Information Diffusion for Few-Shot Learning in Robotic Residual Errors Compensation*

Zeyuan Yang[1,2], Xiaohu Xu[3], Cheng Li[1], Sijie Yan[1(✉)], Shuzhi Sam Ge[2], and Han Ding[1]

[1] State Key Lab of Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan, China 430074
[2] Department of Electrical and Computer Engineering, National University of Singapore, Singapore 117576
[3] The Institute of Technological Sciences, Wuhan University, Wuhan, China 430072
sjyan@hust.edu.cn

Abstract. In this work, a novel model-free robotic residual errors compensation method is proposed based on the information-diffusion-based dataset enhancement (ID-DE) and the Gradient-Boosted Decision Trees (GBDT). Firstly, the dataset enhancement method is developed by utilizing the normal membership function based on the information diffusion technology. Then, merging it with multiple GBDTs, the multi-output residual errors learning model (ID-GBDTs) is constructed, and the grid search is used to determine the optimal hyper-parameters to accomplish the accurate prediction of residual errors. Finally, the compensation of robotic residual errors is realized by using the calibrated kinematic model. Experiments show that ID-DE can significantly improve the generalization ability of various learning models on the few-shot dataset. The R-squared of ID-GBDTs is improved from 0.58 to 0.77 along with the MAE decreased from 0.23 to 0.16, compared to original GBDT. Through the compensation of the residual errors, the mean/maximum absolute positioning error of the UR10 robot are optimized from 4.51/9.42 mm to 0.81/2.65 mm, with an accuracy improvement of 82.03%.

Keywords: Robot Calibration · Residual errors · Dataset enhancement.

1 Introduction

The articulated serial robots have recently been widely employed for accuracy applications such as robot-based machining, robot-based measurement, robot-assisted surgery, and so on [1,2]. However, inadequate absolute position accuracy (APC) caused by manufacturing tolerances, assembly errors, deformation errors, etc. continues to be a significant barrier to its development in high-precision

* Supported by the National Key Research and Development Program of China (No. 2019YFA0706703), the National Nature Science Foundation of China (Nos. 52075204, 52105514) and the China Scholarship Council (No. 202106160036)

operations. Therefore, calibration and compensation of the position errors are the critical steps to enhance the robot's APC before putting it into service.

Currently, much work has been conducted using geometric models and data-driven approaches. The former primarily adopt model-based approaches such as M-DH model [3], POE formula [4], CPC model [5], and other error models to calibrate robot geometric parameter errors caused by the non-ideal geometry of structural elements of the robot. However, non-geometric parameter errors due to gearing, part wear and temperature variations are difficult to establish using the model-based approach. The data-driven approach is a good solution to this challenge. Li et al. [6,7] measured a robot end position dataset containing 1,042 sample data based on a drawstring displacement sensor and implemented a data-driven calibration process using the neural network. Landgraf et al. [8] measured four different series of datasets with 16,811 representative sample data for various scenarios and developed a hybrid neural network for improving the APC. Chen et al. [9] proposed a global compensation method combining a genetic algorithm with and deep neural network (DNN) prediction model based on the more than 3,500 poses dataset being grid-wise distributed in the workspace. In order to obtain enough high-quality sample data, Zhao et al. [10] first calibrated the geometric errors based on the DH model and proposed an end-pose measurement algorithm to avoid the measurement process from being obscured, then achieved the calibration of residual errors by DNN. There are already a growing number of research employing model-free strategies to improve APC, including deep belief network [11], radial basis function neural network [12], and so on [13,14]. However, these approaches have stringent dataset requirements, relying on the small measurement errors and much sample data to assure learning more information about the real sample space. These requirements are challenging to achieve because of the complicated measurement process along with the vast number of robot error sources.

This work seeks to bridge this gap by developing a novel few-shot learning method to realize the prediction and compensation of the residual errors after the geometric errors being compensated. The rest of this paper is organized as follows: Section 2.1 introduces ID-DE that can enhance the generalization ability of the learning models based on the information diffusion. Section 2.2 presents the information-diffusion-based GBDT (ID-GBDTs) model with the few-shot dataset for residual errors prediction and compensation. Section 3 evaluates the ID-DE and ID-GBDTs, as well as the compensation results. The conclusions are finally presented in Section 4.

2 Residual Errors Estimation and Compensation Method

The measuring steps for the position of the robot end are generally done manually and required extensive human intervention, limiting the number of measurements and complicating data processing. Simultaneously, data errors caused by environmental interference and measurement errors make it difficult to ensure the quality of the acquired data. The collected data in this way is referred to

as an incomplete dataset since it does not always fully reflect the real sample space [15], making model-free robot calibration challenging. The following part presents an information-diffusion-based dataset enhancement (ID-DE) and the ID-GBDTs compensation model to address this problem.

2.1 Information-Diffusion-based Dataset Enhancement

Define the incomplete residual errors dataset $\boldsymbol{X}$ as

$$\boldsymbol{X}=\left\{\boldsymbol{x}_i=\left(\boldsymbol{\theta}^{(i)},\boldsymbol{e}^{(i)}\right)\middle|\boldsymbol{\theta}_i^{(i)}=\left[\theta_1^{(i)}\cdots\theta_n^{(i)}\right]^T,\boldsymbol{e}^{(i)}=\left[e_x^{(i)}\quad e_y^{(i)}\quad e_z^{(i)}\right]^T\right\} \tag{1}$$

where, $\boldsymbol{x}_i=\left(\boldsymbol{\theta}^{(i)},\boldsymbol{e}^{(i)}\right)$ denotes the i-th sample pair used for residual errors calibration, $\boldsymbol{\theta}^{(i)}=\left[\theta_1^{(i)}\ \cdots\ \theta_n^{(i)}\right]^T$ denotes the joint angle sequence, where $1\sim n$ are indexes of the joints, respectively. $\boldsymbol{e}^{(i)}=\left[e_x^{(i)}\quad e_y^{(i)}\quad e_z^{(i)}\right]^T$ denotes the corresponding position errors of the end-effector along the $\mathrm{x},\mathrm{y},\mathrm{z}$ direction. And $i=1,\ldots,m$, m denotes the sample number of $\boldsymbol{X}$.

Definition 1. *For a given incomplete dataset $\boldsymbol{X}$, $\hat{R}\left(\vartheta;\boldsymbol{X}\right)$ denotes the estimation of the real relation R obtained by the operator ϑ. if $\exists\ \ell$ generates the new dataset $\mathcal{D}\left(\boldsymbol{X}\right)=\left\{\boldsymbol{x}_{new}\left|\boldsymbol{x}_{new}=\ell\left(\boldsymbol{x},\mu_\Gamma\left(\boldsymbol{x}\right)\right),\boldsymbol{x}=\left(\boldsymbol{\theta},\boldsymbol{e}\right)\in\boldsymbol{X},\mu_\Gamma\left(\boldsymbol{x}\right)\in U\right.\right\}$ satisfies that*

$$\left\|\widetilde{R}\left(\vartheta;\mathcal{D}(\boldsymbol{X})\right)-R\right\|\leq\left\|\hat{R}\left(\vartheta;\boldsymbol{X}\right)-R\right\| \tag{2}$$

then ℓ is called a diffusion function about $\boldsymbol{X}$. μ_Γ is the membership function of the fuzzy set Γ of X, and U is the corresponding membership values set. The new dataset $\mathcal{D}(\boldsymbol{X})$ is called the diffusion set. $\widetilde{R}\left(\vartheta;\mathcal{D}\left(\boldsymbol{X}\right)\right)=\left\{\vartheta\left(\boldsymbol{x}_{new}\right)\left|\boldsymbol{x}_{new}\in\mathcal{D}\left(\boldsymbol{X}\right)\right.\right\}$ is called the diffusion estimation about R.

Definition 1 illustrates that, for any incomplete sample set, the original dataset can be processed by information diffusion to bring it closer to the true sample space, thus making the diffusion estimate closer to the true relationship, as shown in Fig. 1. Consequently, to improve the estimation accuracy, it is crucial to find the suitable diffusion function ℓ to achieve the sample enhancement. Considering the randomness of the robotic residual errors dataset $\boldsymbol{X}$, the normalized normal membership functions are chosen for each sample (the triangular distribution, trapezoidal distribution, etc. can also be selected according to the actual situation). Then, the membership values can be related to the new samples:

$$\begin{cases}\mu_\Gamma\left(\boldsymbol{\theta}^{(i)}\right)=\exp\left(-\dfrac{1}{2}\left(\boldsymbol{\theta}^{(i)}-\boldsymbol{\theta}_{new}^{(i)}\right)^T\Sigma_{\boldsymbol{\theta}}^{-1}\left(\boldsymbol{\theta}^{(i)}-\boldsymbol{\theta}_{new}^{(i)}\right)\right)\\ \mu_\Gamma\left(\boldsymbol{e}^{(i)}\right)=\exp\left(-\dfrac{1}{2}\left(\boldsymbol{e}^{(i)}-\boldsymbol{e}_{new}^{(i)}\right)^T\Sigma_{\boldsymbol{e}}^{-1}\left(\boldsymbol{e}^{(i)}-\boldsymbol{e}_{new}^{(i)}\right)\right)\end{cases} \tag{3}$$

where, $\Sigma_{\boldsymbol{\theta}}$ and $\Sigma_{\boldsymbol{e}}$ are the diffusion coefficient diagonal matrices about $\boldsymbol{\theta}$ and $\boldsymbol{p}$, respectively. As an approach to incorporate the distribution information of the

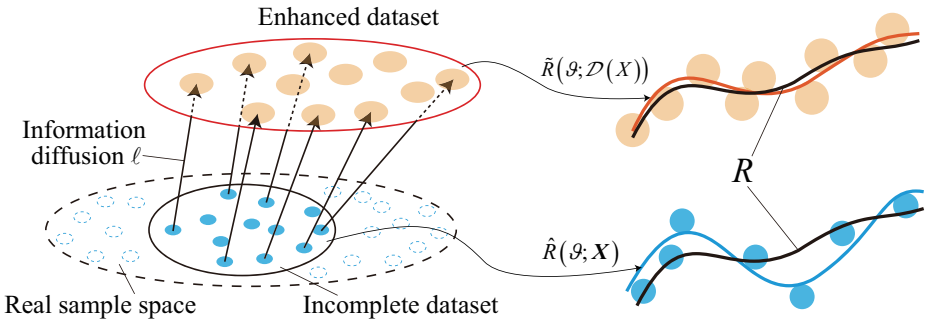

Fig. 1: Information diffusion principle for incomplete dataset.

original dataset $\boldsymbol{X}$ into the diffusion set $\mathcal{D}(\boldsymbol{X})$, the diffusion coefficients matrices associated with the variance of $\boldsymbol{X}$ are designed and can be expressed as

$$\begin{cases} \Sigma_{\boldsymbol{\theta}} = \operatorname{diag}\left[\dfrac{\sum_{i=1}^{m}\left(\theta_1^{(i)}-\bar{\theta}_1\right)^2}{m\,(m-1)},\ldots,\dfrac{\sum_{i=1}^{m}\left(\theta_n^{(i)}-\bar{\theta}_n\right)^2}{m\,(m-1)}\right] \\ \Sigma_{\boldsymbol{e}} = \operatorname{diag}\left[\dfrac{\sum_{i=1}^{m}\left(e_x^{(i)}-\bar{e}_y\right)^2}{m\,(m-1)},\dfrac{\sum_{i=1}^{m}\left(e_y^{(i)}-\bar{e}_y\right)^2}{m\,(m-1)},\dfrac{\sum_{i=1}^{m}\left(e_z^{(i)}-\bar{e}_z\right)^2}{m\,(m-1)}\right] \end{cases} \tag{4}$$

Bring Eq. (4) into Eq. (3), there has

$$\begin{cases} \sum\limits_{j=1}^{n}\dfrac{\left(\theta_j^{(i)}-\theta_{j,new}^{(i)}\right)^2\sum\limits_{i=1}^{m}\left(\theta_j^{(i)}-\bar{\theta}_j\right)^2}{m\,(m-1)} = -2\ln\mu_{\Gamma}\left(\boldsymbol{\theta}^{(i)}\right) \\ \sum\limits_{j=x}^{z}\dfrac{\left(e_j^{(i)}-e_{j,new}^{(i)}\right)^2\sum\limits_{i=1}^{m}\left(e_j^{(i)}-\bar{e}_j\right)^2}{m\,(m-1)} = -2\ln\mu_{\Gamma}\left(\boldsymbol{e}^{(i)}\right) \end{cases} \tag{5}$$

Equation (5) describes that the expanded new samples $\boldsymbol{x}_{new}$ can be derived from the normal diffusion when given the membership values. Since the joint angle vectors have n features (the corresponding residual errors have 3 features in the Cartesian space), the equation has infinite solutions. An effective approach is to minimize the overall error by constructing a loss function associated with the learning model. Considering here that the features of the joint angle samples are independent of each other, $\theta_j \sim \mathcal{N}\left(\theta_j\,|\theta_{j,new},h_j\right), e_j \sim \mathcal{N}\left(e_j\,|e_{j,new},\varsigma_j\right)$, $\theta_{j,new}^{(i)}$ and $e_{j,new}^{(i)}$ correspond to their membership values $\mu_{\Gamma}(\theta_j^{(i)})$ and $\mu_{\Gamma}(e_j^{(i)})$

one-to-one, respectively. There has

$$\begin{cases} \left(\theta_j^{(i)} - \theta_{j,new}^{(i)}\right)^2 \dfrac{\sum_{i=1}^{m}\left(\theta_j^{(i)} - \bar{\theta}_j\right)^2}{m\,(m-1)} = -2\ln\mu_\Gamma\left(\theta_j^{(i)}\right), j = 1,\ldots,n \\ \left(e_j^{(i)} - e_{j,new}^{(i)}\right)^2 \dfrac{\sum_{i=1}^{m}\left(e_j^{(i)} - \bar{e}_j\right)^2}{m\,(m-1)} = -2\ln\mu_\Gamma\left(e_j^{(i)}\right), j = x,y,z \end{cases} \tag{6}$$

Therefore, the new dataset can be obtained by rectifying Eq. (6) as

$$\begin{cases} \boldsymbol{\theta}_{new} = \boldsymbol{\theta} \pm \mathrm{sqrt}\left(-\boldsymbol{I}_{n\times 1} \otimes \mathrm{diag}^{\vee}\left(\Sigma_{\boldsymbol{\theta}}\right) \circ \ln\mu_\Gamma\left(\boldsymbol{\theta}\right)\right) \\ \boldsymbol{e}_{new} = \boldsymbol{e} \pm \mathrm{sqrt}\left(-\boldsymbol{I}_{n\times 1} \otimes \mathrm{diag}^{\vee}\left(\Sigma_{\boldsymbol{e}}\right) \circ \ln\mu_\Gamma\left(\boldsymbol{e}\right)\right) \end{cases} \tag{7}$$

where, $\mathrm{diag}^{\vee}(\bullet)$ denotes the vectorization of the diagonal matrix to a $1 \times n$ row vector. $\otimes$ denotes the Kronecker product and $\circ$ denotes the Hadamard product. $\boldsymbol{I}_{n\times 1}$ denotes the $n \times 1$ unit column vector.

Additionally, $\mathrm{Skew}_L = N_L/(N_L + N_U)$ and $\mathrm{Skew}_U = N_U/(N_L + N_U)$ are used as the left and right skewness magnitudes of the original dataset to characterize the asymmetric diffusion [16], where N_L and N_U denote the number of samples smaller and larger than the mean values of samples, respectively. Then, the diffusion function $\ell\left(\boldsymbol{\theta}, \boldsymbol{e}\right)$ for deriving the new dataset with the skewness information can be expressed as

$$\ell: \begin{cases} \boldsymbol{\theta}_{new,L} = \boldsymbol{\theta} - \mathrm{Skew}_{\theta,L} \cdot \mathrm{sqrt}\left(-\boldsymbol{I}_{n\times 1} \otimes \mathrm{diag}^{\vee}\left(\Sigma_{\boldsymbol{\theta}}\right) \circ \ln\mu_\Gamma\left(\boldsymbol{\theta}\right)\right) \\ \boldsymbol{\theta}_{new,U} = \boldsymbol{\theta} + \mathrm{Skew}_{\theta,U} \cdot \mathrm{sqrt}\left(-\boldsymbol{I}_{n\times 1} \otimes \mathrm{diag}^{\vee}\left(\Sigma_{\boldsymbol{\theta}}\right) \circ \ln\mu_\Gamma\left(\boldsymbol{\theta}\right)\right) \\ \boldsymbol{e}_{new,L} = \boldsymbol{e} - \mathrm{Skew}_{p,L} \cdot \mathrm{sqrt}\left(-\boldsymbol{I}_{n\times 1} \otimes \mathrm{diag}^{\vee}\left(\Sigma_{\boldsymbol{e}}\right) \circ \ln\mu_\Gamma\left(\boldsymbol{e}\right)\right) \\ \boldsymbol{e}_{new,U} = \boldsymbol{e} + \mathrm{Skew}_{p,U} \cdot \mathrm{sqrt}\left(-\boldsymbol{I}_{n\times 1} \otimes \mathrm{diag}^{\vee}\left(\Sigma_{\boldsymbol{e}}\right) \circ \ln\mu_\Gamma\left(\boldsymbol{e}\right)\right) \end{cases} \tag{8}$$

Obviously, the new sample pairs $\boldsymbol{x}_{new} = (\boldsymbol{\theta}_{new}, \boldsymbol{e}_{new})$ are derived according to the membership values in domain U with respect to $\boldsymbol{x} = (\boldsymbol{\theta}, \boldsymbol{e})$. Although the new dataset created by Eq. (8) can only be twice the size of the original dataset, the additional new samples can be generated by taking different membership values or several iterations. Then, the new dataset $\mathcal{D}\left(\boldsymbol{X}\right)$ can be expressed as

$$\begin{aligned} \mathcal{D}\left(X\right) = \big\{&\left(\boldsymbol{\theta}, \boldsymbol{e}, 1\right), \left(\left(\boldsymbol{\theta}_{new,U}, \boldsymbol{e}_{new,U}\right)_1, \left(\boldsymbol{\theta}_{new,L}, \boldsymbol{e}_{new,L}\right)_1, \mu_{\Gamma,1}\right), \\ &\ldots, \left(\left(\boldsymbol{\theta}_{new,U}, \boldsymbol{e}_{new,U}\right)_g, \left(\boldsymbol{\theta}_{new,L}, \boldsymbol{e}_{new,L}\right)_g, \mu_{\Gamma,g}\right)\big\} \end{aligned} \tag{9}$$

where, $\mu_\Gamma\left(\boldsymbol{\theta}\right) = \mu_\Gamma\left(\boldsymbol{e}\right)$ is taken to ensure the correspondence of the input-output relationship of the new samples, so it is abbreviated as μ_Γ. The numerical subscripts $(1, ..., g)$ are used to distinguish the use of different membership values.

2.2 Residual Errors Prediction and Compensation based on ID-GBDTs

Following the enhancement of the original dataset, the Gradient-Boosted Decision Trees (GBDT), a machine learning framework based on decision trees [17],

is utilized as an optional regression model to estimate residual errors. It's composed of numerous subtree models, where the latter subtree is built depending on the outcome of the former subtree, and the ultimate output is the total of all subtree's predictions. Using the weak decision trees as the base learners and merging them through iteration as a powerful ensemble learning model allows GBDT to combine the interpretability and fast computing of decision trees with the strong generalization performance of gradient boosting. The loss function of the GBDT can be expressed as

$$\vartheta_t(\boldsymbol{\theta}) = \vartheta_{t-1}(\boldsymbol{\theta}) + \arg\min Loss\left(\boldsymbol{e}_{err,i}, \vartheta_{t-1}(\boldsymbol{\theta}_i) + h(\boldsymbol{\theta}_i)\right) \tag{10}$$

where, ϑ_t denotes the accumulated models, $h(\boldsymbol{\theta}_i)$ denotes the new decision trees. Eq. (10) illustrates that in the function space, a weak learner $h(\boldsymbol{\theta}_i)$ should be chosen such that the loss function $Loss\left(\boldsymbol{e}_{err,i}, \vartheta_{t-1}(\boldsymbol{\theta}_i) + h(\boldsymbol{\theta}_i)\right)$ is minimized once this weak learner is added.

However, the GBDT can only be used for the single-output system. The proposed ID-GBDTs can cascade several GBDTs, the regressor of which is fitted for each axial residual errors, allowing knowledge of every axial residual errors to be gained by inspecting its corresponding regressor, thereby eliminating the problem that GBDT cannot accomplish multi-target output. Fig. 2 illustrates the principle of the ID-GBDTs.

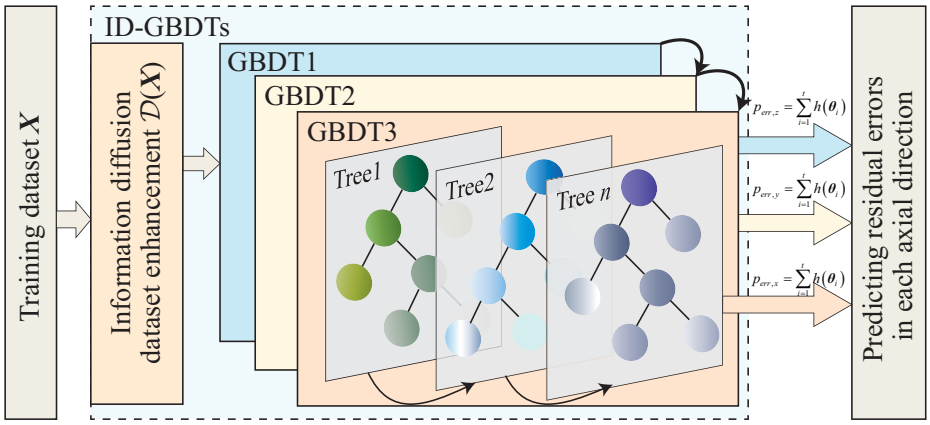

Fig. 2: The learning process of the ID-GBDTs.

After the prediction of the residual errors, the compensated end positions can be obtained by adding the theoretical end positions to their corresponding residual errors.

3 Experimental Setup and Verification

3.1 Experimental Setup

Figure 3 illustrates the experimental setup of the robot calibration process. The experimental setup consists of a desk-mounted Universal Robot UR10, a Leica AT960-LR laser tracker with the measurement accuracy of $\left(\pm 15\mu\mathrm{m} + 6\frac{\mu\mathrm{m}}{\mathrm{m}}\right)$, and a 6-DOF (degrees of freedom) T-MAC target attached to the flange.

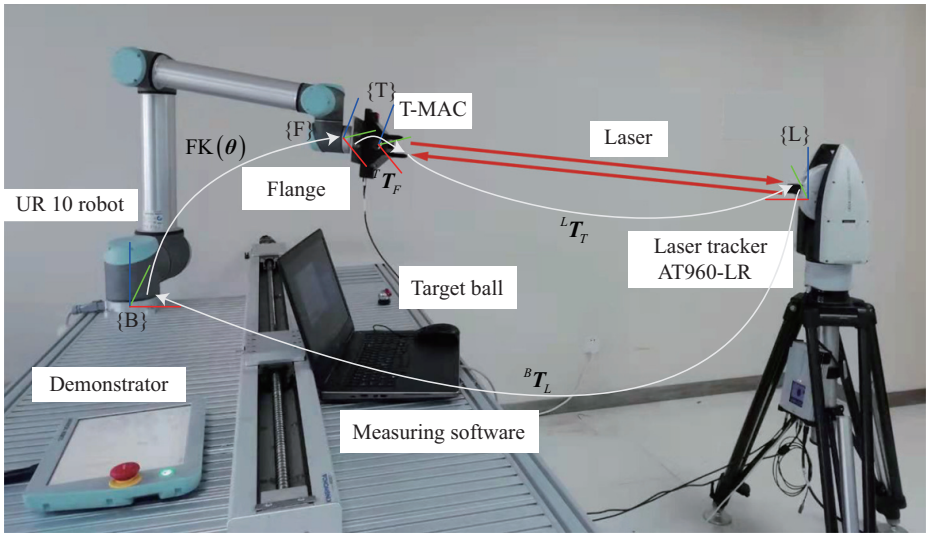

Fig. 3: UR10 robot calibration based on the AT960-LR laser tracker.

The pose transformation matrix ${}^{L}\boldsymbol{T}_{B}$ of laser tracker coordinate system $\{L\}$ with respect to the robot base coordinate system $\{B\}$ can be fitted by measuring the T-MAC poses when rotating the robot's 1-th and 3-th joints, respectively. Then, the robot is programmed offline to move in space (avoiding T-MAC being obscured), while the laser tracker measures in real-time to get T-MAC poses in $\{L\}$. The robot end position errors ${}^{B}\boldsymbol{e}$ are calculated by converting the measured T-MAC position ${}^{L}\boldsymbol{p}$ in $\{L\}$ into $\{B\}$ through ${}^{L}\boldsymbol{T}_{B}$ and comparing it to the T-MAC position acquired using the compensated forward kinematics $\mathrm{FK}(\boldsymbol{\theta})$. ${}^{B}\boldsymbol{e}$ can be expressed as

$$ {}^{B}\boldsymbol{e} = {}^{B}\boldsymbol{T}_{L} \cdot {}^{L}\boldsymbol{p} - {}^{T}\boldsymbol{T}_{F} \cdot \mathrm{FK}(\boldsymbol{\theta}) \tag{11} $$

Based on the above method, 60 sets of residual error data were randomly collected in the range of $800\times500\times800$ mm^3 of the robot's common workspace, of which 42 samples were selected as the original training sample set and 18 samples as the test set. The dataset has been made publicly available at https://github.com/Alvin8584/RobotCalibationDataforUR10.

3.2 ID-DE Evaluation and Residual Errors Compensation

Figure 4 depicts the influence of the proposed ID-based dataset enhancement method on various learning methods. In this case, membership value of the ID-DE is 0.99, and the number of created samples is 65. The selected models are Bayesian Regression (Bayes), Stochastic Gradient Descent (SGD), Support Vector Regression (SVR), RandomForest (RF), Gradient Boosted Decision Trees (GBDT), eXtreme Gradient Boosting (XGBoost) and Back Propagation Neural Network (BPNN), respectively. And the selected evaluation metrics are Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared (R^2), respectively. By using ID-DE, the MAE and MSE of the above learning methods on the test set were decreased by the mean/maximum of 13.62/27.43% and 25.9/52.82% respectively, while R^2 was significantly improved by mean/maximum of 26.67/53.15%. Among them, the scores of BPNN have the greatest noticeable improvement, whereas GBDT has the best overall performance. This indicates that the ID-DE has good compatibility with different learning models and has a significant enhancement effect, especially for few-shot learning.

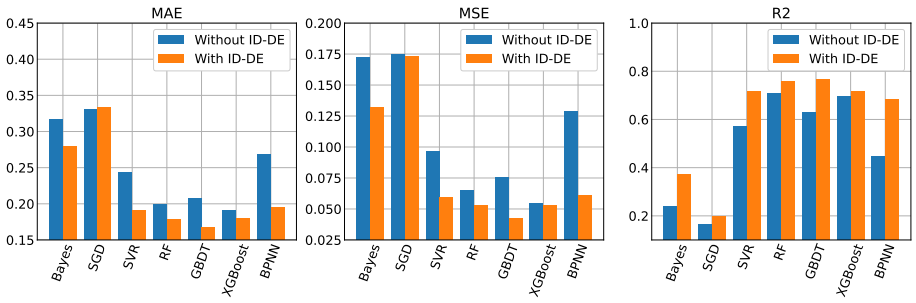

Fig. 4: Performance of learning models with and without ID-DE on test set.

Then, the generalizability ability of the GBDT with different ID-DE parameters is further discussed, as shown in Fig. 5. Generally, as the number of generated samples increases, the performance of the learning model first is improved significantly and then held constant or slightly reduced, while its decrease rate gradually decreases to zero. This is due to the fact that the samples generated by ID-DE can effectively enhance the original dataset and fully utilize the real sample space information; however, when too many samples are generated or the lower membership values are chosen, the information diffusion error increases, which would result in a stagnant or even counterproductive effect. According to a large number of experiments, the ID-DE can significantly improve the model generalization ability when the membership values generally take the value of 0.97-0.99 and the number of the generated new samples is within 0.4-2 times of the original number of samples. In addition, the ID-DE can reduce the MSE by 48.1% at most when its membership value is 0.98 and the number of the generated samples is 55.

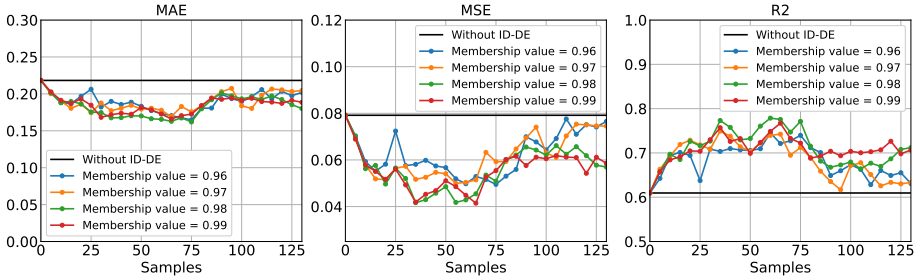

Fig. 5: Performance of ID-GBDTs with different ID-DE parameters on test set.

Based on the optimal parameters obtained above, the learning process and prediction results using the ID-GBDTs are shown in Fig. 6. To further improve the model prediction accuracy, grid search and cross-validation are used to find the hyper-parameters of ID-GBDTs [18]. The MSEs of the final ID-GBDTs along each axis on the test set are 0.06, 0.03 and 0.02, respectively. Comparing to GBDT with only 42 real samples in the training case, ID-GBDTs enhance the R-squared from 0.58 to 0.77, and reduce MAE from 0.23 to 0.16 on the test set with the generalization performance improved by 51.7%.

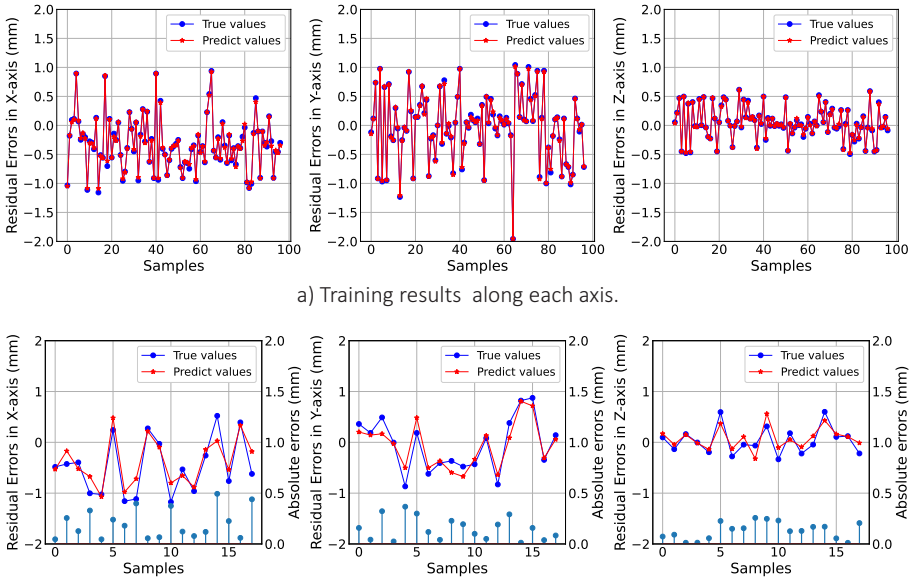

Fig. 6: Training results of ID-GBDTs along each axis.

Figure 7 illustrates the APC comparison of each axial direction before and after the compensation. The initial absolute positioning errors of UR10 are shown to be unevenly distributed, with negative bias in each axis and the largest errors along the y-axis, followed by errors along the x-axis, and finally along the z-axis. After the compensation, the absolute mean/maximum positioning errors of the

UR10 robot along the x, y, and z axes are decreased from 2.37/4.91, 3.39/8.93, and 1.19/2.5 mm to 0.23/2.18, 0.17/1.83, and 0.017/0.83 mm, respectively. The mean/maximum absolute positioning error is decreased from 4.51/9.42 mm to 0.81/2.65 mm, demonstrating an 82.03% improvement in APC.

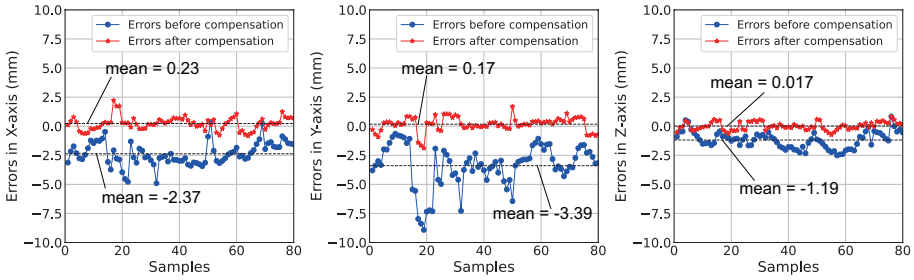

Fig. 7: APC comparison before and after the compensation of residual errors.

4 Conclusions

This paper presents a novel ID-GBDTs learning model for enhancing the few-shot residual errors dataset and improving robot's APC. The following conclusions are achieved:

(1) ID-DE can boost the few-shot dataset and hence improve the generalization capacity of the learning models with great compatibility, which is suitable for regression/classification with the few-show dataset.
(2) The ID-GBDTs enable accurate prediction and compensation of the residual errors, and the mean/maximum absolute positioning errors are reduced from 4.51/9.42 mm to 0.81/2.65 mm, with the APC improved by 82.03%.

References

1. Xie, H. L., Wang, Q. H., Ong, S. K., Li, J. R., Chi, Z. P.: Adaptive Human-Robot Collaboration for Robotic Grinding of Complex Workpieces. CIRP Annals. https://doi.org/10.1016/j.cirp.2022.04.015 (2022)
2. Yang, Z., Chu, Y., Xu, X., Huang, H., Zhu, D., Yan, S., Ding, H.: Prediction and Analysis of Material Removal Characteristics for Robotic Belt Grinding based on Single Spherical Abrasive Grain Model. International Journal of Mechanical Sciences, **190** 106005 (2021)
3. Peng, J., Ding, Y., Zhang, G., Ding, H.: An Enhanced Kinematic Model for Calibration of Robotic Machining Systems with Parallelogram Mechanisms. Robotics and Computer-Integrated Manufacturing **59** pp. 92–103 (2019)
4. Sun, T., Liu, C., Lian, B., Wang, P., Song, Y.: Calibration for Precision Kinematic Control of an Articulated Serial Robot. IEEE Transactions on Industrial Electronics **68**(7) 6000–6009 (2021)
5. Zhuang, H., Wang, L. K., Roth, Z. S.: Error-model-based Robot Calibration Using a Modified CPC Model. Robotics and Computer-Integrated Manufacturing **10**(4) pp. 289–299 (1993)

6. Li, Z., Li, S., Luo, X.: Data-driven Industrial Robot Arm Calibration: A Machine Learning Perspective. In 2021 IEEE International Conference on Networking, Sensing and Control (ICNSC) **1**, pp. 1–6 (2021)
7. Li, Z., Li, S., Bamasag, O. O., Alhothali, A., Luo, X.: Diversified Regularization Enhanced Training for Effective Manipulator Calibration. IEEE Transactions on Neural Networks and Learning Systems (2022)
8. Landgraf, C., Ernst, K., Schleth, G., Fabritius, M., Huber, M. F.: A Hybrid Neural Network Approach for Increasing the Absolute Accuracy of Industrial Robots. In 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE) pp. 468–474 (2021)
9. Chen, X., Zhang, Q., Sun, Y.: Evolutionary robot calibration and nonlinear compensation methodology based on GA-DNN and an extra compliance error model. Mathematical Problems in Engineering **2020** 3981081 (2020)
10. Zhao, G., Zhang, P., Ma, G., Xiao, W.: System Identification of the Nonlinear Residual Errors of an Industrial Robot Using Massive Measurements. Robotics and Computer-Integrated Manufacturing **59**, pp. 104–114 (2019)
11. Wang, W., Tian, W., Liao, W., Li, B.: Pose Accuracy Compensation of Mobile Industry Robot with Binocular Vision Measurement and Deep Belief Network. Optik **238** 166716 (2021)
12. Chen, D., Wang, T., Yuan, P., Sun, N., Tang, H.: A Positional Error Compensation Method for Industrial Robots Combining Error Similarity and Radial Basis Function Neural Network. Measurement Science and Technology **30**(12) 125010 (2019)
13. Gao, G., Liu, F., San, H., Wu, X., Wang, W.: Hybrid Optimal Kinematic Parameter Identification for an Industrial Robot based on BPNN-PSO. Complexity **2018** 4258676 (2018)
14. Gadringer, S., Gattringer, H., Müller, A., Naderer, R.: Robot Calibration Combining Kinematic Model and Neural Network for Enhanced Positioning and Orientation Accuracy. IFAC-PapersOnLine **53**(2) pp. 8432–8437 (2020)
15. Huang, C.: Principle of Information Diffusion. Fuzzy sets and Systems **91**(1), pp. 69-90 (1997)
16. Li, D. C., Wu, C. S., Tsai, T. I., Lina, Y. S.: Using Mega-Trend-Diffusion and Artificial Samples in Small Data Set Learning for Early Flexible Manufacturing System Scheduling Knowledge. Computers and Operations Research **34**(4), pp. 966–982 (2007)
17. Zhang, Z., Jung, C.: GBDT-MO: Gradient-Boosted Decision Trees for Multiple Outputs. IEEE Transactions on Neural Networks and Learning Systems **32**(7), pp. 3156–3167 (2020)
18. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., et al: Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research **12**, pp. 2825–2830 (2011)