Integrated Inspection on PCB Manufacturing in Cyber–Physical–Social Systems

Yutong Wang, *Member, IEEE*, Jiangong Wang, *Student Member, IEEE*, Yansong Cao, Shixing Li, and Oliver Kwan

***Abstract*—The printed circuit boards (PCBs) industry is one of the fastest-growing industries in recent decades. The PCB manufacturing process is highly complicated and severely affected by social factors, which makes it very important to conduct integrated inspection, assuring and improving the production quality. In this article, we propose an artificial systems, computational experiments, and parallel execution-based integrated inspection method in cyber–physical–social systems (CPSS) to realize smart manufacturing. In this inspection system, rather than simply performing modeling, analysis, and control, we perform descriptive intelligence to construct production processes with limited multimodal information, perform predictive intelligence to conduct defect detection and defect prediction, and perform prescriptive intelligence to achieve defect diagnosis and defect management. In this way, our inspection system could offer a learning and training platform for workers to master professional inspection skills, provide an experimentation and evaluation platform for product defect monitoring and early warnings, and supply guidance about defect management and control to improve manufacturing processes. For technical implementation, we leverage a Transformer-based foundation model to achieve knowledge reasoning and human–computer interaction. As a result, we provide an innovative solution to cope with the challenges of quality inspection in current smart manufacturing, and expect its further applications in the PCB industry.**

***Index Terms*—Artificial systems, computational experiments, and parallel execution (ACP) method, cyber–physical–social systems (CPSS), foundation models (FMs), integrated inspection.**

I. Introduction

WITH the rapid development of the global electronic information industry, the printed circuit boards (PCBs) industry has shown explosive growth. PCB is characterized by high precision, high product customization, complex manufacturing processes, intensive human labor, and high end-of-life costs [1]. And this fuels the drive for integrated inspection of the entire PCB manufacturing process and participating entities, including operators, equipment, materials, and environments.

However, there are two major problems that hinder the further development of PCB inspection. The first is current inspection devices automated optical inspection (AOI) that simply focus on product defect inspection, ignoring defect diagnosis. For defect diagnosis, relating defect types to defect causes could make workers aware of potential problems and take precautions the error-prone manufacturing processes, which is very advantageous for manufacturers to improve product quality and production process. However, defect diagnosis highly relies on a holistic knowledge system of the whole manufacturing process, which is difficult to obtain in physical PCB systems. The second is, there is a lack of professional guidance for inspection workers. In spite of the quite high detection accuracy of current AOI devices, there is still a necessity for human workers to complete defects filtering tasks. However, this work requires professional skills and consistent judgment, which cannot be conducted nicely and quickly enough at present.

Therefore, in this article, we propose an artificial systems, computational experiments, and parallel execution (ACP)-based integrated inspection [2], [3], [4] method in cyber–physical–social spaces. In this inspection system, rather than simply performing modeling, analysis, and control, we perform descriptive intelligence to construct the whole PCB production process with limited multimodal information, and iteratively optimize the models through physical and cyber world interactions. For predictive intelligence, based on the massive data generated in artificial systems, we perform sufficient model training and complete defect detection and defect prediction. For prescriptive intelligence, we construct knowledge graphs to build relationships between wrong behaviors and defect types, achieve knowledge representation, and conduct knowledge reasoning for defect diagnosis.

In this way, our inspection system could offer a learning and training platform for workers to master professional inspection skills, provide an experimentation and evaluation platform for product monitoring and early warnings, and supply guidance about defect management and control to improve manufacturing processes. For knowledge reasoning and human–computer

Manuscript received 26 November 2022; accepted 7 December 2022. Date of publication 20 December 2022; date of current version 17 March 2023. This work was supported in part by the Optima Collaborative Research Project of Defect Detection Algorithm for Automated Optical Inspection—Phase II and in part by the National Natural Science Foundation of China under Grant 62103411. This article was recommended by Associate Editor F.-Y. Wang. *(Corresponding author: Jiangong Wang.)*

Yutong Wang and Jiangong Wang are with the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: yutong.wang@ia.ac.cn; wangjiangong2018@ia.ac.cn).

Yansong Cao is with the Department of Engineering Science, Faculty of Innovation Engineering, Macau University of Science and Technology, Macau, China (e-mail: yscao@maverickvc.com).

Shixing Li is with North Automatic Control Technology Institute, Taiyuan, China (e-mail: shixing.li@qaii.ac.cn).

Oliver Kwan is with Motion G, Inc., Singapore 068898 (e-mail: oliver.kwan@motiong.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TSMC.2022.3229096.

Digital Object Identifier 10.1109/TSMC.2022.3229096

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

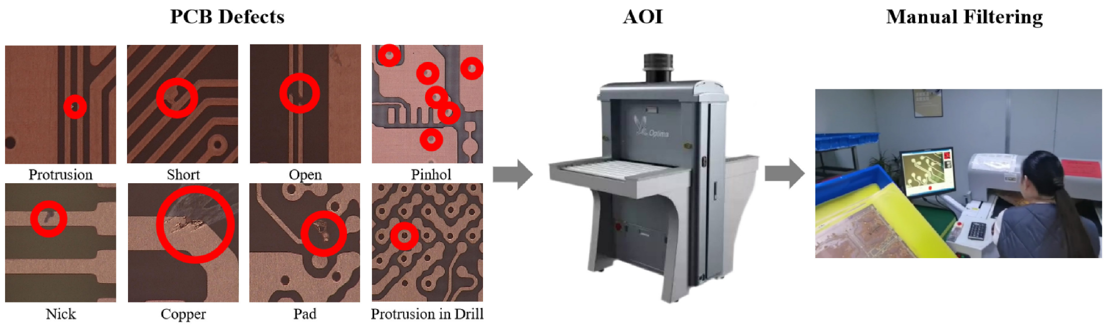

Fig. 1. PCB manufacturing process and AOI.

interaction [5], [6] (HCI), we propose a vision–language-task (ViLT) Transformer-based foundation model (FM) perform multimodal data fusion, data analysis, and results generation. Thereby, in addition to defect detection, we could realize defect prediction, diagnosis, and inspection auxiliary capabilities for integrated inspection, which is an innovative quality inspection for smart PCB manufacturing.

The main contributions of our proposed integrated inspection are as follows.

1) We perform an integrated inspection on the entire PCB manufacturing process in the artificial system, taking participating entities and social factors into full consideration and provide defect prediction, defect diagnosis, and defect management guidance for the actual manufacturing system.
2) We exploit digital humans to facilitate biological workers as personal assistant to complete skills training and inspection filtering, thereby improving work performance, and reducing workload.
3) We propose a ViLT Transformer-based FM to realize knowledge automation and drive the behavior of digital workers.

The remainder of this article is organized as follows. In Section II, we position the related work. In Section III, we propose the overall framework of integrated inspection and its three important components. In Section IV, we elaborate on the key functions of integrated inspection. In Section V, we describe the technical underpinning of integrated inspection. In Section VI, we draw a conclusion about this article.

II. Related Work

In this section, we first give a brief introduction to PCB manufacturing and AOI. Then, we introduce the theoretical support of our method, ACP and cyber–physical–social system (CPSS) approach. Finally, we discuss the technical support of our method, Transformer-based FMs.

A. PCB Manufacturing and AOI

PCBs are laminated sandwich structures of conductive and insulating layers, which affix electronic components and provide reliable electrical connections. As the core element in all electronic units, PCBs are commonly used in consumer electronics, military applications, medical equipment, etc. In recent decades, the high demand for consumer electronics like smartphones has boosted the growth of the PCB industry.

PCBs are characterized by high precision, high product customization, complex manufacturing processes, intensive human labor, and high end-of-life costs [1]. The manufacturing process of PCBs involves multiple complicated processes, and during the manufacturing process of PCB, there may exist a variety of defects, such as open, short, nick, pinhole, and copper. These defects can arise for a variety of reasons, ranging from operator error to pollutants. And defects will directly impact the performance of electronic devices. Therefore, defect detection should be an indispensable process in the whole production line.

Traditional defect inspection is carried out by human workers. Due to the inefficiency and error-prone shortages of manual inspection, AOI equipment [7], [8] is designed and used in manufacturing that largely replaces human labor in defect defection. AOI is an automated noncontact visual inspection device, which autonomously scans PCB with cameras, compares collected PCB images with their templates, and exploits deep neural networks (DNNs)-based defect detection models [9], [10], [11], and produces detection results. In spite of current AOI devices, there is still a necessity for human workers to perform defects filtering task [12], assuring 100% appearance detection accuracy, as shown in Fig. 1.

B. ACP and CPSS

To model, analyze, and control complex systems, Wang first proposed the ACP approach [13], [14]. Basically, A denotes artificial systems for modeling, C denotes computational experiments for analysis, and P denotes parallel execution for control. During parallel execution, the models in the artificial systems are used to guide the physical systems, in turn, the performance of the physical systems is used as feedback for the artificial systems to further optimize the modeling process.

Wang first proposed CPSS [15], [16], [17] as an extension of cyber–physical systems (CPS) [18], [19] that integrates human-related social information. Practical complex systems often involve human behaviors. However, in traditional “Newton systems,” there is no need to consider human or social factors. Since the outputs can be precisely predicted given the inputs. In CPSS-oriented “Merton” societal systems, human factors, including behavior and mentality must be fully considered. Since these factors will definitely influence the outputs, given the same inputs. Combined with ACP, CPSS could find solutions with agility, focus, and convergence to address human-related issues with uncertainty, diversity, and complexity. Up till now, the combination of ACP and CPSS has been successfully applied to transportation [20], [21], agriculture [22], and healthcare [23].

C. Foundation Models

Vaswani et al. first proposed the encoder–decoder Transformer architecture in 2017 [24] to replace long short-term memory (LSTM) [25], [26] and recurrent neural network (RNN) on machine translation tasks. Thanks to the multiheaded self-attention mechanism [27], [28], Transformers could draw global dependencies between input and output, thereby achieving superior performance over existing best models. Furthermore, based on the Transformer architecture, Devlin et al. first proposed bidirectional encoder representations from transformers (BERTs) [29], which is pretrained on broad unlabeled text, then fine-tuned on task-specific labeled data. This training mechanism creates state-of-the-art models for 11 downstream natural language processing (NLP) tasks, including question answering (QA), language inference, etc.

Inspired by the strong representation of Transformer, researchers turn their attention to computer vision (CV) tasks [30]. Abandoning convolutional neural networks (CNNs), Dosovitskiy et al. [31] split an image into patches, feed the resulting sequence of vectors to a standard Transformer encoder, and perform image classification. Transformers have also been applied to a variety of other CV tasks, including object detection [32], [33], semantic segmentation [34], etc. Because of the versatility of Transformers, the concept of FM [35] is proposed, which makes full use of unlabeled data on Transformers and is adapted to achieve better performance on downstream tasks. In addition, FM also achieves strong performance in joint vision-and-language reasoning tasks such as visual QA [36], [37], [38] relying on massive multimodal data.

III. Overall Framework

The overall framework of our ACP-based integrated inspection in CPSS is shown in Fig. 2. It comprises artificial manufacturing systems and actual manufacturing systems providing learning and training, experimentation and evaluation, and management and control capabilities. These three sets of capabilities are iteratively optimized in both artificial and actual systems, influencing and gradually reinforcing each other. In integrated inspection, rather than simply performing modeling, analysis, and control, we need to actively guide the co-evolution of actual and artificial manufacturing systems based on descriptive intelligence, predictive intelligence, and prescriptive intelligence. It is important to emphasize that only by building the virtual–real interaction and closed-loop iteration based on ACP in cyber, physical, and social systems at the same time, we can realize the parallel intelligence and finally solve the integrated inspection on PCB manufacturing.

A. Descriptive Intelligence

PCBs with defects prone to be subpar and malfunctioning, which need to be repaired or scrapped, and this will cause a waste of production costs. Therefore, ensuring product quality and eliminating production defects are the topmost priorities of PCB manufacturers. However, conventional inspection only focuses on product defect detection, ignoring the inspection of participating entities, such as human operators, production equipment, and production processes. Information about these defect sources is not recorded and analyzed by manufacturing systems, thereby this kind of inspection could not provide enough information about defect prediction, defect diagnosis, and production improvement.

For descriptive intelligence, we construct artificial PCB manufacturing systems leveraging digital twin (DT) [39], [40] technology and model-based design (MBD) [41], [42]. Specifically, we model the whole PCB manufacturing process involving materials, equipment, processes, operators, and environments. For three-dimensional (3-D) architecture modeling, we could utilize software, such as SolidWorks, computer-aided design (CAD), and 3ds Max if we have entity prototypes. Without these prototypes, we could generate 3-D models by a data-driven method. Based on RGB-D images, or point clouds data collected from physical manufacturing systems, we could use state-of-the-art deep learning methods [43], such as generative adversarial network (GAN) [44] or structure from motion (SfM) 3-D reconstruction method [45] to achieve automatic modeling.

For the kinematics, dynamics, mechanics, and chemistry characteristics modeling, we could utilize simulation software, such as MATLAB, Multiphysics, and Unity 3-D to build the exact mathematical model if we have the full knowledge about the mechanism. With incomplete information, we could also integrate knowledge-driven and data-driven modeling approaches to conduct modeling [46], [47]. For the knowledge-driven method, we exploit human expert knowledge [48] as a form of imprecise natural language. Thereby we could construct a knowledge base and use knowledge graph [49] to perform knowledge representation and knowledge reasoning [50].

Meanwhile, we iteratively optimize the models through the feedback from the physical world. Specifically, we rectify the models by minimizing the errors between the physical world output and the artificial system output under the same input, making artificial models converge to physical entity models over time.

B. Predictive Intelligence

In this artificial manufacturing system, we perform an integrated inspection of the entire PCB manufacturing process, taking participating entities and human workers into full consideration. With the broad multimodal data (images, videos, text, and audio) generated in the artificial system, we complete multitask learning (defect detection, defect prediction, and defect diagnosis) for predictive intelligence, which plays a significant role in smart manufacturing [51], [52].

Currently, visual quality inspection is carried out by defect detection using supervised or unsupervised methods to obtain

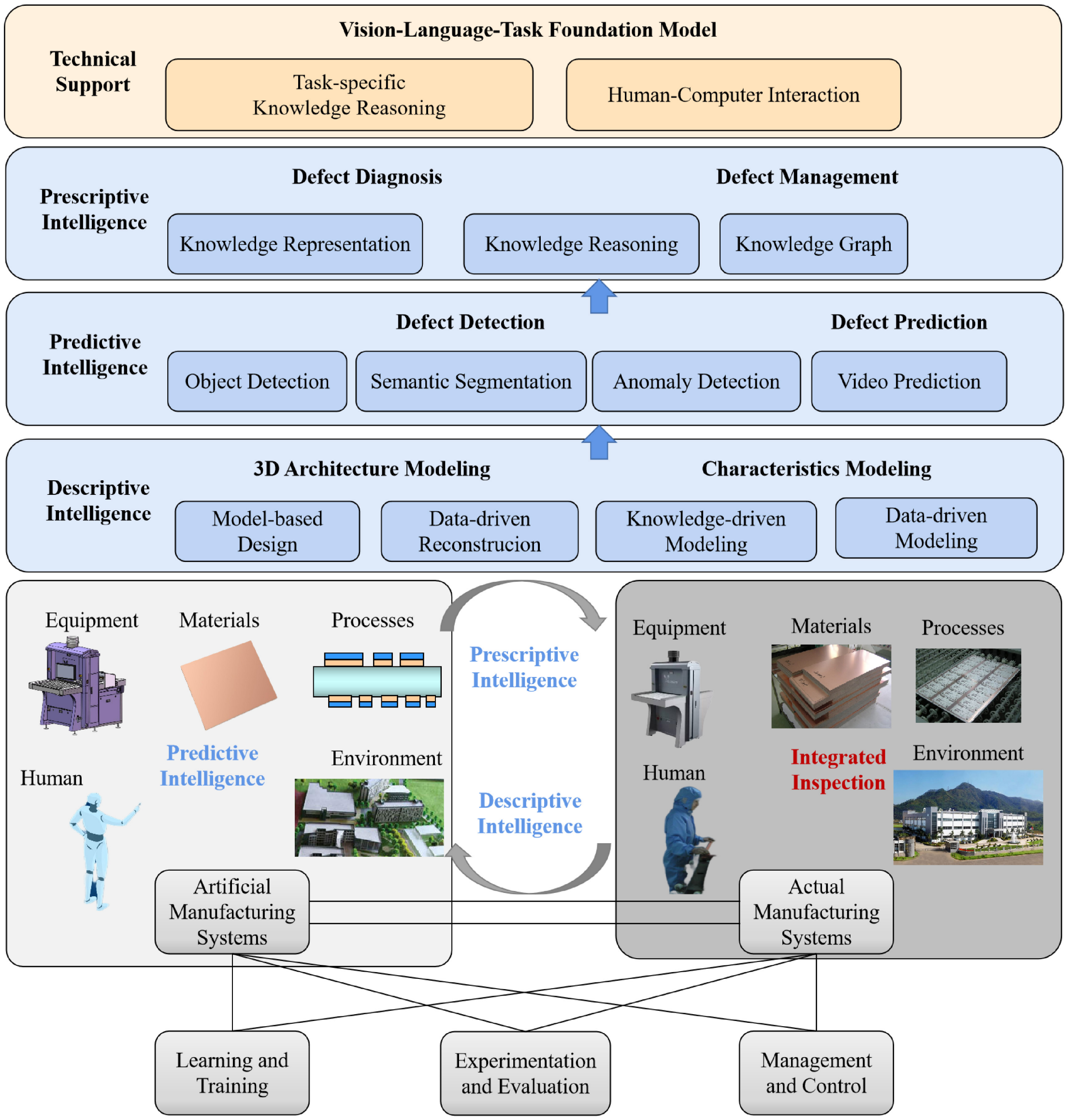

Fig. 2. Framework of the ACP-based integrated inspection for PCB industry in CPSS.

defect type and defect localization. Common supervised methods are object detection [53] and semantic segmentation [54] methods. Major unsupervised method is anomaly detection [55], [56], [57], also known as outlier detection [58] or novelty detection [59]. An anomaly detection model is trained on defect-free images and tested on images with defects. However, since defects in the physical world may have large variations in size, orientation, and shape, the detection results of these methods could not satisfy the demand for PCB manufacturing. In spite of the high accuracy achieved by the state-of-the-art supervised and unsupervised methods on existing public datasets, these methods have poor generalization performance in physical manufacturing systems.

With the artificial PCB manufacturing process, all of the settings are controllable, thus, we could adjust the settings as we need and generate corresponding multimodal data. By this means, we could enrich the training data, especially, defect data. Based on these multimodal data, we could perform a large number of computational experiments to train the model for multitasks and exploit this model to make analyses and predictions in the physical world. The model we use to achieve multitasks is discussed in Section V.

Given the product image information of each manufacturing process, we could perform defect prediction and defect diagnosis like fault prediction [60] and fault diagnosis [61], [62], [63], [64]. For defect prediction, we could perform video prediction [65], [66] on a sequence of product frames, learn spatiotemporal information, and generate future frames. In this way, utilizing sufficient and accessible data in the artificial manufacturing system, we could perform defect monitoring by defect prediction.

C. Prescriptive Intelligence

Achieving intelligent defect management [67] and control could be the ultimate goal for PCB manufacturers. Current

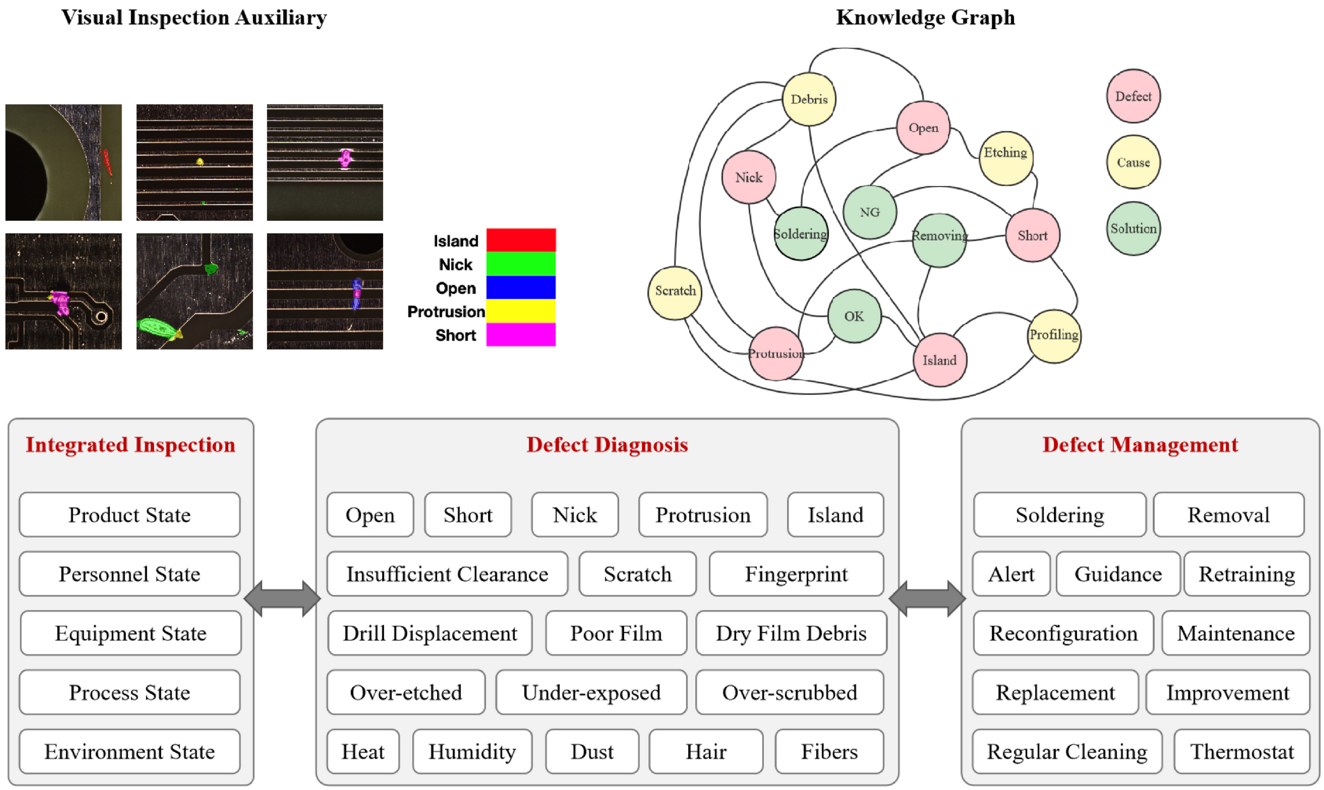

Fig. 3. Training and learning, experimentation and evaluation, and management and control for integrated inspection.

inspection only focuses on product defects, ignoring the defect reasons appearing in the manufacturing process. Wrong operations, aging, corrosion, and wear operations could cost massive defects, and lead to enormous production costs. For prescriptive intelligence, we aim to perform defect management and control in our integrated inspection system. To be specific, we conduct defect diagnosis to determine the defect locations, defect reasons, and defect severity. Considering different operations could cause the same defect type, we construct a knowledge graph to realize knowledge representation. Different manufacturing processes and defects are the entities. Components settings are the attributes of the manufacturing process, and type and severity are the attributes of defects. By this means, we build the relationships between manufacturing processes and defects, and complete defect diagnosis through knowledge reasoning, i.e., deducing improper manufacturing behaviors from defects.

With the diagnosis information, we could perform predictive maintenance, wrong operation alert, and targeted repair of the manufacturing process. In this way, we could realize intelligent management and control of all entities participating manufacturing process, including equipment and operators. Thereby, we could minimize defects and even provide improvement suggestions for better PCB production.

IV. Key Functions

In this section, we discuss the key enabling functions provided by integrated inspection, i.e., learning and training, experimentation and evaluation, and management and control. As shown in Fig. 3, visual inspection auxiliary and knowledge graph are two important tools to achieve integrated inspection. Through visual aids, we help biological workers to visualize and filter defects, while gradually integrating this knowledge into the knowledge graph, which helps subsequent workers' learning and training. In addition, through the reasoning and analysis of the knowledge graph, it can give production guidance at the macro level and manage to reduce the defect rate.

A. Learning and Training

Inspection verification (i.e., defects filtering) workers are one of the most precious resources for a PCB enterprise. These workers are equipped with professional knowledge and practical experience. Therefore, regular training is a basic necessity, which is also cubersome work and consumes massive time and cost. In addition, traditional training is only dependent on defect images with low resolution collected by field engineers. Accurate defect diagnosis always requires multiview observation, but current training methods only provide a very limited perception of defect images.

To deal with these problems, we take full advantage of the artificial manufacturing system we built in descriptive intelligence and spatial immersive extended reality (XR) [68] technology, which provides a virtual learning and training platform for biological workers. With the help of metaverse interaction devices, such as helmets, headsets, and glasses, virtual training is promoted through the modalities of vision, sound, touch, and movement. Moreover, active interaction devices, such as grips and thumbsticks, could also support touching and manipulating virtual entities.

In addition, we also exploit digital humans as personal teachers and mentors for each biological worker. The modeling of digital humans requires the use of 3-D modeling, graphic rendering, NLP, speech synthesis, and other technologies to achieve character generation, character expression, recognition and perception, analysis, and decision making. Therefore, digital humans have the ability to "see," "hear," "understand," and "speak," which is discussed in detail in Section V.

During learning and training, digital workers also care about personal learning ability and mental health. By observing and analyzing the facial expression, working performance, and sentiment [69] of biological humans, digital humans could identify distraction, fatigue, and burnout conditions, and develop targeted interventions [70], [71]. Instead of replacing biological humans, digital humans will release biological humans from tedious and laborious work, largely reduce workload and provide pertinent guidance.

B. Experimentation and Evaluation

Integrated inspection of the entire PCB manufacturing process provides an experimentation and evaluation platform. Specifically, we conduct massive experiments in the cyber world to train FMs and achieve high-accuracy defect prediction with defect type and severity information. Current AOI only outputs approximate defect localization information, verification workers should reclassify the exact defect localization and verify its severity. Based on these FMs, we have developed a visual auxiliary tool for biological workers to complete defect filtering by providing defect type and severity information, as shown in Fig. 3. Common defect types include open (break in the conductor), short (copper bridges between conductors), nick (partial open of conductor), protrusion (copper spurs), and island (spurious metal) [72], [73]. For verification step, we visualize the defect type and segmentation information with different colors, help workers determine the repair measures, which largely saves work time, and improve filtering quality.

C. Management and Control

Integrated inspection could provide guidance about intelligent defect management and control. PCB manufacturing involves massive entities, which means it is difficult to conduct defect management in the physical world. However, with the artificial manufacturing system integrating all production-related information as a whole, integrated inspection could conduct defect diagnosis, thereby achieving intelligent defect management, as shown in Fig. 3. Given the real-time data of product, personnel, equipment, process, and environment in artificial systems, we construct a knowledge graph to represent the relationships between defect type, defect cause, and solution to the defect. And the defect causes are always human misoperation, equipment misconfiguration, process underperformance, and environmental pollutants. According to the defect causes, we perform corresponding remedial measures to achieve defect management, including defect repair, personnel retraining, equipment maintenance, process improvement, regular cleaning, etc.

V. Multimodal Multitask Foundation Model

In this section, we discuss Transformer-based FM as the most important technical underpinnings in integrated inspection to realize knowledge reasoning as well as HCI. The Transformer has an encoder–decoder structure based on the global attention mechanism. And it has the capability of unified representations extracted from multimodal data.

To handle the vision and language data collected both in the physical and cyber world, we need to perform multimodal data fusion [74], data analysis [75], and results generation. To achieve this multimodal multitask learning, we propose a ViLT FM, as shown in Fig. 4. Based on the Transformer encoder architecture, we take images, videos, text, and audio as inputs, encode each modality with a separate encoder, concatenate encoded modalities, and feed modalities into task-specific output heads. And the entire ViLT model is trained end-to-end with losses from each task. These tasks include HCI tasks and knowledge reasoning tasks. HCI contains video classification, text classification, speech recognition, QA, sentiment recognition [76], etc. Knowledge reasoning contains defect detection and defect prediction.

For HCI, conventional chatbot [77], [78], [79] follows a general pipeline: speech recognition, natural language understanding (NLU), dialog management, and natural language generation (NLG) to complete conversation. However, as is known to us all, images speak louder than words. Without visual information, there is a tendency to miss key cues and make an incorrect judgment. Therefore, acting like a visual chatbot, ViLT considers nonverbal information, such as facial expressions, gestures, and actions [80], [81], to better understand worker behavior and generate more appropriate responses.

Equipped with ViLT, digital workers possess visual perception, NLU, and reasoning capabilities within a single model, therefore, realizing knowledge automation as well as human–digital workers interaction. And instead of simply handling queries and maintaining conversations, digital workers own a knowledge base and could provide real-time professional guidance for each human worker as personal assistance. They make personalized strategies about work schedules, operation guidance, and distraction overcoming, thereby increasing the work efficiency of human workers.

VI. Conclusion

Stepping into the smart manufacturing era, a brand new inspection system is urgently needed for PCB manufacturing. In this article, we propose an ACP-based integrated inspection method in CPSS for quality assurance and quality improvement. In this inspection system, we perform descriptive intelligence to construct an artificial PCB manufacturing system, perform predictive intelligence to conduct defect detection and defect prediction, and perform prescriptive intelligence to achieve defect diagnosis and defect management.

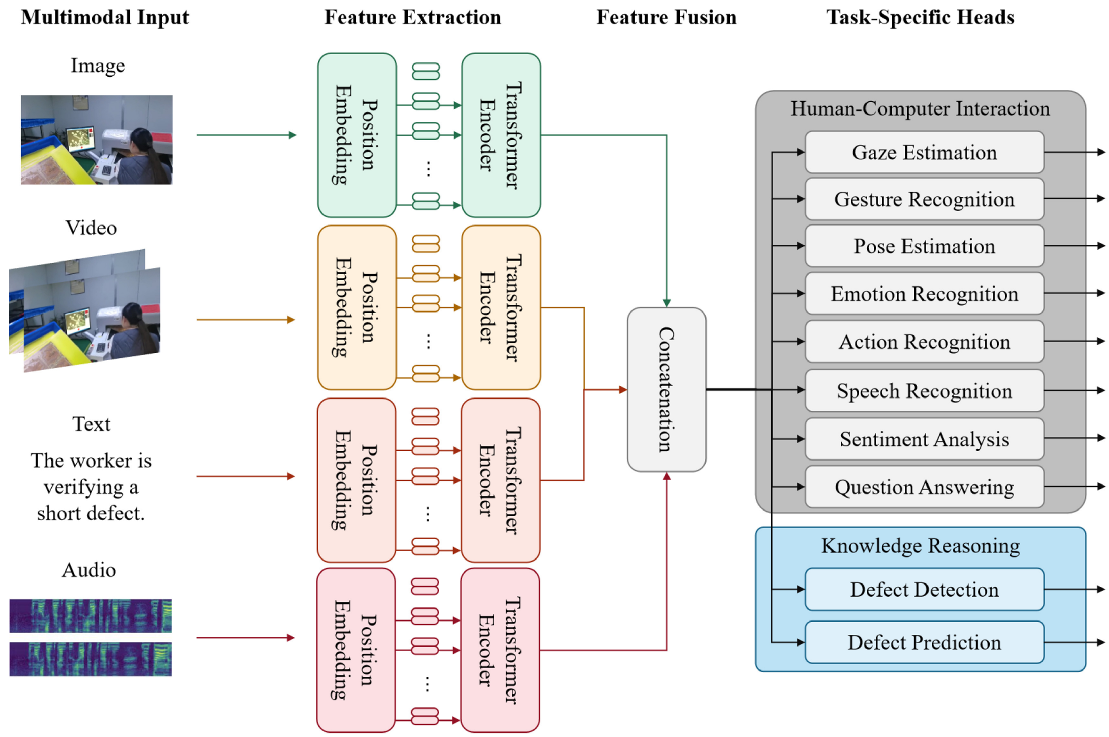

Fig. 4. Multimodal FM for knowledge automation and human–digital workers interaction.

In this inspection system, we could provide a learning and training platform for workers to master professional inspection skills, provide an experimentation and evaluation platform for product defect monitoring and early warnings, and supply guidance about defect minimization. For knowledge reasoning and HCI, we leverage a ViLT FM utilizing multimodal data simultaneously. As a result, stepping toward intelligent inspection, we extract inspection knowledge from the artificial system, perform knowledge reasoning to obtain comprehensive defect information, and provide guidance about diagnosis and improvement for human workers to enhance efficiency and improve productivity.

ACKNOWLEDGMENT

The idea of QoM, QoP, and QoS proposed in this article originated from Prof. Fei-Yue Wang in a meeting in July 2021. On the basis of Prof. Wang's parallel theory, this article has taken shape after the authors finished their field study at Optima Optical Technology (Shenzhen) Company. They would like to take this opportunity to thank Prof. Wang for his insightful ideas and valuable comments on this article.

Also, they would like to extend their sincere gratitude to Dr. Juanjuan Li for her useful help in this article. She spent a lot of time checking for grammatical writing errors and polishing up this article.

REFERENCES

[1] J. LaDou, "Printed circuit board industry," *Int. J. Hyg. Environ. Health*, vol. 209, no. 3, pp. 211–219, 2006.

[2] Y. Wang, J. Wang, Y. Tian, and F.-Y. Wang, "ACP-based automated quality inspection for PCB industry," *Int. J. Intell. Control Syst.*, vol. 1, no. 3, pp. 6–11, 2021.

[3] Y. Wang, J. Wang, Y. Tian, and F.-Y. Wang, "Digital workers for PCB inspection in cyber-physical-social systems," *J. Intell. Sci. Technol.*, vol. 2, no. 1, pp. 12–17, 2022.

[4] Y. Wang, Y. Tian, J. Wang, Y. Cao, and F.-Y. Wang, "Integrated inspection of QoM, QoP, and QoS for AOI industries in metaverses," *IEEE/CAA J. Automatica Sinica*, vol. 9, no. 12, pp. 2071–2078, Dec. 2022.

[5] Z. Shen, A. Elibol, and N. Y. Chong, "Understanding nonverbal communication cues of human personality traits in human-robot interaction," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 6, pp. 1465–1477, Nov. 2020.

[6] X. Wang, X. Zheng, W. Chen, and F.-Y. Wang, "Visual human–computer interactions for intelligent vehicles and intelligent transportation systems: The state of the art and future directions," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 1, pp. 253–265, Jan. 2021.

[7] M. Abd Al Rahman and A. Mousavi, "A review and analysis of automatic optical inspection and quality monitoring methods in electronics industry," *IEEE Access*, vol. 8, pp. 183192–183271, 2020.

[8] H. C. Liao, Z. Y. Lim, Y. X. Hu, and H. W. Tseng, "Guidelines of automated optical inspection (AOI) system development," in *Proc. IEEE 3rd Int. Conf. Signal Image Process.*, 2018, pp. 362–366.

[9] N. Patel, A. N. Saridena, A. Choromanska, P. Krishnamurthy, and F. Khorrami, "Learning-based real-time process-aware anomaly monitoring for assured autonomy," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 4, pp. 659–669, Dec. 2020.

[10] M. Maaref and Z. M. Kassas, "Measurement characterization and autonomous outlier detection and exclusion for ground vehicle navigation with cellular signals," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 4, pp. 670–683, Dec. 2020.

[11] P. Tan et al., "Multialgorithm fusion image processing for high speed railway dropper failure–Defect detection," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 7, pp. 4466–4478, Jul. 2021.

[12] S. Kumar, C. Savur, and F. Sahin, "Survey of human–Robot collaboration in industrial settings: Awareness, intelligence, and compliance," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 1, pp. 280–297, Jan. 2021.

[13] F.-Y. Wang, "Parallel system methods for management and control of complex systems," *Control Decis.*, vol. 19, no. 5, pp. 485–489, 2004.

[14] F.-Y. Wang, "Parallel control and management for intelligent transportation systems: Concepts, architectures, and applications," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 3, pp. 630–638, Sep. 2010.

[15] F.-Y. Wang, "The emergence of intelligent enterprises: From CPS to CPSS," *IEEE Intell. Syst.*, vol. 25, no. 4, pp. 85–88, Jul./Aug. 2010.

[16] J. J. Zhang et al., "Cyber-physical-social systems: The state of the art and perspectives," *IEEE Trans. Comput. Social Syst.*, vol. 5, no. 3, pp. 829–840, Sep. 2018.

[17] F.-Y. Wang, Y. Yuan, X. Wang, and R. Qin, "Societies 5.0: A new paradigm for computational social systems research," *IEEE Trans. Comput. Social Syst.*, vol. 5, no. 1, pp. 2–8, Mar. 2018.

[18] Z. Zhou, B. Wang, M. Dong, and K. Ota, "Secure and efficient vehicle-to-grid energy trading in cyber physical systems: Integration of blockchain and edge computing," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 1, pp. 43–57, Jan. 2020.

[19] D. Ding, Q.-L. Han, X. Ge, and J. Wang, "Secure state estimation and control of cyber-physical systems: A survey," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 1, pp. 176–190, Jan. 2021.

[20] F.-Y. Wang, N.-N. Zheng, D. Cao, C. M. Martinez, L. Li, and T. Liu, "Parallel driving in CPSS: A unified approach for transport automation and vehicle intelligence," *IEEE/CAA J. Automatica Sinica*, vol. 4, no. 4, pp. 577–587, Nov. 2017.

[21] F.-Y. Wang and J. J. Zhang, "Transportation 5.0 in CPSS: Towards ACP-based society-centered intelligent transportation," in *Proc. IEEE Int. Conf. Intell. Transp. Syst.*, 2017, pp. 762–767.

[22] M. Kang, X.-R. Fan, J. Hua, H. Wang, X. Wang, and F.-Y. Wang, "Managing traditional solar greenhouse with CPSS: A just-for-fit philosophy," *IEEE Trans. Cybern.*, vol. 48, no. 12, pp. 3371–3380, Dec. 2018.

[23] F.-Y. Wang, "Parallel healthcare: Robotic medical and health process automation for secured and smart social healthcares," *IEEE Trans. Comput. Social Syst.*, vol. 7, no. 3, pp. 581–586, Jun. 2020.

[24] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 5998–6008.

[25] A. Graves, "Long short-term memory," in *Supervised Sequence Labelling With Recurrent Neural Networks*, Springer, 2012, pp. 37–45.

[26] N. Khairdoost, M. Shirpour, M. A. Bauer, and S. S. Beauchemin, "Real-time driver maneuver prediction using LSTM," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 4, pp. 714–724, Dec. 2020.

[27] K. Messaoud, I. Yahiaoui, A. Verroust-Blondet, and F. Nashashibi, "Attention based vehicle trajectory prediction," *IEEE Trans. Intell. Vehicles*, vol. 6, no. 1, pp. 175–185, Mar. 2021.

[28] P. Liu, Y. Zhou, D. Peng, and D. Wu, "Global-attention-based neural networks for vision language intelligence," *IEEE/CAA J. Automatica Sinica*, vol. 8, no. 7, pp. 1243–1252, Jul. 2021.

[29] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," 2018, *arXiv:1810.04805*.

[30] K. Han et al., "A survey on vision transformer," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 87–110, Jan. 2023.

[31] A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," in *Proc. Int. Conf. Learn. Represent.*, 2021.

[32] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 213–229.

[33] Y. Liu, S. Cao, P. Lasang, and S. Shen, "Modular lightweight network for road object detection using a feature fusion approach," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 8, pp. 4716–4728, Aug. 2021.

[34] S. Zheng et al., "Rethinking semantic segmentation from a sequence-to-sequence perspective with transformers," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2021, pp. 6881–6890.

[35] R. Bommasani et al., "On the opportunities and risks of foundation models," 2021, *arXiv:2108.07258*.

[36] J. Lu, D. Batra, D. Parikh, and S. Lee, "Vilbert: Pretraining task-agnostic visiolinguistic representations for vision-and-language tasks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, 2019, pp. 13–23.

[37] L. H. Li, M. Yatskar, D. Yin, C.-J. Hsieh, and K.-W. Chang, "Visualbert: A simple and performant baseline for vision and language," 2019, *arXiv:1908.03557*.

[38] R. Hu and A. Singh, "UniT: Multimodal multitask learning with a unified transformer," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2021, pp. 1419–1429.

[39] F. Tao, H. Zhang, A. Liu, and A. Y. Nee, "Digital twin in industry: State-of-the-art," *IEEE Trans. Ind. Informat.*, vol. 15, no. 4, pp. 2405–2415, Apr. 2019.

[40] Q. Wang, W. Jiao, P. Wang, and Y. Zhang, "Digital twin for human-robot interactive welding and welder behavior analysis," *IEEE/CAA J. Automatica Sinica*, vol. 8, no. 2, pp. 334–343, Feb. 2021.

[41] G. Franceschini and S. Macchietto, "Model-based design of experiments for parameter precision: State of the art," *Chem. Eng. Sci.*, vol. 63, no. 19, pp. 4846–4872, 2008.

[42] J. C. Jensen, D. H. Chang, and E. A. Lee, "A model-based design methodology for cyber-physical systems," in *Proc. Int. Wireless Commun. Mobile Comput. Conf.*, 2011, pp. 1666–1671.

[43] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015.

[44] H. Li, Y. Zheng, X. Wu, and Q. Cai, "3D model generation and reconstruction using conditional generative adversarial network," *Int. J. Comput. Intell. Syst.*, vol. 12, no. 2, pp. 697–705, 2019.

[45] G. Vacca, G. Furfaro, and A. Dessì, "The use of the UAV images for the building 3D model generation," *Int. Arch. Photogram. Remote Sens. Spatial Inf. Sci.*, vol. 42, no. 4W8, pp. 217–223, 2018.

[46] A. Porwal, E. Carranza, and M. Hale, "Knowledge-driven and data-driven fuzzy models for predictive mineral potential mapping," *Natural Resour. Res.*, vol. 12, no. 1, pp. 1–25, 2003.

[47] L. Todorovski and S. Džeroski, "Integrating knowledge-driven and data-driven approaches to modeling," *Ecol. Model.*, vol. 194, nos. 1–3, pp. 3–13, 2006.

[48] J. Mei and H. Zhao, "Incorporating human domain knowledge in 3-D LiDAR-based semantic segmentation," *IEEE Trans. Intell. Vehicles*, vol. 5, no. 2, pp. 178–187, Jun. 2020.

[49] X. Chen, S. Jia, and Y. Xiang, "A review: Knowledge reasoning over knowledge graph," *Expert Syst. Appl.*, vol. 141, Mar. 2020, Art. no. 112948.

[50] A. V. Kannan et al., "Multimodal knowledge graph for deep learning papers and code," in *Proc. ACM Int. Conf. Inf. Knowl. Manage.*, 2020, pp. 3417–3420.

[51] J. Leng et al., "Blockchain-secured smart manufacturing in industry 4.0: A survey," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 1, pp. 237–252, Jan. 2021.

[52] J. Leng et al., "ManuChain: Combining Permissioned blockchain with a holistic optimization model as bi-level intelligence for smart manufacturing," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 1, pp. 182–192, Jan. 2020.

[53] J. Zhang, Z. Zhang, H. Su, W. Zou, X. Gong, and F. Zhang, "Quality inspection based on quadrangular object detection for deep aperture component," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 10, pp. 5938–5948, Oct. 2021.

[54] D. Jin, X. Bai, and Y. Wang, "Integrating structural symmetry and local homoplasy information in Intuitionistic fuzzy clustering for infrared pedestrian segmentation," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 7, pp. 4365–4378, Jul. 2021.

[55] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM Comput. Surveys*, vol. 41, no. 3, pp. 1–58, 2009.

[56] T. Wang, X. Xu, F. Shen, and Y. Yang, "A cognitive memory-augmented network for visual anomaly detection," *IEEE/CAA J. Automatica Sinica*, vol. 8, no. 7, pp. 1296–1307, Jul. 2021.

[57] S. Tadepalli, N. Akhter, D. Barbará, and A. Shehu, "Anomaly detection-based recognition of near-native protein structures," *IEEE Trans. Nanobiosci.*, vol. 19, no. 3, pp. 562–570, Jul. 2020.

[58] J. Li, J. Zhang, N. Pang, and X. Qin, "Weighted outlier detection of high-dimensional categorical data using feature grouping," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 11, pp. 4295–4308, Nov. 2020.

[59] M. A. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, "A review of novelty detection," *Signal Process.*, vol. 99, pp. 215–249, Jun. 2014.

[60] K. Zhong, M. Han, and B. Han, "Data-driven based fault prognosis for industrial systems: A concise overview," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 2, pp. 330–345, Mar. 2020.

[61] P. M. Papadopoulos, V. Reppa, M. M. Polycarpou, and C. G. Panayiotou, "Scalable distributed sensor fault diagnosis for smart buildings," *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 3, pp. 638–655, May 2020.

[62] X. Xu, X. Yan, C. Sheng, C. Yuan, D. Xu, and J. Yang, "A belief rule-based expert system for fault diagnosis of marine diesel engines," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 2, pp. 656–672, Feb. 2020.

[63] J. Li, Z. Wang, C. K. Ahn, and Y. Shen, "Fault detection for Lipschitz nonlinear systems with restricted frequency-domain specifications," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 12, pp. 7486–7496, Dec. 2021.

[64] T. Sun, D. Zhou, Y. Zhu, and M. V. Basin, "Stability, l_2 -gain analysis, and parity space-based fault detection for discrete-time switched systems under dwell-time switching," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 50, no. 9, pp. 3358–3368, Sep. 2020.

[65] S. Oprea et al., "A review on deep learning techniques for video prediction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 6, pp. 2806–2826, Jun. 2022.

[66] A. Rasouli, “Deep learning for vision-based prediction: A survey,” 2020, *arXiv:2007.00095*.
[67] R. Isermann, *Fault-Diagnosis Systems: An Introduction From Fault Detection to Fault Tolerance*, Springer, 2005.
[68] Å. Fast-Berglund, L. Gong, and D. Li, “Testing and validating extended reality (xR) technologies in manufacturing,” *Procedia Manuf.*, vol. 25, pp. 31–38, May 2018.
[69] Y. Zhang, B. Xu, and T. Zhao, “Convolutional multi-head self-attention on memory for aspect sentiment classification,” *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 4, pp. 1038–1044, Jul. 2020.
[70] C. Maslach, S. E. Jackson, and M. P. Leiter, *MBI: Maslach Burnout Inventory*. Palo Alto, CA, USA: Consulting Psychol. Press, 1996.
[71] C. Maslach, “Job burnout: New directions in research and intervention,” *Current Direct. Psychol. Sci.*, vol. 12, no. 5, pp. 189–192, 2003.
[72] K. Rambabu, D. M. Kumar, V. K. Kumar, S. Nath, and R. Y. Jun, “Optical pattern inspection for flex PCB—Challenges & solution,” *IFAC Proc. Vol.*, vol. 41, no. 2, pp. 8196–8202, 2008.
[73] V. U. Sankar, G. Lakshmi, and Y. S. Sankar, “A review of various defects in PCB,” *J. Electron. Testing*, vol. 38, no. 7, pp. 481–491, 2022.
[74] T. Zhou, M. Chen, and J. Zou, “Reinforcement learning based data fusion method for multi-sensors,” *IEEE/CAA J. Automatica Sinica*, vol. 7, no. 6, pp. 1489–1497, Nov. 2020.
[75] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli, “Multimodal fusion for multimedia analysis: A survey,” *Multimedia Syst.*, vol. 16, no. 6, pp. 345–379, 2010.
[76] R. Socher et al., “Recursive deep models for semantic compositionality over a sentiment treebank,” in *Proc. Conf. Empirical Methods Natural Lang. Process.*, 2013, pp. 1631–1642.
[77] E. Adamopoulou and L. Moussiades, “An overview of chatbot technology,” in *Proc. IFIP Int. Conf. Artif. Intell. Appl. Innovat.*, 2020, pp. 373–383.
[78] B. Luo, R. Y. Lau, C. Li, and Y.-W. Si, “A critical review of state-of-the-art chatbot designs and applications,” *Wiley Interdiscipl. Rev. Data Mining Knowl. Discov.*, vol. 12, no. 1, 2022, Art. no. e1434.
[79] W. Huang, K. F. Hew, and L. K. Fryer, “Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning,” *J. Comput. Assisted Learn.*, vol. 38, no. 1, pp. 237–257, 2022.
[80] A. Calado, P. Roselli, V. Errico, N. Magrofuoco, J. Vanderdonckt, and G. Saggio, “A geometric model-based approach to hand gesture recognition,” *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 52, no. 10, pp. 6151–6161, Oct. 2022.
[81] P. M. Ferreira, D. Pernes, A. Rebelo, and J. S. Cardoso, “DeSIRe: Deep signer-invariant representations for sign language recognition,” *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 51, no. 9, pp. 5830–5845, Sep. 2021.

Yutong Wang (Member, IEEE) received the Ph.D. degree in control theory and control engineering from the University of Chinese Academy of Sciences, Beijing, China, in 2021.

She is currently an Assistant Professor with the State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing. Her research interest covers defect detection.

Jiangong Wang (Student Member, IEEE) received the Bachelor of Engineering degree in electronic information and engineering from Tongji University, Shanghai, China, in 2018. He is currently pursuing the Ph.D. degree in pattern recognition and intelligent systems with the Institute of Automation, Chinese Academy of Sciences, Beijing, China, and the University of Chinese Academy of Sciences, Beijing.

His research interests include parallel vision, unsupervised learning, traffic scene understanding, and medical image processing.

Yansong Cao received the M.B.A. degree from the Yale School of Management, New Haven, CT, USA, in 2007. He is currently pursuing the Ph.D. degree with the Department of Engineering Science, Faculty of Innovation Engineering, Macau University of Science and Technology, Macau, China.

His research interests include smart manufacturing and defect detection.

Shixing Li is an Associate Researcher with North Automatic Control Technology Institute, Taiyuan, China. His main research interest is computer vision. His main technical directions include equipment intelligence and parallel fields.

Oliver Kwan is the Founder and the Chairman of Motion G, Inc., Singapore. His research focuses on the relationships between statistics, inference, and knowledge transfer. He also works on industrial applications of digital twin, foundation model, and intelligent control.