Computer Vision Techniques in Manufacturing

This paper was downloaded from TechRxiv (<https://www.techrxiv.org>).

LICENSE

CC BY 4.0

SUBMISSION DATE / POSTED DATE

05-12-2021 / 11-12-2021

CITATION

Zhou, Longfei; Zhang, Lin; Konz, Nicholas (2021): Computer Vision Techniques in Manufacturing. TechRxiv. Preprint. <https://doi.org/10.36227/techrxiv.17125652.v1>

DOI

[10.36227/techrxiv.17125652.v1](https://doi.org/10.36227/techrxiv.17125652.v1)

Computer Vision Techniques in Manufacturing

Longfei Zhou, *Member, IEEE*, Lin Zhang, *Senior Member, IEEE*, Nicholas Konz

Abstract—Computer vision techniques have played an important role in promoting the informatization, digitization and intelligence of industrial manufacturing systems. Considering the rapid development of computer vision techniques, we present a comprehensive review of the state-of-the-art of these techniques and their applications in manufacturing industries. We survey the most common methods, including feature detection, recognition, segmentation and 3D modeling. A system framework of computer vision in the manufacturing environment is proposed, consisting of a lighting module, a manufacturing system, a sensing module, computer vision algorithms, a decision-making module, and an actuator. Applications of computer vision to different stages of the entire product life cycle are then explored, including product design, modeling and simulation, planning and scheduling, the production process, inspection and quality control, assembly, transportation, and disassembly. Challenges include algorithm implementation, data pre-processing, data labeling, and benchmarks. Future directions include building benchmarks, developing methods for non-annotated data processing, developing effective data pre-processing mechanisms, customizing computer vision models, and opportunities aroused by 5G.

Index Terms—Computer vision, machine intelligence, machine learning, deep learning, manufacturing, production, inspection, robot, assembly, review.

I. INTRODUCTION

Computer vision (CV) is a central subfield of artificial intelligence, and has been developed very rapidly in recent years due to the meteoric rise of deep learning. Computer vision techniques have been applied to different manufacturing industries since the early 1970s, such as food, pharmaceutical, automotive, aerospace, railway, semiconductor, electronic component, plastic, rubber, paper, and forestry-related fields [1]. Researchers and practitioners have been asking questions for decades about how computer vision techniques can be applied to manufacturing industries [2], and vision-based industrial inspection and assembly has gained the most attention [3].

In the early days, only a few computer vision methods were practically used in commercial manufacturing because of limited computing capabilities, until the 1990s [4]. Opportunities for computer vision in manufacturing can be grouped into three broad categories including image-based metrology, manufacturing process interpretability, and material structure analysis [5]. From the perspective of application scenarios, industrial applications of computer vision technologies can be classified into the following: visual inspection, process

control, parts identification, and robotic guidance and control mechanisms [6]. In this survey, our classification framework is based on different stages of the product life cycle in the entire manufacturing process.

Existing review papers on this topic usually specialize in just one of the manufacturing industries, such as the steel industry. For example, [7] reviewed some available computer vision-based froth imaging systems supported by machine vision technologies. In [8], vision-based defect detection and classification of steel surfaces were surveyed for steel mill systems. The survey of [9] reviewed the literature of weld pool state sensing, including conventional sensing, vision sensing, and multi-sensor information fusion technologies, emphasizing the analysis of three-dimensional vision sensing approaches. It is important to consider that these reviews have since become outdated because computer vision techniques have continued to develop at an increasingly rapid pace in recent years. Newly proposed technologies regularly encourage the updating of manufacturing systems. Therefore, it is necessary to make a new survey of the state-of-the-art of computer vision techniques and their applications in different manufacturing tasks. We conduct a literature review because we try to present a critical analysis of existing research on the particular topic of computer vision techniques for manufacturing. Hence, we only include existing study results, but not any answers to a specific question, or any new data, experiments, or unpublished material in any form. A systematic review would indeed focus on providing an in-depth and detailed review of existing literature on a specific topic, and would address a specific, clearly formulated question with suitable responses. However, in a systematic review, some unpublished studies and reports may also be included. Therefore, a literature review is more suitable for our goal which is to help researchers stay updated about the latest research in this field and to identify gaps in the existing literature.

It is difficult to include all literature related to this topic because of the sheer size of the research community. As such, here we make an effort to cover the majority of important methods proposed in the past few decades, and discuss the latest results. Our literature search was mainly conducted based on three publication libraries: Scopus ¹, Google Scholar ², and IEEE Xplore Digital Library ³, and also from additional important journals within this scope. The search range of publication years is from 1970 to 2020. The keywords and inclusion/exclusion criteria we applied can be found in the following search query.

Longfei Zhou is with the Computer Science and Artificial Intelligence Laboratory at Massachusetts Institute of Technology, Cambridge, MA, 02139 USA.

L. Zhang (corresponding author) is with the School of Automation Science and Electrical Engineering at Beihang University, Beijing, 100191 China. e-mail: zhanglin@buaa.edu.cn

Nicholas Konz is with the Department of Electrical and Computer Engineering at Duke University, Durham, NC, 27708.

¹<https://www.scopus.com/>

²<https://scholar.google.com/>

³<https://ieeexplore.ieee.org/>

Title-Abstract-Keywords (manufacturing OR industr* OR production OR machining OR inspection OR 'quality control' OR design* OR test* OR 'modeling and simulation' OR planning OR scheduling OR assembl* OR alignment OR disassembl* OR AGV) AND Title-Abstract-Keywords ('computer vision' OR 'machine vision' OR 'robot vision' OR 'convolutional neural network')

In this paper, the methodology of the literature review is: 1) reviewing the important computer vision techniques including feature detection, recognition, segmentation, and 3D modeling; 2) discussing the system framework of computer vision in the manufacturing environment including a lighting module, a manufacturing system, a sensing module, computer vision algorithms, a decision-making module, and an actuator; and 3) surveying the newest and widely-implemented CV applications for different stages in the entire product life cycle, including product design, modeling and simulation, planning and scheduling, production process, inspection and quality control, assembly, transportation, and disassembly. In addition, we present a critical analysis of challenges and an outlook of future directions.

The rest of this paper is organized as follows. Section II reviews the recent results of important CV techniques including feature detection, recognition, segmentation, and 3D modeling. Section III discusses how computer vision function can be embedded into complex manufacturing systems by proposing a manufacturing-oriented computer vision system framework. In Section IV, computer vision applications for different stages in the entire product life cycle are surveyed. Section V presents a critical analysis of challenges and an outlook of future directions. Finally, conclusions are given in Section VI.

II. COMPUTER VISION TECHNIQUES

The purpose of a computer vision system is to generate a symbolic description of what is being imaged in a scene [10]. This description includes understandings of the scene and then be applied to guide the next operations of the robot system. There are multiple kinds of tasks and algorithms in the computer vision field such as detection, recognition, segmentation, and 3D reconstruction. In this section, we review the state-of-the-art of several important computer vision techniques.

A. Feature Detection

Visual feature detection (e.g., point, edge, and line detection) is the basis of many other computer vision algorithms. In some cases, we are more interested in a specific region of the image, such as human eyes, license plates and corner shapes, which are called keypoint features of the image. In some other cases, we care more about the edge features of objects to achieve identification, segmentation, or other goals. For both keypoint features and edge features, some CV algorithms are designed to identify these two kinds of features, and to match these features between different images. Commonly used evaluation metrics for feature detection include mean

absolute error (MAE), mean squared error (MSE), peak signal-to-noise ratio (PSNR) and structural similarity index measure (SSIM).

1) *Keypoint Detection*: There are mainly two types of methods for keypoint detection: local search detection and global search detection. Common local search detection methods include correlation methods, least squares methods, and learning-based methods [11], [12]. Local search detection methods are more suitable for scenes where images are taken continuously in a high frequency, such as video sequences. Different from local search methods, global search detection methods search the entire image and then match features based on feature appearance. Therefore, global search methods are more suitable for scenes where there are a lot of movements or appearance changes [13], [14].

2) *Edge Detection*: Edges in images often appear at boundaries between different objects, resulting in sudden changes in color and intensity. So, one method of edge detection is by the gradient through the image [15], even though the gradient is easily affected by noise. Hence, a low-pass filter (e.g., Gaussian filter) is needed to filter the image before gradient calculation. Horizontal and vertical convolution operations can be used for edge detection tasks. The zero-crossing method can also be used for edge detection in which zero points of a second-order derivative expression are searched to find the edges [16]. In recent years, deep learning methods, especially convolutional neural networks (CNNs), have been widely applied to edge detection and have obtained exciting results [17], [18].

B. Recognition

Recognition is another important task for computer vision techniques. From the perspective of target objects, recognition problems can be grouped into three categories, including instance recognition, class recognition and general category recognition. We also discussed action recognition in videos which is the current challenging problem and trend for the recognition task. Commonly used evaluation metrics for recognition are accuracy, recall, precision, F1 Score, and ROC/AUC curves.

1) *Instance Recognition*: In instance recognition, the goal is to identify a specific known object. Feature matching strategies can be used for this recognition problem [19]. Other instance recognition methods include viewpoint-invariant feature-based strategies [20], and sparse feature matching [21]. A popular application of instance recognition is face recognition [22]. Some learning-based approaches have also been widely applied to this problem, such as support vector machines (SVM) [23], boosting [24] and neural networks [25].

2) *Class Recognition*: Different from instance recognition, class recognition does not have a specific object as the target. In class recognition problems, the goal is usually to recognize the presence of an instance of a specific category of objects, such as cars or pedestrians. Class recognition problems can be considered as a specific classification problem in which the input is an image and output is the classification of that image.

Throughout the development history of computer vision, the rise of CNNs has been one of the most important breakthroughs for the recent rapid development of computer vision technologies. In recent years, numerous CNNs have been proposed and applied in classification problems, and their models are getting deeper and deeper. Sorted from earlier to recent works, popular CNNs include LeNet-5 [26], AlexNet [27], VGG-16 [28], R-CNN [29], Fast R-CNN [30], Inception Networks [31], [32], ResNet-50 [33], Xception [34], and ResNeXt-50 [35]. Before ResNet-50, most of the proposed CNNs only innovated by adding an increased number of layers in the network, alongside any other engineering changes needed for good performance. The “bottleneck” of this trend was that the model’s accuracy gets saturated and then rapidly decreases as the network became deeper and deeper. The seminal Residual Network firstly inserted shortcut connections in its deep model to deal with this problem [33]. The object detection neural network Fast R-CNN trained the deep VGG16 network 9x faster than R-CNN and 213x faster at test-time. Ren et al. proposed the Faster R-CNN [36] in which a Region Proposal Network was used to share full-image convolutional features with the detection network. For the very deep VGG-16 model, Faster R-CNN had a frame rate of 5fps with 0.732 mAP accuracy on PASCAL VOC 2007 and 0.704 mAP accuracy on PASCAL VOC 2012.

3) *General Category Recognition*: General category recognition is the most challenging recognition problem because we need to identify not only the locations of different objects in the image, but also what category each object belongs to. In this task, all different kinds of objects in the image need to be recognized. Common approaches of general category recognition include bag-of-words models and part-based models. In bag-of-words models, the distribution of visual words in the target image is compared to the training data [37]. In the part-based models, different parts of an image are considered separately so as to determine whether and where an object of interest exists [38]. Recent advances in this task have taken advantage of deep residual learning [33] and deep neural networks [39].

4) *Action Recognition*: The current challenging problem for the recognition task is action recognition. It is not difficult for humans to recognize actions in videos, but it is challenging for machines. Accurately recognizing the actions and behaviors in videos is of great significance for different scenarios, such as pedestrian motion monitoring in autonomous driving [40], and elderly fall detection [41]. In the manufacturing industry, identifying the behaviors of workers in the workshop has also been shown to be a very valuable approach to ensure production safety [42].

C. Segmentation

As one of the most classic tasks in the computer vision field, image segmentation aims to label pixels into different groups according to the objects that each pixel belongs to. Image segmentation can essentially be considered as a clustering problem. Early segmentation techniques usually used region division and merging methods [43]. Later segmentation algorithms applied indicators of consistency, such as

intra-regional consistency and inter-regional dissimilarity [44]. Other segmentation approaches include mean shift [45], graph-based merging [46], graph cuts-based Markov [47], and level sets [48]. In the learning-based segmentation algorithms, one commonly used loss functions is Dice loss which is based on the Dice coefficient. Dice coefficient can be defined as twice the true positive divided by the sum of twice the true positive, false positive and false negative. Another commonly used loss function for deep-learning segmentation algorithms is the Intersection Over Union (IoU).

The latest segmentation algorithms based on machine learning in recent years include Mask RCNN [49] and Dual Attention Network [50]. Recently, U-Net has shown good performance in the segmentation tasks of medical images [51]. There are some variants of U-Net such as Attention U-Net [52], U-Net++ [53], ResUNet++ [54], and TransUNet [55]. The performance of these algorithms not only depends on the algorithm design but also depends on the datasets. Current challenges and trends in segmentation include 3D segmentation and 4D segmentation problems. The goal of 3D segmentation is to segment 3D images in three spatial directions, while 4D segmentation is to segment 4D data which also includes the time dimension in addition to spatial dimensions. The 3D U-Net has been applied to 3D segmentation problems in medical imaging [56] and additive manufacturing defects in X-ray Computed Tomography images [57].

D. 3D Modeling

3D modeling in computer vision can be categorized into two problems: stereo correspondence and 3D reconstruction. Stereo correspondence is the process of generating a 3D model of an object from two or more images of the same object or scene, while 3D reconstruction is to generate a 3D model of an object from only one image [58]. It is a challenge to design a good loss function to evaluate the predicted 3D point cloud and ground-truth. One option is evaluating how well the projections of predicted 3D point clouds cover the ground-truth object’s silhouette [59].

1) *Stereo Correspondence*: A common stereo correspondence method is to find matching pixels in multiple images and map their position in the 2D images to 3D positions in the 3D model. Popular methods for stereo correspondence include epipolar geometry [60], sparse correspondence [61] and dense correspondence [62].

2) *3D Reconstruction*: The earliest approach for 3D reconstruction is to predict object shape from visual shading, which was first proposed by Horn in 1970 [63]. Later, other “shape from X” methods were proposed such as shape from texture [64] and shape from focus [65]. Other 3D reconstruction methods include active rangefinding [66], and model-based reconstruction which has been widely applied to architectural 3D modeling [67]. Recently, deep learning-based algorithms promoted significant improvement in system performance of 3D reconstruction [68], [69].

III. MANUFACTURING-ORIENTED COMPUTER VISION SYSTEM

In this section, we aim to discuss the role of CV in manufacturing systems and the closed-loop from manufactur-

Fig. 1. Manufacturing-oriented computer vision system framework.

ing environments to CV to manufacturing environments. The system framework of computer vision in the manufacturing environment is proposed as shown in Fig. 1. There are optical devices (i.e., the lighting module, the sensing module), hardware (i.e., the manufacturing system, actuators), software (i.e., the CV system, the decision-making module) and data (e.g., optical images, feature descriptions and decision signals) in this framework.

The lighting module is the light source to provide lighting for the manufacturing system so that the sensing module is able to capture images of the manufacturing system. One problem is how to design and set a suitable lighting module to improve the performance of the image sensing process. An important goal of the lighting module is to provide uniform illumination for the scene. Using multiple light sources showed better performance than a single light source for obtaining uniform illumination on the target scene [70]. To obtain uniform near-field irradiance, the angled LED ring array was applied through simplifying the non-rotational symmetric irradiance distribution [71].

In most cases, the sensing module of a computer vision system is one or multiple cameras, which can be categorized into fixed cameras and mobile cameras. Fixed cameras are usually placed on production lines, while mobile cameras are usually placed on robots, such as assembly robots and automated guided vehicles (AGVs). In order to acquire high-quality images even when there is an obstacle along the camera line-of-sight on production lines, an automatic camera placement strategy was presented [72]. Application of handheld computer vision devices can avoid maintaining a consistent camera distance and light source [73].

The CV system takes digital images captured through the sensing module as inputs, and outputs detected features and descriptions of the images. Those popular CV algorithms which are discussed in Section II such as feature detection, recognition, segmentation, and 3D modeling algorithms can be applied here in the CV system.

Detection results of these CV algorithms are then applied to support the decision-making process. The decision-making action is actually the execution process of different rules and strategies in the decision-making module. Typical decision algorithms include types of priority-based rules [74], heuristic algorithms [75] and intelligent optimization algorithms [76].

Decision signals generated by the decision-making module then controls the following actions of the actuator. The action sequence of the actuator reflects interactions between the actuator and the manufacturing system.

Computer vision techniques have also given significant support to the development of digital twin technologies in manufacturing systems [77]. Taking advantage of modeling and simulation technology, the digital twin is playing an increasingly important role in the informatization and intelligentization of manufacturing systems. Computer vision techniques are used to give support to the operation and maintenance of digital twins in manufacturing environments. Using image processing and computer vision technologies, we are able to capture the imaging information of manufacturing systems in real-time (such as the locations, movements, and defects of production parts), and extract abstract information of system states based on these images. The system state information helps to keep the state consistency between the digital twin and the real system. In addition to computer vision, there are other kinds of sensors such as distance sensors, thermal sensors, and motion sensors in the manufacturing systems. Feedback from all these sensors and actuators is given to manufacturing and to be used on-demand.

IV. COMPUTER VISION APPLICATIONS IN MANUFACTURING

From the perspective of the entire product lifecycle, the manufacturing process can be briefly divided into multiple stages including product design, modeling and simulation, planning & scheduling, production process, inspection and quality control, assembly, transportation, and disassembly. In this section, we not only survey the state-of-the-art of all these different application directions but also discuss the important challenges of each direction at the end of each subsection. There is surely overlap between these stages, such as the inspection operations in production and assembly stages, and assembly operations in production processes, depending on different manufacturing industries. We show statistics of the occurrences of keywords of different manufacturing stages in computer vision research papers in recent decades, as shown in Fig. 2. It can be seen that computer vision techniques have been widely applied to inspection, production, and assembly in the manufacturing industry. In this section, we review computer vision applications in these different manufacturing stages.

A. Product Design, Modeling and Simulation

As the first stage of the product life cycle, product design aims to create a new product or a new version of an existing product. Computer-Aided Design (CAD) and Computer-Aided Manufacturing (CAM) techniques have been widely used in the manufacturing industry for decades [78]. In CAD, computers are applied to support the creation, modification, analysis, and optimization of product design. Product design results by CAD are in the form of 3D models. These 3D models are electronic files that can be used for design operations and other manufacturing processes.

Fig. 2. Statistics of keywords of different manufacturing stages in computer vision research papers from 1970 to 2020.

One of the common applications of computer vision in product design is reconstructing 3D models from 2D images of existing products. For example, range data can be used to generate the geometric surface models of viewpoint invariant in CAD/CAM systems [79]. Reverse engineering innovative design was applied to generate the feature-based parametric solid models from scanned data [80]. The constructed models can be applied to the downstream analysis and rapid product prototyping. 3D poses of objects can also be estimated by using a single 2D image and spatial information in CAD models [81].

Another important application of computer vision in product design is the simulation and validation of product designs, especially for vehicle instrument clusters. Modeling and simulation are large topics that can exist in different stages, especially when we consider a digital model as a “digital twin”. Also, it is hard to separate these stages because they do overlap with each other. Hence, the stages here in this paper are relatively separated. In this section, we group together product design, modeling and simulation considering two reasons. Firstly, modeling and simulation techniques are still mostly used in product design activities for now, even though simulation can also be used in subsequent stages. Secondly, there are relatively fewer references about CV techniques in product design, modeling and simulation as we found. Therefore, we merged them together to discuss here. Hardware-in-the-loop simulation and machine vision can be combined to achieve automated design testing and validation of vehicle instrument clusters [82]. Vision-based machine learning systems can be applied in monitoring vehicle instrument clusters in Hardware-in-the-loop simulation and identifying faults in clusters’ design [83].

With the rapid development of CAD software, multi-disciplinary modeling-based product design has become the trend. Therefore, more and more products have their own 3D digital models at earlier stages, which leads to a reduction in

the demand for 3D model reconstruction of products. How to find new applications of CV technology in product design and modeling is an interesting question.

B. Planning and Scheduling

After product design is validated, a production plan is made to identify what to do next, followed by schedules of when and how to execute the planning [84]. One important application of computer vision in production planning and scheduling is aiding the generation of lumber production planning because the internal texture of lumber significantly affects the quality of sawing plans. For example, internal defects can be detected and localized through analyzing Computed Tomography (CT) images of lumber [85]. Then, these defect detections can be used to formulate optimal sawing strategies [86]. Fig. 3 shows some results of detection and localization of knots and holes through a sequence of CT image slices.

Fig. 3. Results of detection and localization of knots and holes in lumber [86], including: (a) results through a continuous sequence of CT images; (b) semi-transparent view of the detection results.

Additionally, computer vision has also been used for production planning in additive manufacturing and computer numerical control (CNC). One option is nesting irregular 3D printing parts in the printing space through representing 3D printing models with pixels [87]. Path plans for golf-club head robotic welding can be generated by recognizing the weld seam in images [88].

There are usually different types of uncertain events in real manufacturing workshops, such as machine breakdown and product quality issues. Hence, a big challenge in applying CV to planning and scheduling is how to update production plans and schedules dynamically according to vision information in real manufacturing workshops.

C. Production Process

The production process is an important stage in the entire product life cycle. Although different industries have different forms of production processes, the basic concept is the same that is transforming raw materials into products through a series of operations. Applications of computer vision in production processes mainly include: production process monitoring, robot guidance, part recognition and classification, 3D position measurement, and production safety monitoring.

1) *Production process monitoring*: To control the trajectory of molten rock in the industrial production of mineral wool, detection algorithms were developed to detect variations of the jet trajectory with time, and the detection results were then used to generate control strategies [89]. An automatic tool was presented to monitor the raising process of a raised cloth by measuring the height and the density of the fibers [90]. In the iron industry, machine vision systems were also applied on flotation cells in iron flotation plants to monitor different features of bubbles such as sizes, numbers, velocity and stability. Froth phases under various production conditions were also recognized through analyzing froth images [91]. Both the wavelet-based and the histogram matching approach in the spatial domain were applied to extract pattern features of a multi-crystalline solar wafer [92]. The feature data was then used to control the flotation process and to improve system performance [93]. In order to overcome the difficulty in labeling and tracking steel materials caused by high temperature and inaccessibility, an online label detection-tracking method was used to label and track steel materials [94].

2) *Robot guidance*: CV-based robot guidance approaches include stereo vision and photogrammetry, projected texture stereo vision, time of flight, structured white light, structured blue LED light, light coding, laser triangulation, and etc. These different visual robot guidance methods can be compared from aspects of accuracy, range, weight, safety, processing time and environmental influences [95]. The fuzzy logic-based controlling method was also applied to intelligent robotics with feedback through a vision module [96]. Building a safe stable collaboration between humans and robots is an important task in robot control. Virtual 3D models of robots and real images of human operators from depth cameras are combined to detect and avoid collisions between humans and robots [97]. Fig. 4 shows a scenario of human-robot collaborative assembly where the robot follows the operator's hand to offer assistance during a collaborative assembly operation.

Fig. 4. A scenario of human-robot collaborative assembly where the robot follows the operator's hand to offer assistance during a collaborative assembly operation [97].

3) *Part classification*: Classification is also an important application of computer vision in the production process for both parts and products, generally based on industrial robotics for recognition and classification tasks [98]. There are different classification methods such as the Mahalanobis-Taguchi system and principal component feature overlap measure [99]. The principal component feature overlap measure showed higher signal-noise ratios and better classification performance than the Mahalanobis-Taguchi system with a lower number of classification features. Color and texture features can also be used in classifying industrial products such as wood and roofing shingles [100]. In addition, vision-based classification has been applied to other different types of products such as olives for oil production when the milling process starts [101], and the batch measurement of sheet-like products [102].

4) *3D position measurement*: Another important task of computer vision systems in the production process is 3D position measurement of objects such as parts and tools. Early techniques mainly achieved the recognition and measurement of specific parts or features of different products (e.g., the locations of screw holes) [103]. Later, some contour-based approaches were used to recognize the parts in random positions and orientation of parts [104]. Besides, the Expectation-Maximization algorithm was applied to estimate the pose of the target and build the dynamics model of the target so that the slow dynamics of sensors could be compensated [105]. More recently, laser dynamic triangulation was applied to determine 3D coordinates of an observed object in the manufacturing robot vision system [106].

5) *Production safety monitoring*: Computer vision systems were also used to monitor the safety of humans and devices in the production process. An intelligent vision-based approach for helmet identification for work safety was applied to ensure the personal safety of the workers in workshops [107]. A color-based hybrid descriptor based on SVM was used for the helmet identification task. For the safety of machining, a chain-processing-based system was developed to automatically avoid collision between tools and machining set-up components by checking whether the actual machining set-up is in conformity with the desired 3D CAD model [108].

One of the challenges in the production process is how to quickly and accurately generate control instructions for the next action of actuators in an uncertain production environment. One possible way is to build a set of possible situations and possible actions and then use computer vision algorithms to recognize the types of situations. Applications include strain control instructions and safety alarms of industrial robots.

D. Inspection

Computer vision technologies have been widely applied to inspection, testing and quality in different manufacturing industries. Inspection is an essential part of quality control in manufacturing, including measurement, examination and testing. The process of inspection is assessing some features of a part or product to determine whether this part or product meets the requirements of quality control through comparing the measuring results with specified requirements [109]. Some

CV-based inspection results in different manufacturing industries are shown in Fig. 5, including mechanical manufacturing, automotive industry, textile industry and 3D printing industry.

Fig. 5. Inspection results in different manufacturing industries based on computer vision techniques, including: (a) mechanical manufacturing (original image from a laser powder bed); (b) mechanical manufacturing (inspection results) [?]; (c) automotive industry (original image of the vehicle surface); (d) automotive industry (inspection results) [110]; (e) textile industry (region of interest in an image of textile materials); (f) textile industry (inspection results) [111]; (g) 3D printing industry (region of interest in a 3D printing image); (h) 3D printing industry (inspection results) [112].

1) *Mechanical manufacturing*: For mechanical machining, machine vision-based approaches have been applied to inspect turned surface in finishing operations [113], to detect defects for spring clamps [114], and to detect damaged parts on assembly lines [115]. Web-based quality control systems were applied to remotely monitor the quality inspection process in production floors [116]. Different machine vision classification methods, including blob analysis, optical flow and a running average method, were tested to detect and classify multiple faults of automated assembly machines [117]. For automatic

localization of exceptions of powder spreading, a CNN was used for pixel-wise semantic segmentation of layer-wise powder bed images. [118].

2) *Automotive industry*: In the automotive industry, computer vision techniques have been widely used in different tasks such as surface quality detection [119] and wheel alignment [120]. The image fusion algorithm and background extraction based on a local directional blurring method were applied to detect defects on specular surfaces not only in areas with smooth slope changes but in edges, corners and deep concavities [110]. Candidate defect regions from images of five sides of automobiles were extracted by a multi-scale matrix fusion method. These candidate defect regions were classified into pseudo-defects, dents and scratches by features such as shape, size, statistics and divergence [121].

3) *Textile industry*: Different CV algorithms have also been applied to quality inspection in textile industry, such as acrylic fiber production [122]. To ensure the highest performance of safety belts, a CV-based system was developed to detect defects such as fiber breaks, knots, and thickness variations [123]. CV algorithms have also been applied to detect fiber orientation under industrial conditions in deterministic time conditions [111].

4) *3D printing*: With the recent rapid development of additive manufacturing technologies, 3D printing has gained more and more attention. Some CV-based systems have also been applied to printing quality inspection in 3D printing process. A multi-material 3D printing platform enabling self-calibration of printheads was developed including a vision module [124]. Augmented reality techniques were also used to monitor the 3D printing process [125] in which a virtual 3D model of the object to be printed was superimposed to the real object. The shape of the object was able to be viewed and compared with its virtual 3D model during the printing process which could be stopped if any printing errors were detected. Multi-view vision detection was used to detect defects in the 3D printing process by changing the field of view in the printing progress [112]. Future development directions for CV-based 3D printing applications are to developing metrics to obtain a better classification accuracy and to develop benchmark databases to include quality scores [126].

5) *Other industries*: Besides, CV-based inspection techniques have also been applied to other industries such as electronics manufacturing [127], [128], display manufacturing [129]. The development of novel smart imaging devices and more efficient learning algorithms is essential for improve the performance of industrial inspection tasks in manufacturing. CV techniques have also been used to alignment tasks in other industries such as precise alignment between optical components in modern corona-graphic systems [130], tile alignment inspection [131], and alignment of raw parts before the machining process starts [132]. The local matching algorithm based on adaptive weight was applied in checking the logistics product anti-counterfeiting identity so that forgery inventory in logistics products can be prevented [133].

The vision-based inspection usually generates large amounts of imaging data. Not all the data is utilized directly but it is very expensive to store this large amount of data. A very

important challenge is how to effectively manage and utilize the large amount of inspection data collected during product inspection stage.

E. Assembly

Assembly, also referred to as progressive assembly, is an important process in product manufacturing, especially for discrete manufacturing such as within the automotive and aviation industries. Parts are assembled into semi-finished products while semi-finished products are then assembled into finished products through assembly lines.

1) *Automatic assembly*: Automatic assembly is an important task in the assembly process in manufacturing. Generally, vision systems are applied to provide assistance and guidance for automatic assembly. The possibility of a totally automatic assembly of motor stators based on CV has been studied [134]. In the assembly process of printed circuit boards, shape-based recognition can be used to guide assembling flexible printed circuit cable onto hard disk drives [135]. Combining 3D object registration and mixed reality can improve the assembly efficiency and quality of cabin products [136]. For automatically picking and placing parts on assembly lines, a flexible assembly system consisting of a vision module and a robotic module was designed and developed [137]. Besides, the graphical user interface based on accurate object recognition and simulation was presented to assist operators to perform path planning of parallel manipulators in packaging and assembling products [138].

2) *Assembly error detection*: Assembly error detection is another important application of CV in the assembly process. Statistical pattern recognition methods can be applied to predict assembly errors when the geometric trajectory of a mated part and the relational position of a base part are out of the allowable tolerance in a robotic assembly system [139]. Vision-based identification algorithms can also be applied to recognize unique features around fasteners so that the assemble results of fasteners are able to be verified [140]. This kind of CV application can also be found in automotive factory assembly lines such as identification of human errors involved in the bolt securing process [141]. This technique has been applied to the actual automotive manufacturing industry (e.g., Mercedes-Benz).

3) *Other assembly applications*: Augmented reality (AR) techniques have been combined with CV for decades to guide assembly sequences [142], such as interactive tools for evaluating assembly sequences [143]. CV can also be applied to automatically segment and recognize continuous human motions in mechanical assembly operations [144]. Another interesting problem in the assembly process is the single-image-based 3D part assembly problem which involves challenges such as ambiguity among parts and 3D pose prediction [145]. In the applied two-pipeline network, 2D-3D correspondences were used to infer relationships between parts. The results of the single-image-based 3D part assembly problem through different methods are shown in Fig. 6. There are some open imaging datasets of the assembly process such as the dataset of pixel-level labeled images of hands performing different assembly tasks [146].

Fig. 6. Results of single-image-based 3D part assembly problem through different methods [145].

A very important challenge in the assembly process is how to achieve safe, flexible and intelligent human-robot cooperation in executing assembly tasks. The latest machine learning-based methods should be helpful for us to make the machines be trained closer to the ideal situation.

F. Transportation

Transportation activities in manufacturing can be material or part transportation through automated guided vehicles (AGV) within a single workshop and can also be logistics between factories through various transportation services. AGV is an autonomous transportation system applied in manufacturing workshops to deliver materials, parts or products around an industrial environment. AGV generally consists of a navigation system, a power system, and a controlling system. The path planning of AGV is implemented through programming inside the controlling system according to specific requirements in workshops. Earlier AGV usually followed along marked lines or wires on the floor in factories and warehouses. Recent AGV have been designed to use radio waves, vision cameras, magnets, or lasers for their navigation functions.

Computer vision techniques have been mainly applied in the navigation module of AGV to provide vision-related information for the AGV control [147], such as obstacle detection [148], and path planning in transportation tasks [149]. Navigation systems in AGV can be grouped into locally guided navigation and remotely guided navigation systems. Locally guided navigation mainly includes model-based approaches [150], downward vision system, laser ranging system, and stacking vision [151]. In addition, deep learning algorithms have also used within computer vision modules of AGV robots to improve their navigation reasoning capabilities [152]. For remotely guided AGV, the web-controlled vehicle system is an important direction in which AGV can be driven from

anywhere through the Internet over a secured server [153]. Support vector machine-based segmentation methods have also been used to distinguish the original color features of path images from their illumination artifacts [154].

In addition to AGV, computer vision technologies have also been applied to large-size industrial autonomous transportation systems such as forklift trucks and logistics vehicles. As the main sensing module for the autonomous guidance system of a forklift, computer vision systems can support both the navigation function and the load recognition function of the forklift truck [155]. A covariance matrix algorithm in the computer vision system was improved to detect and track moving objects in logistics systems [156]. Self-driving vehicles have also been designed for industrial logistics using autonomous driving technologies [157], [158].

A big challenge of the vision techniques for transportation systems in manufacturing is the variation between images such as shadows and highlights in the captured images caused by complex illumination conditions. More effective denoising and normalization techniques need to be combined with CV algorithms in navigation modules to achieve more robust transportation systems.

G. Disassembly

As a key step for the effective disposal of end-of-life products, disassembly should be included in the entire product life cycle. However, automatic product disassembly did not get much developed in actual manufacturing industries due to cost and technical limitations until robotics and computer-aided manufacturing techniques started to be rapidly improved recently [159]. Traditional disassembly was usually done manually by workers, which is a labor-intensive task and expensive. Robotic automated disassembly systems can greatly reduce disassembly costs [160]. According to the degree of automation of the disassembly process, disassembly systems can be classified into semi-automatic disassembly systems and fully-automated disassembly systems. Computer vision is one of the necessary technologies to realize an automatic disassembly system, not only for the generation of the disassembly sequence plan, but also for the navigation and control of automatic disassembly robots.

According to the difference of prior knowledge, product disassembly can be classified into the product structure-based disassembly and disassembly without product structure knowledge. An important task in product disassembly is to generate an effective disassembly sequence plan. This generation task is always difficult due to complicated constraints between components, especially for complex products. For the disassembly sequence generation problem, computer vision techniques have been mainly applied to identify possible assembly structures of the product according to images of the product, to support the follow-up disassembly sequence generation tasks [161]. Another application of computer vision technologies in product disassembly is being used as the visual recognition and navigation system of disassembly robots to support the motion control of disassembly robots. The robotics automatic disassembly technology has been widely used in

the disassembly tasks, especially in the automotive industry, electronics industry [162], [163], and screen products [164]. Some characteristics of screws such as gray-scale, color-scale and depth can be used in image processing algorithms of screw detection to support automatic disassembly tasks [165]. Most current automatic disassembly systems are not fully automated, which means human participation is needed in the disassembly process. Therefore, how to achieve effective and safe human-robot interaction in the disassembly process is an important goal for current automatic disassembly systems [166].

V. CRITICAL ANALYSIS AND FUTURE DIRECTIONS

A. Critical Analysis

Although computer vision techniques have been widely applied to almost every stage in the entire manufacturing process of different industries, there are still some long-standing issues and new challenges to be considered to achieve a better place. Based on the survey, we here give a critical analysis of the state-of-art and challenges in computer vision applications for manufacturing industries in terms of implementation, data collection, data pre-processing, data labeling and benchmarks.

1) *Challenge of implementation*: The rapid development of computer vision technologies (e.g., CNNs and deep learning models) has brought a lot of the latest exciting results to the community. However, the computer vision algorithms which are currently used in actual manufacturing systems are still those relatively classic algorithms, such as support vector machines (SVM) and k-nearest neighbors algorithms (KNN). The latest achievements of the computer vision community are not being quickly applied to the thorny problems facing real manufacturing environments. One possible reason is that large-scale companies usually have sufficient resources to implement the latest research results, but meanwhile, the manufacturing systems of these large-scale companies are more complicated, involving different software systems, hardware systems in different departments (e.g., R&D, Design, Production, Testing, Sales, etc.). This complexity brings more challenges to the application of the latest CV research results in the actual manufacturing environments. Another possible reason is that research activities conducted by the computer vision community are usually based on ideal problem models with a large amount of labeled data, which is different from the issues in the actual manufacturing systems. This also leads to the difficulty of applying the latest research results.

2) *Challenge of data collection*: Data is always one of the most important parts for machine learning and computer vision tasks. Even though more data can be obtained with the rapid development of the Internet of Things (IoT) and sensor technologies, it is still a challenge to collect high-quality data especially for 3D surfaces and reflecting surfaces in some complex manufacturing environments with lighting problems. There are various reasons why data collection is challenging. One reason is the low-quality lighting problem in some complex manufacturing scenarios. Another possible reason is that the surface of the object is reflective resulting in the bias between the obtained images and the real object.

3) *Challenge of data pre-processing*: As more and more data sensing devices are deployed in manufacturing systems, a greater volume of structured and unstructured data (e.g., high-resolution images and videos) are collected from manufacturing production sites. However, not all the collected images and videos are worthy to be sent to the computer vision system for further processing. Concerning this demand, the challenge is the lack of effective data pre-processing mechanisms for a large amount of original manufacturing data. Some large-scale companies choose to temporarily store all collected data in databases for a period of time and delete the early data such as the data from three months ago at a regular interval. This undoubtedly leads to much more expensive data storage and lower data processing efficiency.

4) *Challenge of data labeling*: Although more and more high-quality visual data can be collected from the manufacturing sites, the collected raw data usually lacks necessary labels which are important to supervised learning algorithms. Manually labeling large amounts of raw data is expensive. The current challenge is a lack of effective algorithms for handling non-labeled data, as well as lacking methods to automatically label original visual data. In order to apply deep learning-based vision technologies to different manufacturing scenarios where a large amount of image data can be obtained, more effort needs to be made in this regard. One possibility is recent self-supervised deep learning methods that perform automatic labeling of unlabeled data such as SimCLR [167].

5) *Challenge of benchmarks*: There are already some task-oriented benchmarks such as COCO ⁴, the Multiple Object Tracking Benchmark ⁵, and the UA-DETRAC Benchmark ⁶ to compare and evaluate newly proposed algorithms. But it is difficult to apply these benchmarks to specific manufacturing cases because these benchmarks are mostly designed for particular tasks such as detection of vehicles and pedestrians. Hence, more benchmarks for manufacturing applications are needed to be built for computer vision techniques to be further continuously applied in manufacturing industries. MVTec is a recent commonly used dataset for benchmarking anomaly detection methods with a focus on industrial inspection [168], [169].

B. Future Directions

Based on the above analysis of challenges, some future directions are proposed here as follows to accelerate the application of computer vision techniques in manufacturing industries.

1) *Benchmarking*: One direction is to build benchmark standard datasets for specific tasks in the manufacturing environment to evaluate the performance of different applied computer vision algorithms. There are some steps needed to be taken to build the benchmarks for different CV tasks in manufacturing systems, such as deciding which CV task and which companies to be included, collecting and analyzing data, making the metrics of performance measurement for CV algorithms and methods.

2) *Big data pre-processing*: It is also a trend to develop effective pre-processing mechanisms for specific image data, including data cleaning, and automatic or semi-automatic data annotation methods for original image data. It is difficult to establish fully automated equipment for different fields and different departments to achieve automated data cleaning and labeling. One strategy is to use the domain knowledge from experts in different departments to help with processing data in different structures and types. Another very important strategy is to establish the connection between the upstream production stages and the downstream data processing stage.

3) *Non-supervised learning*: Another direction is to develop effective methods that can process non-annotated manufacturing image data, so that deep learning-based computer vision algorithms can be then applied to those manufacturing scenarios where a large amount of image data can be obtained. Possible directions to deal with this issue include newly-developed machine learning approaches such as one-shot learning, transfer learning, semi-supervised learning.

4) *Task-oriented models*: It is also important to develop deep learning-based computer vision models (e.g., CNNs) for future specific image processing tasks in manufacturing to improve the usefulness of manufacturing data and algorithm efficiency. This work generally requires significant training computations of different network structures and parameters to obtain satisfactory results.

5) *5G-involved CV*: The 5G communication technology, with its rapid development [170], will serve as one of the catalysts to provide new opportunities for the application of computer vision in manufacturing and bring new development directions for solving long-standing bottlenecks. How to optimize existing manufacturing computer vision systems considering 5G and how to design new computer vision applications based on the new 5G architecture to improve system performance are also interesting directions.

VI. CONCLUSIONS

Considering the rapid development of computer vision techniques, we presented a comprehensive review of several important computer vision techniques relevant to manufacturing, as well as their latest applications in different stages of the product life cycle within the entire manufacturing process. These surveyed computer vision techniques include feature detection, recognition, segmentation, and 3D modeling. A system framework of computer vision in the manufacturing environment was proposed consisting of lighting module, manufacturing system, sensing module, computer vision algorithms, decision-making module, and actuator. Applications of computer vision in different stages in the product life cycle were then surveyed, including product design, modeling and simulation, planning and scheduling, production process, inspection and quality control, assembly, transportation, and disassembly. Although computer vision techniques have been widely applied to almost every stage in the entire manufacturing process of different industries, there are still some long-standing issues and new challenges that are discussed in the critical analysis. Future development directions include building benchmarks for specific manufacturing image

⁴<http://cocodataset.org/>

⁵<https://motchallenge.net/>

⁶<https://detrac-db.rit.albany.edu/>

processing tasks, developing effective methods of processing non-annotated data, developing effective data pre-processing mechanisms (e.g., data cleaning, automatic or semi-automatic data annotation methods), developing computer vision models (e.g., CNNs) for specific manufacturing tasks to improve the usefulness of manufacturing data and the efficiency of CV algorithms, and new opportunities kindled by 5G.

REFERENCES

- [1] Z. Liu, H. Ukida, K. Niel, and P. Ramuhalli, "Industrial inspection with open eyes: Advance with machine vision technology," in *Integrated Imaging and Vision Techniques for Industrial Inspection*. Springer, 2015, pp. 1–37.
- [2] J. K. West, "Machine vision in the real world of manufacturing," *Computer Design*, vol. 22, no. 5, p. 89, 1983.
- [3] G. J. Agin, "Computer vision systems for industrial inspection and assembly," *Computer*, no. 5, pp. 11–20, 1980.
- [4] A. Soini, "Machine vision technology take-up in industrial applications," in *ISPA 2001. Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis. In conjunction with 23rd International Conference on Information Technology Interfaces (IEEE Cat. IEEE, 2001, pp. 332–338.*
- [5] J. A. Noble, "From inspection to process understanding and monitoring: a view on computer vision in manufacturing," *Image and vision computing*, vol. 13, no. 3, pp. 197–214, 1995.
- [6] H. Golnabi and A. Asadpour, "Design and application of industrial machine vision systems," *Robotics and Computer-Integrated Manufacturing*, vol. 23, no. 6, pp. 630–637, 2007.
- [7] C. Aldrich, C. Marais, B. Shean, and J. Cilliers, "Online monitoring and control of froth flotation systems with machine vision: A review," *International Journal of Mineral Processing*, vol. 96, no. 1–4, pp. 1–13, 2010.
- [8] N. Neogi, D. K. Mohanta, and P. K. Dutta, "Review of vision-based steel surface inspection systems," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, p. 50, 2014.
- [9] X. Wang, "Three-dimensional vision-based sensing of gtaw: a review," *The International Journal of Advanced Manufacturing Technology*, vol. 72, no. 1–4, pp. 333–345, 2014.
- [10] B. K. P. Horn, *Robot vision*. MIT press, 1986.
- [11] M. Ozuysal, M. Calonder, V. Lepetit, and P. Fua, "Fast keypoint recognition using random ferns," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 3, pp. 448–461, 2009.
- [12] S. Hinterstoisser, S. Benhimane, N. Navab, P. Fua, and V. Lepetit, "Online learning of patch perspective rectification for efficient object detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*. IEEE, 2008, pp. 1–8.
- [13] M. Brown and D. G. Lowe, "Automatic panoramic image stitching using invariant features," *International journal of computer vision*, vol. 74, no. 1, pp. 59–73, 2007.
- [14] R. Fergus, P. Perona, and A. Zisserman, "Weakly supervised scale-invariant learning of models for visual recognition," *International journal of computer vision*, vol. 71, no. 3, pp. 273–303, 2007.
- [15] S. Jain, "Edge detection using fuzzy gradient information," Aug. 11 2020, uS Patent 10,740,903.
- [16] G. Shrivakshan and C. Chandrasekar, "A comparison of various edge detection techniques used in image processing," *International Journal of Computer Science Issues (IJCSI)*, vol. 9, no. 5, p. 269, 2012.
- [17] Y. Liu, M.-M. Cheng, X. Hu, K. Wang, and X. Bai, "Richer convolutional features for edge detection," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 3000–3009.
- [18] C. Jeong, H. S. Yang, and K. Moon, "A novel approach for detecting the horizon using a convolutional neural network and multi-scale edge detection," *Multidimensional Systems and Signal Processing*, vol. 30, no. 3, pp. 1187–1204, 2019.
- [19] J. Ponce, M. Hebert, C. Schmid, and A. Zisserman, *Toward category-level object recognition*. Springer, 2007, vol. 4170.
- [20] J. Sivic and A. Zisserman, "Efficient visual search of videos cast as text retrieval," *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 4, pp. 591–606, 2008.
- [21] K. Lai, L. Bo, X. Ren, and D. Fox, "Sparse distance learning for object recognition combining rgb and depth information," in *2011 IEEE International Conference on Robotics and Automation*. IEEE, 2011, pp. 4007–4013.
- [22] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," *Electronics*, vol. 9, no. 8, p. 1188, 2020.
- [23] G. Guo, S. Z. Li, and K. Chan, "Face recognition by support vector machines," in *Proceedings fourth IEEE international conference on automatic face and gesture recognition*. IEEE, 2000, pp. 196–201.
- [24] P. Viola and M. J. Jones, "Robust real-time face detection," *International journal of computer vision*, vol. 57, no. 2, pp. 137–154, 2004.
- [25] Y. Sun, D. Liang, X. Wang, and X. Tang, "Deepid3: Face recognition with very deep neural networks," *arXiv preprint arXiv:1502.00873*, 2015.
- [26] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [27] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.
- [28] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
- [29] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2014, pp. 580–587.
- [30] R. Girshick, "Fast r-cnn," in *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.
- [31] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818–2826.
- [32] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-first AAAI conference on artificial intelligence*, 2017.
- [33] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [34] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.
- [35] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated residual transformations for deep neural networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1492–1500.
- [36] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in *Advances in neural information processing systems*, 2015, pp. 91–99.
- [37] L. Wu, S. C. Hoi, and N. Yu, "Semantics-preserving bag-of-words models and applications," *IEEE Transactions on Image Processing*, vol. 19, no. 7, pp. 1908–1920, 2010.
- [38] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object detection with discriminatively trained part-based models," *IEEE transactions on pattern analysis and machine intelligence*, vol. 32, no. 9, pp. 1627–1645, 2009.
- [39] Z. Wu, C. Shen, and A. Van Den Hengel, "Wider or deeper: Revisiting the resnet model for visual recognition," *Pattern Recognition*, vol. 90, pp. 119–133, 2019.
- [40] L. Chen, N. Ma, P. Wang, J. Li, P. Wang, G. Pang, and X. Shi, "Survey of pedestrian action recognition techniques for autonomous driving," *Tsinghua Science and Technology*, vol. 25, no. 4, pp. 458–470, 2020.
- [41] X. Wang, J. Ellul, and G. Azzopardi, "Elderly fall detection systems: A literature survey," *Frontiers in Robotics and AI*, vol. 7, p. 71, 2020.
- [42] C.-L. Yang, S.-C. Hsu, Y.-W. Hsu, and Y.-C. Kang, "Human action recognition on exceptional movement of worker operation," in *International Conference on Applied Human Factors and Ergonomics*. Springer, 2021, pp. 376–383.
- [43] S. L. Horowitz and T. Pavlidis, "Picture segmentation by a tree traversal algorithm," *Journal of the ACM (JACM)*, vol. 23, no. 2, pp. 368–388, 1976.
- [44] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
- [45] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.
- [46] P. F. Felzenszwalb and D. P. Huttenlocher, "Efficient graph-based image segmentation," *International journal of computer vision*, vol. 59, no. 2, pp. 167–181, 2004.

- [47] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient nd image segmentation," *International journal of computer vision*, vol. 70, no. 2, pp. 109–131, 2006.
- [48] D. Cremers, M. Rousson, and R. Deriche, "A review of statistical approaches to level set segmentation: integrating color, texture, motion and shape," *International journal of computer vision*, vol. 72, no. 2, pp. 195–215, 2007.
- [49] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask r-cnn," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969.
- [50] J. Fu, J. Liu, H. Tian, Y. Li, Y. Bao, Z. Fang, and H. Lu, "Dual attention network for scene segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 3146–3154.
- [51] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.
- [52] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz et al., "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.
- [53] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, 2018, pp. 3–11.
- [54] D. Jha, P. H. Smedsrud, M. A. Riegler, D. Johansen, T. De Lange, P. Halvorsen, and H. D. Johansen, "Resunet++: An advanced architecture for medical image segmentation," in *2019 IEEE International Symposium on Multimedia (ISM)*. IEEE, 2019, pp. 225–2255.
- [55] J. Chen, Y. Lu, Q. Yu, X. Luo, E. Adeli, Y. Wang, L. Lu, A. L. Yuille, and Y. Zhou, "Transunet: Transformers make strong encoders for medical image segmentation," *arXiv preprint arXiv:2102.04306*, 2021.
- [56] Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: learning dense volumetric segmentation from sparse annotation," in *International conference on medical image computing and computer-assisted intervention*. Springer, 2016, pp. 424–432.
- [57] V. W. H. Wong, M. Ferguson, K. H. Law, Y.-T. T. Lee, and P. Witherell, "Automatic volumetric segmentation of additive manufacturing defects with 3d u-net," *arXiv preprint arXiv:2101.08993*, 2021.
- [58] R. Szeliski, *Computer vision: algorithms and applications*. Springer Science & Business Media, 2010.
- [59] N. Zubić and P. Liò, "An effective loss function for generating 3d models from single 2d image without rendering," in *IFIP International Conference on Artificial Intelligence Applications and Innovations*. Springer, 2021, pp. 309–322.
- [60] G. Xu and Z. Zhang, *Epipolar geometry in stereo, motion and object recognition: a unified approach*. Springer Science & Business Media, 2013, vol. 6.
- [61] J. Maciel and J. P. Costeira, "A global solution to sparse correspondence problems," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 2, pp. 187–199, 2003.
- [62] C. Liu, J. Yuen, and A. Torralba, "Sift flow: Dense correspondence across scenes and its applications," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 5, pp. 978–994, 2010.
- [63] B. K. P. Horn, "Shape from shading: A method for obtaining the shape of a smooth opaque object from one view," Ph.D. dissertation, Massachusetts Institute of Technology, 1970.
- [64] J. Aloimonos, "Shape from texture," *Biological cybernetics*, vol. 58, no. 5, pp. 345–360, 1988.
- [65] S. K. Nayar and Y. Nakagawa, "Shape from focus," *IEEE Transactions on Pattern analysis and machine intelligence*, vol. 16, no. 8, pp. 824–831, 1994.
- [66] M. Hebert, "Active and passive range sensing for robotics," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, vol. 1. IEEE, 2000, pp. 102–110.
- [67] S. N. Sinha, D. Steedly, R. Szeliski, M. Agrawala, and M. Pollefeys, "Interactive 3d architectural modeling from unordered photo collections," *ACM Transactions on Graphics (TOG)*, vol. 27, no. 5, pp. 1–10, 2008.
- [68] T. Groueix, M. Fisher, V. G. Kim, B. Russell, and M. Aubry, "AtlasNet: A Papier-Mâché Approach to Learning 3D Surface Generation," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018.
- [69] L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A. Geiger, "Occupancy networks: Learning 3d reconstruction in function space," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4460–4470.
- [70] S. K. Kopparapu, "Lighting design for machine vision application," *Image and Vision Computing*, vol. 24, no. 7, pp. 720–726, 2006.
- [71] J.-t. Dong, R.-S. Lu, Y.-Q. Shi, R.-X. Xia, Q. Li, and Y. Xu, "Optical design of color light-emitting diode ring light for machine vision inspection," *Optical Engineering*, vol. 50, no. 4, p. 043001, 2011.
- [72] M. Montironi, P. Castellini, L. Stroppa, and N. Paone, "Adaptive autonomous positioning of a robot vision system: Application to quality control on production lines," *Robotics and Computer-Integrated Manufacturing*, vol. 30, no. 5, pp. 489–498, 2014.
- [73] R. Crossley and S. Ratchev, "Aerospace assembly gap measurement using low cost smart tools with machine vision," in *International Precision Assembly Seminar*. Springer, 2018, pp. 158–168.
- [74] L. Zhou, L. Zhang, L. Ren, and J. Wang, "Real-time scheduling of cloud manufacturing services based on dynamic data-driven simulation," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 9, pp. 5042–5051, 2019.
- [75] L. Zhou, L. Zhang, B. R. Sarker, Y. Laili, and L. Ren, "An event-triggered dynamic scheduling method for randomly arriving tasks in cloud manufacturing," *International Journal of Computer Integrated Manufacturing*, vol. 31, no. 3, pp. 318–333, 2018.
- [76] L. Zhou, L. Zhang, C. Zhao, Y. Laili, and L. Xu, "Diverse task scheduling for individualized requirements in cloud manufacturing," *Enterprise Information Systems*, vol. 12, no. 3, pp. 300–318, 2018.
- [77] L. Zhang, L. Zhou, and B. K. Horn, "Building a right digital twin with model engineering," *Journal of Manufacturing Systems*, vol. 59, pp. 151–164, 2021.
- [78] M. Sarcar, K. M. Rao, and K. L. Narayan, *Computer aided design and manufacturing*. PHI Learning Pvt. Ltd., 2008.
- [79] M. Milanova, L. Nikolov, and S. Fotev, "Three dimensional computer vision for computer aided design and manufacturing applications," in *International Workshop on Structural and Syntactic Pattern Recognition*. Springer, 1996, pp. 279–288.
- [80] X. Ye, H. Liu, L. Chen, Z. Chen, X. Pan, and S. Zhang, "Reverse innovative design—an integrated product design methodology," *Computer-aided design*, vol. 40, no. 7, pp. 812–827, 2008.
- [81] M. Ulrich, C. Wiedemann, and C. Steger, "Combining scale-space and similarity-based aspect graphs for fast 3d object recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 34, no. 10, pp. 1902–1914, 2011.
- [82] Y. Huang, R. McMurrin, G. Dhadyalla, R. P. Jones, and A. Mouzakitis, "Model-based testing of a vehicle instrument cluster for design validation using machine vision," *Measurement Science and Technology*, vol. 20, no. 6, p. 065502, 2009.
- [83] M. D. Raj and V. S. Kumar, "Vision based feature diagnosis for automobile instrument cluster using machine learning," in *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*. IEEE, 2017, pp. 1–5.
- [84] L. Zhou, L. Zhang, Y. Laili, C. Zhao, and Y. Xiao, "Multi-task scheduling of distributed 3d printing services in cloud manufacturing," *The International Journal of Advanced Manufacturing Technology*, vol. 96, no. 9–12, pp. 3003–3017, 2018.
- [85] S. M. Bhandarkar, T. D. Faust, and M. Tang, "Design and prototype development of a computer vision-based lumber production planning system," *Image and Vision Computing*, vol. 20, no. 3, pp. 167–189, 2002.
- [86] S. M. Bhandarkar, X. Luo, R. F. Daniels, and E. W. Tollner, "Automated planning and optimization of lumber production using machine vision and computed tomography," *IEEE transactions on automation science and engineering*, vol. 5, no. 4, pp. 677–695, 2008.
- [87] Y. Wang, P. Zheng, X. Xu, H. Yang, and J. Zou, "Production planning for cloud-based additive manufacturing—a computer vision-based approach," *Robotics and Computer-Integrated Manufacturing*, vol. 58, pp. 145–157, 2019.
- [88] M. J. Tsai, H.-W. Lee, and N.-J. Ann, "Machine vision based path planning for a robotic golf club head welding system," *Robotics and Computer-Integrated Manufacturing*, vol. 27, no. 4, pp. 843–849, 2011.
- [89] F. Trdič, B. Širok, P. R. Bullen, and D. R. Philpott, "Monitoring mineral wool production using real-time machine vision," *Real-Time Imaging*, vol. 5, no. 2, pp. 125–140, 1999.
- [90] M. Carfagni, R. Furferi, and L. Governi, "A real-time machine-vision system for monitoring the textile raising process," *Computers in Industry*, vol. 56, no. 8–9, pp. 831–842, 2005.

- [91] J. Liu, W. Gui, Z. Tang, H. Hu, and J. Zhu, "Machine vision based production condition classification and recognition for mineral flotation process monitoring," *International Journal of Computational Intelligence Systems*, vol. 6, no. 5, pp. 969–986, 2013.
- [92] D.-M. Tsai and M.-C. Lin, "Machine-vision-based identification for wafer tracking in solar cell manufacturing," *Robotics and Computer-Integrated Manufacturing*, vol. 29, no. 5, pp. 312–321, 2013.
- [93] A. Mehrabi, N. Mehrshad, and M. Massinaei, "Machine vision based monitoring of an industrial flotation cell in an iron flotation plant," *International Journal of Mineral Processing*, vol. 133, pp. 60–66, 2014.
- [94] Q.-J. Zhao, P. Cao, and D.-W. Tu, "Toward intelligent manufacturing: label characters marking and recognition method for steel products with machine vision," *Advances in Manufacturing*, vol. 2, no. 1, pp. 3–12, 2014.
- [95] L. Pérez, Í. Rodríguez, N. Rodríguez, R. Usamentiaga, and D. F. García, "Robot guidance using machine vision techniques in industrial environments: A comparative review," *Sensors*, vol. 16, no. 3, p. 335, 2016.
- [96] D. D. Ligutan, L. J. S. Cruz, M. C. D. Del Rosario, J. N. S. Kudhal, A. C. Abad, and E. P. Dadios, "Design and implementation of a fuzzy logic-based joint controller on a 6-dof robot arm with machine vision feedback," in *2017 Computing Conference*. IEEE, 2017, pp. 249–257.
- [97] A. Mohammed, B. Schmidt, and L. Wang, "Active collision avoidance for human–robot collaboration driven by vision sensors," *International Journal of Computer Integrated Manufacturing*, vol. 30, no. 9, pp. 970–980, 2017.
- [98] M. Peña-Cabrera, I. Lopez-Juarez, R. Rios-Cabrera, and J. Corona-Castuera, "Machine vision approach for robotic assembly," *Assembly Automation*, vol. 25, no. 3, pp. 204–216, 2005.
- [99] J. Häcker, F. Engelhardt, and D. D. Freyw, "Robust manufacturing inspection and classification with machine vision," *International journal of production research*, vol. 40, no. 6, pp. 1319–1334, 2002.
- [100] M. A. Akhloufi, W. B. Larbi, and X. Maldague, "Framework for color-texture classification in machine vision inspection of industrial products," in *2007 IEEE International Conference on Systems, Man and Cybernetics*. IEEE, 2007, pp. 1067–1071.
- [101] D. M. Gila, D. A. Puerto, J. G. García, and J. G. Ortega, "Automatic classification of olives for oil production using computer vision," in *2015 IEEE International Conference on Industrial Technology (ICIT)*. IEEE, 2015, pp. 1651–1656.
- [102] H. Zhao, R. Dai, and C. Xiao, "A machine vision system for stacked substrates counting with a robust stripe detection algorithm," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 11, pp. 2352–2361, 2019.
- [103] M. Yachida and S. Tsuji, "A versatile machine vision system for complex industrial parts," *IEEE Transactions on Computers*, no. 9, pp. 882–894, 1977.
- [104] D.-M. Tsai and M.-F. Chen, "A fast machine vision approach for automatic recognition of industrial parts," *International journal of production research*, vol. 34, no. 3, pp. 687–699, 1996.
- [105] C.-Y. Lin, C. Wang, and M. Tomizuka, "Pose estimation in industrial machine vision systems under sensing dynamics: A statistical learning approach," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2014, pp. 4436–4442.
- [106] L. Lindner, O. Sergiyenko, J. C. Rodríguez-Quinonez, M. Rivas-Lopez, D. Hernandez-Balbuena, W. Flores-Fuentes, F. N. Murrieta-Rico, and V. Tyrsa, "Mobile robot vision system using continuous laser scanning for industrial application," *Industrial Robot: An International Journal*, 2016.
- [107] H. Wu and J. Zhao, "An intelligent vision-based approach for helmet identification for work safety," *Computers in Industry*, vol. 100, pp. 267–277, 2018.
- [108] B. Karabagli, T. Simon, and J.-J. Orteu, "A new chain-processing-based computer vision system for automatic checking of machining set-up application for machine tools safety," *The International Journal of Advanced Manufacturing Technology*, vol. 82, no. 9-12, pp. 1547–1568, 2016.
- [109] D. Nelsen and S. E. Daniels, "Quality glossary," *Quality Progress*, vol. 40, no. 6, p. 39, 2007.
- [110] J. Molina, J. E. Solanes, L. Arnal, and J. Tornero, "On the detection of defects on specular car body surfaces," *Robotics and Computer-Integrated Manufacturing*, vol. 48, pp. 263–278, 2017.
- [111] R. Schmitt, T. Fürtjes, B. Abbas, P. Abel, W. Kimmelman, P. Kosse, and A. Buratti, "Real-time machine-vision-system for an automated quality monitoring in mass production of multiaxial non-crimp fabrics," in *Automation, Communication and Cybernetics in Science and Engineering 2015/2016*. Springer, 2016, pp. 769–782.
- [112] H. Shen, W. Sun, and J. Fu, "Multi-view online vision detection based on robot fused deposit modeling 3d printing technology," *Rapid Prototyping Journal*, 2019.
- [113] X. Li, L. Wang, and N. Cai, "Machine-vision-based surface finish inspection for cutting tool replacement in production," *International journal of production research*, vol. 42, no. 11, pp. 2279–2287, 2004.
- [114] X. Zhu, R. Chen, and Y. Zhang, "Automatic defect detection in spring clamp production via machine vision," in *Abstract and Applied Analysis*, vol. 2014. Hindawi, 2014.
- [115] P. Arjun and T. Mirnalinee, "Machine parts recognition and defect detection in automated assembly systems using computer vision techniques," *Rev. Téc. Ing. Univ. Zulia*, vol. 39, no. 1, pp. 71–80, 2016.
- [116] R. Chiou, P. Mookiah, and Y. Kwon, "Manufacturing e-quality through integrated web-enabled computer vision and robotics," *The International Journal of Advanced Manufacturing Technology*, vol. 43, no. 7-8, pp. 720–730, 2009.
- [117] V. Chauhan and B. Surgenor, "Fault detection and classification in automated assembly machines using machine vision," *The International Journal of Advanced Manufacturing Technology*, vol. 90, no. 9-12, pp. 2491–2512, 2017.
- [118] L. Scime, D. Siddel, S. Baird, and V. Paquit, "Layer-wise anomaly detection and classification for powder bed additive manufacturing processes: A machine-agnostic algorithm for real-time pixel-wise semantic segmentation," *Additive Manufacturing*, vol. 36, p. 101453, 2020.
- [119] J. M. Parker, "A robust machine vision system design to facilitate the automation of surface appearance inspections," in *2001 IEEE/ASME International Conference on Advanced Intelligent Mechanisms. Proceedings (Cat. No. 01TH8556)*, vol. 1. IEEE, 2001, pp. 87–92.
- [120] R. Furferi, L. Governì, Y. Volpe, and M. Carfagni, "Design and assessment of a machine vision system for automatic vehicle wheel alignment," *International Journal of Advanced Robotic Systems*, vol. 10, no. 5, p. 242, 2013.
- [121] Q. Zhou, R. Chen, B. Huang, C. Liu, J. Yu, and X. Yu, "An automatic surface defect inspection system for automobiles using machine vision methods," *Sensors*, vol. 19, no. 3, p. 644, 2019.
- [122] P. Heleno, R. Davies, B. A. B. Correia, and J. Dinis, "A machine vision quality control system for industrial acrylic fibre production," *EURASIP Journal on Advances in Signal Processing*, vol. 2002, no. 7, p. 298501, 2002.
- [123] D. Ponsa, R. Benavente, F. Lumbreras, J. Martinez, and X. Roca, "Quality control of safety belts by machine vision inspection for real-time production," *Optical Engineering*, vol. 42, no. 4, pp. 1114–1120, 2003.
- [124] P. Sitthi-Amorn, J. E. Ramos, Y. Wang, J. Kwan, J. Lan, W. Wang, and W. Matusik, "Multifab: a machine vision assisted platform for multi-material 3d printing," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, pp. 1–11, 2015.
- [125] A. Ceruti, A. Liverani, and T. Bombardi, "Augmented vision and interactive monitoring in 3d printing process," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 11, no. 2, pp. 385–395, 2017.
- [126] K. Okarma and J. Fastowicz, "Computer vision methods for non-destructive quality assessment in additive manufacturing," in *International Conference on Computer Recognition Systems*. Springer, 2019, pp. 11–20.
- [127] D.-B. Perng, C.-C. Chou, and S.-M. Lee, "Design and development of a new machine vision wire bonding inspection system," *The International Journal of Advanced Manufacturing Technology*, vol. 34, no. 3-4, pp. 323–334, 2007.
- [128] Y. Yang, Z.-J. Zha, M. Gao, and Z. He, "A robust vision inspection system for detecting surface defects of film capacitors," *Signal Processing*, vol. 124, pp. 54–62, 2016.
- [129] A. Yousefian-Jazi, J.-H. Ryu, S. Yoon, and J. J. Liu, "Decision support in machine vision system for monitoring of tft-lcd glass substrates manufacturing," *Journal of Process Control*, vol. 24, no. 6, pp. 1015–1023, 2014.
- [130] D. Savransky, S. J. Thomas, L. A. Poyneer, and B. A. Macintosh, "Computer vision applications for coronagraphic optical alignment and image processing," *Applied Optics*, vol. 52, no. 14, pp. 3394–3403, 2013.
- [131] K.-L. Lin and J.-L. Fang, "Applications of computer vision on tile alignment inspection," *Automation in construction*, vol. 35, pp. 562–567, 2013.
- [132] A. Mendikute and M. Zatarain, "Automated raw part alignment by a novel machine vision approach," *Procedia Engineering*, vol. 63, pp. 812–820, 2013.

- [133] L. Zhu and L. Bai, "Design and research of logistics product anti-counterfeiting algorithm based on 3d computer vision," in *Applied Mechanics and Materials*, vol. 644. Trans Tech Publ, 2014, pp. 1778–1782.
- [134] N. Heraković and D. Noe, "Experimental analysis of conditions for machine vision control in em stator assembly process," *IFAC Proceedings Volumes*, vol. 40, no. 2, pp. 96–101, 2007.
- [135] J.-Y. Chang and J.-Y. Shen, "Machine vision assistance for flex cable-actuator comb manufacturing in hard disk drives," in *2008 15th International Conference on Mechatronics and Machine Vision in Practice*. IEEE, 2008, pp. 195–200.
- [136] Y. Liu, S. Li, J. Wang, H. Zeng, and J. Lu, "A computer vision-based assistant system for the assembly of narrow cabin products," *The International Journal of Advanced Manufacturing Technology*, vol. 76, no. 1–4, pp. 281–293, 2015.
- [137] P. Nerakae, P. Uangpairoj, and K. Chamniprasart, "Using machine vision for flexible automatic assembly system," *Procedia Computer Science*, vol. 96, pp. 428–435, 2016.
- [138] H.-d. Chen, Y.-f. Wang, Z. Guo, W.-x. Chen, and P. Zhao, "A gui software for automatic assembly based on machine vision," in *2018 IEEE International Conference on Mechatronics, Robotics and Automation (ICMRA)*. IEEE, 2018, pp. 105–111.
- [139] S. Okumura, N. Take, and N. Okino, "Error prevention in robotic assembly tasks by a machine vision and statistical pattern recognition method," *International journal of production research*, vol. 43, no. 7, pp. 1397–1410, 2005.
- [140] L. Rusli and A. Luscher, "Fastener identification and assembly verification via machine vision," *Assembly Automation*, 2018.
- [141] E. Turner, L. Newberry, S. Santinga, J. Gray, S. Gopu, J. Peoples, J. Hobbs, and S. White, "Applying computer vision to track tool movement in an automotive assembly plant," in *Proceedings of the 2019 ACM Southeast Conference*, 2019, pp. 214–217.
- [142] R. Sharma and J. Molineros, "Computer vision-based augmented reality for guiding manual assembly," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 3, pp. 292–317, 1997.
- [143] J. Molineros, V. Raghavan, and R. Sharma, "Computer vision based augmented reality for guiding and evaluating assembly sequences," in *Proceedings. IEEE 1998 Virtual Reality Annual International Symposium (Cat. No. 98CB36180)*. IEEE, 1998, p. 214.
- [144] Q. Jiang, M. Liu, X. Wang, M. Ge, and L. Lin, "Human motion segmentation and recognition using machine vision for mechanical assembly operation," *SpringerPlus*, vol. 5, no. 1, p. 1629, 2016.
- [145] Y. Li, K. Mo, L. Shao, M. Sung, and L. Guibas, "Learning 3d part assembly from a single image," *arXiv preprint arXiv:2003.09754*, 2020.
- [146] R. Shilkrot, Z. Chai, and M. Hoai, "Busyhands: A hand-tool interaction database for assembly tasks semantic segmentation," *arXiv preprint arXiv:1902.07262*, 2019.
- [147] J. Kang, J. Lee, H. Eum, C.-H. Hyun, and M. Parks, "An application of parameter extraction for agv navigation based on computer vision," in *2013 10th International Conference on Ubiquitous Robots and Ambient Intelligence (URAI)*. IEEE, 2013, pp. 622–626.
- [148] J. Varagul and T. Ito, "Simulation of detecting function object for agv using computer vision with neural network," *Procedia Computer Science*, vol. 96, pp. 159–168, 2016.
- [149] S. Sheth, A. Ajmera, A. Sharma, S. Patel, and C. Kathrecha, "Design and development of intelligent agv using computer vision and artificial intelligence," in *Soft Computing: Theories and Applications*. Springer, 2018, pp. 337–349.
- [150] P.-S. Lee, Y.-E. Shen, and L.-L. Wang, "Model-based location of automated guided vehicles in the navigation sessions by 3d computer vision," *Journal of robotic systems*, vol. 11, no. 3, pp. 181–195, 1994.
- [151] A. Kelly, B. Nagy, D. Stager, and R. Unnikrishnan, "Field and service applications—an infrastructure-free automated guided vehicle based on computer vision—an effort to make an industrial robot vehicle that can operate without supporting infrastructure," *IEEE Robotics & Automation Magazine*, vol. 14, no. 3, pp. 24–34, 2007.
- [152] C. Chuixin and C. Hanxiang, "Agv robot based on computer vision and deep learning," in *2019 3rd International Conference on Robotics and Automation Sciences (ICRAS)*. IEEE, 2019, pp. 28–34.
- [153] M. R. T. Hossai, M. A. Shahjalal, and N. F. Nuri, "Design of an iot based autonomous vehicle with the aid of computer vision," in *2017 International Conference on Electrical, Computer and Communication Engineering (ECCE)*. IEEE, 2017, pp. 752–756.
- [154] X. Wu, C. Sun, T. Zou, L. Li, L. Wang, and H. Liu, "Svm-based image partitioning for vision recognition of agv guide paths under complex illumination conditions," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101856, 2020.
- [155] G. Garibotto, S. Masciangelo, P. Bassino, C. Coelho, A. Pavan, and M. Marson, "Industrial exploitation of computer vision in logistic automation: autonomous control of an intelligent forklift truck," in *Proceedings. 1998 IEEE International Conference on Robotics and Automation (Cat. No. 98CH36146)*, vol. 2. IEEE, 1998, pp. 1459–1464.
- [156] S. S. Zhang and J. P. Wachs, "The improvement and application of intelligence tracking algorithm for moving logistics objects based on machine vision sensor," *Sensor Letters*, vol. 11, no. 5, pp. 862–869, 2013.
- [157] L. Neuweiler and P. V. Riedel, "Autonomous driving in the logistics industry: A multi-perspective view on self-driving trucks, changes in competitive advantages and their implications," 2017.
- [158] B. Van Meldert and L. De Boeck, "Introducing autonomous vehicles in logistics: a review from a broad perspective," *FEB Research Report KBI_1618*, 2016.
- [159] M. Kerin and D. T. Pham, "A review of emerging industry 4.0 technologies in remanufacturing," *Journal of Cleaner Production*, p. 117805, 2019.
- [160] S. Vongbunyong and W. H. Chen, "Disassembly automation," in *Disassembly Automation*. Springer, 2015, pp. 25–54.
- [161] H. Poschmann, H. Brüggemann, and D. Goldmann, "Disassembly 4.0: A review on using robotics in disassembly tasks as a way of automation," *Chemie Ingenieur Technik*, vol. 92, no. 4, pp. 341–359, 2020.
- [162] A. ElSayed, E. Kongar, S. M. Gupta, and T. Sobh, "A robotic-driven disassembly sequence generator for end-of-life electronic products," *Journal of Intelligent & Robotic Systems*, vol. 68, no. 1, pp. 43–52, 2012.
- [163] E. Yildiz and F. Wörgötter, "Dcnv-based screw detection for automated disassembly processes," in *2019 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*. IEEE, 2019, pp. 187–192.
- [164] S. Vongbunyong, S. Kara, and M. Pagnucco, "A framework for using cognitive robotics in disassembly automation," in *Leveraging technology for a sustainable world*. Springer, 2012, pp. 173–178.
- [165] M. Bdiwi, A. Rashid, and M. Putz, "Autonomous disassembly of electric vehicle motors based on robot cognition," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2016, pp. 2500–2505.
- [166] J. Huang, D. T. Pham, Y. Wang, M. Qu, C. Ji, S. Su, W. Xu, Q. Liu, and Z. Zhou, "A case study in human-robot collaboration in the disassembly of press-fitted components," *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*, vol. 234, no. 3, pp. 654–664, 2020.
- [167] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A simple framework for contrastive learning of visual representations," in *International conference on machine learning*. PMLR, 2020, pp. 1597–1607.
- [168] P. Bergmann, M. Fauser, D. Sattlegger, and C. Steger, "Mvtec ad—a comprehensive real-world dataset for unsupervised anomaly detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 9592–9600.
- [169] P. Bergmann, K. Batzner, M. Fauser, D. Sattlegger, and C. Steger, "The mvtec anomaly detection dataset: a comprehensive real-world dataset for unsupervised anomaly detection," *International Journal of Computer Vision*, vol. 129, no. 4, pp. 1038–1059, 2021.
- [170] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. Soong, and J. C. Zhang, "What will 5g be?" *IEEE Journal on selected areas in communications*, vol. 32, no. 6, pp. 1065–1082, 2014.