



Research  
Laser Micro/Nano-Manufacturing—Article

# Real-Time Machine Learning-Based Position Recognition in Laser Nanofabrication with Sub-Half-Wavelength Precision



Hao Zhang<sup>a</sup>, Jinchuan Zheng<sup>a,\*</sup>, Guiyuan Cao<sup>b</sup>, Han Lin<sup>c,\*</sup>, Baohua Jia<sup>c,\*</sup>

<sup>a</sup> Swinburne University of Technology, John Street, Hawthorn, VIC 3122, Australia

<sup>b</sup> Nanophotonics Research Center, Shenzhen Key Laboratory of Micro-Scale Optical Information Technology, Shenzhen University, Shenzhen 518060, China

<sup>c</sup> Centre for Atomaterials and Nanomanufacturing (CAN), RMIT University, Melbourne, VIC 3000, Australia

## ARTICLE INFO

### Article history:

Received 23 May 2024

Revised 6 January 2025

Accepted 14 March 2025

Available online 16 June 2025

### Keywords:

Laser nanofabrication

Machine learning

Image-based position recognition

Neural network

Image processing

## ABSTRACT

Laser nanofabrication with tightly focused ultrafast laser pulses enable versatile fabrication of arbitrary two-dimensional (2D)/three-dimensional (3D) micro/nanostructures. Accurate positioning of the laser focal spot is crucial, especially for high-resolution integrated devices in 2D materials, requiring precise placement on atomically thin surfaces. However, uneven surfaces and surface tilt poses significant challenges. Existing methods to detect focal positions often involve complex setups with additional optical components or sensors, achieving limited accuracy. This study introduces a machine learning-based method to accurately detect the focal position during laser nanofabrication by analyzing the shape and intensity of the laser focal spot. We compare four machine learning methods: rational quadratic Gaussian process regression, kernel approximation least square, quadratic support vector machine, and trilayered neural network. Our experiments show that trilayered neural network (TNN) method achieves a detection accuracy of 257 nm (half the fabrication laser wavelength), surpassing the required accuracy for tightly focused laser beam (typically larger than one wavelength). This method can map focal positions along the fabrication trajectory to compensate for surface roughness or tilt. Moreover, it can be directly implemented in any laser nanofabrication system with a camera for *in situ* monitoring, without requiring additional optical components, indicating broad applicability.

© 2025 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The advent of nanomachining has significantly integrated high-performance optical components into daily life applications [1], such as camera lenses and eyeglasses, and critical domains such as aeronautics [2], electronics [3], and chemical and biological fields [4]. Recent advance [5] in laser nanofabrication technology have revolutionized the manufacturing process for optical devices requiring surface pattern generation. Operating at a submicrometer scale, laser nanofabrication enables the production of miniaturized three-dimensional (3D) photonic devices [6]. Known for its precision and programmability, 3D laser nanoprinting offers meticulous control and positional selectivity [7]. This technique excels in producing intricate patterns and complex geometries with unparalleled precision, environmental friendliness [8], and flexibility [9].

The adoption of laser processing methodologies has become a pivotal scientific approach, enhancing functional attributes and uncovering novel properties in complex materials [10]. The versatility of femtosecond laser technology spans micro/nanofabrication methodologies [11,12], surface treatment modalities, structural modifications [13], and controlled surface deformation [14]. Laser nanofabrication precisely focuses controllable laser fluence into a nanometric high-intensity focal spot to modify material properties [15]. Accurately positioning this high-resolution focal spot is crucial for the optimal laser beam and sample interaction, and is essential for the quality, resolution, performance, and reliability of fabricated patterns and devices. This is particularly critical for photonic devices, whose performance is highly sensitive to minor fabrication defects, requiring submicrometer precision.

Previous studies have developed methodologies for detecting focal positions and calibrating focusing systems during laser writing processes. Implementations include focus-adjusting mechanisms such as deformable mirrors [16], remote sensing [17], laser focal point scanning [18], and laser micromachining [19]. However,

\* Corresponding authors.

E-mail addresses: [jzheng@swin.edu.au](mailto:jzheng@swin.edu.au) (J. Zheng), [han.lin2@rmit.edu.au](mailto:han.lin2@rmit.edu.au) (H. Lin), [baohua.jia@rmit.edu.au](mailto:baohua.jia@rmit.edu.au) (B. Jia).

these systems are often excessively complex or costly, making them impractical for large-scale industrial production. Advanced astigmatic principles such as dynamic scan detection [20], extracted focusing error signal [21], and confocal point sensors [22] have been explored; however, their high cost and inconsistent results hinder their applicability in manufacturing environments requiring large sample volumes and complex structural geometries. Adjusting optical setups to prevent defocus during the laser nanofabrication process is also considered unfeasible [23]. Recently, Xu et al. [24] introduced a machine vision technique, using a charged-coupled device (CCD) camera to capture diffraction from light interacting with the workpiece surface to discern focusing states. Despite its precision, this method faced drawbacks such as low speed, non-real-time focus detection, and reliance on a high numerical aperture (NA) objective. Recent advances in laser manufacturing have increasingly incorporated neural networks to enhance precision and efficiency. Deep neural networks (DNNs) now detect fine-grained flaws in laser-fabricated materials by learning layer-wise imaging profiles, improving defect identification [25]. Additionally, artificial neural networks (ANNs) optimize laser processing parameters [26] and predict outcomes in laser micro-grooving [27] and engraving, providing accurate estimations of qualitative characteristics during fabrication [28]. Moreover, recent advancements have incorporated machine learning techniques in processing high-resolution images from scanning electron microscopy (SEM), improving the fabrication fidelity of integrated nanophotonic devices [29,30]. Polat et al. [31] applied an advanced convolutional neural network (CNN) technique to discern surface focus, surpassing the conventional Gaussian fitting approach. However, their detection accuracy ( $\pm 5 \mu\text{m}$ ) does not meet the stringent requirements of nanofabrication processes. While these methods detect focal positions, some rely on complex setups with additional optical components or sensors without real-time processing [18,21,24]. Other works are in real-time but with limited accuracy [17,31]. Therefore, there is an urgent need for a cost-effective focus distance detection system that enhances production efficiency and quality in laser nanofabrication, combining high accuracy and fast speed.

This study addresses the high sensitivity of the focal spot shape to the focal position during laser nanofabrication by proposing a machine learning-based method to accurately identify the focal position through analysis of the laser focal spot shape. Four machine learning techniques are examined and compared: rational quadratic Gaussian process regression (RQGPR), kernel approximation least square (KALS), quadratic support vector machine (QSVM), and trilayered neural network (TNN). The final results demonstrate detection accuracy of a tightly focused laser beam that exceeds fabrication accuracy. This method allows for the mapping of focal positions along the fabrication trajectory, enabling accurate compensation for surface roughness or tilt during laser nanofabrication. Importantly, this approach can be integrated into any laser nanofabrication system with a monitoring camera, without requiring additional optical components. Thus, the proposed methods can be broadly applied in the laser nanofabrication process.

## 2. Experimental setup and data acquisition

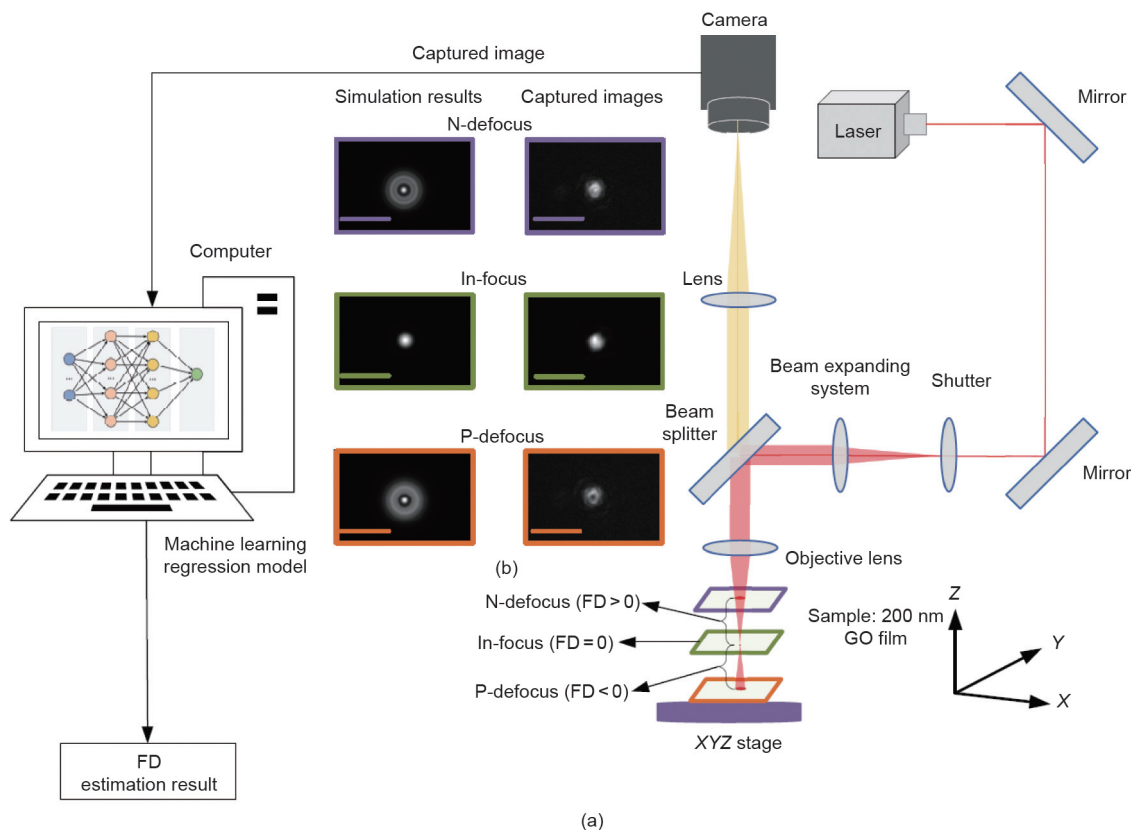
Fig. 1 illustrates the schematic of a laser nanofabrication system, while Fig. 1(b) presents simulated and experimentally captured images that exhibit distinguishable shape changes at various focal positions (detailed simulation and experimental images are provided in Figs. S1 and S2 in Appendix A). These images offer real-time feedback on the position of the focal plane relative to the sample. Although laser-written features serve as post-process verification, the primary method for determining

the in-focus position depends on the shape and intensity of the focal spot observed in the captured images, ensuring proper sample alignment before fabrication. Fig. 1 also illustrates the focal distance (FD), defined as the difference between the current focus position and the ideal in-focus position. The in-focus region is identified based on the captured images of the laser focal spot reflected from the workpiece surface. Conversely, the P-defocus region indicates that the focal spot is above the sample surface, while the N-defocus region signifies that the focus point is below the sample surface. In either defocus state, the laser energy cannot be fully transferred to the sample surface, thereby compromising writing quality or causing failure. The simulated results based on vectorial Debye diffraction theory are depicted in Fig. S3 in Appendix A.

Fig. 1 illustrates the experimental setup of the laser nanofabrication system under investigation, which includes a laser source, beam splitter, beam expansion apparatus, shutter mechanism, mirrors, objective lens, and XYZ positioning stage. The laser beam is expanded via a beam expansion system to produce a substantially planar wavefront. This expanded beam is then directed through an objective lens with an NA of 0.8 and projected onto a complementary metal-oxide-semiconductor (CMOS) camera (XIMEA MC050MG-SY, Ximea, Germany). This study employs a commercial laser nanoprinting system (NanoPrint3D, Innofocus, Australia). The femtosecond fabrication laser operates at a wavelength of 800 nm, a pulse repetition frequency (PRF) of 1000 Hz, and a pulse width (PW) of 400 femtoseconds. Key fabrication parameters, such as output power and scanning speed, were set at  $0.2 \mu\text{W}$  and  $30 \mu\text{m}\cdot\text{s}^{-1}$ , respectively. The fabrication lens used in this study was the Olympus plan N 50 $\times$  oil immersion objective (Olympus, Japan). Fabrication procedures were recorded as videos using a CMOS camera, which features both color and monochromatic sensors and operates at a frame rate of 23 frames per second (FPS) with a resolution of 5 megapixel (MP) (i.e.,  $2464 \times 2056$  pixels).

In laser nanofabrication systems, a CMOS camera is extensively used for real-time monitoring of the writing process. The camera captures live images as the laser beam scans across the material, facilitating the fabrication of 3D structures. Its primary functions include tracking the position of the laser beam and assessing the writing quality. During the writing procedure, the camera continuously captures images of the material under the laser beam exposure. Fig. S2 illustrates the measured focal intensity distribution at different positions. The figure shows that the adjacent images are highly similar, making them difficult to distinguish. The control software analyzes these captured images to evaluate the planar position of the laser beam, enabling fine-tuning of the movement of the positioning stage to construct the desired structure. This iterative feedback loop ensures accuracy and precision in the fabricated structure.

In this research, graphene oxide (GO) is utilized as the monoelement 2D material for fabrication. GO, a graphene-based material with oxygen-containing functional group [32], has a thickness of 200 nm for various integrated photonic applications. The thickness of the GO film can influence the accuracy of position recognition during laser nanofabrication due to light-matter interaction. Since the thickness of the GO film is smaller than the full width at half maximum (FWHM) of the focal spot along the Z-axis, reflections from the interfaces between the cover glass and GO, and GO and air, are expected. The resulting intensity is the sum of the intensities from these two interfaces, affecting the detected intensity distributions. The high absorption of GO at the detection wavelength means thicker film results in higher absorption and lower light reflection. A thickness of 200 nm is chosen as it is optimal for fabricating GO lenses [33], allowing accurate position detection. The structural fabrication profile is an Archimedean spiral with 499 fabrication points, creating a pattern on the GO film. To generate



**Fig. 1.** Schematic of a laser nanofabrication system with machine learning-enabled real-time position recognition during fabrication. (a) Schematic of the laser nanofabrication system. (b) Simulated and experimentally captured images. N-defocus indicates the focal point below the sample surface, In-focus indicates the sample at the optimal focal position for clear laser pattern delineation on the fabrication material, and P-defocus indicates the focal spot above the sample surface. GO denotes graphene oxide.

comprehensive data covering both in-focus and defocus situations, the sample is mounted on a surface with a 7° inclination.

### 3. Methodology

This section discusses the methods of image processing, data labeling and data splitting, which may affect the performance metrics of the FD detection model.

Image resizing is crucial for discarding unnecessary information from raw images or maintaining consistent image sizes from cameras with varying resolutions. It significantly reduces processing time. However, no universal resizing method guarantees optimal performance; trials are often necessary. Occorsio et al. [34] demonstrated that the bicubic interpolation (BIC) method offers a satisfactory trade-off between image quality and processing time compared to non-adaptive methods such as bilinear and nearest neighbor approaches. The BIC method calculates a weighted average of the four nearest pixels to determine one unknown pixel value, as depicted in Fig. S4 in Appendix A. Hence, our study employs the BIC resizing method and conducts a comparative study to identify the optimal resizing factor.

Data labeling involves assigning relevant and descriptive labels to data sets, a crucial annotation process that provides contextual information for machine learning models to learn effectively. In this research, the labels of interest are FDs measured in micrometers. Approximately half of the labels can be obtained using position sensors through careful experimental design during data collection. However, the number of collected data sets exceeds that of the unknown labels. To address this, a data interpolation technique is used to construct the unknown labels based on the two adjacent known labels.

Data interpolation is a fundamental technique in both machine learning and computer vision, primarily involving the use of existing data to reconstruct missing data. This facilitates the creation of a comprehensive data set for subsequent analysis or training of machine learning models. This proactive step helps mitigate biases from incomplete data sets and potentially enhances model accuracy. Among interpolation methods, linear interpolation estimates values for points lying between two known data points on a straight line. This method calculates an estimated value at an intermediate point by computing a weighted average of the two known data points and is simple to implement. Consequently, this study employs the linear interpolation method to compute the unknown labels.

Data splitting involves partitioning a data set into distinct subsets for training and testing a machine learning model. There is no universally optimal split ratio for training and testing data; the appropriate split ratio depends on the specific nature of the problems and data sets. Therefore, conducting experiments with various split ratios, along with different configurations such as image resizing and region of interest (ROI) size is crucial, to determine the most suitable combination of image pre-processing.

### 4. Results and discussions

This section presents the results of data processing and comparisons of four regression models. The subsequent sections follow the procedural flow depicted in Fig. S5 in Appendix A. The initial image processing step involves image resizing and selecting the ROI to extract image pixel intensity values as input features for a regression model to detect the FD. In addition, the split ratio between training and testing data are studied to optimize the regression

models for the best detection performance. This research compares four selected machine learning regression algorithms to evaluate the best-performing model in Fig. S6 in Appendix A, and finally, the trained models are applied for real-time FD detection.

The data set consists of 1516 images acquired during a fabrication process, with the sample placed on a surface at a 7° inclination to generate varying, pre-determined FDs for data labeling. This setup indicates variations in FD across both in-focus and defocus region.

Fig. S7 in Appendix A examines various resizing factors and ROI-area percentages about detection accuracy and time to determine the optimal resizing factor and ROI size for regression models. Specifically, Fig. S7(a) compares detection accuracy among different resizing factors and ROI-area percentages. The  $x$ -axis represents the ROI-area percentage, defined as the selected ROI area divided by the entire image area, while the  $y$ -axis denotes the root mean squared error (RMSE) values of detection. Smaller ROIs result in underfitting and low accuracy, as the model may fail to capture important features in the images. For instance, with a 1% ROI-area percentage, the RMSE value for detection accuracy is as high as 1.8  $\mu\text{m}$ . Conversely, excessively large ROI-area percentage can lead to overfitting, and deteriorating detection accuracy because the model fit only the training data rather than generalizing to new data. Additionally, a larger ROI area increases detection time due to expanded observation features, as shown in Fig. S7(b).

Based on the above comparisons and analysis, a resizing factor of 0.2 provides the best balance between detection accuracy and processing time. Consequently, the input images were resized from their original dimension of  $1024 \times 1024$  to  $205 \times 205$  pixels. In addition, an ROI-area percentage of 6.2% resulted in the lowest RMSE value for detection accuracy. Thus, an ROI size of  $51 \times 51$  pixels was selected and is shown in Fig. S8 in Appendix A. For comparison, Fig. S8 also illustrates various ROI sizes to visualize their coverage of the laser spot feature relative to the entire image area.

As previously mentioned, 1516 data sets were collected, but only 499 have corresponding from the known fabrication profile. The remaining data sets were labeled with their corresponding FDs using linear interpolation, as shown in Fig. S9 in Appendix A. The subsequent subsection will present the cases of data split ratios studied in this research and the detection results to determine the optimal split ratio between training and testing.

To evaluate model performance, four metrics are employed: mean absolute error (MAE), RMSE, coefficient of determination ( $R^2$ ), and mean prediction speed (MPS). Specifically, MPS refers to the average time taken to process real-time observations by a trained model. MPS will be used to assess the efficiency and speed of FD detection, which is crucial for real-time defocus calibration in the control system.

$$\text{MAE} = \frac{\sum_{i=1}^N |y_i - x_i|}{N} \quad (1)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (x_i - y_i)^2}{N}} \quad (2)$$

$$R^2 = \frac{\sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^N (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^N (y_i - \bar{y})^2}} \quad (3)$$

$$\text{MPS} = \frac{N}{\sum_{i=1}^N t_i} \quad (4)$$

where  $N$  is the number of observations,  $x_i$  is observation values,  $y_i$  is predicted values, and  $t_i$  is the prediction time.  $i \in \mathbf{N}^+$ ,  $\bar{x}$  denotes the mean of the observation values, and,  $\bar{y}$  denotes the mean of the

predicted values. RMSE is measured in micrometers ( $\mu\text{m}$ ), and MPS is expressed in observation per second ( $\text{obs}\cdot\text{s}^{-1}$ ).

The processed data sets were used to train the four regression models: RQGPR, KALS, QSVM, and TNN (detailed explanation in Note S1 and Table S1 in Appendix A). To mitigate overfitting or underfitting and enhance model robustness, the  $K$ -fold cross-validation technique was applied, where  $K$  denotes the number of data partitions. With  $K = 5$ , this technique randomly partitions the data sets into five equally sized folds, using one subset as the test set and the remaining four as training sets. The training and testing processes for each regression model were then conducted. The experimental results are compared and summarized in Fig. 2.

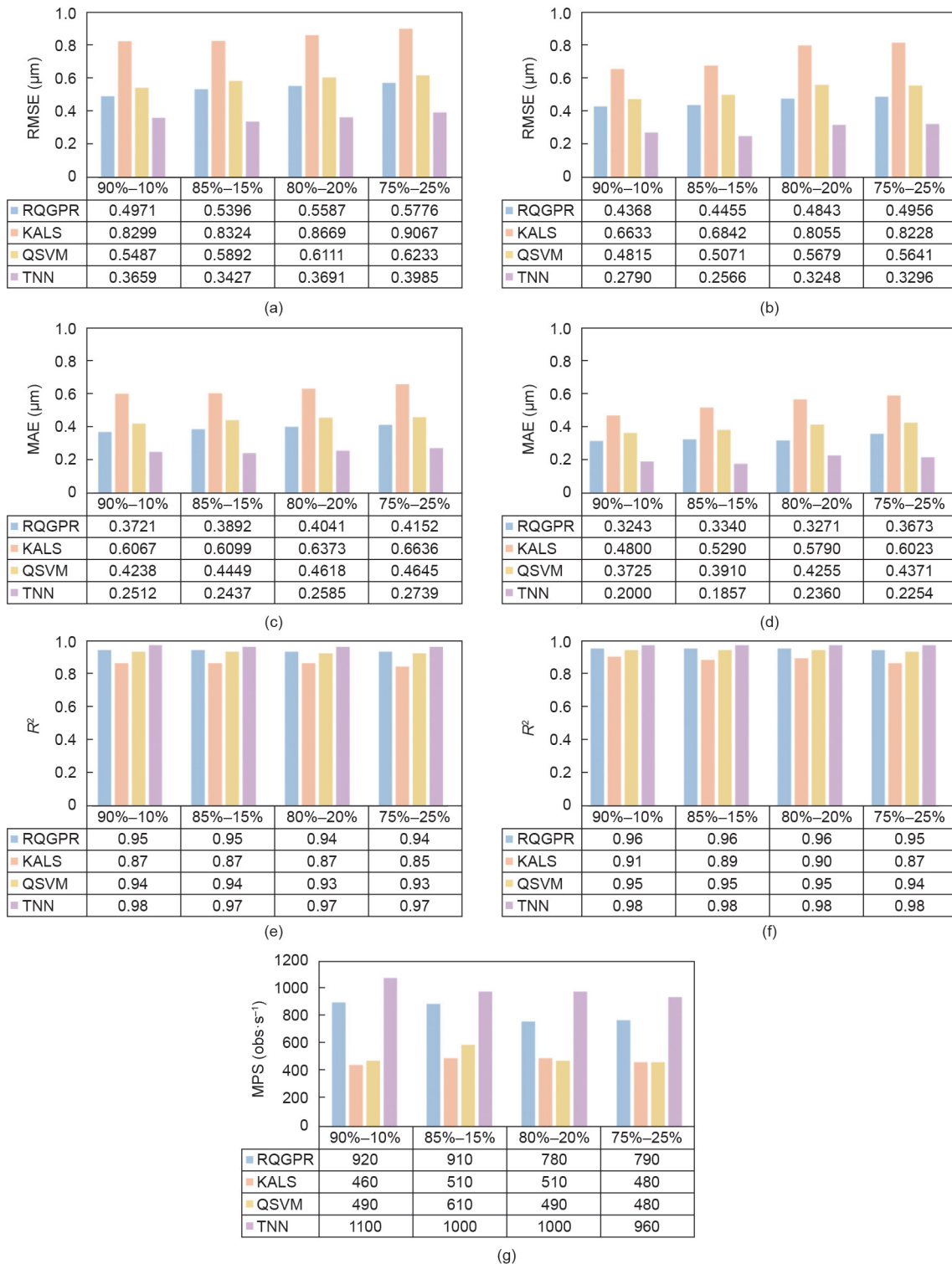
First, the data split ratios between training and testing, listed in Table S2 in Appendix A, are studied. For training accuracy, Figs. 2(a), (c), and (e) shows that RMSE and MAE values increase as the ratio of training data decreases. Similarly,  $R^2$  values approach unity as the volume of training data increases. Conversely, the testing results in Figs. 2(b), (d), and (f) indicate that RMSE and MAE metrics reach their minimum with approximately 10%–15% testing data. Increasing the training data ratio results in a decrease in  $R^2$  values. Notably, the 15% testing data ratio under the TNN model yields the best RMSE, MAE, and  $R^2$  results. For detection speed, the 85%–15% data split ratio exhibits the fastest processing speed in the KALS and QSVM models. Consequently, the 85%–15% data split ratio is selected, corresponding to 1289 images for training and 227 images for testing.

Next, the results in Fig. 2 summarize the training and testing accuracy achieved by each regression model. For training results, the KALS model has the largest RMSE (0.8324) and MAE (0.6099), while the TNN model shows superior performance with the smallest RMSE (0.3427) and MAE (0.2437). The TNN model also exhibits an  $R^2$  value of 0.97 (close to the ideal value of 1), indicating robust detection accuracy. Regarding detection speed, which is crucial for real-time systems, the KALS method is time- and memory-intensive, resulting in relatively slow detection. In contrast, the TNN model achieves the highest detection speed with an MPS of  $1000 \text{ obs}\cdot\text{s}^{-1}$ . According to the testing results, the KALS model presents the largest RMSE (0.6842) and MAE (0.5290). Conversely, the TNN model again outperforms others, achieving the smallest RMSE (0.2566) and MAE (0.1857), and the best  $R^2$  value of 0.98. Notably, the detection accuracy of our method surpasses the Rayleigh length (1.269  $\mu\text{m}$ ) of the focused beam (detailed calculation in Note S2 in Appendix A). These results are based on 15% of images from the data set as testing data. In conclusion, based on the evaluation of performance metrics, the TNN regression model performs the best in this study.

The trained TNN model comprises fully connected layers with 10 nodes each, represented by a  $1 \times 3$  row vector. ReLU functions [35] are applied as activation functions for these layers. Predictor means and standard deviations are indicated as  $1 \times 1024$  row vectors. The learned layer weights are encapsulated in a  $1 \times 4$  cell array. The first input layer contains a  $10 \times 2601$  matrix, while the second and third layers feature  $10 \times 10$  matrices. The final output layer comprises a  $1 \times 10$  vector for calculating the FD detection. Additionally, the TNN model includes layer biases, with hidden layer biases represented by  $10 \times 1$  vectors and the final layer bias as a scalar.

To compare and verify the robustness of the four trained regression models under real conditions, new data sets comprising 1517 images were acquired using the same working configuration. These images were collected from data sets where the sample was placed on a surface at a 5° inclination.

A statistical analysis of the data set using a box plot is depicted in Fig. 3, where the FDs on the  $x$ -axis are segmented into 13 groups with 0.5  $\mu\text{m}$  intervals within the in-focus region of  $\pm 3 \mu\text{m}$ . The  $y$ -axis indicates the variation of detection errors concerning the FD.



**Fig. 2.** Comparisons of performance metrics versus data split ratio between training and testing: (a, c, e) training accuracy; (b, d, f) testing accuracy; (g) detection speed. The ratio 90%–10%, 85%–15%, 80%–20%, and 75%–25% represents the percentage splits between training and testing data, such as 90%–10% indicating 90% training and 10% testing data.

The plot reveals that the accuracy of the KALS model decreases as the FD increases. When the FD exceeds  $\pm 2 \mu\text{m}$ , the whiskers extend, indicating a wider spread of the middle 50% of detection errors. Additionally, both the RQGPR and QSVM models exhibit a wider span of detection errors when the FD is between  $+0.5 \mu\text{m}$  to  $+2.5 \mu\text{m}$ . Conversely, the TNN model achieves the smallest span

of detection errors within  $\pm 1.5 \mu\text{m}$  when the FDs vary within  $\pm 3.5 \mu\text{m}$ . Furthermore, the detection errors in the TNN model display more condensed boxes, suggesting more uniform and reliable performance compared to the others.

Nevertheless, performance limitations are evident in all regression models, particularly when the FD deviates in P-defocus (below

-2  $\mu\text{m}$ ) or N-defocus (above +2  $\mu\text{m}$ ), leading to laser defocus and fabrication failures. Poor fabrication quality can damage the target sample. At these extreme FDs, the intensity of the focal spot is significantly low and blurred, limiting the information that can be

extracted. Therefore, it is more meaningful to consider detection accuracy within the in-focus region. Fig. 4 shows detection error counts for the in-focus region within  $\pm 2 \mu\text{m}$  among 1075 images. The x-axis denotes detection error, and the y-axis indicates the

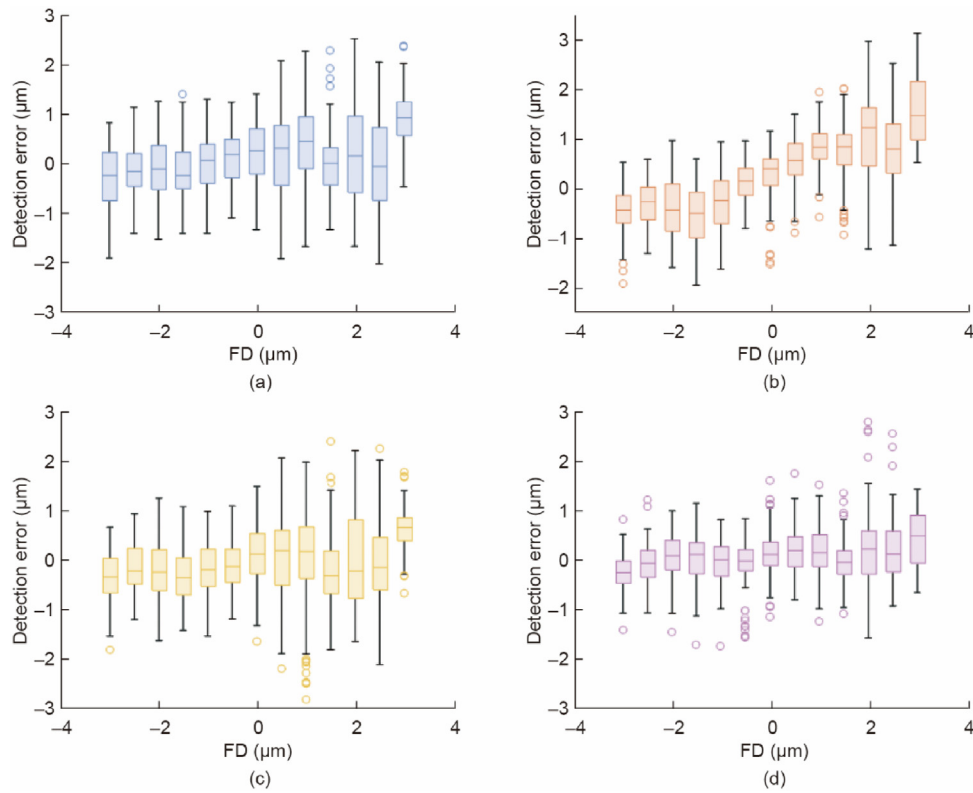


Fig. 3. Comparison of FD detection errors: (a) RQGPR, (b) KALS, (c) QSVM, and (d) TNN.

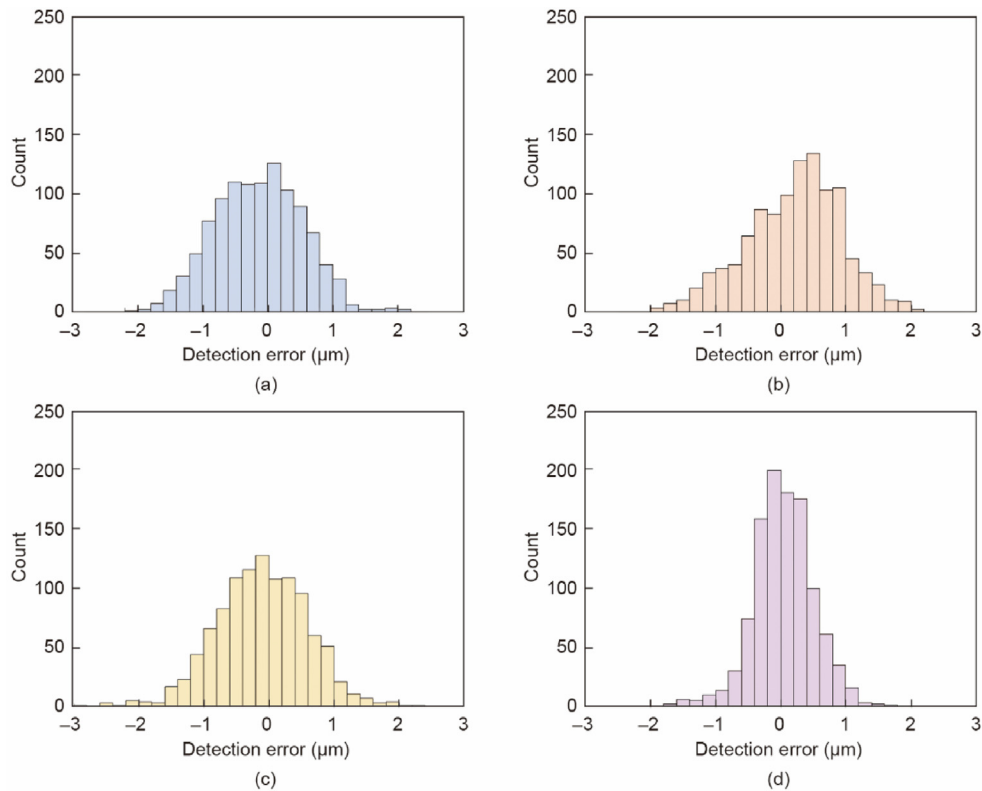


Fig. 4. Detection error counts within the in-focus region ( $\pm 2 \mu\text{m}$ ): (a) RQGPR, (b) KALS, (c) QSVM, and (d) TNN.

count of detection error in each bin. Symmetric distributions centered at 0  $\mu\text{m}$  are observed across all figures except KALS, which is skewed to the right with its peak count at +0.6  $\mu\text{m}$ . The RQGPR and QSVM models display lower detection error counts compared to TNN in the in-focus region. In contrast, TNN provides the highest counts, with more than 200 counts at zero detection error. Additionally, most detection errors fall within  $\pm 0.4 \mu\text{m}$ , which is much

smaller than the required fabrication accuracy of a tightly focused laser beam, indicating high precision and repeatability. Therefore, TNN FD detection is suitable for the nanofabrication of femtosecond lasers.

Fig. 5 presents the FD detection results, where asterisk lines indicate the detected FD and green lines indicate the actual one. KALS exhibits the largest detection error amplitude, with an RMSE

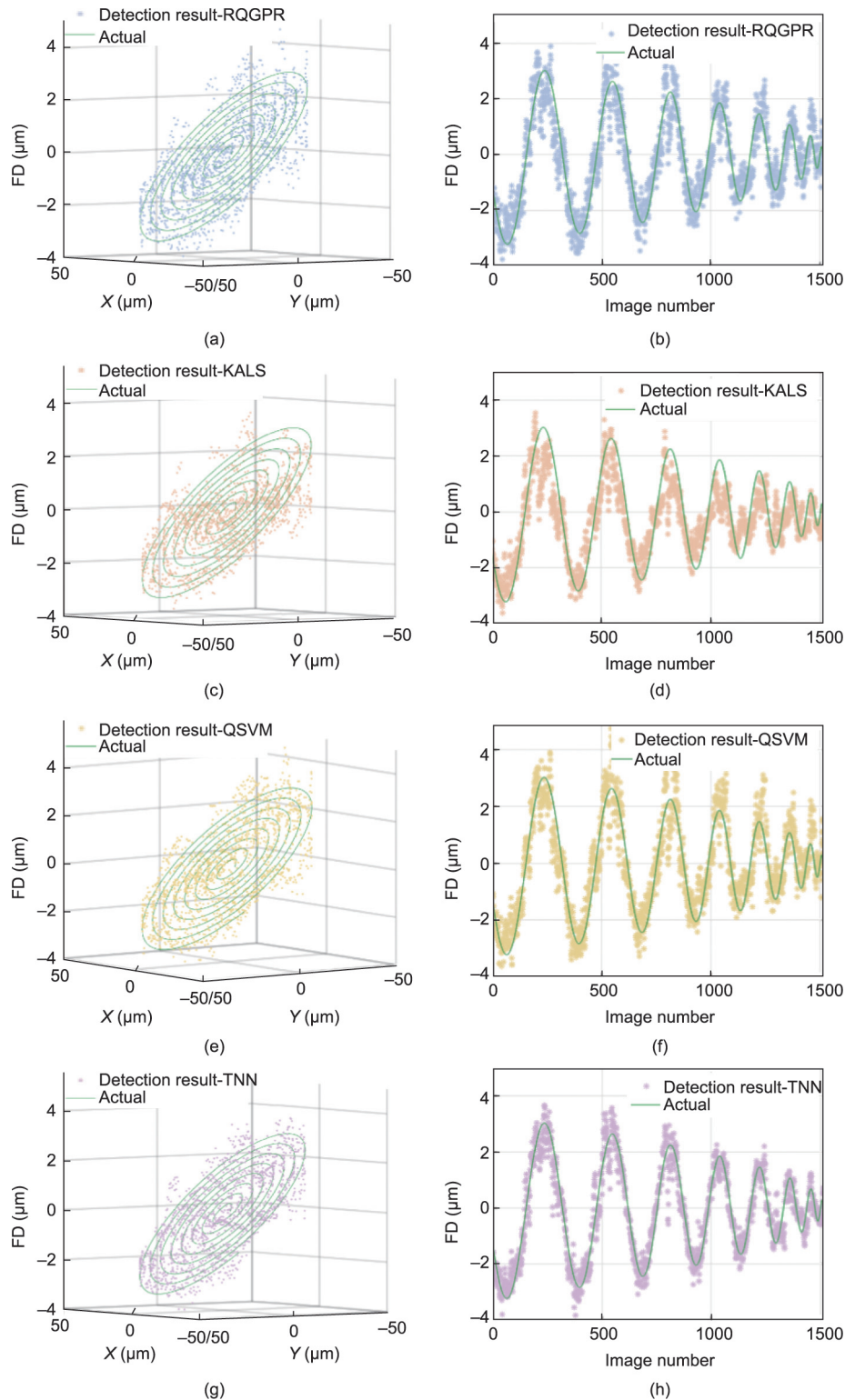
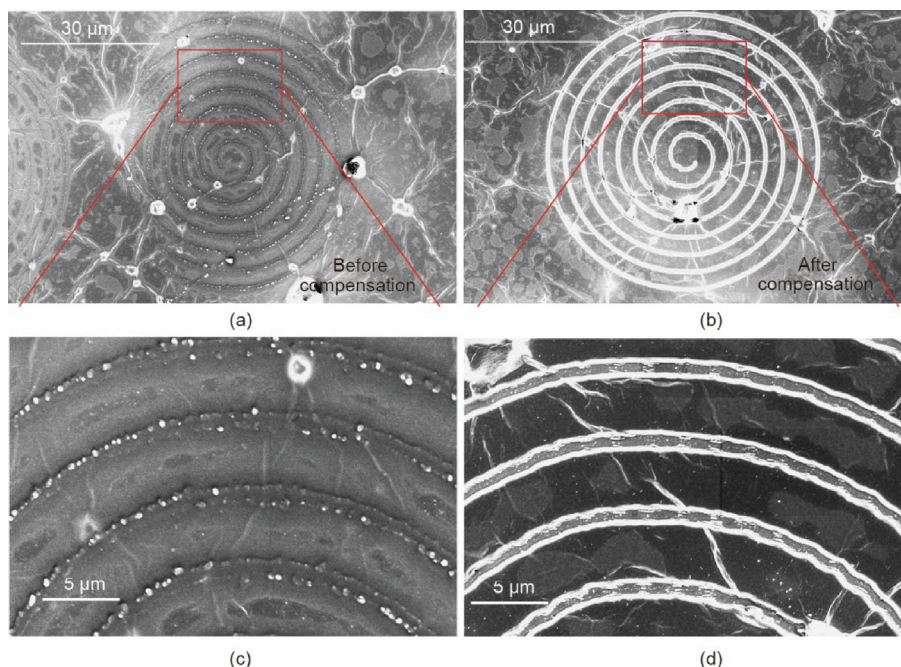


Fig. 5. FD detection results under four regression models. The sample is placed on a surface at a 5° inclination (green line), and the detected FD is indicated by asterisk lines. The left images are 3D plots, and the right images are 2D plots: (a, b) RQGPR, (c, d) KALS, (e, f) QSVM, and (g, h) TNN.



**Fig. 6.** SEM images of machine learning prediction for data compensation at low magnification and high magnification. (a) Before compensation at low magnification; (b) after compensation at low magnification; (c) before compensation at high magnification; (d) After compensation at high magnification.

**Table 1**  
Comparisons of detection accuracy between RQGPR and TNN.

Data set	Angle of inclination (°)	Number of images	RMSE		Improvement by TNN
			RQGPR	TNN	
1	3	1987	0.8792	0.6917	21%
2	5	1938	1.0225	0.6462	37%
3	5	1301	1.1051	0.7833	29%
4	5	1316	0.8209	0.7651	7%
5	5	1517	0.7133	0.4888	31%
6	5	1220	0.7623	0.6083	20%
7	8	1315	1.0138	0.7580	25%

value of 0.8435. Both RQGPR and QSVM display larger detection errors within  $\pm 2 \mu\text{m}$  compared to TNN, particularly when the FD deviates within the N- or P-defocus regions. The RMSE values for RQGPR and QSVM are 0.7133 and 0.7162  $\mu\text{m}$ , respectively. Conversely, TNN demonstrates the best performance, with the smallest RMSE of 0.5034  $\mu\text{m}$ .

In conclusion, TNN exhibits the highest detection accuracy, followed by RQGPR, which slightly outperforms QSVM, while KALS shows the poorest performance. The SEM images of data compensation using machine learning in a GO sample are presented in Fig. 6. Fig. 6(a) generated without compensation, shows an incomplete structure with out-of-focus left and right sides, rendering it unable to interact with materials. This structure is considered a failure in device fabrication. Conversely, Fig. 6(b) shows a more uniform and high-contrast structure produced with compensation due to the strong interaction between the laser focus and the material. The fabricated area is properly reduced, displaying high conductivity (bright part in the SEM image), resulting in very high contrast. This structure can function as a functional orbital angular momentum (OAM) GO lens [36], highlighting the improvement in accuracy when utilizing machine learning prediction for data compensation technology during the patterning process. Additionally, Table 1 outlines all the detection accuracy results from 10 594 images between the two closely performing models, RQGPR and

TNN, where the data sets are collected under additional working conditions with different angles of inclination of the sample. It is concluded that the TNN model outperforms the RQGPR model by an average improvement of 24% in FD detection accuracy concerning RMSE values. Therefore, the trained TNN model is regarded as the most suitable FD estimator.

### 5. Conclusion

This study offers a comprehensive analysis of machine learning regression models that incorporate image processing algorithms with fabrication data in the laser nanofabrication process. The research examines the impact of key parameters designs, such as the resizing factor of the input image, ROI size, and the data split ratio between training and testing, to optimize detection accuracy and efficiency. Comparative analysis indicated that a resizing factor of 0.2 and an ROI size of  $51 \times 51$  pixels yielded the best performance, characterized by the lowest RMSE and detection time. Additionally, the optimal models evaluated—RQGPR, KALS, QSVM, and TNN—TNN demonstrated superior performance. Specifically, the TNN split ratio of 85%–15% between training and testing data resulted in superior performance metrics. Among the four regression model achieved the smallest RMSE of 0.3427  $\mu\text{m}$  and MAE of 0.2437  $\mu\text{m}$  for training data, and RMSE of 0.2566  $\mu\text{m}$  and MAE

of 0.1857  $\mu\text{m}$  for testing data. Furthermore, the TNN model attained the highest  $R^2$  value of 0.98 for testing data and the highest MPS of 1000 ( $\text{obs}\cdot\text{s}^{-1}$ ). Statistical analysis confirmed that the TNN model provided the most reliable detection within the in-focus region of  $\pm 2 \mu\text{m}$ .

Based on the findings, the TNN model emerges as a promising FD estimator for future research, particularly regarding the developing automatic focus calibration systems in laser nanofabrication. This advancement could substantially enhance the monitoring and control of laser FDs, thereby ensuring the fabrication quality across a broad spectrum of applications in laser nanofabrication.

In addition, this high-precision FD detection technique can be applied to other 2D material nanofabrication systems and large-scale laser fabrication, which demand high accuracy and repeatability [37]. Furthermore, it can be utilized for various types of laser machining, including picosecond (PS) and nanosecond (NS) machining. Without requiring additional optical components, this method can be directly integrated into laser nanofabrication systems equipped with cameras for in situ process monitoring, aiding in dynamic focus adjustment and enabling exceptionally rapid decisions. The next step involves feeding the motorized stage with FD detection results and adjusting the focal spot through feedback control, thereby creating a real-time autofocusing mechanism for the laser nanofabrication setup.

### CRedit authorship contribution statement

**Hao Zhang:** Writing – review & editing, Writing – original draft, Validation, Investigation, Formal analysis, Conceptualization. **Jinchuan Zheng:** Writing – review & editing, Supervision, Methodology, Conceptualization. **Guiyuan Cao:** Methodology. **Han Lin:** Writing – review & editing, Supervision, Resources. **Baohua Jia:** Writing – review & editing, Supervision, Resources.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

This work was supported by the Australia Research Council (DP220100603, FT210100806, and FT220100559), the Industrial Transformation Training Centres scheme (IC180100005), the Linkage Project scheme (LP210200345, LP210100467, and LP240100504), the LIEF Project (LE250100078), and the Centre of Excellence Program (CE230100006). The authors would like to thank Innofocus Photonics Technology Pty Ltd. for providing the experimental platform and Frank Yao and Pei-chun Kao for helping set up the platform.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.eng.2025.03.037>.

### References

- Lin H, Zhang Z, Zhang H, Lin KT, Wen X, Liang Y, et al. Engineering van der Waals materials for advanced metaphotonics. *Chem Rev* 2022;122(19):15204–355.
- Vercillo V, Tonnichia S, Romano JM, García-Girón A, Aguilar-Morales AI, Alamri S, et al. Design rules for laser-treated icephobic metallic surfaces for aeronautic applications. *Adv Funct Mater* 2020;30(16):1910268.
- Wu J, Lin H, David J, Loh KP, Jia B. Graphene oxide for photonics, electronics and optoelectronics. *Nat Rev Chem* 2023;7(3):162–83.
- Cao G, Lin H, Fraser S, Zheng X, Del Rosal B, Gan Z, et al. Resilient graphene ultrathin flat lens in aerospace, chemical, and biological harsh environments. *ACS Appl Mater Interfaces* 2019;11(22):20298–303.
- Zhang Y, Wu J, Jia L, Qu Y, Yang Y, Jia B, et al. Graphene oxide for nonlinear integrated photonics. *Laser Photonics Rev* 2023;17(3):2200512.
- Kim YG, Rhee HG, Ghim YS. Real-time method for fabricating 3D diffractive optical elements on curved surfaces using direct laser lithography. *Int J Adv Manuf Technol* 2021;114:1497–504.
- Wang W, Liu YQ, Liu Y, Han B, Wang H, Han DD, et al. Direct laser writing of superhydrophobic PDMS elastomers for controllable manipulation via marangoni effect. *Adv Funct Mater* 2017;27(44):1702946.
- Wu T, Yin K, Pei J, He Y, Duan JA, Arnusch CJ. Femtosecond laser-textured superhydrophilic coral-like structures spread agnws enable strong thermal camouflage and anti-counterfeiting. *Appl Phys Lett* 2024;124(16):161602.
- Ma M, Wang Z, Wang D, Zeng X. Control of shape and performance for direct laser fabrication of precision large-scale metal parts with 316L stainless steel. *Opt Laser Technol* 2013;45:209–16.
- Scarisoareanu ND, Craciun F, Dinescu M, Ion V, Andrei A, Moldovan A, et al. Laser processing of nanostructures: enhancing functional properties of lead-free perovskite nanostructures through chemical pressure and epitaxial strain. In: Dinca V, Suchea MP, editors. *Micro and nano technologies, functional nanostructured interfaces for environmental and biomedical applications*. Amsterdam: Elsevier; 2019. p. 113–52.
- Yang P, Yin K, Song X, Wang L, Deng Q, Pei J, et al. Airflow triggered water film self-sculpturing on femtosecond laser-induced heterogeneously wetted micro/nanostructured surfaces. *Nano Lett* 2024;24(10):3133–41.
- He Y, Yin K, Wang L, Wu T, Chen Y, Arnusch CJ. Femtosecond laser structured black superhydrophobic cork for efficient solar-driven cleanup of crude oil. *Appl Phys Lett* 2024;124(17):171601.
- Weng W, Deng Q, Yang P, Yin K. Femtosecond laser-chemical hybrid processing for achieving substrate-independent superhydrophobic surfaces. *J Cent South Univ* 2024;31(1):1–10.
- Li L, Hong M, Schmidt M, Zhong M, Malshe A, Huis In'Tveld B, et al. Laser nano-manufacturing—state of the art and challenges. *CIRP Ann* 2011;60(2):735–55.
- Dinh VH, Hoang LP, Vu YNT, Cao XB. Auto-focus methods in laser systems for use in high precision materials processing: a review. *Opt Lasers Eng* 2023;167:107625.
- Agafonov VV, Safronov AG. Efficiency of objectives with deformable mirrors. 1. Controlling the focal length and the position of the focal spot. *J Opt Technol* 2005;72(6):448–54.
- Wang D, Ding X, Zhang T, Kuang H. A fast auto-focusing technique for the long focal lens TDI CCD camera in remote sensing applications. *Opt Laser Technol* 2013;45:190–7.
- Chen TH, Fardel R, Arnold CB. Ultrafast z-scanning for high-efficiency laser micro-machining. *Light Sci Appl* 2018;7(4):17181.
- Alexeev I, Strauss J, Gröschl A, Cvecek K, Schmidt M. Laser focus positioning method with submicrometer accuracy. *Appl Opt* 2013;52(3):415–21.
- Luo J, Liang Y, Yang G. Dynamic scan detection of focal spot on nonplanar surfaces: theoretical analysis and realization. *Opt Eng* 2011;50(7):073601.
- Bai Z, Wei J. Focusing error detection based on astigmatic method with a double cylindrical lens group. *Opt Laser Technol* 2018;106:145–51.
- Antti M, Ville H, Jorma V. Precise online auto-focus system in high speed laser micromachining applications. *Phys Procedia* 2012;39:807–13.
- Luo J, Liang Y, Yang G. Realization of autofocusing system for laser direct writing on non-planar surfaces. *Rev Sci Instrum* 2012;83(5):053102.
- Xu SJ, Duan YZ, Yu YH, Tian ZN, Chen QD. Machine vision-based high-precision and robust focus detection for femtosecond laser machining. *Opt Express* 2021;29(19):30952–60.
- Imani F, Chen R, Diewald E, Reutzel E, Yang H. Deep learning of variant geometry in layerwise imaging profiles for additive manufacturing quality control. *J Manuf Sci Eng* 2019;141(11):111001.
- Vo C, Zhou B, Yu X. Optimization of laser processing parameters through automated data acquisition and artificial neural networks. *J Laser Appl* 2021;33(4):042025.
- Subramonian S, Khalim AZ, Yusoff Y, Pujari S, Malingam SD, Amran MA. Optimization and prediction of laser micro-grooving by artificial neural network. *Int J Eng Technol* 2019;7(4):6481–7.
- Rahimi MH, Shayganmanesh M, Noorossana R, Pazhuheian F. Modelling and optimization of laser engraving qualitative characteristics of Al-SiC composite using response surface methodology and artificial neural networks. *Opt Laser Technol* 2019;112:65–76.
- Mohanavel V, Gandhimathi G, Bhardwaj D, Kavitha M, Ramkumar G, Ishwarya MV, et al. Deep learning-guided femtosecond laser processing in optical materials and devices for nano fabrication advancements. *Opt Quantum Electron* 2024;56(2):210.
- Gostimirovic D, Grinberg Y, Xu DX, Liboiron-Ladouceur O. Improving fabrication fidelity of integrated nanophotonic devices using deep learning. *ACS Photonics* 2023;10(6):1953–61.
- Polat C, Yapici GN, Elahi S, Elahi P. High-precision laser focus positioning of rough surfaces by deep learning. *Opt Lasers Eng* 2023;168:107646.
- Lin H, Sturmberg BCP, Lin KT, Yang Y, Zheng X, Chong TK, et al. A 90-nm-thick graphene metamaterial for strong and extremely broadband absorption of unpolarized light. *Nat Photonics* 2019;13(4):270–6.
- Cao G, Lin H, Jia B, Yuan X, Somekh M, Wei S. Design of a dynamic multi-topological charge graphene orbital angular momentum metalens. *Opt Express* 2023;31(2):2102–11.

- [34] Occorsio D, Ramella G, Themistoclakis W. Image scaling by de la Vallée-Poussin filtered interpolation. *J Math Imaging Vis* 2022;65:513–41.
- [35] Fukushima K. Visual feature extraction by a multilayered network of analog threshold elements. *IEEE Trans Syst Sci Cybern* 1969;5(4):322–33.
- [36] Cao G, Lin H, Jia B. Broadband diffractive graphene orbital angular momentum metalens by laser nanoprinting. *Ultrafast Sci* 2023;3:0018.
- [37] Lin K, Nian X, Li K, Han J, Zheng N, Lu X, et al. Highly efficient flexible structured metasurface by roll-to-roll printing for diurnal radiative cooling. *eLight* 2023;3:22.