Journal of Engineering Mechanics and Machinery (2026)
Clausius Scientific Press, Canada

DOI: 10.23977/jemm.2026.110118
ISSN 2371-9133 Vol. 11 Num. 1

Research on Intelligent Identification of Defocus Distance in Laser Powder Feeding Additive Manufacturing Based on PCA-CNN

Jiayu Wang, Jiali Gao*, Pengfei Qiu

University of Shanghai for Science and Technology, Shanghai, China

Keywords: Laser additive manufacturing, defocus distance, principal component analysis, online monitoring, convolutional neural network

Abstract: Precise control of defocus distance is key to ensuring the cladding quality in laser powder feeding additive manufacturing. To achieve online real-time monitoring and intelligent identification of the defocus state, this study proposes an intelligent identification framework that integrates principal component analysis (PCA) with deep learning. Three models—multilayer perceptron (MLP), autoencoder (AE), and convolutional neural network (CNN)—were systematically constructed and compared. The results show that, through its inherent convolution and pooling operations, the PCA-CNN model can adaptively capture the gradual changes in the two-dimensional spatial morphology of the molten pool caused by minor variations in defocus distance, achieving an overall accuracy of 99.28% on an independent test set. This performance is significantly superior to that of PCA-MLP (95.34%) and PCA-AE (93.19%). The proposed PCA-CNN intelligent identification framework enables rapid and accurate determination of the defocus state during the cladding process, providing a reliable technical foundation for subsequent real-time compensation of process parameters and closed-loop quality control.

1. Introduction

Laser powder feeding additive manufacturing technology has been widely applied in fields such as high-performance fabrication of complex components, repair of damaged parts, and preparation of functionally graded materials (2024)[1]. The final forming quality of this process heavily depends on the stability of the cladding process. The working distance, defined as the vertical distance between the outlet of the powder feeding nozzle and the surface of the substrate (or the deposited layer), directly affects the laser energy density and the convergence state of the powder stream on the substrate; therefore, it is a key parameter influencing process stability and forming quality.

However, during actual processing, especially when manufacturing large or complex curved parts, factors such as substrate unevenness or thermal deformation can easily cause dynamic deviation of the working distance, leading to defocus. When the working distance is too short, the laser processing head operates under negative defocus, which tends to cause mechanical interference. At the same time, spatter from the molten pool can contaminate the nozzle of the processing head, hindering powder delivery and, in severe cases, interrupting the process and ruining the forming of the entire

component (2023)[2]. When the working distance is too long, the positive defocus state of the laser processing head causes increased attenuation of laser energy during transmission. Powder particles cannot be sufficiently preheated before reaching the molten pool, resulting in a reduced powder capture rate, which in turn leads to insufficient single-track forming height or defects such as lack of fusion in the part (2020)[3].

Therefore, real-time and accurate monitoring and control of the defocus distance are crucial for achieving high-quality and stable forming in laser powder feeding additive manufacturing. Current monitoring methods are mainly divided into two categories: offline and online techniques. Offline monitoring (e.g., calipers, probes) cannot be used for dynamic process control. Existing online techniques, such as scanning systems based on laser triangulation or structured light, can achieve non-contact measurement, but typically require scanning before processing or after process interruption; due to spatial offsets, they are difficult to realize true synchronous in-situ monitoring. Furthermore, some researchers have attempted to use laterally mounted CCD cameras or acoustic signals for monitoring, but these methods have limitations in terms of application scenario universality or accuracy (2018)[4], (2021)[5]. In recent years, optical coherence tomography (OCT), owing to its strong anti-interference capability and coaxial integration potential, has been successfully applied to measure molten pool depth or distance in additive manufacturing processes, demonstrating its potential as an advanced process monitoring tool (2022)[6]; however, it is costly and involves complex optical system integration. In contrast, coaxial vision-based monitoring methods offer advantages such as low cost, rich information content, and ease of integration. By analyzing molten pool images, visual features reflecting the defocus state can be directly extracted [7]–[9]. Nevertheless, high-dimensional image data often contain substantial redundant information, and direct processing faces challenges of low computational efficiency and insufficient real-time performance.

This study proposes an intelligent identification framework for the defocus distance of the laser processing head in laser powder feeding additive manufacturing, which integrates principal component analysis (PCA) with deep learning models. First, by fabricating a series of single tracks with different preset defocus distances, a dataset pairing defocus distance with coaxial images was synchronously constructed. Subsequently, PCA was performed on all raw images to achieve unified dimensionality reduction and feature extraction. On this basis, a multilayer perceptron (MLP) model and an autoencoder (AE) model were constructed to train and identify the reduced one-dimensional feature vectors. A convolutional neural network (CNN) was also constructed to further train and identify the defocus distance after reconstructing the reduced principal components into denoised two-dimensional images. This study systematically compares these three models—all derived from the same PCA feature source but differing in input format and network architecture—aiming to explore an efficient, low-cost approach for intelligent identification of the defocus distance of the laser processing head in laser powder feeding additive manufacturing, thereby laying a foundation for ultimately achieving closed-loop quality control in the additive manufacturing process.

2. Experiment and Molten Pool Image Acquisition

In this study, molten pool image acquisition experiments were carried out during the laser powder feeding additive manufacturing process. A coaxial vision system was used to capture images of the molten pool region under different defocus distances, providing raw visual data for subsequent research. The core acquisition device of this experiment is a coaxial image acquisition system. The image acquisition unit of this system is a Basler ace 2 GigE series industrial camera, which is equipped with a Sony IMX392 color CMOS image sensor. Owing to the use of a global shutter, the system can significantly reduce motion blur when capturing the dynamic molten pool. The image acquisition resolution is set to 1920 × 1200 pixels.

The defocus distance refers to the vertical distance between the laser focal plane and the actual processing surface of the workpiece. In Figure 1(a), the laser focus is exactly on the processing plane, and the defocus distance is zero. In Figure 1(b), the processing head is in a negative defocus state (–4 mm), meaning the processing plane is above the focus. In Figure 1(c), the processing head is in a positive defocus state (+8 mm), meaning the processing plane is below the focus. The influence of defocus distance on forming quality is mainly related to the laser energy distribution. At the focal point, the laser spot is the smallest and the unit energy is the highest. As the defocus distance increases, the laser energy becomes more dispersed and the spot becomes larger.

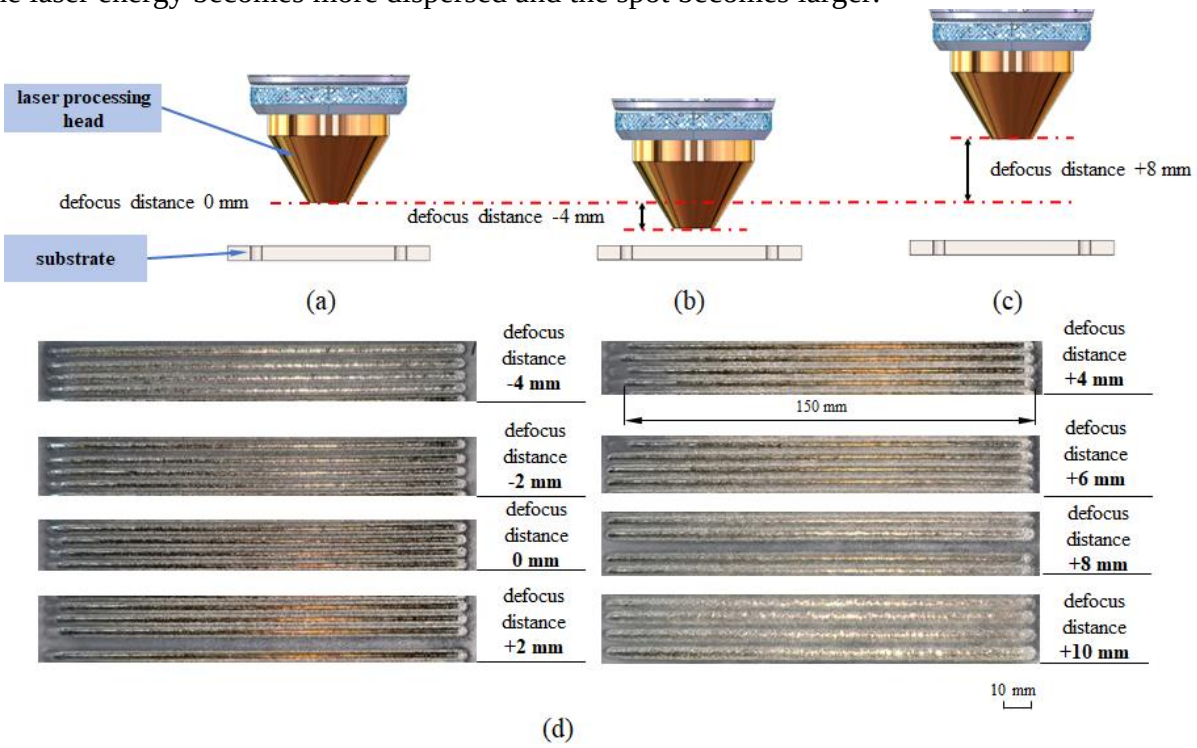


Figure 1 Schematic diagram of defocus states: (a) defocus distance of 0; (b) negative defocus state (–4 mm); (c) positive defocus state (+8 mm); (d) fabrication results for different defocus distances.

To achieve online real-time identification and monitoring of the defocus distance, this study conducted Ti6Al4V single-track deposition experiments on a Ti6Al4V substrate. The design of the defocus distance and the single-track deposition results are shown in Figure 1(d). The defocus distance ranged from –4 mm to +10 mm, with a total of eight levels (at intervals of 2 mm). The other process parameters were kept constant: laser power 1200 W, scanning speed 8 mm/s, and powder feed rate 6.78 g/min. The length of each single track was designed to be 150 mm. For each defocus distance level, deposition was repeated four times, resulting in a total of 32 experimental single tracks. During the deposition of each single track, the coaxial image acquisition system was triggered to capture molten pool images. To construct a dataset for model training and testing, all captured molten pool images were organized and partitioned.

The dataset partitioning followed the principle of ensuring independence between the training set and the test set. For the four repeated single tracks under each defocus distance level, all images corresponding to three of the single tracks were selected as the training set, and all images corresponding to the remaining single track were selected as the test set. Thus, the image samples of the entire dataset were divided into a training set and a test set at a ratio of 3:1. This partitioning scheme aims to evaluate the model based on new single-track data that have not appeared in training, thereby testing its generalization performance.

3. Construction of Intelligent Identification Model for Defocus Distance

3.1 PCA Feature Extraction

Principal component analysis (PCA) is a linear unsupervised dimensionality reduction and feature extraction method (2024)[10]. Its core objective is to find a set of new orthogonal bases for high-dimensional data, namely principal components (PCs), such that when the original data is projected onto a low-dimensional subspace spanned by the first k principal components, the most significant variation information is retained. For image data, each grayscale image of size $H \times W$ is treated as a $D = H \times W$ dimensional vector $\mathbf{x}$. Given a training set containing n samples, a data matrix $\mathbf{x} \in \mathrm{R}^{n \times D}$ is constructed. PCA first centers the data as follows:

$$\mathbf{X}_c = \mathbf{X} - \bar{\mathbf{X}} \tag{1}$$

where $\bar{\mathrm{X}}$ is the sample mean vector. The covariance matrix is then computed as:

$$\mathbf{C} = \frac{1}{n-1}\mathbf{X}_c^{\mathrm{T}}\mathbf{X}_c \tag{2}$$

and the principal component directions $\mathbf{v}$ along with their corresponding variance contributions (eigenvalues λ) are obtained by solving the eigenvalue decomposition problem:

$$\mathbf{C}\mathbf{v} = \lambda\mathbf{v} \tag{3}$$

By selecting the eigenvectors corresponding to the first k largest eigenvalues to form the projection matrix $\mathbf{V}_k \in \mathrm{R}^{D \times k}$, the original high-dimensional data can be reduced to k dimensions:

$$\mathbf{Z} = \mathbf{X}_c\mathbf{V}_k \tag{4}$$

where $\mathbf{z} \in \mathrm{R}^{n \times k}$ is the reduced-dimensional feature representation.

In this study, PCA was applied to uniformly preprocess the coaxial molten pool images acquired during the laser powder feeding additive manufacturing process. First, all training set images were converted to grayscale and resized to a uniform dimension. Subsequently, each two-dimensional image matrix was flattened into a one-dimensional vector, and all sample vectors together formed the initial high-dimensional data matrix. To ensure comparability across pixel dimensions, the data matrix was standardized by performing zero-mean and unit-variance normalization for each pixel position (feature dimension). Based on the standardized data, PCA was performed to extract a specified number of principal components. This process determined the linear mapping from the original high-dimensional pixel space to the low-dimensional principal component space.

In this study, the molten pool images after preprocessing and cropping were uniformly resized to 350 × 350 pixels, meaning each image formed a 122,500-dimensional vector after flattening. Fifty principal components were extracted using the PCA algorithm, reducing the dimensionality from 122,500 to 50, a reduction ratio of approximately 99.96%. With only 50 principal components, about 95% of the core information of the original images was retained. This effectively preserves the key features of the molten pool while achieving a substantial reduction in data dimensionality, thereby simplifying the computational complexity of subsequent models.

For each subsequent new input image, the preprocessing procedure includes the following steps: first, the same grayscale conversion, resizing, and standardization operations are performed; then, through the established mapping relationship, the image is projected onto the principal component space to obtain a low-dimensional feature vector. In addition, this low-dimensional feature vector is inversely mapped back to the original image pixel space to generate a reconstructed image. This reconstructed image can filter out some of the random noise in the original image while preserving

its main structural information. Figure 2 shows the images reconstructed after PCA processing. Figure 2(a) is the original molten pool image, and Figure 2(b) is the molten pool image reconstructed after PCA feature extraction.

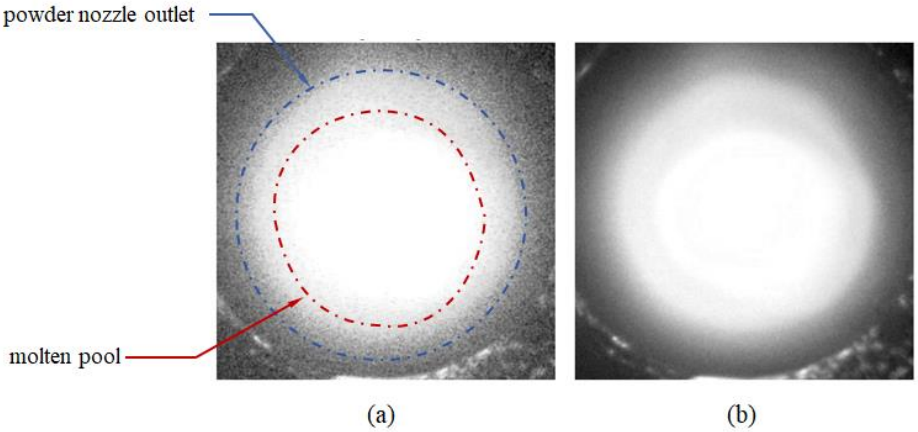


Figure 2 Comparison of PCA-reconstructed images: (a) original molten pool image; (b) molten pool image reconstructed after PCA feature extraction.

In the molten pool images acquired by the coaxial vision system, two bright and dark regions originating from different physical sources can be observed. Among them, the diameter of the largest bright ring in the image (indicated by the blue dashed line) roughly corresponds to the outlet aperture of the actual powder feeding nozzle. This region is mainly formed by diffuse reflection of laser light from the inner wall or outlet edge of the nozzle; it is not the molten pool itself, but its contour clearly indicates the real-time position and working distance of the nozzle. In addition, the smaller and brightest part inside the annular bright region (indicated by the red dashed line) is the molten pool itself. Its high brightness originates from the thermal radiation of the metal molten pool and the direct reflection of the laser, making it the core observation area reflecting the morphology, size, and stability of the molten pool. As can be seen from the comparison in the figure, the original molten pool image contains obvious granular random noise, with rough background texture and randomly distributed noise points. In the image reconstructed by PCA, these high-frequency random noises are effectively filtered out, resulting in a smoother overall visual effect, while key information such as the core shape structure and brightness gradient distribution of the molten pool is fully preserved. By performing inverse reconstruction from the low-dimensional PCA features, the interference of irrelevant noise is suppressed while retaining the core structural information, providing a more robust and representative feature input for subsequent identification of the defocus distance.

3.2 PCA-MLP Model

Based on the 50-dimensional low-dimensional feature vectors obtained from PCA preprocessing, this study constructed a multilayer perceptron (MLP) model to identify the defocus distance in laser powder feeding additive manufacturing. As a classic feedforward artificial neural network, the MLP can effectively learn complex mapping relationships between input features and target categories through multi-layer nonlinear transformations [11]–[13]. The input layer of this model aligns with the PCA feature dimension, receiving 50 feature values. The network backbone consists of three fully connected hidden layers, with neuron counts of 256, 128, and 64, respectively. After each hidden layer, a batch normalization layer, a ReLU activation function, and a dropout operation with a rate of 0.5 are sequentially introduced. As a regularization technique, dropout suppresses overfitting by randomly masking a subset of neurons during training, thereby improving the generalization ability

of the model. Finally, the output layer consists of 8 neurons, corresponding to the 8 defocus distance categories to be identified, and uses the Softmax function to output a normalized class probability distribution. This model contains approximately 128,000 trainable parameters.

During the model training phase, the cross-entropy loss function was used to measure the difference between the predicted probability distribution of the model and the true labels. The optimization process employed the AdamW optimizer with an initial learning rate of 0.001, combined with a weight decay strategy to further control model complexity. A learning rate scheduler was applied to halve the learning rate when the validation accuracy did not improve for 10 consecutive training epochs.

Model training and validation were performed on the feature dataset processed by PCA, with the training set and validation set split in a ratio of 8:2. The training process lasted for 50 epochs. Figure 3 shows the training history curves. The training loss decreased rapidly from an initial value of 3.42 to approximately 0.14, while the validation loss also decreased from 2.11 to about 0.10 without significant rebound, indicating that the model did not overfit. In terms of accuracy, the training accuracy steadily increased from 19.20% to 96.02%, while the validation accuracy increased from 34.97% to a higher value of 96.72% and stabilized in the later stage of training. Notably, starting from the 26th epoch, the validation accuracy consistently exceeded the training accuracy, demonstrating good generalization ability of the model.

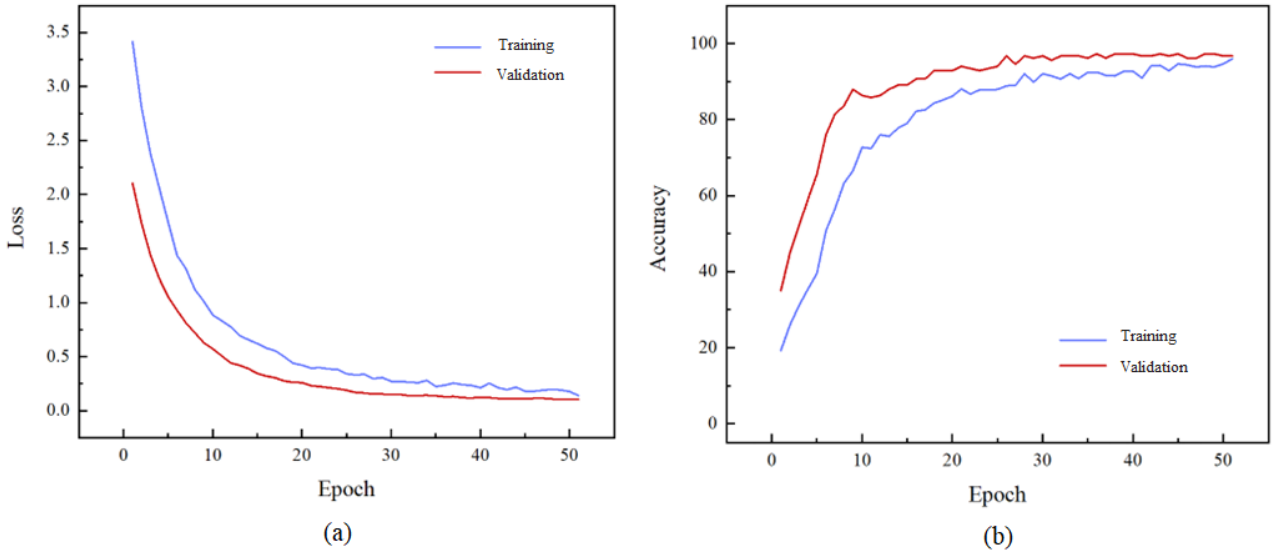


Figure 3 Training history curves of PCA-MLP: (a) loss curve; (b) accuracy curve.

3.3 PCA-AE Model

Based on the 50-dimensional low-dimensional feature vectors obtained from PCA preprocessing, this study constructed an end-to-end learning model that integrates an autoencoder and a classifier. The model accepts the 50-dimensional PCA feature vector as input, and its core is a shared encoder network. The encoder consists of three fully connected layers stacked sequentially, with neuron counts of 64, 32, and 16, respectively. Each layer includes batch normalization, a ReLU activation function, and a dropout operation with a rate of 0.3, ultimately compressing the input into a 16-dimensional latent space feature representation.

The model adopts a dual-branch architecture and a joint training strategy. The latent features output by the encoder are simultaneously fed into two parallel branches: one is a structurally symmetric decoder, whose objective is to reconstruct the original 50-dimensional feature space from the latent features; the other is a classifier, responsible for predicting the corresponding defocus distance

category based on the latent features. During training, the model is optimized using a weighted combination loss function, which consists of the cross-entropy loss for the identification task and the mean squared error loss for autoencoder reconstruction, with weights set to 0.7 and 0.3, respectively. The reconstruction loss is introduced as an auxiliary task to impose a strong constraint on the latent features, forcing them to preserve as much of the complete information structure of the original input as possible while focusing on classification discrimination, thereby enabling the learning of more robust and generalizable feature representations.

The model was trained using the AdamW optimizer on a training set and validation set split in a ratio of 8:2. The training history loss curves show that the total loss function value steadily decreased and eventually converged on both the training and validation sets. Figure 4 shows the training history curves. The total loss function value exhibited a stable decline and eventual convergence on both the training and validation sets. Analysis of the accuracy curves reveals that the validation accuracy gradually increased from an initial value of approximately 20.77%, entered a steady rising phase in the middle and later stages, and finally fluctuated within the range of 95% to 98% in the late training stage, reaching a maximum of 97.81%. The validation accuracy curve and the training accuracy curve did not show significant divergence, reflecting good learning stability and generalization ability of the model under the joint optimization objective.

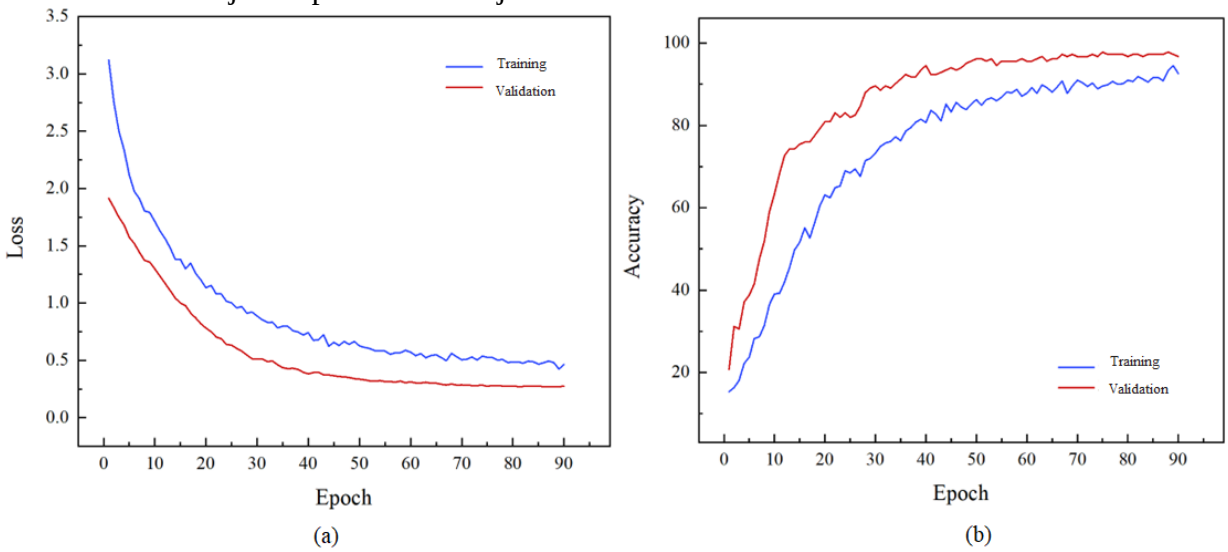


Figure 4 Training history curves of PCA-AE: (a) loss curve; (b) accuracy curve.

3.4 PCA-CNN Model

Based on the denoised images reconstructed by principal component analysis (PCA), this study constructed a convolutional neural network (CNN) model to identify the defocus distance in laser powder feeding additive manufacturing. By leveraging its properties of local connectivity, weight sharing, and spatial downsampling, the CNN can automatically and efficiently extract hierarchical visual features from two-dimensional images. The input of this model is a single-channel grayscale image reconstructed by PCA. The network backbone consists of four sequentially stacked convolutional blocks. Each convolutional block comprises two convolutional layers, each followed by batch normalization and a ReLU activation function, and finally a downsampling layer. The first three convolutional blocks use 2×2 max pooling for downsampling, while the fourth convolutional block uses adaptive average pooling to unify the feature map size to 4×4. The numbers of output channels for the four convolutional blocks are designed to be 32, 64, 128, and 256, respectively. All convolutional operations use 3×3 kernels with spatial padding. After flattening the output feature map

of the last convolutional block, it is fed into a classifier consisting of several fully connected layers, in which a dropout operation with a rate of 0.5 is introduced to prevent overfitting. Finally, the output layer consists of 8 neurons corresponding to the 8 defocus distance categories to be identified, and uses the Softmax function to output a normalized class probability distribution. During the model training phase, the cross-entropy loss function was adopted, and the Adam optimizer was used for parameter updates with an initial learning rate of 0.001, combined with a weight decay strategy.

The model was trained and validated on the PCA-reconstructed image dataset. The dataset was divided into a training subset and a validation subset at a ratio of 8:2. The training process lasted for a total of 100 epochs, and the changes in loss and accuracy are shown in Figure 5. The training loss generally showed a downward trend with fluctuations, steadily converging from 1.86 at the first epoch to 0.23 at the 100th epoch. Correspondingly, the training accuracy gradually increased from 25.38% to 92.73%, indicating that the model's ability to fit the training data continuously improved. The performance on the validation set exhibited staged fluctuations. In the early stage of training, the validation accuracy rapidly climbed to over 80%, reaching a maximum of 86.89%, demonstrating the model's ability to quickly capture general features. However, in the middle stage of training, the validation accuracy reached peaks multiple times but was interspersed with several significant performance regressions, accompanied by sharp spikes in the validation loss. This phenomenon may be related to the optimizer's exploration behavior on a complex non-convex loss surface or temporary instability when the model jumps out of local optima. As training progressed, the model gradually stabilized in the later stage, with reduced fluctuations in validation accuracy, ultimately maintaining a high range between 93% and 98%. Finally, the model achieved its best validation accuracy of 98.36% at the 87th epoch, indicating excellent classification and discrimination capability.

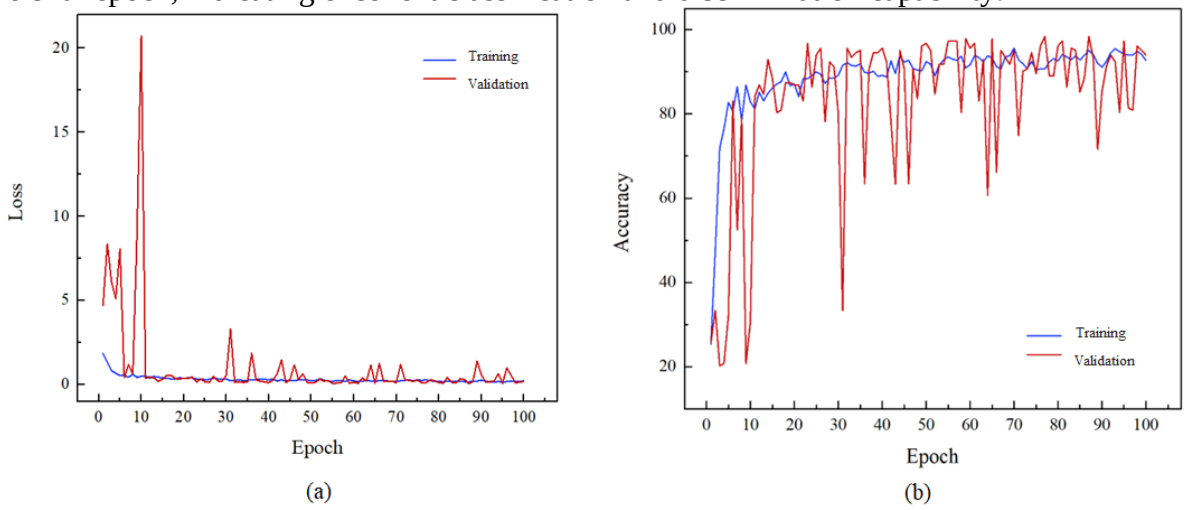
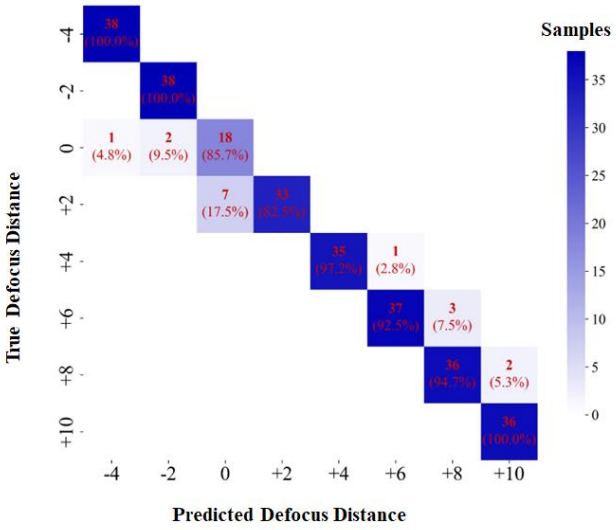


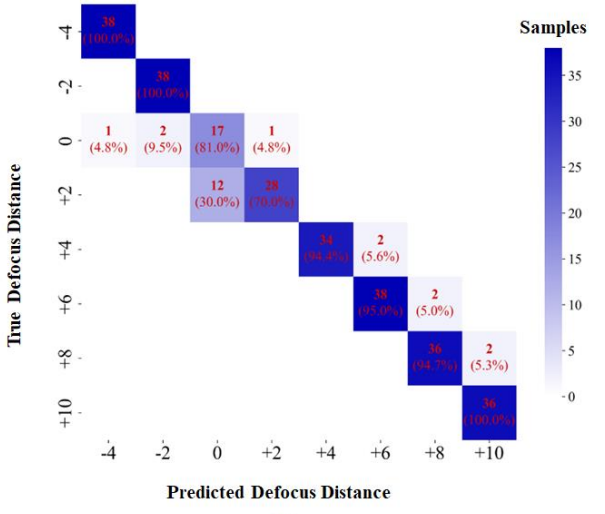
Figure 5 Training history curves of PCA-CNN: (a) loss curve; (b) accuracy curve.

4. Results and Analysis

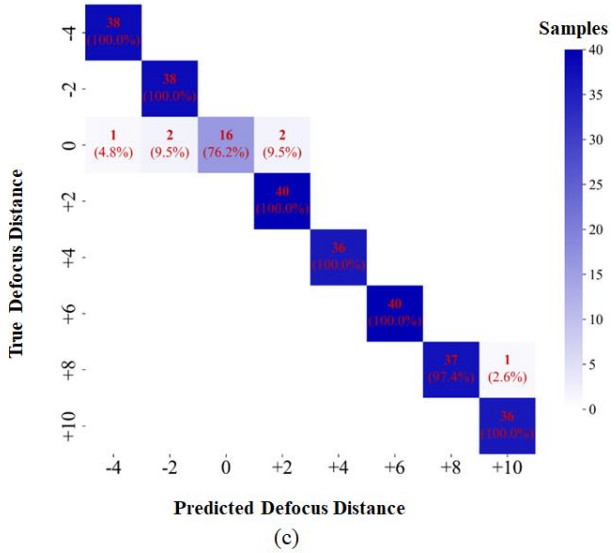
To further evaluate the performance of the proposed models, this section presents a detailed analysis of the identification results on the independent test set for the three models (PCA-MLP, PCA-AE, PCA-CNN). The test set consists of the images from the reserved fourth complete single track for each defocus distance category, ensuring that the evaluation data are completely unseen during model training, thereby truly reflecting the generalization ability of the models. After inputting the test set data into the models for identification, the confusion matrices of all models on the test set are shown in Figure 6.



(a)



(b)



(c)

Figure 6 Confusion matrix of defocus distance identification on the test set: (a) PCA-MLP model; (b) PCA-AE model; (c) PCA-CNN model.

Table 1 shows the comparison results of evaluation metrics for the three identification models.

Table 1. Evaluation results of the three identification models

Defocus distance	PCA-MLP				PCA-AE				PCA-CNN			
/mm	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score	Accuracy	Precision	Recall	F1-score	Accuracy
-4	1.000	1.000	1.000		1.000	1.000	1.000		1.000	1.000	1.000	
-2	1.000	1.000	1.000		1.000	1.000	1.000		0.974	1.000	0.987	
0	0.696	1.000	0.821		0.556	0.938	0.698		1.000	0.938	0.968	
+2	1.000	0.811	0.896		0.962	0.676	0.794		1.000	1.000	1.000	
+4	1.000	0.972	0.986		1.000	0.944	0.971		1.000	1.000	1.000	
+6	0.974	0.925	0.949		0.950	0.950	0.950		1.000	1.000	1.000	
+8	0.923	0.947	0.935		0.947	0.947	0.947		1.000	0.974	0.987	
+10	0.947	1.000	0.973		0.947	1.000	0.973		0.973	1.000	0.986	
Overall	0.942	0.957	0.945	0.953	0.920	0.932	0.917	0.932	0.994	0.989	0.991	0.993

For the 50-dimensional one-dimensional feature vectors extracted by PCA, the PCA-MLP model achieved excellent overall performance on the test set. The overall accuracy of this model on all test samples reached 95.34%, with macro-average precision, recall, and F1-score of 0.942, 0.957, and 0.945, respectively, indicating strong defocus distance identification capability. Further analysis of the performance for each category shows that the PCA-MLP model achieved nearly perfect recognition for the four distance categories of –4 mm, –2 mm, +4 mm, and +10 mm. However, the PCA-MLP model exhibited confusion between some adjacent defocus distance categories, which was its main source of error. Specifically:

For the 0 mm category, the recall of PCA-MLP reached 100%, but the precision was only 69.57%. Detailed analysis of the prediction results revealed that multiple samples belonging to +2 mm were incorrectly predicted as 0 mm.

For the +2 mm category, the precision was 100%, but the recall was 81.08%. The missed samples were exactly those misclassified as 0 mm.

For the +6 mm and +8 mm categories, there was slight mutual misclassification between the two.

For the vast majority of correctly identified samples, the PCA-MLP model assigned very high prediction probabilities to the true categories, exhibiting very confident judgments. For the few misidentified samples, the highest probability output by the PCA-MLP model was often significantly lower than that for correct samples, and the probability gap between the true category and the predicted category was small, intuitively reflecting the model's uncertainty on these boundary cases.

The PCA-AE model achieved an overall accuracy of 93.19% on the test set, with macro-average precision, recall, and F1-score of 0.920, 0.932, and 0.917, respectively. Analyzing its identification performance, this model also encountered significant difficulties with the 0 mm and +2 mm categories. The precision for the 0 mm category was as low as 0.5556, meaning nearly half of the samples predicted as 0 mm actually belonged to other categories; meanwhile, the recall for the +2 mm category was only 0.6757, with a large number of missed detections. Consequently, the F1-scores for these two categories (0.6977 and 0.7937) were the lowest among the three models. Although the model achieved perfect or near-perfect identification for categories such as –4 mm, –2 mm, and +10 mm, its serious errors on the core difficult categories reduced the overall performance, resulting in a lower overall accuracy than the PCA-MLP model.

The relatively limited performance of the PCA-AE model may be related to objective conflicts in multi-task learning. This model simultaneously optimizes category discrimination and accurate feature reconstruction, and the learned latent representation must trade off between retaining

sufficient detail to reconstruct the 50-dimensional input and abstracting the most discriminative features. In this study, the mandatory high-fidelity reconstruction constraint may have biased the latent space too much toward preserving all variations in the input, rather than focusing on extracting the most critical subtle features that distinguish adjacent defocus distance categories.

In contrast, the PCA-CNN model is directly oriented toward the identification objective and fully utilizes the spatial structural information of images, resulting in more task-specific feature representations. The PCA-CNN model takes the two-dimensional denoised images reconstructed by PCA as input and performs defocus distance identification using the inherent spatial feature extraction capability of convolutional neural networks. This model achieved extremely outstanding performance on the same independent test set as the PCA-MLP model, with a significant leap in identification accuracy. Its overall accuracy reached 99.28%, an improvement of nearly 4 percentage points over the PCA-MLP model (95.34%). Its macro-average F1-score was as high as 0.991, indicating a nearly perfect balance between precision and recall. A detailed comparison of category-wise performance shows that the PCA-CNN model achieved improvements in the weak areas where the PCA-MLP and PCA-AE models exhibited confusion:

For the weakest categories, 0 mm and +2 mm, the PCA-CNN achieved fundamental improvement. The precision for the 0 mm category improved from 0.6957 to a perfect 1.0000, while maintaining a high recall of 93.75%, and the F1-score increased significantly from 0.8205 to 0.9677. More importantly, the recall for the +2 mm category increased from 0.8108 to 100%, achieving an F1-score of 1.0000. The model clarified the complex feature boundary between these two adjacent categories.

The PCA-CNN model continued to deliver top-tier performance on the remaining categories. The model not only reduced the overall error rate by approximately 60% (from 4.66% to 1.72%), but also exhibited very small variance in performance across categories, demonstrating highly balanced and robust generalization ability.

The significant advantage of the PCA-CNN model is related to the deep integration of its architecture and input form. The PCA-reconstructed images suppress random noise while fully preserving the two-dimensional spatial structure of the molten pool. Through its stacked convolutional and pooling operations, the CNN can adaptively extract and combine local and global features from this spatial structure, forming a hierarchical, structured understanding of the molten pool morphology. This capability enhances its sensitivity and resolution to visual patterns caused by small changes in defocus distance. Therefore, it can learn more discriminative feature representations, enabling more accurate decisions on category boundaries where physical distances are adjacent and visual differences are subtle.

In summary, the model based on PCA-reconstructed images and the CNN architecture achieved a test accuracy of 99.28% on the defocus distance identification task in laser powder feeding additive manufacturing, validating the effectiveness and superiority of this technical approach.

5. Conclusion

This study focuses on the accurate identification of defocus distance during the laser powder feeding additive manufacturing process and proposes an intelligent monitoring framework that integrates principal component analysis (PCA) preprocessing with deep learning models. Through systematic processing and analysis of coaxially acquired molten pool images, and by comparing three models based on the same PCA feature source but with different architectures, the following main conclusions are drawn:

Principal component analysis (PCA) can effectively address the feature redundancy problem in high-dimensional molten pool image data. Through standardization, dimensionality reduction, and feature reconstruction, it not only preserves the core structural information of the molten pool images

but also filters out some random noise. While retaining approximately 95% of the core information, it reduces the data dimensionality by about 99.96%, significantly improving the processing efficiency of subsequent models. The resulting one-dimensional feature vectors and two-dimensional reconstructed images can be respectively adapted to different types of deep learning models, achieving unified and efficient utilization of the feature source.

The performance of different model architectures varies significantly, with the PCA-CNN model achieving an overall identification accuracy of 99.28%, which is notably superior to PCA-MLP (95.34%) and PCA-AE (93.19%). This model extracts hierarchical spatial local features from PCA-reconstructed images and is highly sensitive to subtle gradual changes in molten pool morphology and brightness distribution. Consequently, it has achieved fundamental improvements on the two key classification difficulties of 0 mm defocus distance and +2 mm defocus distance (the average F1-score for these two categories increased from 0.858 for MLP to 0.984).

Acknowledgments

This paper was supported by National Natural Science Foundation of China (Grant No. 52475479).

References

[1] Zhou Y, Jiang D, Al-Akailah A, et al. Understanding the formation of laser-induced melt pools with both wire and powder feeding in directed energy deposition[J]. Additive Manufacturing, 2024, 89: 104312.

[2] Kittel J, Wendt F, Hoelters S, et al. Approach for advanced working distance monitoring and control capability in laser metal deposition processing for additive manufacturing[J]. Journal of Laser Applications, 2023, 35.

[3] Gong X, You W, Li X, et al. Modeling the influence of injection parameters on powder efficiency in laser cladding[J]. Welding in the World, 2020, 64(8): 1437-1448.

[4] Shi T, Lu B, Shen T, et al. Closed-loop control of variable width deposition in laser metal deposition[J]. The International Journal of Advanced Manufacturing Technology, 2018, 97(9): 4167-4178.

[5] Chabot A, Rauch M, Hascoet J. Novel control model of Contact-Tip-to-Work Distance (CTWD) for sound monitoring of arc-based DED processes based on spectral analysis[J]. The International Journal of Advanced Manufacturing Technology, 2021, 116: 1-10.

[6] Stehmar C, Gipperich M, Kogel-Hollacher M, et al. Inline Optical Coherence Tomography for Multidirectional Process Monitoring in a Coaxial LMD-w Process[J]. Applied Sciences, 2022, 12(5).

[7] Ye J, Bab-Hadiashar A, Hoseinnezhad R, et al. Predictions of in-situ melt pool geometric signatures via machine learning techniques for laser metal deposition[J]. International Journal of Computer Integrated Manufacturing, 2023, 36(9): 1345-1361.

[8] Hsu H W, Lo Y L, Lee M H. Vision-based inspection system for cladding height measurement in Direct Energy Deposition (DED)[J]. Additive Manufacturing, 2019, 27: 372-378.

[9] Da Silva A, Frostevarg J, Kaplan A F H. Melt pool monitoring and process optimisation of directed energy deposition via coaxial thermal imaging[J]. Journal of Manufacturing Processes, 2023, 107: 126-133.

[10] Costa A P de A, Choren R, Pereira D A de M, et al. Integrating multicriteria decision making and principal component analysis: a systematic literature review[J]. Cogent Engineering, 2024, 11(1).

[11] Khan T A, Tulsi J, Alam M, et al. Analysis and visualization of fraud detection patterns through data mining and classification using MLP and hybrid deep learning model[J]. Egyptian Informatics Journal, 2025, 32: 100829.

[12] Liu W, Gai M. PV-MLP: A lightweight patch-based multi-layer perceptron network with time-frequency domain fusion for accurate long-sequence photovoltaic power forecasting[J]. Renewable Energy, 2025, 251.

[13] Mysliwiec P, Kubit A, Szawara P. Optimization of 2024-T3 Aluminum Alloy Friction Stir Welding Using Random Forest, XGBoost, and MLP Machine Learning Techniques[J]. Materials, 2024, 17(7).