On-line part deformation prediction based on deep learning

Zhiwei Zhao[1], Yingguang Li*[2], Changqing Liu[2], James Gao[3]

1. College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, 29 Yudao Street, Nanjing, 210016, China.
2. National Key Laboratory of Science and Technology on Helicopter Transmission, Nanjing University of Aeronautics and Astronautics, Nanjing, 210016, China. (E-mail: liyingguang@nuaa.edu.cn, Tel: 0086-025-84895835)
3. School of Engineering, University of Greenwich, Chatham Maritime, Kent, ME4 4TB, UK

Abstract: Deformation prediction is the basis of deformation control in manufacturing process planning. This paper presents an on-line part deformation prediction method using a deep learning model during numerical control machining process, which is different from traditional methods based on finite element simulation of stress release prior to the actual machining process. A fourth-order tensor model is proposed to represent the continuous part geometric information, process information, and monitoring information, which is used as the input to the deep learning model. A deep learning framework with a Conventional Neural Network and a Recurrent Neural Network has been constructed and trained by monitored deformation data and process information associated with interim part geometric information. The proposed method can be generalised for different parts with certain similarities and has the potential to provide a reference for an adaptive machining control strategy for reducing part deformation. The proposed method was validated by actual machining experiments, and the results show that the prediction accuracy has been improved compared with existing methods**.** Furthermore, this paper shifts the difficult problem of residual stress measurement and off-line deformation prediction to the solution of on-line deformation prediction based on deformation monitoring data.

Key words: Deformation prediction, Monitoring data, Deep learning, Tensor model

1. Introduction

Deformation is one of the main causes of geometric error and scrap parts (Moehring et al. 2018), especially in machining large scale complex aerospace structural parts, where deformation is typically controlled within 0.05mm/m (Heinzel et al. 2017). Deformation mainly results from the release of initial residual stress within the workpiece and machining-induced stress. Previous research in deformation control had been focused on the areas of material, design and manufacturing. In the manufacturing area that this research is devoted to, there were basically two approaches, i.e., machining process optimisation based on deformation prediction, and closed process control based on real time monitoring data (Li and Wang 2017). Previous research was mainly aimed at establishing the relationship between deformation and the impacting factors using numerical simulation (Chantzis et al. 2013) and analytical models (Wu et al. 2016). Many deformation prediction methods were reported, in which prediction results were obtained prior to machining processes based on measured residual stress distribution. In practice, it is very difficult to accurately predict deformation because of the difficulty to accurately measure residual stress, and the challenge of modelling the properties of workpiece materials and the subsequent effect on plastic deformation, chip formation and separation, all of which have influence on machining-induced stress.

2. Literature review

2.1 Deformation prediction approaches

Numerical methods are mainly applied in commercial finite element methods (FEM) software tools to predict workpiece deformation. Chantzis et al. (2013) developed a deformation control method by predicting the deformation caused by initial residual stresses which is obtained by displacement measurements. Guo et al. (2009) measured the initial residual stresses by using a modified layer-removal method and established a finite element model of machining distortion for aluminum alloy multi-frame components. Huang et al. (2015) presented a prediction model based on FEM to study the effect of black initial residual stress on part deformation and a crack compliance method was used to measure original residual stress. Gulpak et al. (2013) proposed a hybrid model consisting of 3 sub-models with a regression model and FE simulation to calculate shape deformation. Regression models were used to calculate specific boundary conditions for FE simulations.

Compared to equivalent FEM models, analytical models present major advantages by significantly reducing the computational time from days to seconds (Arrazola et al. 2013). Shang (1995) presented a two-dimensional analytical model and deduced the formulae of re-distribution and distortion by initial residual stress during the milling process by elasticity theory. Wu et al. (2016) proposed a mathematical prediction model in thin-walled plates by using the finite difference method (FDM), which is simple and efficiency. Nervi et al. (2009) established a third-dimensional mathematical model for the prediction of distortion of airframe components from aluminum plates according to the Navier–Lamè equation. Heinzel et al. (2017) presented an analytical method based on multilayer source stress model to analyze the effects of initial residual stress and machining induced stress on materials modifications. The shortcoming of analytical models is that these methods are only suitable for very simple and regular parts, it is not possible to calculate the deformation of complex parts.

Basically, there are two inducing factors impacting deformation: original residual stress and machining residual stress due to cutting effects, while accurate measurement or calculation of both these factors are worldwide challenges, so accurate deformation prediction is still a difficult problem (Chantzis et al. 2013; Arrazola et al. 2013).

2.2 Deformation monitoring approaches

In-process monitoring technologies have been developed to obtain key physical data such as cutting force, vibration and temperature (Pratama et al. 2017) and interim machining effects of the workpiece which are difficult to predict before machining (Abellan-Nebot and Subirón 2010). Möhring et al. (2010) proposed a process force monitoring approach by integrating a sensory fixture system and a milling spindle so as to analyze workpiece deflection, which supplied a basis for deformation compensation. Yoshioka et al. (2014) presented a method to monitoring the distance between the workpiece surface and the tool, which provides a means to improve surface quality. In order to overcome the difficulty of workpiece deformation monitoring, the authors' research group developed an adaptive machining method using flexible fixtures, where deformation can be monitored when the workpiece is not restrained by the fixtures during the non-cutting intervals (Li et al. 2015; Hao et al. 2018).

In-process monitoring methods can obtain accurate machining data, and it is useful for timely problem such as local deformation compensation based on monitoring data. The monitoring deformation data of a current process provides a very important reference to obtaining the deformation of the following machining processes, and is therefore worth further study.

2.3 Data-driven prediction approach

Data-driven methods have been widely used in manufacturing process applications, such as fault detection (Wang et al. 2018a), machining condition monitoring (Pratama et al. 2017) and machining accuracy prediction (Pimenov et al. 2017; Cheng et al. 2015). Deformation prediction is a time-series prediction problem, because the deformation monitoring data sets are time-dependent. Time series prediction approaches are used to predict future trending based on historical time series data. There are many time series prediction approaches which have been successfully applied in industrial areas. For example, hidden Markov model, Lagrange's interpolation, ARIMA (Autoregressive Integrated Moving Average). Yu (2017) proposed an adaptive hidden Markov model-based online learning framework for faulty bearing detection and performance degradation monitoring. Guiassa and Mayer (2011) proposed a predictive compliance-based model for compensation in multi-pass milling by on-machine probing, where Lagrange's interpolation-based approach is adopted so as to compensate the final cut more effectively. Wang et al. (2018b) improved forecasting compensatory control to guarantee the remaining wall thickness for pocket milling of a large thin-walled part, where ARIMA with Kalman filtering was used to predict the deformation due to cutting force of the next time point for deformation compensation.

Essentially, Lagrange's interpolation is suitable for the modeling of two variables, while ARIMA and Kalman filtering are only suitable for linear systems. The issue of deformation prediction due to residual stress is a nonlinear problem with multiple influencing factors. It is still a challenge for the hidden Markov model to address the high dimensional observation values, because the training of the hidden Markov model is very difficult for high dimensional data. Therefore, existing time series prediction approaches are not suitable for this paper.

2.4 Deep learning approaches

The principle of machining deformation prediction is to establish the relationship between workpiece deformation and its impacting factors. From this perspective a deep learning method has been proofed to be an effective way to establish complex relationship models for the machining process (Wang et al. 2018), which benefits from plentiful accurate monitoring data during the machining process.

Deep learning models are composed of multiple processing layers to learn the relationships of data with multiple levels of feature abstraction (Lecun et al. 2015). Deep learning methods have many successful applications, such as picture recognition (He et al. 2016) and playing games (Silver et al. 2016). It has turned out that it is very good at discovering intricate structures in high-dimensional data and has attracted increasing attention in manufacturing industry (Wang et al. 2017a). Wang et al. (2017a) proposed a data-driven prediction model for material removal rate during chemical mechanical polishing using a deep belief network with the input to the model being process parameters. Fu et al. (2015) employed deep belief networks to build end milling feature spaces for cutting states monitoring based on vibration signal. Lin et al. (2018) proposed a method to inspect LED chips automatically by using a deep convolutional neural network. Wu et al. (2018) proposed a novel approach for fault prognosis of equipment based on a Long Short-Term Memory(LSTM) network.

For part deformation prediction the challenging problems are the complexity of part structure as well as the coupling effects with process information, especially related to the continuous changing of part geometry during the machining process. At present octree-based part geometry representation is deemed as a successful way for deep learning of geometric shape involved problems (Wang et al. 2017b). However, this method cannot represent the geometric state to reflect the stiffness of the workpiece as well as the changes during the machining process. The multi-dimensional data including the non-structural geometric information, process data, and monitoring data makes it a challenge to represent all the deformation factors, and currently there is a lack of a suitable deep learning network for this issue.

3. Overview of the proposed approach

To address the above challenges, this paper proposes an on-line part deformation prediction method driven by monitoring data and process information associated with interim workpiece states. The prediction model is constructed based on a deep learning method, in which the data are represented by a fourth-order tensor model. The deformation resulting from the machining to follow is predicted based on the monitoring data of the current workpiece state, as illustrated in Fig. 1. This method is different from existing prediction models, because it avoids the need to measure the residual stress. The model established the relationship of previous deformation and subsequent deformation by developing a network using combined Conventional Neural Network (CNN) and Recurrent Neural Network (RNN).

The data, which provided an accurate representation of the workpiece state and can be measured accurately, have been recorded or monitored during the machining process. There is significant potential to further mine the internal relationships between the recorded data and the deformation. The complex data contains many physical relationships, especially for the coupling effects amongst the data. Data driven intelligence based on deep learning can model the complex multivariate nonlinear relationships among those data with no in-depth understanding of system physical behaviors being required.

In this paper, the aim of deformation prediction is to establish the relationship between current workpiece states and the deformation resulted by the follow-up machining process. First of all, a prediction model needs to be constructed from a wide range of workpiece states data. However, the recorded data are collected from different manufacturing processes, which are multi-format, multi-dimensional, and multi-modality. This makes it a great challenge for data analysis. The representation of geometry-process-deformation information is the key to the solution. In addition to that, how to construct a deep learning framework to learn the relationships among the data is another challenge. This paper addresses these two issues in the following sections.

4. Representation of geometry-process-deformation information based on tensor model

Workpiece deformation during machining is affected by its stiffness, the machining process performed and residual stress distribution. The workpiece stiffness is reflected by the interim state of workpiece geometry; the machining process can be represented by process parameters; and residual stress is reflected by the monitored deformation. In order to establish the relationship between the current workpiece state and the deformation resulting by the follow-up machining process, all of the geometry-process-deformation data need to be presented appropriately, which is the key to the deep learning model. The data are multi-dimensional, and related to individual workpiece states during machining. The information is un-structured, and continuously changing during the whole machining process. There are also coupling effects, i.e., the geometric information is affected by process information. Efficient data processing needs better expression, and those complexities place higher demands on the data representation.

Inspired by the application of tensor model in computer vision, data mining and signal processing (Kolda and Bader 2009), the authors found the advantages of tensor in this application. It has the ability to represent multi-dimensional data while keeping the structural relationship among the different information elements. Tensor is a multi-dimensional representation of data while maintaining the individual structure of each data element, which is very important for multi-dimensional data analysis. A tensor can be represented as $\mathbb{A} \in \mathbb{R}^{I_1 \times I_2 \times \ldots I_n \ldots \times I_N}$, the order of the tensor is N, the data in $\mathbb{A}$

is $a_{i_1 \dots i_n \dots i_N}$ and $1 \le i_n \le I_n$, I_n is the dimension size of nth order. The tensor can be unfolded in each order to represent information from a special perspective. In the proposed model, a tensor including related information of different dimensions can be represented as:

$$\mathbb{A} = [\boldsymbol{A_1}, \boldsymbol{A_2}, \dots, \boldsymbol{A_s}] \; \mathbb{A} \in \mathbb{R}^{I_1 \times I_2 \times I_3 \times I_4} \tag{1}$$

$$\boldsymbol{A_s} = [A_1, A_2, A_3, A_4] \; , \boldsymbol{A_s} \in \mathbb{R}^{I_1 \times I_2 \times I_3}, \; A_i \in \mathbb{R}^{I_1 \times I_2} \tag{2}$$

where $\mathbb{A}$ is a fourth order tensor representing the whole machining states; $\boldsymbol{A_s}$ is the sub-tensor model unfolded in the fourth order of $\mathbb{A}$, which represents the information in each interim machining state $\boldsymbol{s}$; A_i is a sub-array of the different deformation related data in the tensor model. I_1 and I_2 can be regarded as a plane mapped from the perspective of the X-Y plane of the part. I_3 is the dimension of the deformation related information, whose size is determined by the kinds of different machining information. Tensor $\boldsymbol{A_s}$ is sliced from this order and forms a sub-array in I_1 and I_2; I_4 is the machining state. In each A_i, the data of the plane is the value of the x, y position in the part. As illustrated in Fig. 2, the four types of machining related data, including part geometry, cutting-depth, fixture positions and deformation data, are represented by four sub-arrays.

To address the difficulty in representing the workpiece state, the workpiece geometric state is mapped to order A_1 of the proposed tensor model, each of the elements in order A_1 corresponds to a unit of the workpiece, i.e., each unit represents a certain area of the workpiece, where the element in this sub-array is the material height of the corresponding part in the X-Y region. If there is no material in an element, the value is 0 in the corresponding unit. In order to keep the genericity of the representation, the number of data elements in one dimension of this sub-array can be set according to the workpiece size of the part types. A_2 is the depth of cut in position (x, y). The depth of cut sub-array represents the location where material is removed and the cutting depth of the follow-up process. A_3 and A_4 are fixture information and monitored data.

The deformation data are monitored by responsive fixtures on certain key points(Li et al. 2015), and the fixture positions in A_3 are set as value 1, the other positions are 0. In the same way, deformation data is recorded in the tensor at the monitored point. In this model, geometric-process-deformation information is related, and some other process parameters could be represented to add the number of A_i. The geometric and process information of interim workpiece states can be obtained using the dynamic feature modelling method proposed by the authors (Li et al. 2012). This model maintains the multi-dimensional data structure of individual data elements, and the interdependency and complementarity between them. Tensor representation provides a possibility to reveal the implicit relationships among multi-dimensional data, and a basis to improve the ability of generalising the model in tensor space.

5. Deformation prediction model based on a deep learning network

In order to establish the deformation prediction model, deep learning is applied in the whole prediction process as well as in the process of extracting features and special relationships between different data elements. As illustrated in Fig. 3, the tensor model is the input to the prediction model. In the deep learning process, CNN is used to extract features and relationships in the tensor model by translating the tensor model to vectors with reduced dimensions. Then the vectors are sent to RNN, which establishes the relationship between the input information in the current workpiece state and the deformation resulting from the following machining process.

5.1 Feature extraction process based on CNN

For deformation prediction modeling, a suitable representation of the impacting factors is essential for a better learning of the intricate relationships in the machining data. However, the workpiece state information is modeled by tensor including $I_1 \times I_2 \times I_3 \times I_4$ data, which makes it a challenge to learn the relationship in tensors directly. Generally, the useful information is at a low dimensional space in the tensor model, where useful structural and relational information is distributed sparsely. In order to obtain a low dimensional representation for the tensor, a feature extraction process is essential for the prediction process.

CNN has multiple layers and has been developed to derive information from the data with multiple arrays. It has been used in many aspects, such as relational information extraction and image classification. A deep hierarchy CNN model with its multilayer abstraction ability can map the object into a regularised space. The main module of the CNN network is the convolutional layers, in which a set of weights called filter bank are used to connect the previous layer and the feature maps, and extract the feature with highly correlated data. For non-linear deformation related data, CNN can obtain a deep structure for modeling the data. The deep stacks of CNN enable the model to extract higher level representation of the latent output of the previous layer. As a result, CNN offers a powerful model for learning robust hierarchical implicit representations for structured inputs.

In order to abstract the features effectively, a CNN structure called Res-Net is used here. The Res-Net is easy to optimize, and makes the network structure deeper. It means that the higher level features will be abstracted from the original data.

5.2 Predicting follow-up machining deformation based on RNN

Machining Deformation is a time series process related to the historical machining data of workpiece. The workpiece status changes throughout the machining process, and the workpiece deformation changes in both positive and negative directions. Thus it is difficult to predict prior to the following machining process. For this reason, the deformation information of the following workpiece state in the historical data of machined parts will be used for predicting the deformation of workpiece state following the 'current' state.

RNN is a kind of artificial neural sequence model, where connections between units form a directed cycle. RNN is often used for sequenced tasks. It takes an input sequence as one element at one time from the input layer, maintaining a 'state vector' in the recurrent layer that implicitly contains historical information of all the past elements of the sequence, and then the historical information can be transferred to the current output layer or the next recurrent layer. RNN has a strong capability of capturing contextual information within sequences of different lengths.

Deformation related data are auto-correlation data. The subsequent deformation could be predicted by the data of the current workpiece state and the historical data. In the deformation prediction model, the sequence of the deformation related data should be considered by using the auto-correlation data to construct the model and describe the dynamic process. Therefore, the machining deformation data are 'context-sensitive' data. Under this circumstance, a class of RNN, i.e., Bidirectional Long Short Term Memory (BiLSTM) is adopted to establish the prediction model by considering the entire deformation related information. LSTM is a variant of RNN proposed and applied to solve the problem of vanishing gradient of RNN models for long series.

For the structure depicted in Fig. 4, BiLSTM uses two LSTMs to learn each relationship of the sequence based on both the past and the future context of the input data, and then the hidden states of the two LSTMs are combined together to predict the deformation of the following machining process. One LSTM processes the sequence from t_1 to t_n, and the other is from t_n to t_1. And then the two hidden states h_t^1 and h_t^2 are concatenated into a vector, the combined outputs y is the prediction of data series. $x(t)$ is the input of the network, for each time step, the input is composed of current part states including the geometry and deformation data, and the following machining process parameters. In sequenced workpiece states, for a particular workpiece state, both information of previous and following states is useful and complementary to each other. Therefore, the BiLSTM is more suitable to model the relationships among the sequenced workpiece states.

The structure of the complete prediction model is shown in Fig. 3. It has two modules: a feature extraction module and a deformation prediction module. Part deformation is correlated in the entire part space. In order to propagate from one corner of the feature information to another, Fully Connected (FC) layers are used to handle the information propagation network. Thus, the extracted features of CNN are unfolded into a vector and pass through two FC layers. The output of the propagation network is represented as:

$$F_i = \boldsymbol{CNN}(\boldsymbol{A}_i) \quad F_i \in \boldsymbol{F}(\boldsymbol{s}) = [F_1, F_2, \dots, F_s] \tag{3}$$

$$M_i = f_1(U_2 \cdot ((U_1 \cdot F_i) + b_{u1})) + b_{u2}) \tag{4}$$

where $\boldsymbol{A_i}$ is tensor model of the machining state i; $\boldsymbol{CNN}$ is the CNN network operator; F_i is the feature extracted by CNN; U_1, U_2, b_{u1}, b_{u2} are weights and biases of two FC layers. f_1 is Tanh activation function. M_i is the output result of the FC layers.

Then the vectors are input to the prediction module. The structure of prediction module consists of some LSTM units and four layers in each prediction process, i.e., FC input layer, recurrent layer, multimodal layer and FC regression layer. Firstly, the FC input layer extends the extracted data to a dense representation with the same size of LSTM input:

$$I_\mathrm{i} = f_2(W_1 \cdot M_i + b_1) \tag{5}$$

where W_1, b_1 are the weights and biases of the FC input layer. f_2 is the Tanh activation function. In the LSTM input units, the two hidden states are connected by the multi-modal layer and the result is passed through the last FC output layer to export the result:

$$h_s^1 = \boldsymbol{LSTM_1}(I_1, I_2, \dots, I_s) \qquad h_s^2 = \boldsymbol{LSTM_2}(I_1, I_2, \dots, I_s) \tag{6}$$

$$P_{s+1} = f_4(O(H_3 \cdot (f_3(H_1 \cdot h_t^1 + H_2 \cdot h_t^2) + b_2) + b_3) \tag{7}$$

where $\boldsymbol{LSTM_1}$ and $\boldsymbol{LSTM_2}$ are BiLSTM operator; h_s^1 and h_s^2 are the hidden states of the two LSTMs; H_1, H_2, H_3, O, b_2, and b_3 are the weights and biases of multimodal layer(m-layer) and regression layer(r-layer). f_4 and f_3 are activation functions, P_{s+1} is the output of prediction module.

When predicting the deformation of state $s+1$, the result P_{s+1} could be regarded as the deformation monitoring data after the current machining process, then the data could be used to predict the deformation resulted by the following machining process with associated process information.

5.3 Network training and prediction performance

Totally 480 parts including frames and beams with aluminum alloy material were collected from aircraft manufacturing factories and our machining laboratory, and the geometric, process and deformation information of the interim machining process of each part is included and stored in the proposed four-order tensor model, which is used to train our deformation prediction model. Both samples from real machining and from simulation environments are included in our data set. While the samples are still not enough for training the model effectively, the dataset is expanded to 43200 samples with data augmentation by data translation and rotation approach. The samples were divided into a training set and an evaluation set in the division principle of 80% and 20%.

The selection of the hyper-parameters of the deep learning framework is very important for achieving a good prediction performance. The main hyper-parameters include the number of neurons per layer, the number of layers, the definition of neuron activation function as well as dropout or learning rate.

The parameters of CNN, RNN and other layers' configurations are described in Table 1 and Table 2. To speed up the training rate, the Mini-Batch Normalization was adopted in convolutional layers.

The entire network has deep structures, i.e., the CNN has 22 layers, its parameters consist of filter values and connection weights, along with their biases. In order to get optimized weights of the network, mini-batch stochastic gradient descent was used and backpropagation were applied. The learning rate starts with a value of 1×10^{-5}, and the batch-size is 16, which is limited by the memory size of GPUs. The initial momentum was selected from 0.7 to 0.95. The parameters are randomly initialized using normal distribution. The network is trained by regressing to the true deformation data. The deep learning model is implemented in Python^{T} and CUDA based on CNN and RNN kernels in PyTorch^{T}. The proposed model is trained and tested on a workstation with an Intel Xeon E5 CPU and NVIDIA Tesla P4 GPU. Because of the huge parameters, the training time of the network requires about 14d.

The mean square is used as a loss function for the entire network:

$$MSE = \frac{1}{m}\sum_{i=1}^{m}(d_s - P_s)^2 \tag{8}$$

An evaluation function is given to evaluate the prediction model:

$$\text{Eva} = \frac{1}{N}\sum_{i=1}^{N}\left|d_{predicted_i} - d_{measured_i}\right| \tag{9}$$

The training results illustrated in Fig. 5 suggests that the proposed deep learning framework performs well in deformation prediction. The training loss is less than 9×10^{-6}, and the evaluation result is nearly 0.009 for the evaluation set. The error may be significantly reduced with the increase of the monitoring points and the training set.

Some process data of Res-Net of CNN is visualized in Fig. 6 with some process data of convolutional layers being extracted from the network in the training process. Fig. 6 (a) shows the process data of a training sample (beam part) in Conv-1. There are 32 process data processed by 32 filter banks in Conv-1 on the geometric-process-deformation information of the beam part. For the sake of observation, the process information data are processed as images, the grayscale value of the pixel represents the level of sensitivity of the filter banks for the original information. Since the original input information of the geometric-process-deformation is organized on geometry, the output images are also visualized based on geometry, causing the images to look like the part. The higher value of grayscale means more sensitive.

The process data of Convs-2, Convs-3, Convs-4, and Convs-5 are also showed in Fig. 6. The output images of Fig. 6 (a) and Fig. 6 (b) of Conv-1 and Conv-2 only identifies the primary features, such as edge information in the images. With the more convolutional processing, the higher features are extracted by the network. From the view of Fig. 6 (c), (d) and (e), the filter banks are sensitive to different area features containing more information.

In order to further analyse the network performance, Fig. 7 shows the 64 process information images of Cons-3, which are easily explained. After eight convolution layers, the higher information is extracted from the original data. As illustrated in Fig. 7, the 1-red-frame shows this filter bank is sensitive to the area features looking like long ribs. The other three red frames indicate that the other three filter banks are sensitive to the middle ribs, pockets and the whole ribs of the part. The extracted information is important for deformation prediction. The other images are difficult to translate. With the increase in the number of network layers, the complex representations of the geometric-process-deformation information will be extracted, which is more difficult to understand by humans. The process data shows that the deep learning framework has a powerful capability for information extraction in deformation prediction problems.

6. Case study

In order to verify the proposed method, two typical aircraft structural parts composed of different machining features but with the same dimensions (600mm×120mm×30mm) are machined. The material of the parts is aluminum alloy 7075, which are machined on a DMG 80P machine tool, as shown in Fig. 8. As illustrated in Fig. 9, the deformations of each part are monitored at 4 points by responsive fixtures. The machining parameters are listed in Table 3. As illustrated in Fig. 9, the part is fixed by seven fixtures. The red points represent the fixed fixtures, and the four green points represent the responsive fixtures, which are used to monitoring the deformation of the part. The responsive fixtures are always set at the far end of the long side to get the overall deformation of the parts. All the parts in our samples are long beam parts with a similar structure and shape to the parts in Fig. 9. In order to meet machining and monitoring requirements, at least 4 responsive fixtures are required.

Deformation data were sampled when the part deformation was released during the machining intervals, and the machining process followed a layer-by-layer removal process. In other words, the deformation data were collected after each layer was finished. Both of the two workpieces were machined by 7 layers with 6 roughing layers and the last finishing layer. After each layer was machined, the deformation data were monitored by responsive fixtures, and the machining information including part geometry, process parameters, positions of monitoring points and part deformations, was modeled as a tensor and input to the prediction model to predict the deformation resulted by the following layer of machining.

Prediction results for each part are presented in Fig. 10, and it can be seen that the prediction accuracy is not good in the earlier stages, as there is only small sample data to describe the workpiece state. The prediction accuracy gets better as the monitoring data increase. The max prediction error of the last layer is 0.029 mm for the parts machined, and the relative errors are within 10.61%. The prediction results are listed in Table 4.

A typical deformation prediction method based on FEM (Tang et al. 2013) shows that, the relative error is 20% and prediction error is 0.035 mm with part dimension 260mm×40mm×20mm, while the scale and complexity are both much less than the parts of our paper.

The model presented in our paper improves the prediction accuracy and does not need the measurement of the residual stress. The prediction model based on deep learning has shown its feasibility in predicting deformation, and with the increase of the dataset size the prediction accuracy increases.

7. Conclusions and further work

Traditional offline-based prediction methods require the destruction of the entire material to obtain a distribution of residual stresses. Due to the inaccuracy measurement of residual stress and the different distribution of the same batch of materials, the existing methods can only predict an approximate general trend, which is not enough for deformation control. This paper presents an on-line part deformation prediction method for deformation control using a deep learning method, which is different from existing prediction methods before machining process based on FE simulation. Our proposed deformation prediction method is driven by part deformation monitoring data, which is a very accurate representation of the state of the part. CNN and RNN are trained by monitored deformation and process information including a large amount of historical data associated with interim workpiece states, which are represented by a proposed forth order tensor model. Case studies show a high prediction accuracy, with which suitable machining process strategies can be adopted to control the deformation.

Compared with existing deformation prediction methods, the proposed method avoids the measurement and residual and machining-induced stress inside the workpiece, which is difficult and inaccurate to measure. This method shifts the difficult problem of residual stress measurement and off-line deformation prediction to the solution of on-line deformation prediction based on deformation monitoring data. The proposed method also provides a reference for representing the input information of machining problems using deep learning methods.

At present, the aluminium alloy materials are used in the case studies. Howerer it should also be feasible to apply the approach to other metal materials such as titanium alloys. In the proposed tensor model, it would be feasible to add other process information in addition to cutting-depth, and to extend the tensor dimension of I_3. On-going and further work of the authors is to investigate the above promising application areas.

Acknowledgement

The reported research was funded by the National Natural Science Foundation of China (Ref. 51775278), the National Natural Science Foundation of China – Chinese Aerospace Science and Technology Corporation on Advanced Manufacturing (Ref: U1537209), and the Jiangsu Province Outstanding Youth Fund (Ref: BK20140036).

References

Abellan-Nebot, J. V., & Subirón, F. R. (2010). A review of machining monitoring systems based on artificial intelligence process models. *International Journal of Advanced Manufacturing Technology, 47*(1-4), 237-257.

Arrazola, P. J., Özel, T., Umbrello, D., Davies, M., & Jawahir, I. S. (2013). Recent advances in modelling of metal machining processes. *CIRP Annals - Manufacturing Technology, 62*(2), 695-718.

Chantzis, D., Van-Der-Veen, S., Zettler, J., & Sim, W. M. (2013). An industrial workflow to minimise part distortion for machining of large monolithic components in aerospace industry. *14th CIRP Conference on Modeling of Machining Operations (CIRP CMMO), 8*, 281-286.

Cheng, Q., Zhao, H., Zhao, Y., Sun, B., & Gu, P. (2018). Machining accuracy reliability analysis of multi-axis machine tool based on Monte Carlo simulation. *Journal of Intelligent Manufacturing*, 29(1), 191–209.

Fu, Y., Zhang, Y., Qiao, H., Li, D., Zhou, H., & Leopold, J. (2015). Analysis of Feature Extracting Ability for Cutting State Monitoring Using Deep Belief Networks. *Procedia CIRP, 31*, 29-34.

Guiassa, R., & Mayer, J. R. R. (2011). Predictive compliance based model for compensation in multi-pass milling by on-machine probing. *CIRP Annals - Manufacturing Technology, 60*(1), 391-394.

Gulpak, M., Sölter, J., & Brinksmeier, E. (2013). Prediction of Shape Deviations in Face Milling of Steel. *Procedia CIRP, 8*, 15-20.

Guo, H., Zuo, D. W., Wu, H. B., Xu, F., & Tong, G. Q. (2009). Prediction on milling distortion for aero-multi-frame parts. *Materials Science and Engineering a-Structural Materials Properties Microstructure and Processing, 499*(1-2), 230-233.

Hao, X., Li, Y., Chen, G., & Liu, C. (2018). 6+X locating principle based on dynamic mass centers of structural parts machined by responsive fixtures. *International Journal of Machine Tools & Manufacture, 125*, 112-122.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep Residual Learning for Image Recognition.* 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, 770-778.

Heinzel, C., Sölter, J., Gulpak, M., & Riemer, O. (2017). An analytical multilayer source stress approach for the modelling of material modifications in machining. *CIRP Annals - Manufacturing Technology, 66*(1), 531-534.

Huang, X. M., Sun, J., & Li, J. F. (2015). Finite element simulation and experimental investigation on the residual stress-related monolithic component deformation. *International Journal of Advanced Manufacturing Technology, 77*(5-8), 1035-1041.

Kolda, T. G., & Bader, B. W. (2009). Tensor Decompositions and Applications. *Siam Review, 51*(3), 455-500.

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

Li, J. G., & Wang, S. Q. (2017). Distortion caused by residual stresses in machining aeronautical aluminum alloy parts: recent advances. *International Journal of Advanced Manufacturing Technology, 89*(1-4), 997-1012.

Li, Y., Liu, C., Hao, X., Gao, J. X., & Maropoulos, P. G. (2015). Responsive fixture design using dynamic product inspection and monitoring technologies for the precision machining of large-scale aerospace parts. *CIRP Annals - Manufacturing Technology, 64*(1), 173-176.

Li, Y., Liu, X., Gao, J. X., & Maropoulos, P. G. (2012). A dynamic feature information model for integrated manufacturing planning and optimization. *CIRP Annals - Manufacturing Technology, 61*(1), 167-170.

Lin, H., Li, B., Wang, X., Shu, Y., & Niu, S. (2018). Automated defect inspection of LED chip using deep convolutional neural network. *Journal of Intelligent Manufacturing*, doi:https://doi.org/10.1007/s10845-018-1415-x.

Moehring, H. C., Wiederkehr, P., Gonzalo, O., &Kolar, P. (2018). *Intelligent Fixtures for the Manufacturing of Low Rigidity Components*: Springer International Publishing.

Möhring, H. C., Litwinski, K. M., & Gümmer, O. (2010). Process monitoring with sensory machine tool components. *CIRP Annals - Manufacturing Technology, 59*(1), 383-386.

Nervi, S., Szabó, B. A., & Young, K. A. (2009). Prediction of Distortion of Airframe Components Made from Aluminum Plates. *Aiaa Journal, 47*(7), 1635-1641.

Pimenov, D. Y., Bustillo, A., & Mikolajczyk, T. (2017). Artificial intelligence for automatic prediction of required surface roughness by monitoring wear on face mill teeth. *Journal of Intelligent Manufacturing, 29*(1), 1-17.

Pratama, M., Dimla, E., Lai, C. Y., & Lughofer, E. (2017). Metacognitive learning approach for online tool condition monitoring. *Journal of Intelligent Manufacturing*.Doi. 10.1007%2Fs10845-017-1348-9

Shang, M. E. H. (1995). Prediction of the dimensional instability resulting from machining of residually stressed components. *Texas Tech University*.

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Driessche, G. V. D., et al. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*(7587), 484.

Tang, Z. T., Yu, T., Xu, L. Q., & Liu, Z. Q. (2013). Machining deformation prediction for frame components considering multifactor coupling effects. *International Journal of Advanced Manufacturing Technology, 68*(1-4), 187-196.

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

Wang, P., Gao, R. X., & Yan, R. (2017a). A deep learning-based approach to material removal rate prediction in polishing. *CIRP Annals - Manufacturing Technology, 66*(1), 429-432.

Wang, P. S., Liu, Y., Guo, Y. X., Sun, C. Y., & Tong, X. (2017b). O-CNN: octree-based convolutional neural networks for 3D shape analysis. *Acm Transactions on Graphics, 36*(4), 1-11.

Wang, T., Qiao, M., Zhang, M., Yang, Y., & Snoussi, H. (2018a). Data-driven prognostic method based on self-supervised learning approaches for fault detection. *Journal of Intelligent Manufacturing*(3), 1-9.

Wang, X., Bi, Q., Zhu, L., & Ding, H. (2018b). Improved forecasting compensatory control to guarantee the remaining wall thickness for pocket milling of a large thin-walled part. *International Journal of Advanced Manufacturing Technology, 94*(5-8), 1-12.

Wu, Q., Ding, K., & Huang, B. (2018). Approach for fault prognosis using recurrent neural network. *Journal of Intelligent Manufacturing*(3), doi:https://doi.org/10.1007/s10845-018-1428-5.

Wu, Q., Li, D. P., Ren, L., & Mo, S. (2016). Detecting milling deformation in 7075 aluminum alloy thin-walled plates using finite difference method. *International Journal of Advanced Manufacturing Technology, 85*(5-8), 1291-1302.

Yoshioka, H., Shinno, H., Sawano, H., & Tanigawa, R. (2014). Monitoring of distance between diamond tool edge and workpiece surface in ultraprecision cutting using evanescent light. *CIRP Annals - Manufacturing Technology, 63*(1), 341-344.

Yu, J. (2017). Adaptive hidden Markov model-based online learning framework for bearing faulty detection and performance degradation monitoring. *Mechanical Systems & Signal Processing, 83*, 149-162.

Fig. 1 On-line part deformation prediction model

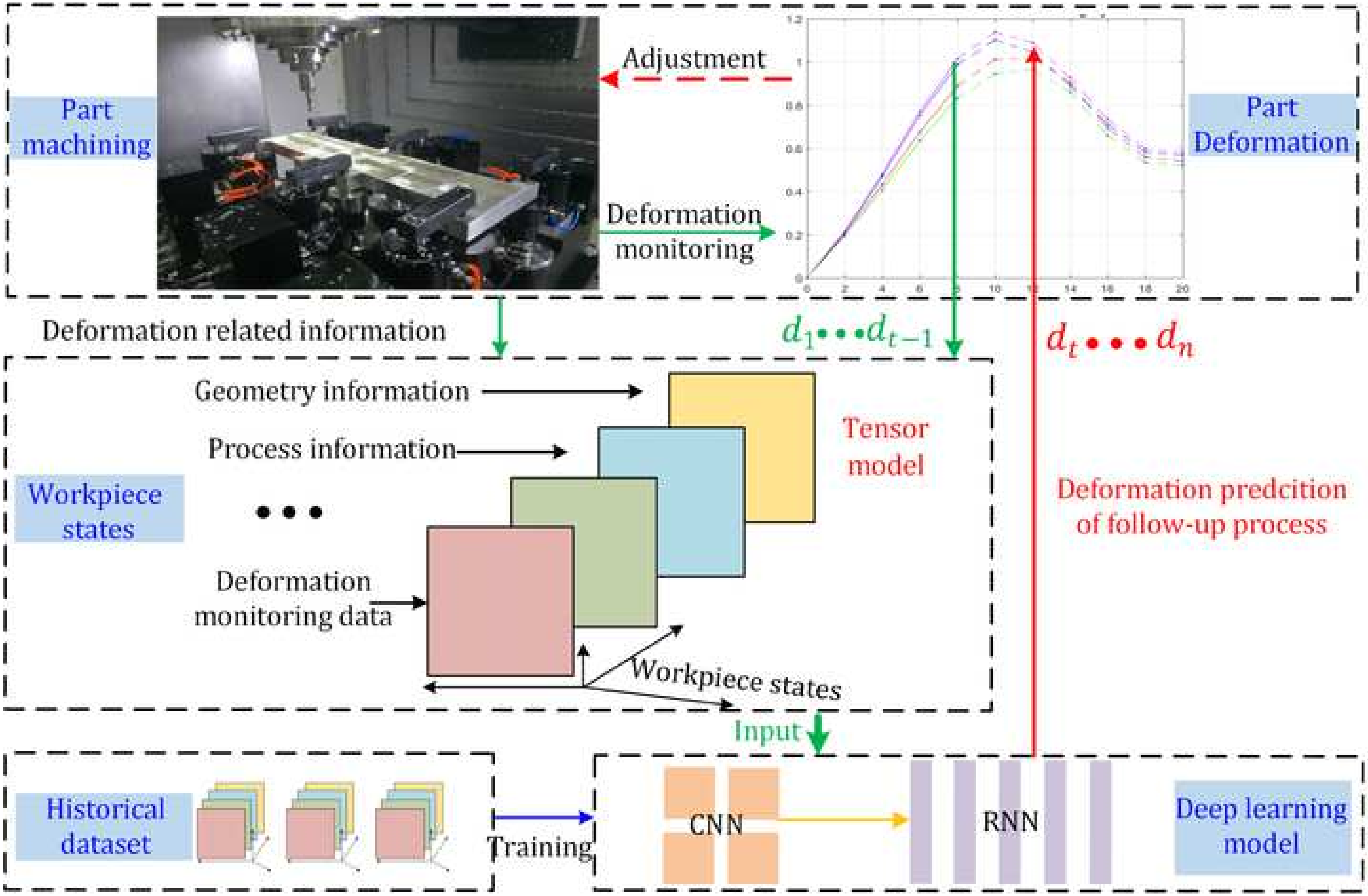


Fig. 2 Tensor model of geometry-process-deformation information

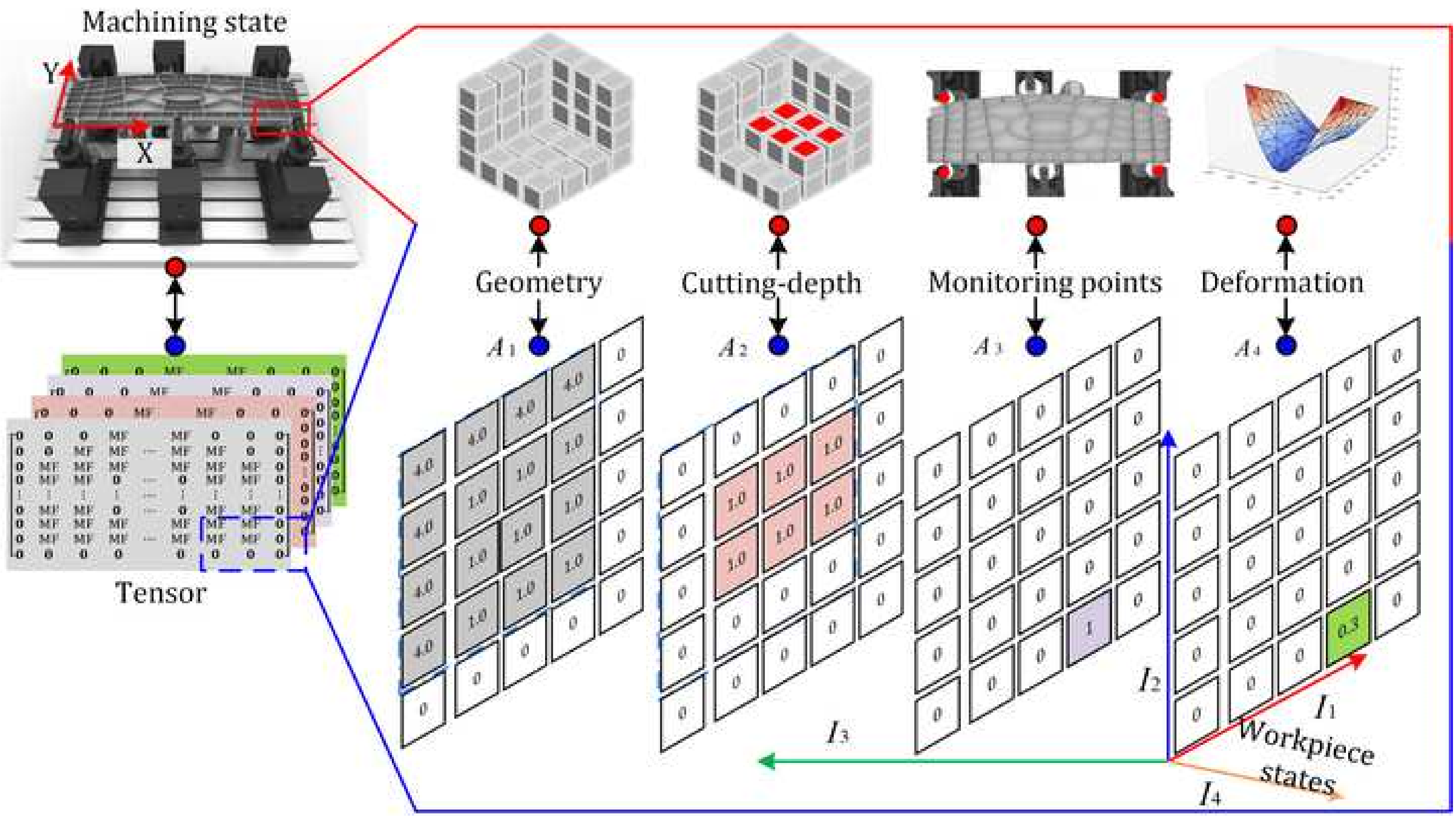


Fig. 3 The structure of deformation prediction model

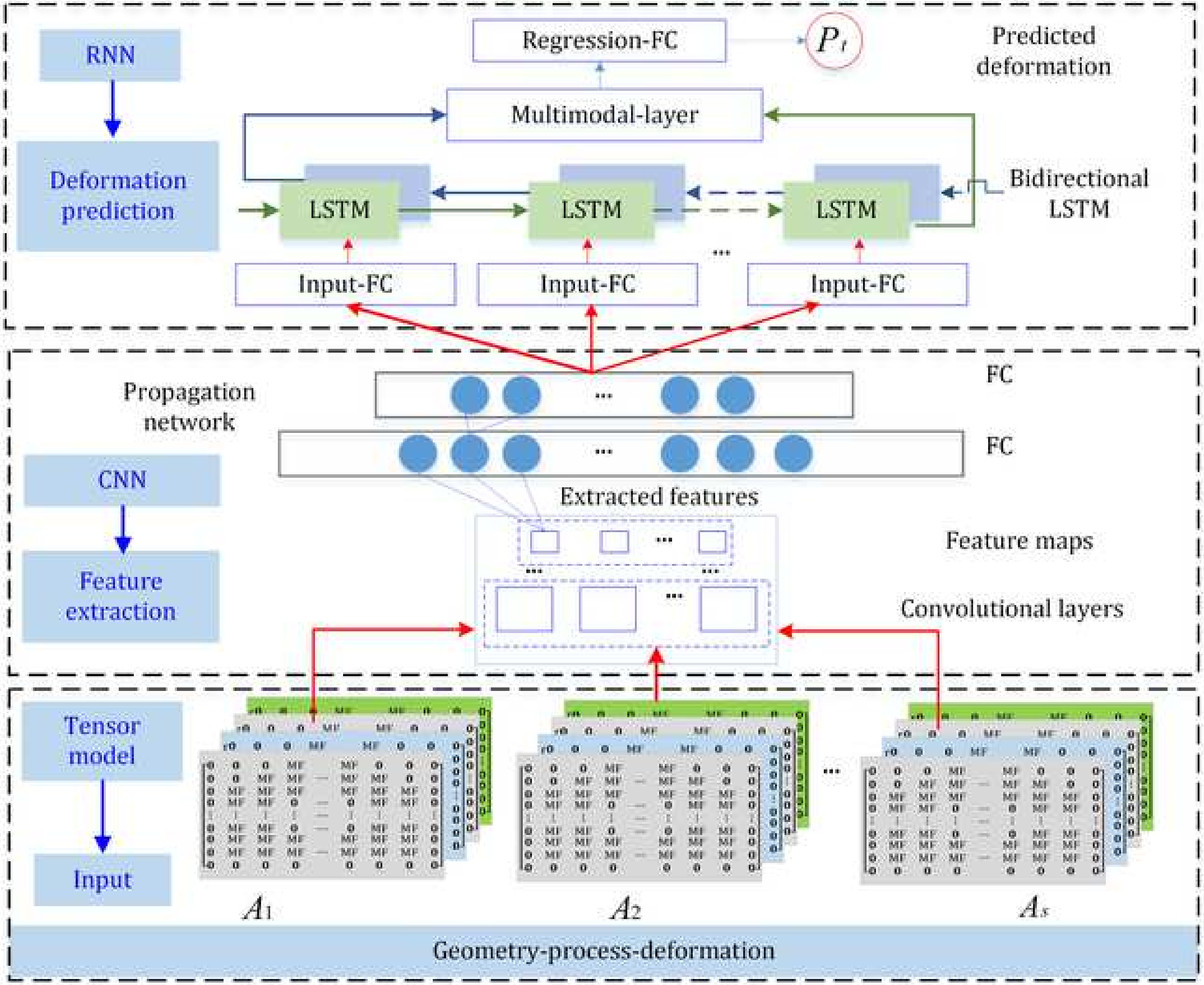


Fig. 4 Bidirectional Long Short Term Memory(BiLSTM)

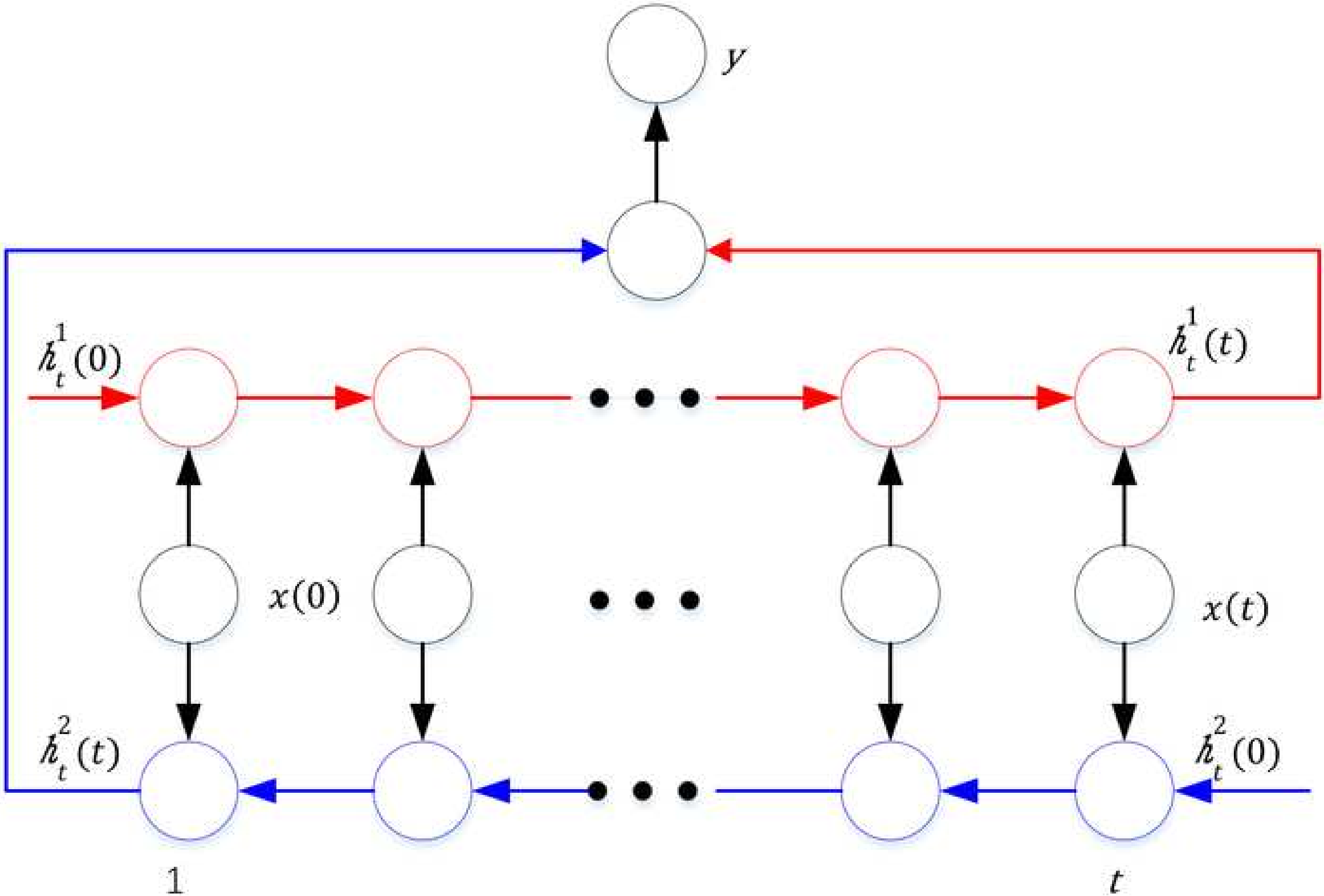


Fig. 5 Training Loss and Evaluation value

Click here to access/download;Figure;FIG. 5.tif

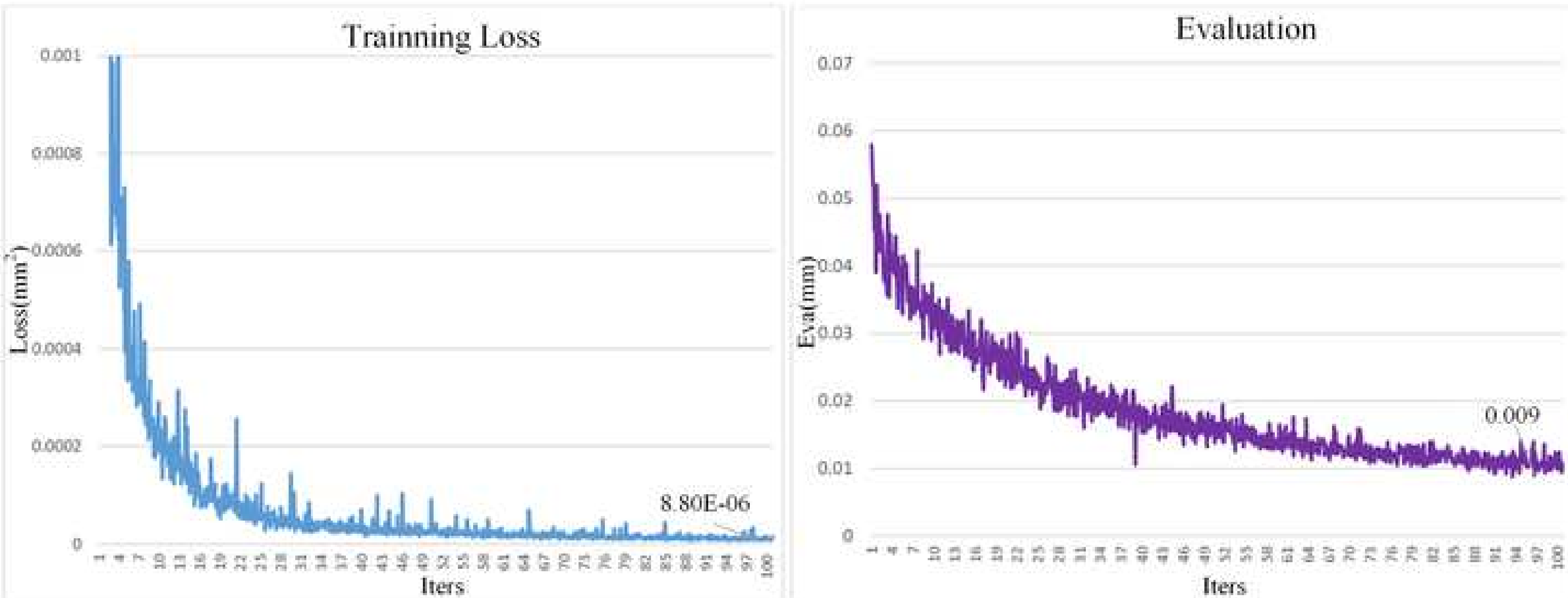


Fig. 6 Process data of Res-Net visualized by the output data of CNN middle layers

Click here to access/download;Figure;FIG. 6.tif

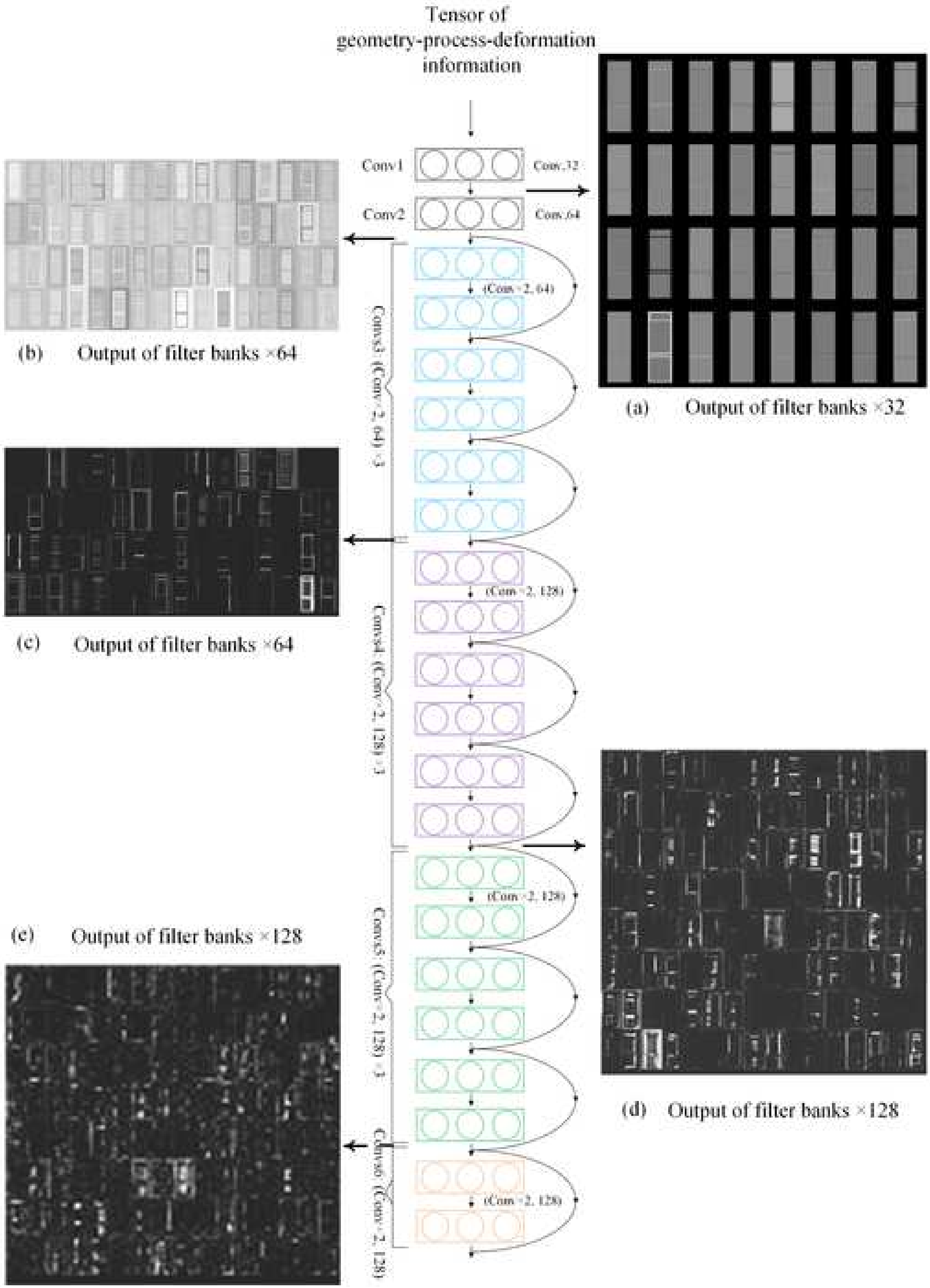


Fig. 7 The visualized process data of Convs-3

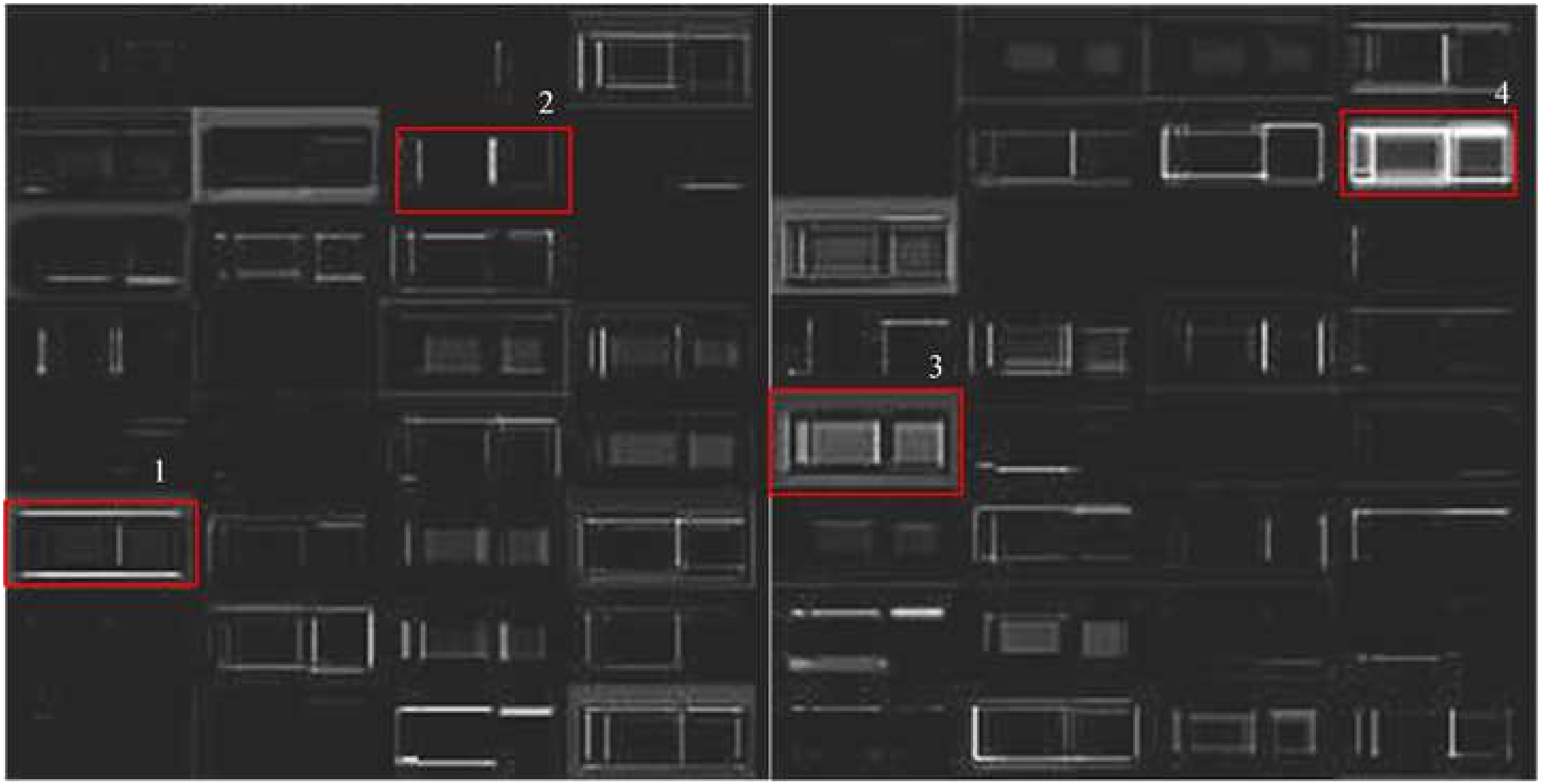
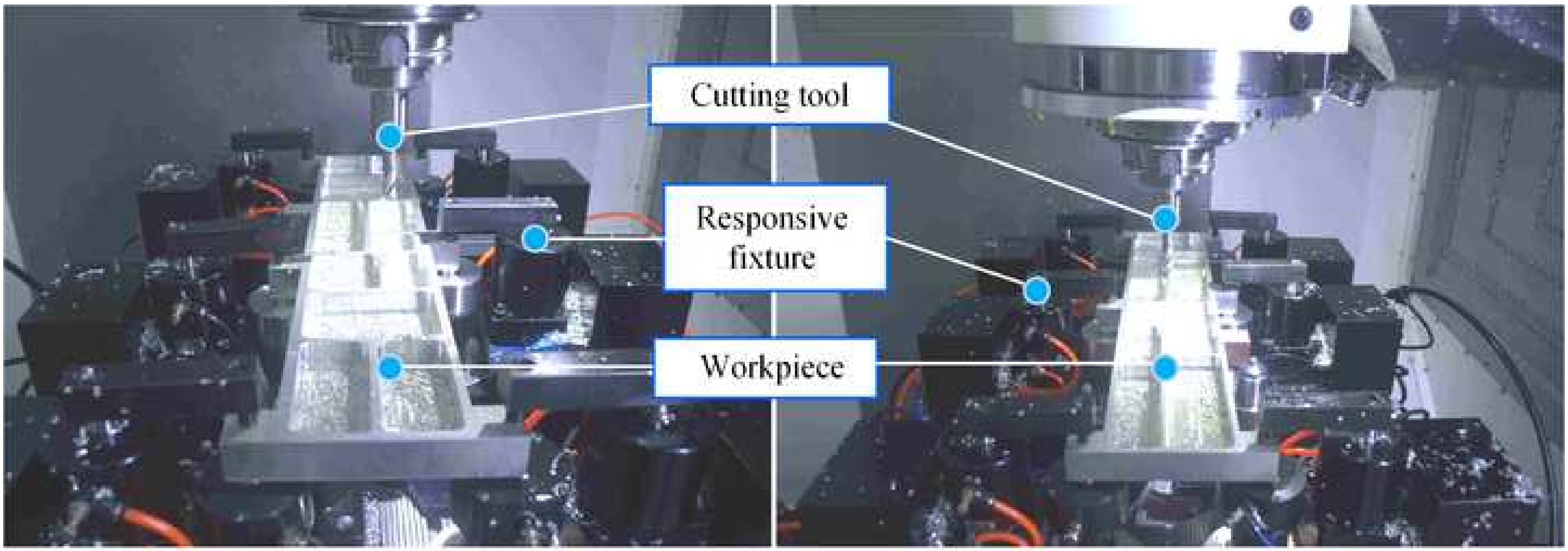


Fig. 8 Machining experiment environment



(a) Maching state of Part-1

(b) Machining state of Part-2

Fig. 9 Part information and monitoring points in plan view

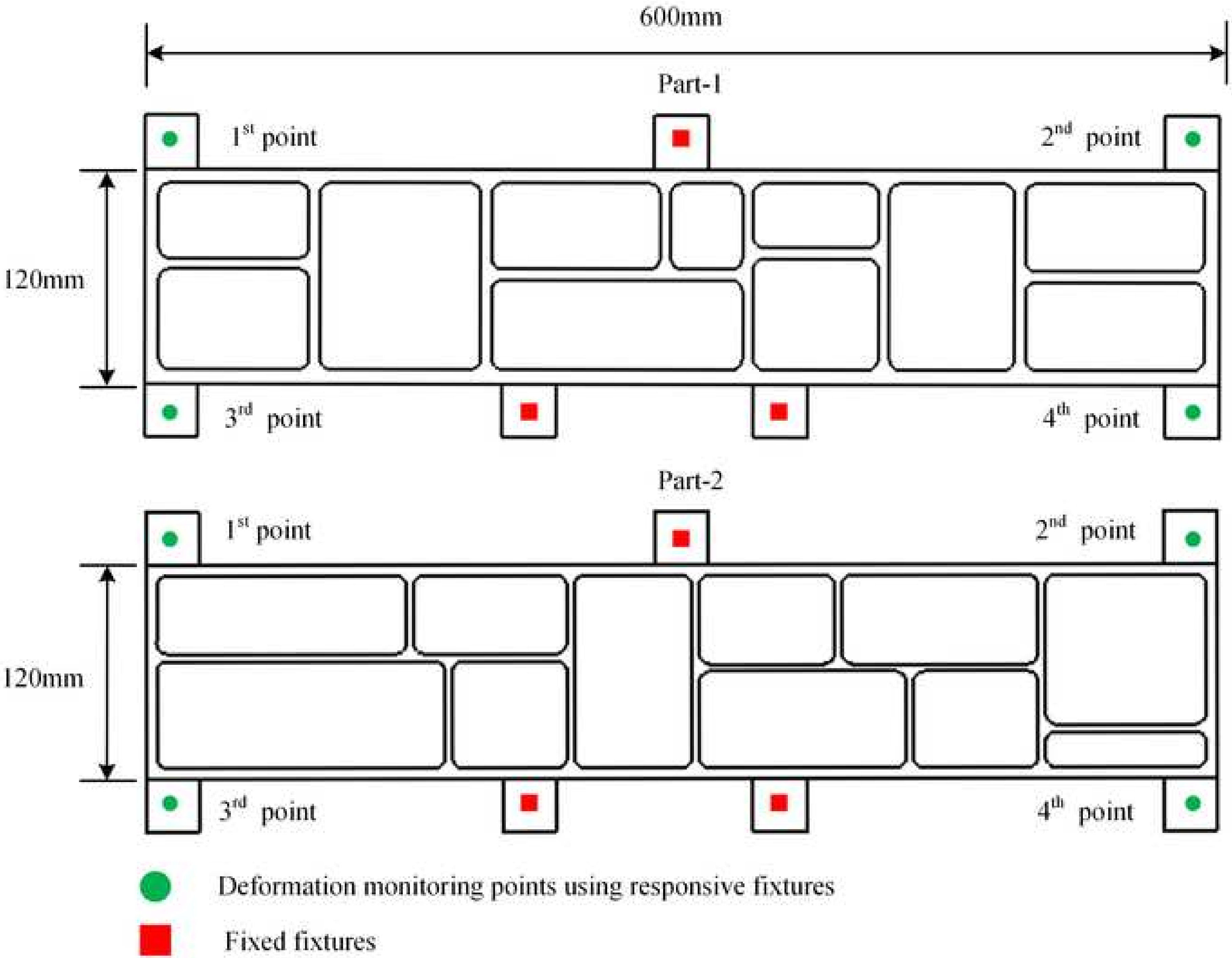


Fig. 10 Prediction results of the deformation for different machining states

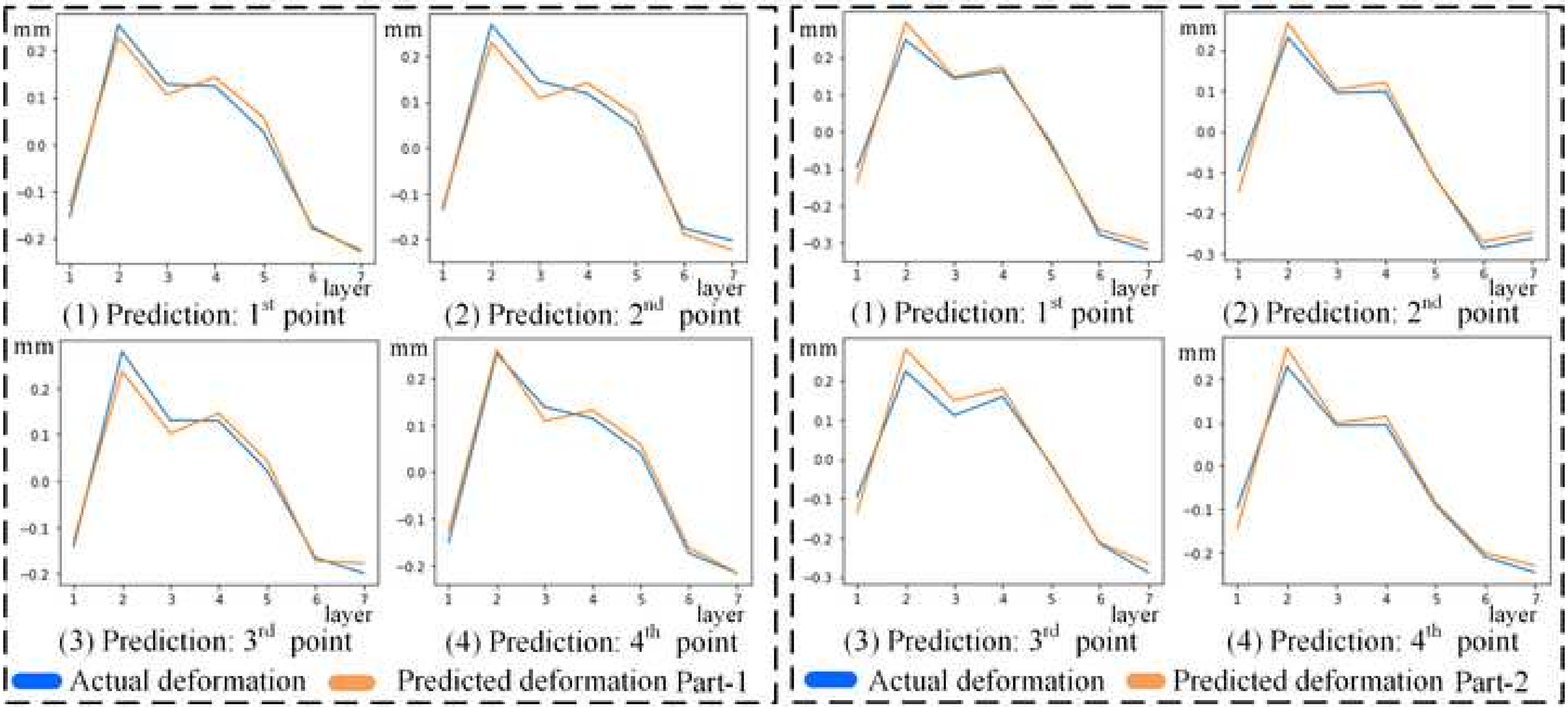


Table 1 Parameters of the CNN and RNN

Layer	Input	Output	Kernel	Stride	Padding	
Conv-1	4	32	(5,5)	(3,3)	(1,1)	
Conv-2	32	64	(3,3)	(2,2)	(1,1)	
Convs-3	64	64	(3,3)	(2,2)	(1,1)	×3
	64	64	(3,3)	(2,2)	(1,1)	
Convs-4	128	128	(3,3)	(2,2)	(1,1)	×3
	128	128	(3,3)	(2,2)	(1,1)	
Convs-5	128	128	(3,3)	(2,2)	(1,1)	×3
	128	128	(3,3)	(2,2)	(1,1)	
Convs-6	128	128	(3,3)	(2,2)	(1,1)	×1
	128	128	(3,3)	(2,2)	(1,1)	
BiLSTM		Input size: 128		Hidden units: 256		

Table 2 Parameters of auxiliary layers in the neural network

Layer	FC1	FC2	Input-FC	m-layer	r-layer
Input	128×4×7	512	256	128	512
Output	512	256	128	512	4

Table 3 Machining process parameters

Spindle Speed:	5000rpm
Cutting Mode	Down Milling
Cutting tool	D20
Cutting Depth (Roughing)	4mm
Cutting Depth (Finishing)	2mm
Cutting Width:	10mm
Feed Rate:	2000mm/min
Cutting Fluid:	Yes

Table 4 Prediction results for the parts machining experiments

Part	Data	Point-1	Point-2	Point-3	Point-4
Part-1	Monitoring data(mm)	-0.228	-0.202	-0.198	-0.216
	Prediction data(mm)	-0.223	-0.223	-0.177	-0.215
	Relative error	2.19%	10.40%	10.61%	0.46%
Part-2	Monitoring data(mm)	-0.318	-0.262	-0.289	-0.246
	Prediction data(mm)	-0.295	-0.247	-0.261	-0.228
	Relative error	7.23%	5.73%	9.69%	7.32%