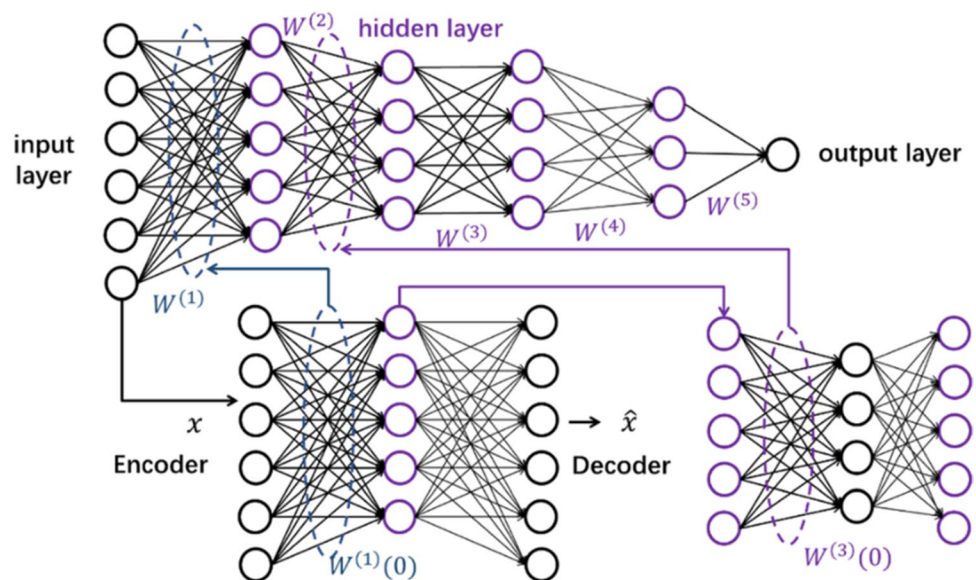

Fig. 14 Graphic depiction of a DBN [173]

Fig. 13 Stacked auto-encoders for deep neural network pre-training [267]. Copyright 2020, Springer Nature

5 Conclusions and Perspectives

Machine vision has significantly improved the scope, efficiency, quality, and reliability of industrial inspection, which has ushered in a series of achievements that cannot be ignored in contemporary industry. However, there are further explorations to be carried out in the application of machine vision.

First, machine vision is a type of real-time in-line detection, which involves large amounts of data, redundant information, and a high-dimensional feature space. The image processing speed is one of the main bottlenecks influencing the real-time performance of vision systems. It is still difficult to achieve real-time in-line detection for objects with complex shape features.

The second issue is the anti-interference aspect of vision detection systems. Visual inspection should be capable of increasing the robustness of detection to reduce the dependence on the image acquisition environment.

The intelligence level of the vision detection system is another bottleneck; whereas a complex interference environment can be identified manually at a glance, it is difficult for a machine to do the same, and it may even make an incorrect judgement.

Although machine vision technology may not be perfect, defect detection based on machine vision is still the main direction for future research and development in this area. Therefore, some important points need to be considered in future development.

5.1 Robust General Algorithm for Balancing Efficiency and Precision

Artificial intelligence, represented by deep learning, has become an important area in industry as a result of the rapid technological developments in recent years. Deep learning marks a significant milestone in visual inspection. Many algorithms can be employed to achieve high accuracy but cannot be used for real-time online detection. In contrast, some algorithms are very fast but cannot reach the ideal accuracy. In addition, some algorithms can work for the detection of products in experimental cases but may not suit practical production. Therefore, it is a meaningful research direction to study a robust algorithm that achieves both efficiency and accuracy.

Moreover, most of the deep learning algorithms are heavily dependent on large-scale sample datasets, which has become a major factor that limits the application of these methods in some areas. Transfer learning is most effective when the source network has been trained with data that is similar to the target network [273]. Weak supervised learning, including incomplete supervision, inexact supervision, inaccurate supervision, or even unsupervised learning, will be an effective way to solve the problem of expensive data acquisition [274–276].

5.2 Fusion of Multiple Detection Technologies

Visual inspection is an image-based detection technology that is mainly aimed at the surface of objects. However, in many cases, industrial inspection concerns not only the surface but also the performance of the whole object.

In the era of Industry 4.0, in order to make the machine more intelligent, comprehensive sensing detection technology should be further studied [277]. Visual inspection can be combined with micro-thermal sensors [278], ultrasonic guided waves [279], eddy current detection [280], laser scanning thermography [281], etc., to achieve a full range of inspection and evaluation of objects.

5.3 Real-Time Performance

Machine vision is mainly used in industrial production line, which requires real-time processing ability. The amount of data involved in visual inspection is very large. However, image processing requires time, thereby leading to a lag in the entire system. The main difficulty in developing real-time detection is the speed of image processing.

Image processing and analysis algorithm should be further optimized to improve the speed of visual inspection system, which is the key technology for further research in the future. Of course, excellent hardware facilities are also very important, such as high-performance computers. Predictably, with the development of 5G communication technology with low delay, using network to upload image data to some powerful cloud servers for processing is also a worth solution [282]. In addition, the next generation of computing technology, represented by quantum computing [283], is expected to contribute fast computing capabilities to the process of visual inspection.

5.4 Extreme Small-Scale Visual Inspection

Manufacturing III takes atomic and close-to-scale manufacturing (ACSM) as the core technology and has become the primary future development trend in manufacturing [79, 272, 284]. To develop ACSM, defect detection will be a very important area. As an example, a neural network can look through a microscope on a sample surface and return information about the atomic structure and lattice defects in real time. On an atomic scale, the size of the datasets grows exponentially; therefore, the application of deep learning could be an effective approach to making great breakthroughs.

References

1. Wang, T., Chen, Y., Qiao, M., & Snoussi, H. (2018). A fast and robust convolutional neural network-based defect detection model in product quality control. *The International Journal of Advanced Manufacturing Technology, 94*(9–12), 3465–3471.
2. Liao, Z., Abdelhafeez, A., Li, H., Yang, Y., Diaz, O. G., & Axinte, D. (2019). State-of-the-art of surface integrity in machining of metal matrix composites. *International Journal of Machine Tools and Manufacture, 143*, 63–91.
3. Kim, D. H., Kim, T. J., Wang, X. L., Kim, M., Quan, Y. J., Oh, J. W., et al. (2018). Smart machining process using machine learning: A review and perspective on machining industry. *International Journal of Precision Engineering and Manufacturing-Green Technology, 5*(4), 555–568.
4. Bulnes, F. G., Usamentiaga, R., Garcia, D. F., & Molleda, J. (2016). An efficient method for defect detection during the manufacturing of web materials. *Journal of Intelligent Manufacturing, 27*(2), 431–445.
5. Song, K., & Yan, Y. (2013). A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects. *Applied Surface Science, 285*, 858–864.
6. Zhang, H., Shen, X., Bo, A., Li, Y., Zhan, H., & Gu, Y. (2017). A multiscale evaluation of the surface integrity in boring trepanning association deep hole drilling. *International Journal of Machine Tools and Manufacture, 123*, 48–56.
7. Rao, X., Zhang, F., Lu, Y., Luo, X., & Chen, F. (2020). Surface and subsurface damage of reaction-bonded silicon carbide induced by electrical discharge diamond grinding. *International Journal of Machine Tools and Manufacture, 154*, 103564.
8. Huang, S. H., & Pan, Y. C. (2015). Automated visual inspection in the semiconductor industry: A survey. *Computers in industry, 66*, 1–10.
9. Ravimal, D., Kim, H., Koh, D., Hong, J. H., & Lee, S.-K. (2020). Image-based inspection technique of a machined metal surface for an unmanned lapping process. *International Journal of Precision Engineering and Manufacturing-Green Technology, 7*(3), 547–557.
10. Mital, A., Govindaraju, M., & Subramani, B. (1998). A comparison between manual and hybrid methods in parts inspection. *Integrated Manufacturing Systems, 9*(6), 344–349.
11. Malamas, E. N., Petrakis, E. G., Zervakis, M., Petit, L., & Legat, J.-D. (2003). A survey on industrial vision systems, applications and tools. *Image and Vision Computing, 21*(2), 171–188.
12. Kopardekar, P., Mital, A., & Anand, S. (1993). Manual, hybrid and automated inspection literature and current research. *Integrated Manufacturing Systems, 4*(1), 18–29.
13. Davies, E. R. (2012). *Computer and machine vision: Theory, algorithms, practicalities*. . Academic.
14. Park, J. K., Kwon, B. K., Park, J. H., & Kang, D. J. (2016). Machine learning-based imaging system for surface defect inspection. *International Journal of Precision Engineering and Manufacturing-Green Technology, 3*(3), 303–310.
15. Kim, J., & Lee, S. (2017). Deep learning of human visual sensitivity in image quality assessment framework. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1676–1684).
16. Batchelor, B. G. (2012). *Machine vision handbook*. . Springer.
17. Penumuru, D. P., Muthuswamy, S., & Karumbu, P. (2019). Identification and classification of materials using machine vision and machine learning in the context of industry 4.0. *Journal of Intelligent Manufacturing, 31*, 1229–1241.
18. Ali, M. A., & Lun, A. K. (2019). A cascading fuzzy logic with image processing algorithm–based defect detection for automatic visual inspection of industrial cylindrical object's surface. *The International Journal of Advanced Manufacturing Technology, 102*(1–4), 81–94.
19. Badmos, O., Kopp, A., Bernthaler, T., & Schneider, G. (2020). Image-based defect detection in lithium-ion battery electrode using convolutional neural networks. *Journal of Intelligent Manufacturing, 31*(4), 885–897.
20. Di Leo, G., Liguori, C., Pietrosanto, A., & Sommella, P. (2017). A vision system for the online quality monitoring of industrial manufacturing. *Optics and Lasers in Engineering, 89*, 162–168.
21. Sun, T. H., Tien, F. C., Tien, F. C., & Kuo, R. J. (2016). Automated thermal fuse inspection using machine vision and artificial neural networks. *Journal of Intelligent Manufacturing, 27*(3), 639–651.
22. Zhang, X. W., Ding, Y. Q., Lv, Y. Y., Shi, A. Y., & Liang, R. Y. (2011). A vision inspection system for the surface defects of strongly reflected metal based on multi-class SVM. *Expert Systems with Applications, 38*(5), 5930–5939.
23. Sun, X., Gu, J., Tang, S., & Li, J. (2018). Research progress of visual inspection technology of steel products—A review. *Applied Sciences, 8*(11), 2195.
24. Liu, H. J., Wang, Y. N., & Duan, F. (2003). Image capture in machine vision. *Computer and Information Technology,* (1), 18–21.
25. Wang, J. L., Qu, X. H., & Zhao, Y. (2009). Design of lighting system in multi vision detection. *Electro-Optic Technology Application, 24*(4), 1–5.
26. Mersch, S. (1987). Overview of machine vision lighting techniques. In *Optics, illumination, and image sensing for machine vision* (Vol. 728, pp. 36–38). International Society for Optics and Photonics.
27. Cowan, C. K. (1991). Automatic camera and light-source placement using CAD models. In *Workshop on directions in automated CAD-based vision* (pp. 22–32). IEEE Computer Society.
28. Sieczka, E. J., & Harding, K. G. (1992). Light source design for machine vision. In *Optics, illumination, and image sensing for machine vision VI* (Vol. 1614, pp. 2–10). International Society for Optics and Photonics.
29. Yi, S., Haralick, R. M., & Shapiro, L. G. (1995). Optimal sensor and light source positioning for machine vision. *Computer Vision and Image Understanding, 61*(1), 122–137.
30. Kopparapu, S. K. (2006). Lighting design for machine vision application. *Image and Vision Computing, 24*(7), 720–726.

31. Li, Y., Wang, S., Tian, Q., & Ding, X. (2015). A survey of recent advances in visual feature detection. *Neurocomputing, 149*, 736–751.

32. Dan, D., Lei, M., Yao, B., Wang, W., Winterhalder, M., Zumbusch, A., et al. (2013). DMD-based LED-illumination super-resolution and optical sectioning microscopy. *Scientific Reports, 3*, 1116.

33. Moreno, I. (2012). Image-like illumination with LED arrays: design. *Optics Letters, 37*(5), 839–841.

34. Chertov, A. N., Gorbunova, E. V., Korotaev, V. V., & Peretyagin, V. S. (2014). Solution of multi-element LED light sources development automation problem. In *Thirteenth international conference on solid state lighting* (Vol. 9190, p. 919015). International Society for Optics and Photonics.

35. Moreno, I., Avendaño-Alejo, M., & Tzonchev, R. I. (2006). Designing light-emitting diode arrays for uniform near-field irradiance. *Applied Optics, 45*(10), 2265–2272.

36. Hou, T.-H.T. (2001). Automated vision system for IC lead inspection. *International Journal of Production Research, 39*(15), 3353–3366.

37. Perng, D. B., Liu, H. W., & Chang, C. C. (2011). Automated SMD LED inspection using machine vision. *The International Journal of Advanced Manufacturing Technology, 57*(9–12), 1065–1077.

38. Albeanu, D. F., Soucy, E., Sato, T. F., Meister, M., & Murthy, V. N. (2008). LED arrays as cost effective and efficient light sources for widefield microscopy. *PLoS ONE, 3*(5), e2146.

39. Gao, F., Li, Z., Xiao, G., Yuan, X., & Han, Z. (2012). An online inspection system of surface defects for copper strip based on computer vision. In *2012 5th international congress on image and signal processing* (pp. 1200–1204). IEEE.

40. Liu, Y. J., Kong, J. Y., Wang, X. D., & Jiang, F. Z. (2010). Research on image acquisition of automatic surface vision inspection systems for steel sheet. In *2010 3rd international conference on advanced computer theory and engineering (ICACTE)* (Vol. 6, pp. 189–192). IEEE.

41. Ryer, A., & Light, V. (1997). *Light measurement handbook*. Newburyport, MA: International Light.

42. Braun, D., & Heeger, A. J. (1991). Visible light emission from semiconducting polymer diodes. *Applied Physics Letters, 58*(18), 1982–1984.

43. Lee, D.-J., Schoenberger, R., Archibald, J., & McCollum, S. (2008). Development of a machine vision system for automatic date grading using digital reflective near-infrared imaging. *Journal of Food Engineering, 86*(3), 388–398.

44. Al-Mallahi, A., Kataoka, T., Okamoto, H., & Shibata, Y. (2010). Detection of potato tubers using an ultraviolet imaging-based machine vision system. *Biosystems Engineering, 105*(2), 257–265.

45. Mery, D., Lillo, I., Loebel, H., Riffo, V., Soto, A., Cipriano, A., et al. (2011). Automated fish bone detection using X-ray imaging. *Journal of Food Engineering, 105*(3), 485–492.

46. Rocha, H., Peretta, I. S., Lima, G. F. M., Marques, L. G., & Yamanaka, K. (2016). Exterior lighting computer-automated design based on multi-criteria parallel evolutionary algorithm: Optimized designs for illumination quality and energy efficiency. *Expert Systems with Applications, 45*, 208–222.

47. Sansoni, G., Biancardi, L., Minoni, U., & Docchio, F. (1994). A novel, adaptive system for 3-D optical profilometry using a liquid crystal light projector. *IEEE Transactions on Instrumentation and Measurement, 43*(4), 558–566.

48. Yang, S., Yang, L., Zhang, G., Wang, T., & Yang, X. (2018). Modeling and calibration of the galvanometric laser scanning three-dimensional measurement system. *Nanomanufacturing and Metrology, 1*(3), 180–192.

49. Zhang, S. (2018). High-speed 3D shape measurement with structured light methods: A review. *Optics and Lasers in Engineering, 106*, 119–131.

50. Li, Y., Wang, Q. L., Li, Y. F., Xu, D., & Tan, M. (2008). On-line visual measurement and inspection of weld bead using structured light. In *2008 IEEE instrumentation and measurement technology conference* (pp. 2038–2043). IEEE.

51. Lilienblum, E., & Al-Hamadi, A. (2015). A structured light approach for 3-D surface reconstruction with a stereo line-scan system. *IEEE Transactions on Instrumentation and Measurement, 64*(5), 1258–1266.

52. Chen, J. H., Shen, K., Wu, X., Zhou, X., Han, F. F., & Li, J. W. (2015). Design of stroboscopic light source used in measurement for high-speed motion object. In *2015 fifth international conference on instrumentation and measurement, computer, communication and control (IMCCC)* (pp. 1135–1138). IEEE.

53. Tian, G. Y., Lu, R. S., & Gledhill, D. (2007). Surface measurement using active vision and light scattering. *Optics and Lasers in Engineering, 45*(1), 131–139.

54. Yang, Y., Miao, C., Li, X., & Mei, X. (2014). On-line conveyor belts inspection based on machine vision. *Optik, 125*(19), 5803–5807.

55. Peng, T. G., & He, Y. H. (2013). Adaptive illumination light source for online machine vision inspection of tin steel strips. *Baosteel Technical Research, 7*(4), 25.

56. Clancy, N. T., Stoyanov, D., Yang, G.-Z., & Elson, D. S. (2012). Stroboscopic illumination scheme for seamless 3D endoscopy. In *Advanced biomedical and clinical diagnostic systems X* (Vol. 8214, pp. 82140M). International Society for Optics and Photonics.

57. Golnabi, H., & Asadpour, A. (2007). Design and application of industrial machine vision systems. *Robotics and Computer-Integrated Manufacturing, 23*(6), 630–637.

58. Bigas, M., Cabruja, E., Forest, J., & Salvi, J. (2006). Review of CMOS image sensors. *Microelectronics Journal, 37*(5), 433–451.

59. Schroder, D. (1974). A two-phase germanium charge-coupled device. *Applied Physics Letters, 25*(12), 747–749.

60. Taylor, S. A. (1998). *CCD and CMOS imaging array technologies: Technology review*. . Xerox Research Centre Europe.

61. White, M. H., Lampe, D. R., Blaha, F. C., & Mack, I. A. (1974). Characterization of surface channel CCD image arrays at low light levels. *IEEE Journal of Solid-State Circuits, 9*(1), 1–12.

62. Dillon, P. L., Lewis, D. M., & Kaspar, F. G. (1978). Color imaging system using a single CCD area array. *IEEE Journal of Solid-State Circuits, 13*(1), 28–33.

63. Beyer, H. A. (1990). Calibration of CCD-cameras for machine vision and robotics. In *Automated inspection and high-speed vision architectures III* (Vol. 1197, pp. 88–98). International Society for Optics and Photonics.

64. Jurkovic, J., Korosec, M., & Kopac, J. (2005). New approach in tool wear measuring technique using CCD vision system. *International Journal of Machine Tools and Manufacture, 45*(9), 1023–1030.

65. Dworkin, S., & Nye, T. (2006). Image processing for machine vision measurement of hot formed parts. *Journal of Materials Processing Technology, 174*(1–3), 1–6.

66. Nehir, M., Frank, C., Aßmann, S., & Achterberg, E. P. (2019). Improving optical measurements: non-linearity compensation of compact charge-coupled device (CCD) spectrometers. *Sensors, 19*(12), 2833.

67. Mehta, S., Patel, A., & Mehta, J. (2015). CCD or CMOS Image sensor for photography. In *2015 international conference on communications and signal processing (ICCSP)* (pp. 0291–0294). IEEE.

68. Akhlaq, M., Sheltami, T. R., Helgeson, B., & Shakshuki, E. M. (2012). Designing an integrated driver assistance system using image sensors. *Journal of Intelligent Manufacturing, 23*(6), 2109–2132.

69. Sun, T. H., Tseng, C. C., & Chen, M.-H. (2010). Electric contacts inspection using machine vision. *Image and Vision Computing, 28*(6), 890–901.

70. Chiou, Y. C., & Li, W. C. (2009). Flaw detection of cylindrical surfaces in PU-packing by using machine vision technique. *Measurement, 42*(7), 989–1000.

71. Shen, H., Li, S., Gu, D., & Chang, H. (2012). Bearing defect inspection based on machine vision. *Measurement, 45*(4), 719–733.

72. Sun, J., & Zhu, J. H. (2008). Calibration and correction for omnidirectional image with a fisheye lens. In *2008 fourth international conference on natural computation* (Vol. 6, pp. 133–137). IEEE.

73. Hansen, P., Alismail, H., Rander, P., & Browning, B. (2015). Visual mapping for natural gas pipe inspection. *The International Journal of Robotics Research, 34*(4–5), 532–558.

74. Chen, Y. H., Chang, C. L., Hwang, C. H., & Wang, W. C. (2013). Omnidirectional image of fish-eye lens for contact lens inspection system. In *2013 IEEE International instrumentation and measurement technology conference (I2MTC)* (pp. 1152–1155). IEEE.

75. Kogumasaka, N., Ohtani, K., & Baba, M. (2017). Surface finishing inspection using a fisheye camera system. In *2017 56th annual conference of the society of instrument and control engineers of Japan (SICE)* (pp. 487–491). IEEE.

76. Scholz-Reiter, B., Weimer, D., & Thamer, H. (2012). Automated surface inspection of cold-formed micro-parts. *CIRP Annals, 61*(1), 531–534.

77. Zhang, S., Zhou, Y., Zhang, H., Xiong, Z., & To, S. (2019). Advances in ultra-precision machining of micro-structured functional surfaces and their typical applications. *International Journal of Machine Tools and Manufacture, 142*, 16–41.

78. Fang, F. Z., Zhang, X. D., Gao, W., Guo, Y. B., Byrne, G., & Hansen, H. N. (2017). Nanomanufacturing-Perspective and applications. *CIRP Annals, 66*(2), 683–705.

79. Mathew, P. T., Rodriguez, B. J., & Fang, F. Z. (2020). Atomic and close-to-atomic scale manufacturing: A review on atomic layer removal methods using atomic force microscopy. *Nanomanufacturing and Metrology*, 3, 167–186.

80. Weimer, D., Thamer, H., Fellmann, C., Lütjen, M., Thoben, K.-D., & Scholz-Reiter, B. (2014). Towards 100% in-situ 2D/3D quality inspection of metallic micro components using plenoptic cameras. *Procedia CIRP, 17*, 847–852.

81. Li, D. J., Wang, S. W., & Fu, Y. (2017). Quality detection system and method of micro-accessory based on microscopic vision. *Modern Physics Letters B, 31*(29), 1750270.

82. Hart, J. M., Resendiz, E., Freid, B., Sawadisavi, S., Barkan, C., & Ahuja, N. (2008). Machine vision using multi-spectral imaging for undercarriage inspection of railroad equipment. In *Proceedings of the 8th world congress on railway research, Seoul, Korea* (Vol. 18).

83. Banterle, F. (2011). Advanced high dynamic range imaging: Theory and practice (1st ed.). A K Peters/CRC Press.

84. Feng, W., Zhang, F. M., Wang, W. J., Xing, W., & Qu, X. H. (2017). Digital micromirror device camera with per-pixel coded exposure for high dynamic range imaging. *Applied Optics, 56*(13), 3831–3840.

85. Zhang, T., Liu, J. H., Liu, S. L., Tang, C. T., & Jin, P. (2017). A 3D reconstruction method for pipeline inspection based on multi-vision. *Measurement, 98*, 35–48.

86. Rongsheng, L., Ang, W., Tengda, Z., & Yonghong, W. (2018). Review on automated optical (visual) inspection and its applications in defect detection. *Acta Optica Sinica, 38*(8), 0815002.

87. Gonzalez, R. C., & Woods, R. E. (2007). *Digital image processing*. (3rd ed.). Prentice-Hall Inc.

88. Ejiri, M., Uno, T., Mese, M., & Ikeda, S. (1973). A process for detecting defects in complicated patterns. *Computer Graphics and Image Processing, 2*(3–4), 326–339.

89. Hara, Y., Akiyama, N., & Karasaki, K. (1983). Automatic inspection system for printed circuit boards. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 623–630.

90. Kang, H. S., Lee, J. Y., Choi, S., Kim, H., Park, J. H., Son, J. Y., et al. (2016). Smart manufacturing: Past research, present findings, and future directions. *International Journal of Precision Engineering and Manufacturing-Green Technology, 3*(1), 111–128.

91. Herrmann, C., Schmidt, C., Kurle, D., Blume, S., & Thiede, S. (2014). Sustainability in manufacturing and factories of the future. *International Journal of Precision Engineering and Manufacturing-Green Technology, 1*(4), 283–292.

92. Rusk, N. (2016). Deep learning. *Nature Methods, 13*(1), 35–35.

93. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.

94. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278–2324.

95. Hinton, G. E., Osindero, S., & Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation, 18*(7), 1527–1554.

96. Bengio, Y. (2009). *Learning deep architectures for AI*. . Now Publishers Inc.

97. Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735–1780.

98. Bai, J., & Feng, X. C. (2007). Fractional-order anisotropic diffusion for image denoising. *IEEE Transactions on Image Processing, 16*(10), 2492–2502.

99. Thakur, K. V., Damodare, O. H., & Sapkal, A. M. (2016). Poisson noise reducing bilateral filter. *Procedia Computer Science, 79*, 861–865.

100. Fukushima, N., Sugimoto, K., & Kamata, S.-I. (2018). Guided image filtering with arbitrary window function. In *2018 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 1523–1527) . IEEE.

101. Torres-Huitzil, C. (2013). Fast hardware architecture for grey-level image morphology with flat structuring elements. *IET Image Processing, 8*(2), 112–121.

102. Brigham, E. O., & Morrow, R. (1967). The fast Fourier transform. *IEEE Spectrum, 4*(12), 63–70.

103. Cooley, J. W., & Tukey, J. W. (1965). An algorithm for the machine calculation of complex Fourier series. *Mathematics of Computation, 19*(90), 297–301.

104. Zhang, Z., Wang, Y., & Wang, K. (2013). Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *Journal of Intelligent Manufacturing, 24*(6), 1213–1227.

105. Shao, H., Shi, X., & Li, L. (2011). Power signal separation in milling process based on wavelet transform and independent component analysis. *International Journal of Machine Tools and Manufacture, 51*(9), 701–710.

106. Pislaru, C., Freeman, J., & Ford, D. G. (2003). Modal parameter identification for CNC machine tools using wavelet transform. *International Journal of Machine Tools and Manufacture, 43*(10), 987–993.

107. Boujelbene, R., Jemaa, Y. B., & Zribi, M. (2019). A comparative study of recent improvements in wavelet-based image coding schemes. *Multimedia Tools and Applications, 78*(2), 1649–1683.

108. Luisier, F., Blu, T., & Unser, M. (2007). A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding. *IEEE Transactions on Image Processing, 16*(3), 593–606.

109. Jain, P., & Tyagi, V. (2015). LAPB: Locally adaptive patch-based wavelet domain edge-preserving image denoising. *Information Sciences, 294*, 164–181.

110. Yan, Z., Xu, W., & Yang, C. (2018). A power thresholding function-based wavelet image denoising method. *Journal of Imaging Science and Technology, 62*(1), 10506–10501.

111. Xu, X., Wang, Y., & Chen, S. (2016). Medical image fusion using discrete fractional wavelet transform. *Biomedical Signal Processing and Control, 27*, 103–111.

112. Daniel, E. (2018). Optimum wavelet-based homomorphic medical image fusion using hybrid genetic–grey wolf optimization algorithm. *IEEE Sensors Journal, 18*(16), 6804–6811.

113. Rein, S. A., Fitzek, F. H., Gühmann, C., & Sikora, T. (2015). Evaluation of the wavelet image two-line coder: A low complexity scheme for image compression. *Signal Processing: Image Communication, 37*, 58–74.

114. Bruylants, T., Munteanu, A., & Schelkens, P. (2015). Wavelet based volumetric medical image compression. *Signal Processing: Image Communication, 31*, 112–133.

115. Mehra, I., & Nishchal, N. K. (2015). Optical asymmetric image encryption using gyrator wavelet transform. *Optics Communications, 354*, 344–352.

116. Yang, Y., Su, Z., & Sun, L. (2010). Medical image enhancement algorithm based on wavelet transform. *Electronics Letters, 46*(2), 120–121.

117. Jung, C., Yang, Q., Sun, T., Fu, Q., & Song, H. (2017). Low light image enhancement with dual-tree complex wavelet transform. *Journal of Visual Communication and Image Representation, 42*, 28–36.

118. Nixon, M., & Aguado, A. (2019). *Feature extraction and image processing for computer vision.* . Academic.

119. Lin, Z., Fu, J., Shen, H., Xu, G., & Sun, Y. (2016). Improving machined surface texture in avoiding five-axis singularity with the acceptable-texture orientation region concept. *International Journal of Machine Tools and Manufacture, 108*, 1–12.

120. Li, M., Wan, S., Deng, Z., & Wang, Y. (2019). Fabric defect detection based on saliency histogram features. *Computational Intelligence, 35*(3), 517–534.

121. Chatzichristofis, S. A., & Boutalis, Y. S. (2008). Fcth: Fuzzy color and texture histogram-a low level feature for accurate image retrieval. In *2008 ninth international workshop on image analysis for multimedia interactive services* (pp. 191–196). IEEE.

122. Hadjidemetriou, E., Grossberg, M. D., & Nayar, S. K. (2004). Multiresolution histograms and their use for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 26*(7), 831–847.

123. Mohanaiah, P., Sathyanarayana, P., & GuruKumar, L. (2013). Image texture feature extraction using GLCM approach. *International Journal of Scientific and Research Publications, 3*(5), 1.

124. Pushpalatha, K., Karegowda, A. G., & Ramesh, D. (2017). Identification of similar looking bulk split grams using GLCM and CGLCM texture features. *International Journal of Computer Applications, 167*(6), 30–36.

125. Zhu, D., Pan, R., Gao, W., & Zhang, J. (2015). Yarn-dyed fabric defect detection based on autocorrelation function and GLCM. *Autex Research Journal, 15*(3), 226–232.

126. Wang, X., Ding, X., & Liu, C. (2005). Gabor filters-based feature extraction for character recognition. *Pattern Recognition, 38*(3), 369–379.

127. Tan, X., & Triggs, B. (2007). Fusing Gabor and LBP feature sets for kernel-based face recognition. In *International workshop on analysis and modeling of faces and gestures* (pp. 235–249). Springer.

128. Raheja, J. L., Kumar, S., & Chaudhary, A. (2013). Fabric defect detection based on GLCM and Gabor filter: A comparison. *Optik, 124*(23), 6469–6474.

129. Ojala, T., Pietikainen, M., & Maenpaa, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*(7), 971–987.

130. Nosaka, R., Ohkawa, Y., & Fukui, K. (2011). Feature extraction based on co-occurrence of adjacent local binary patterns. In *Pacific-rim symposium on image and video technology* (pp. 82–91). Springer.

131. Shan, C. (2012). Learning local binary patterns for gender classification on real-world face images. *Pattern Recognition Letters, 33*(4), 431–437.

132. Zhang, L., Jing, J., & Zhang, H. (2015). Fabric defect classification based on LBP and GLCM. *Journal of Fiber Bioengineering and Informatics, 8*(1), 81–89.

133. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60*(2), 91–110.

134. Lindeberg, T. (2012). *Scale invariant feature transform. Scholarpedia, 7*(5), 10491.

135. Dunderdale, C., Brettenny, W., Clohessy, C., & van Dyk, E. E. (2020). Photovoltaic defect classification through thermal infrared imaging using a machine learning approach. *Progress in Photovoltaics: Research and Applications, 28*(3), 177–188.

136. Dalal, N., & Triggs, B. (2005). Histograms of oriented gradients for human detection. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)* (Vol. 1, pp. 886–893). IEEE.

137. Halfawy, M. R., & Hengmeechai, J. (2014). Automated defect detection in sewer closed circuit television images using histograms of oriented gradients and support vector machine. *Automation in Construction, 38*, 1–13.

138. Bay, H., Ess, A., Tuytelaars, T., & Van Gool, L. (2008). Speeded-up robust features (SURF). *Computer Vision and Image Understanding, 110*(3), 346–359.

139. Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010). Brief: Binary robust independent elementary features. In *European conference on computer vision* (pp. 778–792). Springer.

140. Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011). ORB: An efficient alternative to SIFT or SURF. In *2011 International conference on computer vision* (pp. 2564–2571). Ieee.

141. Guo, Z., Zhang, L., & Zhang, D. (2010). A completed modeling of local binary pattern operator for texture classification. *IEEE Transactions on Image Processing, 19*(6), 1657–1663.

142. Liu, L., Zhao, L., Long, Y., Kuang, G., & Fieguth, P. (2012). Extended local binary patterns for texture classification. *Image and Vision Computing, 30*(2), 86–99.

143. Chen, J., Kellokumpu, V., Zhao, G., & Pietikäinen, M. (2013). RLBP: Robust local binary pattern. In *BMVC*.

144. Zhao, J., Peng, Y., & Yan, Y. (2018). Steel surface defect classification based on discriminant manifold regularized local descriptor. *IEEE Access, 6*, 71719–71731.

145. Suykens, J. A., & Vandewalle, J. (1999). Least squares support vector machine classifiers. *Neural Processing Letters, 9*(3), 293–300.

146. Keller, J. M., Gray, M. R., & Givens, J. A. (1985). A fuzzy k-nearest neighbor algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, 4*, 580–585.

147. Jia, H., Murphey, Y. L., Shi, J., & Chang, T.-S. (2004). An intelligent real-time vision system for surface defect detection. In *Proceedings of the 17th international conference on pattern recognition, 2004. ICPR 2004.* (Vol. 3, pp. 239–242). IEEE.

148. Li, T.-S., & Huang, C.-L. (2009). Defect spatial pattern recognition using a hybrid SOM–SVM approach in semiconductor manufacturing. *Expert Systems with Applications, 36*(1), 374–385.

149. Valavanis, I., & Kosmopoulos, D. (2010). Multiclass defect detection and classification in weld radiographic images using geometric and texture features. *Expert Systems with Applications, 37*(12), 7606–7614.

150. Huang, Y., Wu, D., Zhang, Z., Chen, H., & Chen, S. (2017). EMD-based pulsed TIG welding process porosity defect detection and defect diagnosis using GA-SVM. *Journal of Materials Processing Technology, 239*, 92–102.

151. Zhang, X. W., Gong, F., & Xu, L. Z. (2012). Inspection of surface defects in copper strip using multivariate statistical approach and SVM. *International Journal of Computer Applications in Technology, 43*(1), 44–50.

152. Liang, R., Ding, Y., Zhang, X., & Chen, J. (2008). Copper strip surface defects inspection based on SVM-RBF. In *2008 fourth international conference on natural computation* (Vol. 2, pp. 41–45). IEEE.

153. You, D., Gao, X., & Katayama, S. (2014). WPD-PCA-based laser welding process monitoring and defects diagnosis by using FNN and SVM. *IEEE Transactions on Industrial Electronics, 62*(1), 628–636.

154. Chen, B., Yan, Z., & Chen, W. (2014). Defect detection for wheel-bearings with time-spectral kurtosis and entropy. *Entropy, 16*(1), 607–626.

155. Wang, J., Neskovic, P., & Cooper, L. N. (2006). Neighborhood size selection in the k-nearest-neighbor rule using statistical confidence. *Pattern Recognition, 39*(3), 417–423.

156. Lei, Y. G., & Zuo, M. J. (2009). Gear crack level identification based on weighted K nearest neighbor classification algorithm. *Mechanical Systems and Signal Processing, 23*(5), 1535–1547.

157. Yıldız, K., Buldu, A., & Demetgul, M. (2016). A thermal-based defect classification method in textile fabrics with K-nearest neighbor algorithm. *Journal of Industrial Textiles, 45*(5), 780–795.

158. Cetiner, I., Var, A. A., & Cetiner, H. (2016). Classification of knot defect types using wavelets and KNN. *Elektronika ir elektrotechnika, 22*(6), 67–72.

159. Das, S., & Jena, U. R. (2016). Texture classification using combination of LBP and GLRLM features along with KNN and multiclass SVM classification. In *2016 2nd international conference on communication control and intelligent systems (CCIS)* (pp. 115–119). IEEE.

160. Mjahed, S., El Hadaj, S., Bouzaachane, K., & Raghay, S. (2018). Engine fault signals diagnosis using genetic algorithm and K-means based clustering. In *Proceedings of the international conference on learning and optimization algorithms: theory and applications* (pp. 1–6).

161. Hamdi, A. A., Sayed, M. S., Fouad, M. M., & Hadhoud, M. M. (2018). Unsupervised patterned fabric defect detection using texture filtering and K-means clustering. In *2018 international conference on innovative trends in computer engineering (ITCE)* (pp. 130–144). IEEE.

162. Jiao, L. C., Yang, S. Y., Liu, F., Wang, S. G., & Feng, Z. X. (2016). Seventy years beyond neural networks: retrospect and prospect. *Chinese Journal of Computers, 39*(8), 1697–1716.

163. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM, 60*(6), 84–90.

164. Lin, M., Chen, Q., & Yan, S. (2013). Network in network. arXiv: 1312.4400

165. Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv:arXiv:1409. 1556

166. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1–9).

167. Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv:1502.03167

168. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818–2826).

169. Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 31, p. 1).

170. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770–778).

171. Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 4700–4708).

172. Zhang, Q., Zhang, M., Chen, T., Sun, Z., Ma, Y., & Yu, B. (2019). Recent advances in convolutional neural network acceleration. *Neurocomputing, 323*, 37–51.

173. Voulodimos, A., Doulamis, N., Doulamis, A., & Protopapadakis, E. (2018). Deep learning for computer vision: A brief review. *Computational intelligence and neuroscience, 2018*, 1–13.

174. Wang, J., Ma, Y., Zhang, L., Gao, R. X., & Wu, D. (2018). Deep learning for smart manufacturing: Methods and applications. *Journal of Manufacturing Systems, 48*, 144–156.

175. Gollapudi, S. (2019). Deep learning for computer vision. In *Learn computer vision using OpenCV* (pp. 51–69). Springer.

176. Du, W., Shen, H., Fu, J., Zhang, G., Shi, X., & He, Q. (2020). Automated detection of defects with low semantic information in X-ray images based on deep learning. *Journal of Intelligent Manufacturing, 32*, 141–156.

177. Zhang, Y., You, D., Gao, X., Wang, C., Li, Y., & Gao, P. P. (2019). Real-time monitoring of high-power disk laser welding statuses based on deep learning framework. *Journal of Intelligent Manufacturing, 31*, 799–814.

178. Masci, J., Meier, U., Ciresan, D., Schmidhuber, J., & Fricout, G. (2012) Steel defect classification with max-pooling convolutional neural networks. In *The 2012 international joint conference on neural networks (IJCNN)* (pp. 1–6) . IEEE.

179. Lin, J., Yao, Y., Ma, L., & Wang, Y. (2018). Detection of a casting defect tracked by deep convolution neural network. *The International Journal of Advanced Manufacturing Technology, 97*(1–4), 573–581.

180. Nguyen, T. P., Choi, S., Park, S.-J., Park, S. H., & Yoon, J. (2020). Inspecting method for defective casting products with convolutional neural network (CNN). *International Journal of Precision Engineering and Manufacturing-Green Technology, 8*, 583–594.

181. Kim, M., Lee, M., An, M., & Lee, H. (2020). Effective automatic defect classification process based on CNN with stacking ensemble model for TFT-LCD panel. *Journal of Intelligent Manufacturing, 31*(5), 1165–1174.

182. Imoto, K., Nakai, T., Ike, T., Haruki, K., & Sato, Y. (2018). A CNN-based transfer learning method for defect classification in semiconductor manufacturing. In *2018 international symposium on semiconductor manufacturing (ISSM)* (pp. 1–3). IEEE.

183. Viola, P., & Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001* (Vol. 1). IEEE.

184. Neubeck, A., & Van Gool, L. (2006). Efficient non-maximum suppression. In *18th International Conference on Pattern Recognition (ICPR'06)* (Vol. 3, pp. 850–855). IEEE.

185. Felzenszwalb, P., McAllester, D., & Ramanan, D. (2008). A discriminatively trained, multiscale, deformable part model. In *2008 IEEE conference on computer vision and pattern recognition* (pp. 1–8). IEEE.

186. Van de Sande, K. E., Uijlings, J. R., Gevers, T., & Smeulders, A. W. (2011). Segmentation as selective search for object recognition. In *2011 international conference on computer vision* (pp. 1879–1886). IEEE.

187. Uijlings, J. R., Van De Sande, K. E., Gevers, T., & Smeulders, A. W. (2013). Selective search for object recognition. *International Journal of Computer Vision, 104*(2), 154–171.

188. Zitnick, C. L., & Dollár, P. (2014). Edge boxes: Locating object proposals from edges. In *European conference on computer vision* (pp. 391–405). Springer.

189. Shumin, D., Zhoufeng, L., & Chunlei, L. (2011). AdaBoost learning for fabric defect detection based on HOG and SVM. In *2011 International conference on multimedia technology* (pp. 2903–2906). IEEE.

190. Dou, Y., Huang, Y., Li, Q., & Luo, S. (2014). A fast template matching-based algorithm for railway bolts detection. *International Journal of Machine Learning and Cybernetics, 5*(6), 835–844.

191. He, B., Hou, Y., Xiong, Y., & Li, B. (2019). Railway fastener defects detection using Gaussian mixture deformable part model. In *Journal of physics: Conference series* (Vol. 1302, pp. 022102, Vol. 2). IOP Publishing.

192. Wei, L., Zhang, N., Xue, M., & Huo, J. (2020). Research of express box defect detection based on machine vision. In *Proceedings of the 5th international conference on multimedia and image processing* (pp. 12–17).

193. Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. D. (2019). Object detection with deep learning: A review. *IEEE Transactions on Neural Networks and Learning Systems, 30*(11), 3212–3232.

194. Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580–587).

195. He, K., Zhang, X., Ren, S., & Sun, J. (2015). Spatial pyramid pooling in deep convolutional networks for visual recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 37*(9), 1904–1916.

196. Girshick, R. (2015). Fast r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 1440–1448).

197. Ren, S., He, K., Girshick, R., & Sun, J. (2016). Faster r-cnn: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(6), 1137–1149.

198. Dai, J., Li, Y., He, K., & Sun, J. (2016). R-fcn: Object detection via region-based fully convolutional networks. arXiv:1605.06409

199. Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2117–2125).

200. He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision* (pp. 2961–2969) .

201. Erhan, D., Szegedy, C., Toshev, A., & Anguelov, D. (2014). Scalable object detection using deep neural networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2147–2154).

202. Yoo, D., Park, S., Lee, J.-Y., Paek, A. S., & So Kweon, I. (2015) Attentionnet: Aggregating weak directions for accurate object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2659–2667).

203. Najibi, M., Rastegari, M., & Davis, L. S. (2016). G-cnn: An iterative grid based object detector. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2369–2377).

204. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779–788).

205. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). Ssd: Single shot multibox detector. In *European conference on computer vision* (pp. 21–37). Springer.

206. Redmon, J., & Farhadi, A. (2017). YOLO9000: better, faster, stronger. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 7263–7271) .

207. Lin, T.-Y., Goyal, P., Girshick, R., He, K., & Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision* (pp. 2980–2988).

208. Redmon, J., & Farhadi, A. (2018). Yolov3: An incremental improvement. arXiv:1804.02767

209. Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020). YOLOv4: Optimal speed and accuracy of object detection. arXiv:2004.10934

210. Wu, Y., & Zhang, X. (2020). Automatic fabric defect detection using cascaded mixed feature pyramid with guided localization. *Sensors, 20*(3), 871.

211. Lei, H. W., Wang, B., Wu, H. H., & Wang, A. H. (2018). Defect detection for polymeric polarizer based on faster R-CNN. *Journal of Intelligent Information Hiding and Multimedia Signal Processing, 9*, 1414–1420.

212. Lei, X., & Sui, Z. (2019). Intelligent fault detection of high voltage line based on the faster R-CNN. *Measurement, 138*, 379–385.

213. Sun, X., Gu, J., Huang, R., Zou, R., & Giron Palomares, B. (2019). Surface defects recognition of wheel hub based on improved faster R-CNN. *Electronics, 8*(5), 481.

214. Wang, K.-J., Rizqi, D. A., & Nguyen, H.-P. (2020). Skill transfer support model based on deep learning. *Journal of Intelligent Manufacturing, 32*, 1129–1146.

215. Adibhatla, V. A., Chih, H.-C., Hsu, C.-C., Cheng, J., Abbod, M. F., & Shieh, J.-S. (2020). Defect detection in printed circuit boards using you-only-look-once convolutional neural networks. *Electronics, 9*(9), 1547.

216. Lv, X. M., Duan, F. J., Jiang, J. J., Fu, X., & Gan, L. (2020). Deep active learning for surface defect detection. *Sensors, 20*(6), 1650.

217. Jing, J., Zhuo, D., Zhang, H., Liang, Y., & Zheng, M. (2020). Fabric defect detection using the improved YOLOv3 model. *Journal of Engineered Fibers and Fabrics, 15*, 1558925020908268.

218. Deng, H., Cheng, J., Liu, T., Cheng, B., & Sun, Z. (2020). Research on iron surface crack detection algorithm based on improved YOLOv4 network. In *Journal of physics: Conference series* (Vol. 1631, pp. 012081, Vol. 1). IOP Publishing.

219. Zhai, S., Shang, D., Wang, S., & Dong, S. (2020). DF-SSD: An improved SSD object detection algorithm based on DenseNet and feature fusion. *IEEE Access, 8*, 24344–24357.

220. Haralick, R. M., & Shapiro, L. G. (1985). Image segmentation techniques. *Computer Vision, Graphics, and Image Processing, 29*(1), 100–132.

221. Cuevas, E., Zaldivar, D., & Pérez-Cisneros, M. (2010). A novel multi-threshold segmentation approach based on differential evolution optimization. *Expert Systems with Applications, 37*(7), 5265–5271.

222. Pernkopf, F., & O'Leary, P. (2002). Visual inspection of machined metallic high-precision surfaces. *EURASIP Journal on Advances in Signal Processing, 2002*(7), 650750.

223. Tang, J. (2010). A color image segmentation algorithm based on region growing. In *2010 2nd international conference on computer engineering and technology* (Vol. 6, pp. V6-634–V636-637). IEEE.

224. Chuang, K.-S., Tzeng, H.-L., Chen, S., Wu, J., & Chen, T.-J. (2006). Fuzzy c-means clustering with spatial information for image segmentation. *Computerized Medical Imaging and Graphics, 30*(1), 9–15.

225. Dhanachandra, N., Manglem, K., & Chanu, Y. J. (2015). Image segmentation using K-means clustering algorithm and subtractive clustering algorithm. *Procedia Computer Science, 54*, 764–771.

226. Mia, S., & Rahman, M. M. (2018). An efficient image segmentation method based on linear discriminant analysis and K-means algorithm with automatically splitting and merging clusters. *International Journal of Imaging and Robotics, 18*(1), 62–72.

227. Brejl, M., & Sonka, M. (2000). Object localization and border detection criteria design in edge-based image segmentation: Automated learning from examples. *IEEE Transactions on Medical imaging, 19*(10), 973–985.

228. Bo, T., Jianyi, K., & Shiqian, W. (2017). Review of surface defect detection based on machine vision. *Journal of Image and Graphics, 22*(12), 1640–1663.

229. Celik, T., & Tjahjadi, T. (2010). Unsupervised colour image segmentation using dual-tree complex wavelet transform. *Computer Vision and Image Understanding, 114*(7), 813–826.

230. Lo, E. H., Pickering, M. R., Frater, M. R., & Arnold, J. F. (2011). Image segmentation from scale and rotation invariant texture features from the double dyadic dual-tree complex wavelet transform. *Image and Vision Computing, 29*(1), 15–28.

231. Shih, F. Y. (2009). *Image processing and mathematical morphology: Fundamentals and applications*. . CRC Press.

232. Putera, S. I., & Ibrahim, Z. (2010). Printed circuit board defect detection using mathematical morphology and MATLAB image processing tools. In *2010 2nd international conference on education technology and computer* (Vol. 5, pp. V5-359-V355-363). IEEE.

233. Hammouche, K., Diaf, M., & Siarry, P. (2008). A multilevel automatic thresholding method based on a genetic algorithm for a fast image segmentation. *Computer Vision and Image Understanding, 109*(2), 163–175.

234. Sengur, A., & Guo, Y. (2011). Color texture image segmentation based on neutrosophic set and wavelet transformation. *Computer Vision and Image Understanding, 115*(8), 1134–1144.

235. Shirkhorshidi, A. S., Aghabozorgi, S., Wah, T. Y., & Herawan, T. (2014). Big data clustering: A review. In *International conference on computational science and its applications* (pp. 707–720). Springer.

236. Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics, 3*(3), 32–57.

237. Zhang, T., Ramakrishnan, R., & Livny, M. (1997). BIRCH: A new data clustering algorithm and its applications. *Data Mining and Knowledge Discovery, 1*(2), 141–182.

238. Guha, S., Rastogi, R., & Shim, K. (1998). CURE: an efficient clustering algorithm for large databases. *ACM Sigmod Record, 27*(2), 73–84.

239. Ng, R. T., & Han, J. (2002). CLARANS: A method for clustering objects for spatial data mining. *IEEE Transactions on Knowledge and Data Engineering, 14*(5), 1003–1016.

240. Khan, S. S., & Ahmad, A. (2004). Cluster center initialization algorithm for K-means clustering. *Pattern Recognition Letters, 25*(11), 1293–1302.

241. Zhao, G. F., & Qu, G. Q. (2006). Analysis and implementation of CLARA algorithm on clustering. *Journal of Shandong University of Technology (Science and Technology), 2*, 45–48.

242. Li, J., Wang, K., & Xu, L. (2009). Chameleon based on clustering feature tree and its application in customer segmentation. *Annals of Operations Research, 168*(1), 225–245.

243. Park, H.-S., & Jun, C.-H. (2009). A simple and fast algorithm for K-medoids clustering. *Expert Systems with Applications, 36*(2), 3336–3341.

244. Tran, T. N., Drab, K., & Daszykowski, M. (2013). Revised DBSCAN algorithm to cluster data with dense adjacent clusters. *Chemometrics and Intelligent Laboratory Systems, 120*, 92–96.

245. Maillo, J., Triguero, I., & Herrera, F. (2015). A mapreduce-based k-nearest neighbor approach for big data classification. In *2015 IEEE Trustcom/BigDataSE/ISPA* (Vol. 2, pp. 167–172). IEEE.

246. El Moussawi, A., Giacometti, A., Labroche, N., & Soulet, A. (2020). MAPK-means: A clustering algorithm with quantitative preferences on attributes. *Intelligent Data Analysis, 24*(2), 459–489.

247. Xiong, Z., Li, Q., Mao, Q., & Zou, Q. (2017). A 3D laser profiling system for rail surface defect detection. *Sensors, 17*(8), 1791.

248. Jian, C., Gao, J., & Ao, Y. (2017). Automatic surface defect detection for mobile phone screen glass based on machine vision. *Applied Soft Computing, 52*, 348–358.

249. Melnyk, R., & Tushnytskyy, R. (2020). Detection of defects in printed circuit boards by clustering the etalon and defected samples. In *2020 IEEE 15th international conference on advanced trends in radioelectronics, telecommunications and computer engineering (TCSET)* (pp. 961–964). IEEE.

250. Li, C., Zhang, X., Huang, Y., Tang, C., & Fatikow, S. (2020). A novel algorithm for defect extraction and classification of mobile phone screen based on machine vision. *Computers & Industrial Engineering, 146*, 106530.

251. Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3431–3440).

252. Liu, X. L., Deng, Z. D., & Yang, Y. H. (2019). Recent progress in semantic image segmentation. *Artificial Intelligence Review, 52*(2), 1089–1106.

253. Yu, Z., Wu, X., & Gu, X. (2017). Fully convolutional networks for surface defect inspection in industrial environment. In *International conference on computer vision systems* (pp. 417–426). Springer.

254. Li, S. J., Zhou, H. M., Wang, G. Y., Zhu, X. H., Kong, L. F., & Hu, Z. Y. (2018). Cracked insulator detection based on R-FCN. In *Journal of physics: Conference series* (Vol. 1069, pp. 012147, Vol. 1).

255. Yang, X., Li, H., Yu, Y., Luo, X., Huang, T., & Yang, X. (2018). Automatic pixel-level crack detection and measurement using fully convolutional network. *Computer-Aided Civil and Infrastructure Engineering, 33*(12), 1090–1109.

256. Qiu, L., Wu, X., & Yu, Z. (2019). A high-efficiency fully convolutional networks for pixel-wise surface defect detection. *IEEE Access, 7*, 15884–15893.

257. Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International conference on medical image computing and computer-assisted intervention* (pp. 234–241). Springer.

258. Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(12), 2481–2495.

259. Zhao, H., Shi, J., Qi, X., Wang, X., & Jia, J. (2017). Pyramid scene parsing network. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2881–2890).

260. Yu, H., Li, X., Song, K., Shang, E., Liu, H., & Yan, Y. (2020). Adaptive depth and receptive field selection network for defect

semantic segmentation on castings X-rays. *NDT & E International, 116*, 102345.

261. Tabernik, D., Šela, S., Skvarč, J., & Skočaj, D. (2020). Segmentation-based deep-learning approach for surface-defect detection. *Journal of Intelligent Manufacturing, 31*(3), 759–776.

262. Hu, C., Duan, Y., Liu, S., Yan, Y., Tao, N., Osman, A., et al. (2019). LSTM-RNN-based defect classification in honeycomb structures using infrared thermography. *Infrared Physics & Technology, 102*, 103032.

263. Wang, Q., Liu, Q., Xia, R., Li, G., Gao, J., Zhou, H., et al. (2020). Defect depth determination in laser infrared thermography based on LSTM-RNN. *IEEE Access, 8*, 153385–153393.

264. Liu, T., Bao, J., Wang, J., & Zhang, Y. (2018). A hybrid CNN–LSTM algorithm for online defect recognition of CO2 welding. *Sensors, 18*(12), 4369.

265. Liu, Y., Xu, K., & Xu, J. (2019). Periodic surface defect detection in steel plates based on deep learning. *Applied Sciences, 9*(15), 3127.

266. Jia, F., Lei, Y., Lin, J., Zhou, X., & Lu, N. (2016). Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data. *Mechanical Systems and Signal Processing, 72*, 303–315.

267. Yang, L., & Jiang, H. (2020). Weld defect classification in radiographic images using unified deep neural network with multi-level features. *Journal of Intelligent Manufacturing*, *32*, 459–469.

268. Seker, A., & Yuksek, A. G. (2017). Stacked autoencoder method for fabric defect detection. *Cumhuriyet Science Journal, 38*(2), 342–354.

269. Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science, 313*(5786), 504–507.

270. Chen, Z., Zeng, X., Li, W., & Liao, G. (2016). Machine fault classification using deep belief network. In *2016 IEEE international instrumentation and measurement technology conference proceedings* (pp. 1–6). IEEE.

271. Chen, Z., & Li, W. (2017). Multisensor feature fusion for bearing fault diagnosis using sparse autoencoder and deep belief network. *IEEE Transactions on instrumentation and measurement, 66*(7), 1693–1702.

272. Fang, F. Z. (2020). On atomic and close-to-atomic scale manufacturing—Development trend of manufacturing technology. *Chinese Mechanical Engineering, 31*(09), 1009–1021.

273. Kim, S., Noh, Y.-K., & Park, F. C. (2020). Efficient neural network compression via transfer learning for machine vision inspection. *Neurocomputing, 413*, 294–304.

274. Zhou, Z. H. (2018). A brief introduction to weakly supervised learning. *National Science Review, 5*(1), 44–53.

275. Barlow, H. B. (1989). Unsupervised learning. *Neural Computation, 1*(3), 295–311.

276. Mei, S., Yang, H., & Yin, Z. (2018). An unsupervised-learning-based approach for automated defect inspection on textured surfaces. *IEEE Transactions on Instrumentation and Measurement, 67*(6), 1266–1277.

277. Xu, L. D., Xu, E. L., & Li, L. (2018). Industry 4.0: State of the art and future trends. *International Journal of Production Research, 56*(8), 2941–2962.

278. Shimizu, Y., Matsuno, Y., Chen, Y.-L., Matsukuma, H., & Gao, W. (2018). Design and testing of a micro-thermal sensor probe for nondestructive detection of defects on a flat surface. *Nanomanufacturing and Metrology, 1*(1), 45–57.

279. Liu, S., Zuo, Y., & Zhang, Z. (2018). A new detecting technology for external anticorrosive coating defects of pipelines based on ultrasonic guided wave. *E&ES, 108*(2), 022073.

280. Li, X., Liu, Z., Jiang, X., & Lodewijks, G. (2018). Method for detecting damage in carbon-fibre reinforced plastic-steel structures based on eddy current pulsed thermography. *Nondestructive Testing and Evaluation, 33*(1), 1–19.

281. Hwang, S., An, Y.-K., Yang, J., & Sohn, H. (2020). Remote inspection of internal delamination in wind turbine blades using continuous line laser scanning thermography. *International Journal of Precision Engineering and Manufacturing-Green Technology, 7*, 699–712.

282. Tran, T. X., Hajisami, A., Pandey, P., & Pompili, D. (2017). Collaborative mobile edge computing in 5G networks: New paradigms, scenarios, and challenges. *IEEE Communications Magazine, 55*(4), 54–61.

283. Preskill, J. (2018). Quantum computing in the NISQ era and beyond. *Quantum, 2*, 79.

284. Fang, F. Z. (2020). Atomic and close-to-atomic scale manufacturing: perspectives and measures. *International Journal of Extreme Manufacturing, 2*(3), 030201.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Zhonghe Ren received the B.E. and M.E. degrees in School of Mechanical Engineering from Jiangnan University, China, in 2016 and 2019, respectively. Currently, he is pursuing the Ph.D. degree in School of Precision Instruments and Opto-electronics Engineering at Tianjin University, China. His research interests include computer vision, image processing, and deep learning.

Fengzhou Fang is currently a joint professor in School of Precision Instruments and Opto-electronics Engineering at Tianjin University in China and in School of Mechanical and Materials Engineering at University College Dublin in Ireland. He started his career in teaching and research as an academic staff at university in 1982. His specialist areas of interest include optical freeform design and manufacturing, medical device/implants manufacturing, ultra-precision manufacturing and measurement. He has published over 360 technical papers in peer reviewed journals, along with 12 book chapters. He is currently the editor-in-chief of Nanomanufacturing and Metrology.

Ning Yan received the B.E. degree in School of Precision Instruments and Opto-electronics Engineering from Tianjin University, China, in 2017. Currently, he is pursuing the Ph.D. degree in School of Precision Instruments and Opto-electronics Engineering at Tianjin University, China. His research interests include computer vision and precise measurement.

You Wu received the B.E. degree in School of Mechanical and Aerospace Engineering (SMAE) from Jilin University, China, in 2018. Currently, he is pursuing the M.E. degree in School of Precision Instruments and Opto-electronics Engineering at Tianjin University, China. His research interests include computer vision, machine learning and deep learning.