DOI 10.15302/J-SSCAE-2020.04.017

A Product Process Adaptive Design Method Based on Manufacturing-Related Big Data

Wei Wei, Chen Zheng, Yuan Jun

School of Mechanical Engineering and Automation, Beihang University, Beijing 100191, China

Abstract: As digital and smart production methods are widely applied in manufacturing, enterprises should focus more on the values of manufacturing-related big data, which are important for innovation in product process design. This study proposes a product process adaptive design method based on manufacturing-related big data. This method is proposed based on the requirements of enterprises for data–business deep integration, and it is used to solve the insufficient utilization of manufacturing-related data among enterprises. Hence, a product process adaptive design model named "data + knowledge + decision" is proposed, and the data mining and utilization processes are summarized for the model, i.e., multisource heterogeneous data fusion, data cleaning and preprocessing, data conversion and dimensionality reduction, data mining, data visualization, and design decision. Subsequently, an automobile welding process is presented as an example. A prediction model depicting the relationship between welding parameters and welding defects is established to improve the welding quality and realize adaptive welding process design. This study reveals that manufacturing-related big data contain rich knowledge and patterns; therefore, they can be used to guide product design decisions and support the product process adaptive design under different manufacturing environments. In the future, to enhance the integration of manufacturing-related big data and product process design, the integration of big data with 5G technology should be promoted, and investment should be increased for the development of big data and algorithm design platforms.

Keywords: product process; adaptive design; manufacturing-related big data; data mining; knowledge discovery

1 Introduction

Emerging information technologies, such as big data and cloud computing, integrate with modern industrial technologies rapidly, during which some new manufacturing modes have been introduced, such as smart manufacturing and digital twins [1]. Data-driven methods are effective in managing complex problems. With the wide application of CNC machine tools, sensors, data collectors, and other devices, enterprises have accumulated a significant amount of manufacturing data [2] that shares similar characteristics with big data [3]. The association rules and manufacturing knowledge pertaining to these data have yet to be discovered; hence, data mining is an effective method to guide the iteration and optimization of product process design.

A flow chart of the product process design and data-driven design is shown in Fig. 1. Typical product process design methods are primarily based on simulations and experiments [4–6]. When a new design task is proposed, the designer uses design software to arrange and plan the product process. Subsequently, the trial product is produced and tests are conducted to define the final design before mass production. During the manufacturing process, real-time data of machines, tools, and workpieces are collected to monitor the actual status [7]. Finally, quality inspection is performed on products that generate considerable inspection data [8]. The procedures above are known as "forward design." The data mining process is performed to explore the knowledge and relationships such that the product process design can be optimized and a "backward design" can be realized between product quality and process

Received date: May 26, 2020; **revised date:** June 28, 2020

Corresponding author: Wei Wei, associate professor of School of Mechanical Engineering & Automation of Beihang University. Major research fields include product design and smart manufacturing. E-mail: weiwei@buaa.edu.cn

Funding program: National Key R & D Program (2018YFB1701703); National Natural Science Foundation of China (51675028); CAE Advisory Project "Strategic Studies on the Internet Plus Action Plan (2035)" (2018-ZD-02)

Chinese version: Strategic Study of CAE 2020, 22 (4): 042–049

Cited item: Wei Wei et al. A Product Process Adaptive Design Method Based on Manufacturing-Related Big Data. *Strategic Study of CAE*, https://doi.org/10.15302/J-SSCAE-2020.04.017

parameters by analyzing manufacturing data [9]. Currently, the product process design method based on simulation and experiments is relatively mature in manufacturing enterprises, whereas the "backward design" driven by manufacturing big data is progressing slowly. Problems such as poor adaptation of product process design to the manufacturing environment, as well as the tardiness of product iteration and version update must be solved urgently.

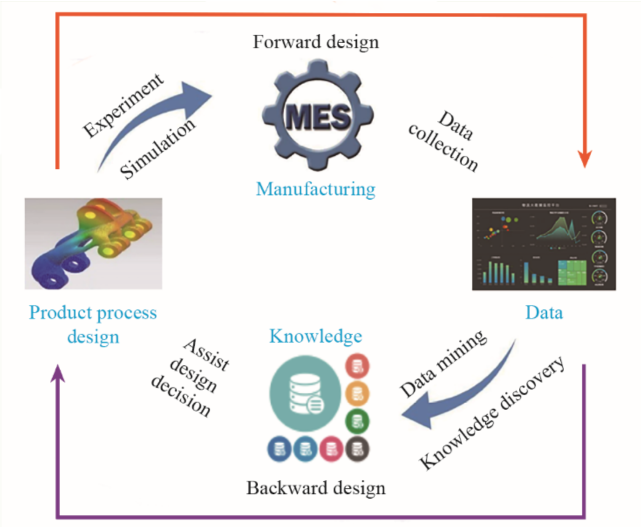

Fig. 1. Product process design and data-driven design flow chart.

Note: MES refers to manufacturing execution system.

Data mining emerged in the 1990s, with the goal of extracting, analyzing, and discovering knowledge from a significant amount of data [10,11]. In recent years, researchers have focused on the application of big data manufacturing. Köksal et al. [8] reviewed the application of data mining to product quality improvement from 1997 to 2007. Ferrerio et al. [7] applied data mining to the optimization of the drilling process of aircraft parts. Zhang [12] analyzed users' comment data to facilitate product innovation and update product markets. Kretschmer [13] investigated the application of data mining in designing the assembly process. Tao [14,15] proposed a product design framework driven by digital twin technology based on big data. Ma [16] proposed a data-driven product family design method and applied it to real product family design examples. Fu [17] provided guidance for data-driven product innovation, and Yu [18] discussed new modes of big data for guiding product design.

Conventional data mining technologies are applied primarily to structured data that differ from massive, multisource, heterogeneous manufacturing data, thereby necessitating data fusion, data processing, and data for managing these manufacturing data. Therefore, a new product process adaptive design mode named "data + knowledge + decision" is proposed herein such that manufacturing big data are indispensable to in product design.

2 Analysis of product process adaptive design method

Owing to the widespread use of CNC machine tools and sensors as well as the rapid development of modern technologies such as the Internet of Things and artificial intelligence, the amount of manufacturing data collected by enterprises has increased, and the rate of data increase is continually increasing. This has become a problem and an opportunity for enterprises to fully utilize manufacturing data. The new product process adaptive design mode, "data + knowledge + decision," is an effective method for investigating the potential value of manufacturing data, driving the adaptive design of the product process, and promoting the improvement in product service and product quality.

The data-mining process for the data-driven product process adaptive design mode is illustrated in Fig. 2. The procedures involved are as follows: (1) Based on data generated by the manufacturing system, manufacturing big data are formed and partitioned into structured and unstructured data, the latter of which are structured through feature and information extractions to combine heterogeneous data. (2) Massive data are selected and sampled such that the sampled data not only reflect the overall distribution, but also retain or reduce the complexity of data analysis. (3) The sampled data are cleaned and preprocessed. (4) The dimensionality of the data is reduced through data transformation. (5) Data mining is performed to discover the laws, patterns, and knowledge pertaining to the data. (6) The data are visualized to assist design decision-making and promote adaptive product process design.

DOI 10.15302/J-SSCAE-2020.04.017

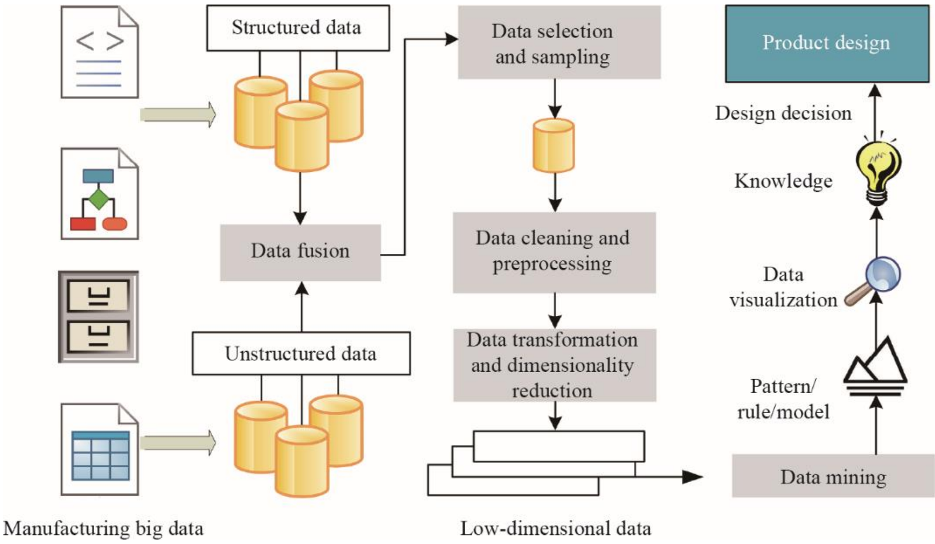

Fig. 2. Data mining process of data-driven product process adaptive design mode.

2.1 Multisource heterogeneous data fusion

As the number and types of sensors used in advanced manufacturing systems increase, the data becomes more complex and various, such as the time series produced by mechanical sensors and speed sensors, image data produced by vision sensor capturing images, and data produced by product data management systems. The distinguishing characteristics of manufacturing big data are diversity, complexity, and uncertainty, which render the unified expression of manufacturing big data key to solving the problem of data fusion [19].

(1) Fusion of time-series data. The states of machine tools and workpieces are typically monitored by collecting time series data of speed, stress, temperature, etc., the frame rate of which is relatively high (40–80 fps). To realize the match and fusion between time series data and the other data, a downsampling operation must be conducted on the time series data, where the mean, variance, etc. represent the state of certain manufacturing stages.

(2) Fusion of image data. For image data in the manufacturing system, the information must be extracted for a structured expression. Conventional image data extraction is primarily manual, which is inefficient. Owing to the application of convolutional neural networks in the field of computer vision, the image perception ability of machines has progressed significantly; hence, convolution neural networks are a favorable alternative for extracting information from image data and converting unstructured image data into structured data.

2.2 Data cleaning and preprocessing

In the data-driven product process adaptive design, both the quantity and quality of data are important. Data quality includes the aspects of accuracy, completeness, consistency, and validity. Accuracy refers to the degree to which the data fit the physical world; completeness refers to the proportion of valid values in the data; consistency refers to the degree to which the data satisfy the specified constraints; and validity refers to the value density of data.

Data cleaning refers to the removal of poverty data for improving data quality, including the processing and denoising of abnormal and missing data. The distance measurement or clustering method can be used to detect outliers in the dataset and remove data points that are distant from the center of the dataset. For missing values in the dataset, interpolation is an effective method of data processing as it yields a complete dataset. For noise in the dataset, a smoothing filtering algorithm can be used for denoising. Manufacturing big data typically contain a significant number of repeated data points; as such, data simplification must be performed to reduce data redundancy.

2.3 Data transformation and dimensionality reduction

Manufacturing big data are the digital expression of manufacturing systems and processes. More data collection in the manufacturing system facilitates the complete description of the system. Data mining, however, can become difficult because of the dimensionality disaster due to the significant amount of data. Dimensionality disasters are one of the most pressing issues encountered when managing high-dimensional data; it not only exacerbates the time and space complexity, but also results in the non-convergence problem of the analysis algorithm.

Various data collected in the manufacturing process typically exhibit certain correlations, such as the voltage and current in the welding process, rotation speed, and cutting speed of machine tools. This correlation causes dimensional redundancy and increases unnecessary calculations, as a result of which data dimensionality reduction is particularly important for retaining the appropriate feature data and reducing redundant data from a high-dimensional data space. The dimension-reduced data can not only retain the original information, but also avoid dimensionality disasters.

2.4 Data mining

The main purpose of mining manufacturing big data is to predict and extract the rules. Data prediction is used to predict the possible value of other variables with known variables, and rule extraction is used to identify hidden rules or knowledge in the data, both of which are beneficial to product design. Typically used data mining methods include classification, regression, clustering, and correlation analysis [20].

(1) Classification and regression are the main methods for data prediction, the former of which maps the feature space to discrete variables, whereas the latter maps the feature space to continuous variables. The classification and regression prediction processes are shown in Fig. 3.

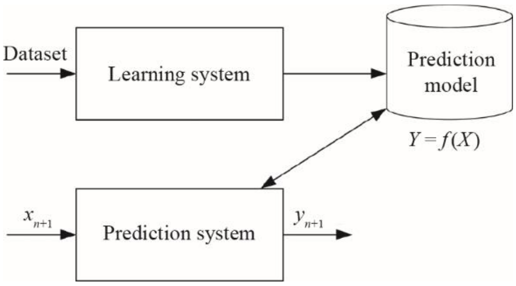

Fig. 3. Prediction process of classification and regression.

For a training dataset T,

$$T = (x_1, y_1), (x_2, y_2), \dots, (x_n, y_n),$$

where $x_i \in R^n$ is the input; $y_i \in R$ is the output, which is a discrete variable in the classification and continuous variable in the regression. The goal is to build a prediction model $Y = \mathrm{f}(X)$, where the output y_{n+1} can be predicted using the new input x_{n+1}.

(2) Clustering analysis refers to the use of an unsupervised learning method to segment instances into natural groups and cluster similar instances into corresponding clusters based on the hidden pattern behind the instance data. Clustering analysis is widely used in manufacturing data mining. For example, a large number of error instances are collected in error prediction, and clustering is performed based on the error characteristics, by which error subcategories are discovered. The typically used clustering algorithms are the K-means method [21], expectation maximum algorithm [22], and Bayesian clustering [23].

(3) Correlation analysis can be expressed in the form of rules. For example, the rule of "if A then B" can be explained as a product that satisfies "condition A likely satisfies condition B." Correlation analysis in manufacturing big data is important for determining the design constraints. Factors that affect product quality can be discovered using mining correlation rules; therefore, design variables can be adjusted during design decision-making to realize adaptive design.

2.5 Data visualization

Data visualization refers to the intuitive visual representation of data that uses image processing, computer graphics, and user interfaces to express, model, and animate data. Data visualization refers to presenting valuable information in large datasets with scatter diagrams, matrix diagrams, curved surfaces, and other intuitive forms to deepen users' understanding of data and accelerate information acquisition [24]. Based on data types, data visualization can be categorized into high-dimensional data visualization, time series data visualization, hierarchical data visualization, and network data visualization [25].

2.6 Design decision

As shown in Fig. 4, the knowledge and correlation rules produced by data mining can be used to assist design decisions, such as for determining design parameters, product architecture, and process flows. When a new

DOI 10.15302/J-SSCAE-2020.04.017

manufacturing plan is executed, the digital twin system collects the data generated from the manufacturing system, which contains two applications in design decisions: one is to be stored in a database for data mining, whereas the other is to provide designers with timely feedback and assist design decision-making to verify the feasibility and rationality of process design decisions of products.

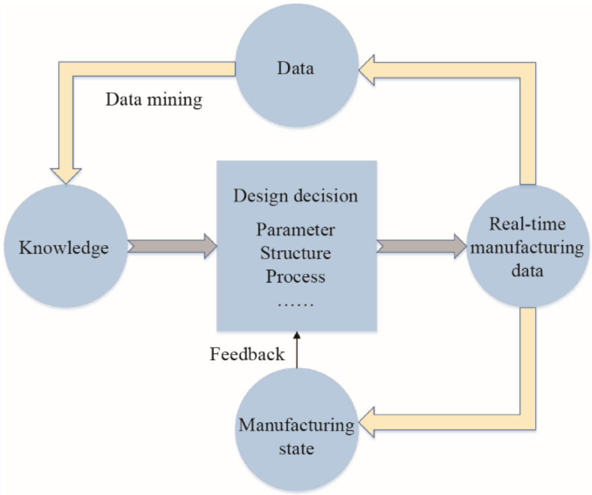

Fig. 4. Data mining assisting design decisions.

3 Application of manufacturing big data in product process adaptive design——data mining as example

To better understand the driving role of manufacturing big data in product process design, the application of data mining in the welding process optimization of automobile products is presented as an example. The data collected during the welding process primarily included process parameters for welding control and X-ray images of welding seams for quality inspection. The X-ray images could not be directly used for data analysis, implying that defect detection must be performed first. Furthermore, the relationship between the defects of welding seams and welding process parameters must be analyzed.

3.1 Defect detection of X-ray image of welding seam

X-ray image detection is a widely used method in welding quality inspection and can accurately detect the position, shape, size, and distribution of defects. For the X-ray images of welding seams generated during X-ray detection (Fig. 5), conventional processing methods used are primarily based on the experience of the professional. Hence, they are inefficient and impose high requirements on the technical capabilities of the inspectors; consequently, large numbers of images are difficult to analyze. Therefore, a semantic segmentation algorithm was used to automatically detect welding defects, where each pixel of the X-ray images is classified into backgrounds, pores, slag inclusions, and cracks.

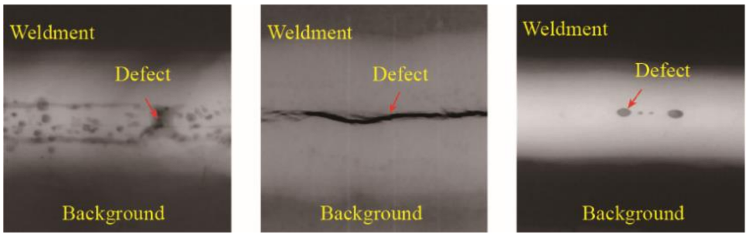

Fig. 5. X-ray image of welding seams.

Based on the convolutional neural network, the encoder–decoder network structure was used for the semantic segmentation of X-ray images (Fig. 6). For the input image, “convolution + batch normalization + activation function” calculation was first performed. Subsequently, a 2 × 2 maximum pooling layer was used for downsampling, rendering the image size 1/32 of the original after the same operation was performed five times. Next, a de-pooling operation was performed for upsampling, where “convolution + batch normalization + activation function” calculation was performed until the image returned to its original size. Finally, a normalized exponential function

(Softmax) was used to classify the output pixel value and count the proportion of each type of defect to the number of all pixels.

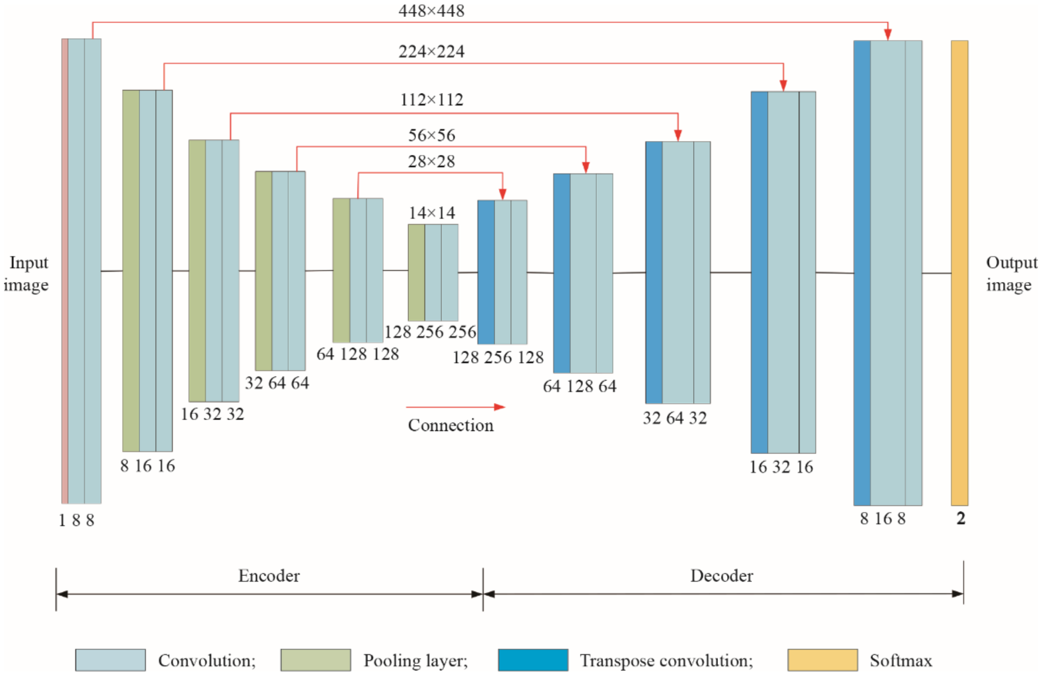

Fig. 6. Semantic segmentation of X-ray images based on "encoder–decoder" structure.

3.2 Correlation model between process parameters and welding quality

A steel plate pre-coating primer followed by butt welding was adopted in the manufacturing of some automobile brands, while the welding quality problem was prominent and required urgent improvement. In this study, 200 X-ray images of welding seams, welding currents, voltages, shielding gas flows, and thicknesses of pre-painted film were recorded during manufacturing for process monitoring. Based on the data above, the processes affecting the welding quality were investigated to optimize the process design.

(1) The X-ray images were detected using a pre-trained semantic segmentation defect detection model by counting the number and proportion of all types of defects. Fig. 7 shows that the pore defect is the most prominent in this product batch; hence, the relationship between pore defects and process parameters must be analyzed.

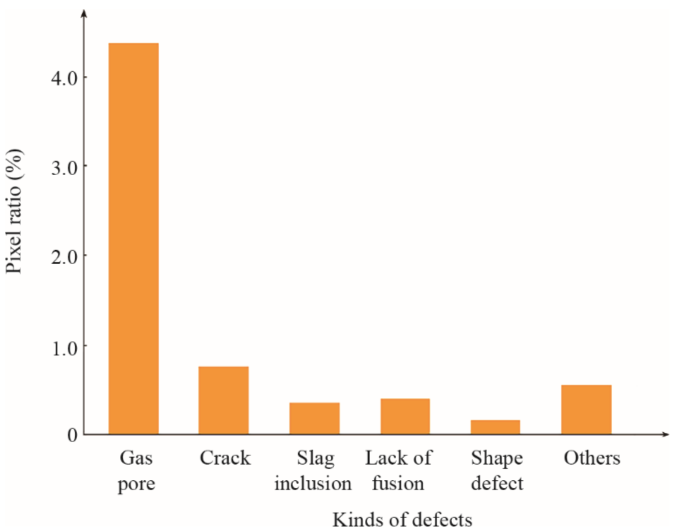

Fig. 7. Detection results of six types of defects.

(2) The time series data of the process parameters related to each image were downsampled to obtain the process parameters corresponding to each welding seam. The correlation between the process parameters and pore defects

DOI 10.15302/J-SSCAE-2020.04.017

was analyzed using the Pearson coefficient method and visualized. As shown in Fig. 8, the welding current, voltage, shielding gas flow, and thickness of the pre-painted film affected the pore defects, and a regression model between the process parameters and number of pores was established using a fully connected neural network (Fig. 9). The input layer of the model included four nodes representing the film thickness, welding speed, welding current, and gas flow, and the output layer represented the proportion of gas pores corresponding to the process parameters in the input layer.

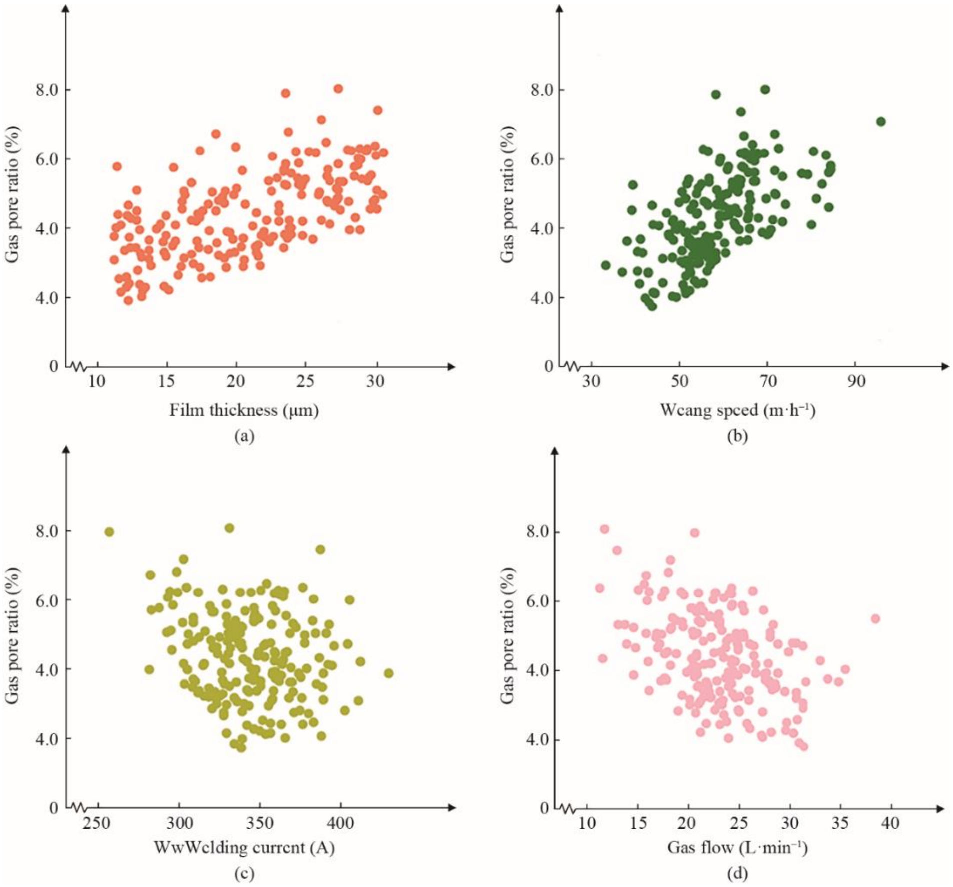

Fig. 8. Correlation between welding parameters and gas pore ratio.

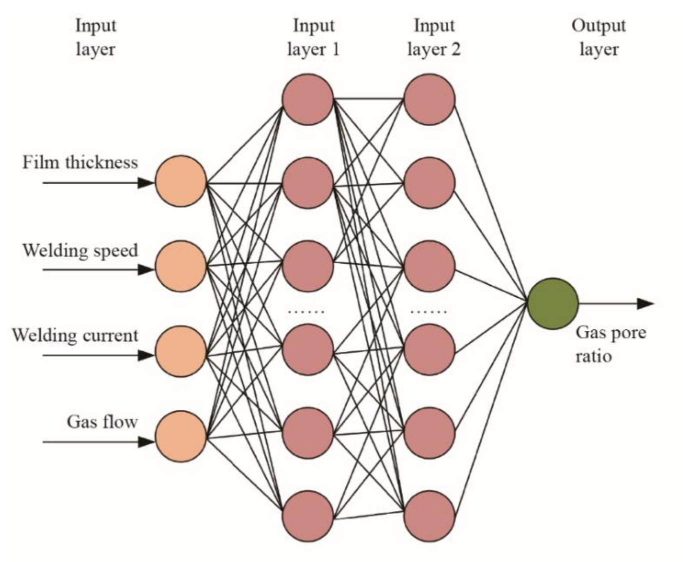

Fig. 9. Regression model with fully connected neural network.

(3) The 200 images were segregated into training and validation sets at a ratio of 6:4, and the mean square error (MSE) loss was selected as the loss function. As shown in Fig. 10, the MSE loss decreased to less than 0.02 after 200 iterations, indicating the good predictive performance of the defect detection regression model that can facilitate designers in adjusting process parameters, reducing welding defects, and improving product quality.

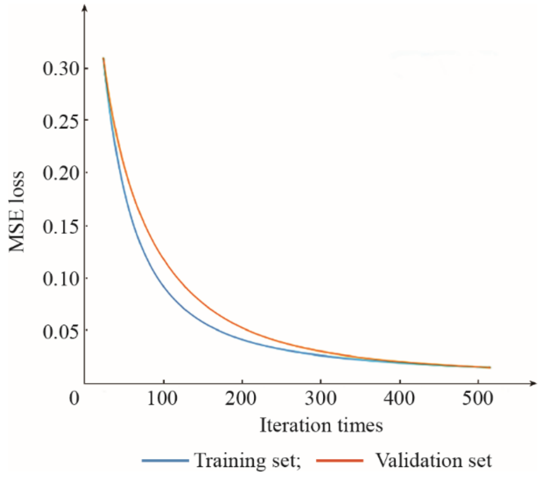

Fig. 10. MSE loss of training and validation sets.

3.3 Application of defect detection model in product process design

The defect detection model can be used to inspect welding products. Conventional manual testing methods are inefficient and require long inspection cycles, rendering it difficult for designers to obtain timely feedback regarding the welding process. The defect detection model can predict whether a product process design plan is reasonable based on the regression relationship between the welding process and welding defects, thereby assisting designers in optimizing design decisions and improving welding quality.

4 Conclusion

A manufacturing big data-driven product process adaptive design method was proposed herein, in which the function of manufacturing data in product process design, the problems of manufacturing data mining, and the application steps of data mining were explained. The application of data mining in the welding process was presented as an example to predict welding defects and promote the optimization of the welding process. Results indicated that the knowledge and rules yielded by manufacturing big data mining can provide effective guidance for product designers and promote the iterative optimization of the product process, thereby achieving the goal of big-data-driven product process adaptive design.

To further promote the application of the manufacturing big-data-driven adaptive design method in product process design, the following are suggested:

(1) Promote the combination of manufacturing big data and 5G technology. Manufacturing big data are the basis for data mining and analysis, and they are the driving force that promotes the transformation of enterprises to innovation. Combined with 5G technology, manufacturing enterprises can obtain more comprehensive data to analyze all manufacturing factors.

(2) Strengthen the development of big data platforms and algorithm design platforms. The big data platform is the basis for the efficient storage, reading, import, and export of manufacturing data. The development of a stable and reliable big data platform can improve the efficiency of data mining, ensure the timeliness of knowledge utilization, and guarantee the information security of enterprises. The algorithm design platform is the basis for the rapid design of data-mining algorithms. The development of a highly integrated algorithm design platform can lower the threshold of the application of data mining, allowing more designers to create new value through manufacturing big data mining.

(3) Promote typical demonstration cases for enterprises The focus is to improve the application demonstration and promotion level, innovate the traditional manufacturing concepts, and stimulate the application of manufacturing big data in the manufacturing industry with cases of successful enterprises.

DOI 10.15302/J-SSCAE-2020.04.017

References

[1] Tao F, Qi Q, Liu A, et al. Data-driven smart manufacturing [J]. Journal of Manufacturing Systems, 2018, 48: 157 169.

[2] Zhang Y, Ren S, Liu Y, et al. A framework for big data driven product lifecycle management [J]. Journal of Cleaner Production, 2017, 159: 229 240.

[3] Wang G C. From 3V to 5V: A further research of big data boosting economic behavior [J]. Tianjin Social Sciences, 2017 (2): 94 99. Chinese.

[4] Thomas W, Oenoki T, Altan T. Process simulation in stamping—Recent applications for product and process design [J]. Journal of Materials Processing Technology, 2000, 98 (2): 232 243.

[5] Zhou F M, Liu W, Wang X, et al. Exhaust valve groove robotic automation welding process design [J]. Journal of Jiangsu University of Science and Technology (Natural Science Edition), 2014, 28(2): 135 139. Chinese.

[6] Qiao Y H, Han J, Zhang C Y. Jig of passenger vehicle chassis design and welding process simulation [C]. Wuhan: 2010 International Conference on Mechanic Automation and Control Engineering, 2010.

[7] Ferreiro S, Sierra B, Irigoien I, et al. Data mining for quality control: Burr detection in the drilling process [J]. Computers & Industrial Engineering, 2011, 60(4): 801 810.

[8] Köksal G, Batmaz I, Testik M C. A review of data mining applications for quality improvement in manufacturing industry [J]. Expert Systems with Applications, 2011, 38(10): 13448 13467.

[9] Tucker C S, Kim H M. Data-driven decision tree classification for product portfolio design optimization [J]. Journal of Computing and Information Science in Engineering, 2011, 9(4): 1 14.

[10] Fayyad U, Stolorz P. Data mining and KDD: Promise and challenges [J]. Future Generation Computer Systems, 1997, 13(2–3): 99 115.

[11] Han J, Kamber M, Pei J. Data mining: Concepts and techniques (third edition) [M]. San Francisco: Morgan Kaufman Publishers, 2012.

[12] Zhang H, Rao H G, Feng J Z. Product innovation based on online review data mining: A case study of Huawei phones [J]. Electronic Commerce Research, 2018, 18(1): 3 22.

[13] Kretschmer R, Pfouga A, Rulhoff S, et al. Knowledge-based design for assembly in agile manufacturing by using data mining methods [J]. Advanced Engineering Informatics, 2017, 33: 285 299.

[14] Tao F, Sui F Y, Liu A, et al. Digital twin-driven product design framework [J]. International Journal of Production Research, 2019, 57(12): 3935 3953.

[15] Tao F, Cheng J F, Qi Q L, et al. Digital twin-driven product design, manufacturing and service with big data [J]. The International Journal of Advanced Manufacturing Technology, 2018, 94(9–12): 3563 3576.

[16] Ma J & Kim H M. Product family architecture design with predictive, data-driven product family design method [J]. Research in Engineering Design, 2016, 27(1): 5 21.

[17] Fu X, Asorey H. Data-driven product innovation [C]. Sydney: The 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2015.

[18] Yu C, Zhu L. Product design pattern based on big data-driven scenario [J]. Advances in Mechanical Engineering, 2016, 8(7): 1 9.

[19] Bleiholder J, Naumann F. Data fusion [J]. ACM Computing Surveys, 2008, 41(1): 1 41.

[20] Wang G H, Jiang P. Survey of data mining g [J]. Journal of Tongji University (Natural Science Edition), 2004, 32 (2): 246 252. Chinese.

[21] Sun J G, Liu J, Zhao L Y. Clustering algorithms research [J]. Journal of Software, 2008, 19(1): 48 61. Chinese.

[22] Xia B, Emilion R, Wang H W. Comparison of EM algorithm and dynamic clustering algorithm of Dirichlet mixtum samples [J]. Journal of Beijing University of Aeronautics and Astronautics, 2019, 45(9): 1805 1811. Chinese.

[23] Zhu J, Chen L F. Bayesian clustering algorithm for categorical data [J]. Journal of Computer Application, 2017, 37(4): 1026 1031. Chinese.

[24] Xu W P. Design and implementation of data visualization system architecture [D]. Beijing: Beijing Jiaotong University (Master's thesis), 2015. Chinese.

[25] Wang Y W. A study on better models for association rule mining in Chinese medicine prescriptions and their applications [D]. Wuhan: Hubei University of Chinese Medicine (Master's thesis), 2017. Chinese.