Sensors and Materials, Vol. 38, No. 3 (2026) 1481–1502
MYU Tokyo

S & M 4389

U-Net-based Framework for Automated Weld Bead Morphology and Cross-section Analysis in Laser-directed Energy Deposition

Ke-Han Su,[1*] Sung-Chu Wang,[1] Yi-Jun Chen,[1]
Keng-Pin Chang,[2] and Haw-Ching Yang[1]

[1]Department of Electrical Engineering, National Kaohsiung University of Science and Technology, No. 1, University Road, Yanchao Dist., Kaohsiung City 82445, Taiwan
[2]Tongtai Machine & Tool Co. Ltd, No. 3, Luke 3rd Road, Luzhu Dist., Kaohsiung City 82151, Taiwan

(Received September 24, 2025; accepted February 16, 2026)

Keywords: laser-directed energy deposition (LDED), metal additive manufacturing, image processing technology, U-Net, machine learning

Laser-directed energy deposition (LDED) has emerged as a promising metal additive manufacturing technique owing to its small heat-affected zone, low material wastage, and minimal environmental constraints. It has been widely applied in aerospace component fabrication, brake disk coating, and high-value part repair. However, the quality and stability of LDED are highly dependent on process parameters such as laser power, powder feed rate, and scanning speed. Conventional parameter optimization largely relies on post-build inspections, in which the manual measurement of bead morphology and cross-sectional geometry is both time-consuming and error-prone. To address these drawbacks, in this study, we propose a deep-learning-based geometric data measurement framework for weld bead characterization. A U-Net-based semantic segmentation model was developed to analyze both appearance morphology and cross-sectional geometry in single-track deposition experiments. The proposed models achieved feature recognition accuracies exceeding 90% on test datasets, demonstrating robust inference performance. Furthermore, the integration of automated image-based inspection software reduced measurement time by approximately 91% compared with manual evaluation. The proposed framework demonstrates the potential of deep-learning-assisted image analysis to improve process analysis efficiency, reduce human-induced errors, and enhance parameter optimization in LDED, thereby contributing to more reliable and intelligent additive manufacturing systems.

1. Introduction

In recent years, with the rapid advancement of technology and the promotion of intelligent manufacturing and digital transformation, additive manufacturing (AM) has gradually become one of the most significant emerging technologies in modern production. Compared with conventional subtractive manufacturing, AM not only enables the fabrication of complex,

*Corresponding author: e-mail: kehansu@nkust.edu.tw
https://doi.org/10.18494/SAM5948

ISSN 0914-4935
https://myukk.org/

multidimensional components, but also offers advantages such as high material utilization, shortened production cycles, flexible design modifications, and digital integration of geometric information. Among the various AM techniques, directed energy deposition (DED) has attracted increasing attention owing to its advantages of a small heat-affected zone, low material waste, and minimal environmental constraints. DED has been widely applied in aerospace component fabrication, mold repair, and the remanufacturing of high-value parts. Although the market share of DED in maintenance, repair, and overhaul (MRO) applications remains limited, several aerospace repair companies have already adopted DED for restoring the critical components of aircraft engines, highlighting its strong industrial potential.

In practice, the quality evaluation and parameter setting of DED processes are often conducted by observing the morphology, geometry, and stability of the deposited tracks. However, in laser-based additive manufacturing, bead morphology is affected by numerous process parameters, resulting in high variability and challenging quality assessment. Aside from requiring significant operator expertise, the clarity of captured images is also highly dependent on imaging conditions. Traditionally, bead inspection has relied on microscopy or other metrology tools to capture weld bead appearance images and cross-sectional features, which are then analyzed manually. Such approaches are labor-intensive, heavily dependent on subjective interpretation, and difficult to standardize, thereby reducing the reliability of quality assessment. With the maturation of digital image processing technologies, process characterization in AM can now be accelerated through automated feature extraction and parameter optimization. The integration of image processing methods to automatically capture bead morphology and cross-sectional features has therefore become a critical research direction for improving process efficiency and consistency.

In previous studies on welding inspection, various image-based approaches were explored. For example, Lee *et al.* proposed an image-projection-based algorithm to analyze bead brightness distribution for quality classification.[1] Lei *et al.* developed a row–column gray-scale segmentation method to extract multilayer bead cross-sectional features such as perimeter and area.[2] Rebouças Filho *et al.* applied the active contour method to estimate dilution during welding.[3] However, these traditional methods often suffer from limitations in robustness and generalization owing to their dependence on feature design and imaging conditions. Active contour approaches, for instance, are highly sensitive to initialization and image quality, and their computational inefficiency constrains practical deployment. The emergence of deep-learning-based semantic segmentation has provided new opportunities for robust and automated welding image analysis. By learning discriminative features directly from data, deep learning methods have outperformed traditional image processing techniques in both accuracy and efficiency. Consequently, conventional contour-based approaches are being replaced by deep learning segmentation models, among which U-Net (Convolutional Networks for Biomedical Image Segmentation) and HRNet (High-Resolution Network) are the most frequently applied.[4–7] For instance, He *et al.*[7] introduced an HRNet-based method for weld cross-sectional feature detection, incorporating curve fitting and geometric transformations to quantify weld quality indicators with sub-millimeter precision. Similarly, Long *et al.*[8] proposed a fully convolutional network (FCN) for semantic segmentation, enabling the pixel-level

classification of welding images. Knaak *et al.*[9] applied CNN-based semantic segmentation to extract keyhole area, melt pool size, and weld width in laser welding, demonstrating the capability of deep learning to replace traditional methods with significantly improved accuracy.

In terms of U-Net applications, Cai *et al.*[10] developed a U-Net-based semantic segmentation method to extract melt pool and keyhole contours while mitigating noise interference. Jin *et al.*[11] combined contrast-limited adaptive histogram equalization with U-Net to predict bead geometry and location. Shang *et al.*[12] proposed AM U-Net, integrating wavelet depthwise separable convolutions to reduce model complexity while preserving detail extraction capability for bead appearance analysis. Yang *et al.*[13] further introduced CvT-UNet, integrating Transformer modules with CNNs to enhance adaptability under varying contrast conditions. Collectively, these studies confirm that U-Net and HRNet architectures are widely adopted in weld bead image analysis and are promising for parameter optimization. For metal additive manufacturing, applications of U-Net in DED remain limited. Yang *et al.*[14] introduced a modified U-Net for monitoring weld pool morphology, yet the focus was restricted to melt pool segmentation. Shah *et al.*[15] employed U-Net for analyzing track geometry in laser powder bed fusion (LPBF) rather than DED. To date, no one has specifically investigated the application of U-Net to L-DED bead cross-sectional geometry analysis. Motivated by this gap, we developed a U-Net-based semantic segmentation model to extract both weld bead appearance images and cross-sectional bead features in L-DED, supported by a geometry measurement module for quantitative evaluation.

In summary, image acquisition, processing, and analysis have become essential tools for improving process quality in laser-based AM. Given the significant variability of bead morphology under different process parameters and imaging conditions, automated and robust segmentation methods are critical for ensuring reliable measurements. As shown in previous studies, U-Net and its variants have demonstrated strong capability for welding image analysis. Accordingly, we established an automated measurement framework that standardizes weld bead geometry measurement and developed deep-learning-based segmentation models for both appearance and cross-sectional features.

The main contributions of this paper can be summarized as follows.

1. Construction of a standardized weld bead image dataset: Both weld bead appearance and cross-sectional images were acquired, preprocessed, and annotated, providing a reliable foundation for geometric feature measurement in laser-directed energy deposition (LDED).
2. Development of a dual-model deep learning framework: Separate architectures were designed for weld bead appearance and cross-sectional images, including a standard U-Net and a VGG16-based U-Net, with comparative analysis of their feature recognition accuracy and performance.
3. Demonstration of practical benefits: Independent dataset evaluation and efficiency assessment confirmed that the proposed system reduces measurement time by 91.6% while maintaining high accuracy, thereby minimizing human error and improving stability, highlighting its potential for real-world deployment.

The remainder of this paper is organized as follows. In Sect. 2, we describe the geometric characteristics of the deposited weld beads and the corresponding measurement methods. In

Sect. 3, we introduce the proposed automated image-based feature measurement system for weld bead analysis. In Sect. 4, we present and discuss the experimental results, and in Sect. 5, we conclude the paper.

2. Geometric Features and Measurement of Weld Beads

Figure 1 shows the schematic of the LDED process, in which a laser serves as the heat source. The system consists of a laser module, a metal powder feeding system, a coaxial nozzle, and a motion platform. Since the choice of powder materials in LDED generally depends on the intended application—particularly with Ni-based superalloys widely adopted in aerospace and energy components—in this study, we employed Inconel 625 as the experimental material to ensure consistency with mainstream industrial practice and relevance to practical engineering applications. During processing, metal powders are carried by shielding gas (e.g., argon) and delivered through the coaxial nozzle into the focal region, where they are instantaneously melted by the laser beam to form a molten pool on the substrate surface. With the relative motion between the laser head and the motion platform, the molten pool advances continuously, generating a weld bead (deposition track) with a distinct geometric profile. The morphology and geometry of the weld bead are strongly affected by process parameters such as laser power, powder feed rate, and scanning speed, often resulting in significant variation in both external appearance and internal cross-sectional features. Consequently, weld bead characterization not only determines deposition quality but also reflects the impact of parameter selection on overall build performance. In this study, weld bead images are categorized into weld bead appearance images and cross-sectional images, which serve as inputs for geometric feature measurement and semantic segmentation model training. To enable effective model development, the geometric features are first clearly defined and their relationships with process parameters and bead quality are analyzed.

This section is organized as follows. In Sect. 2.1, we define the weld bead features; in Sect. 2.2, we describe the measurement equipment; in Sect. 2.3, we introduce the image preprocessing methods; and in Sect. 2.4, we outline the dataset construction procedure.

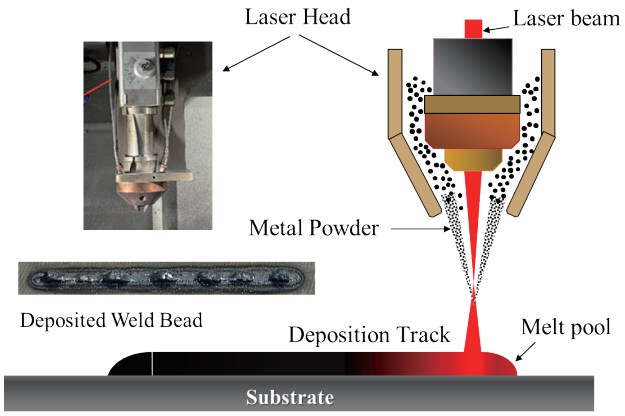


Fig. 1. (Color online) Schematic diagram of deposition process in LDED.

2.1 Definition of weld bead features

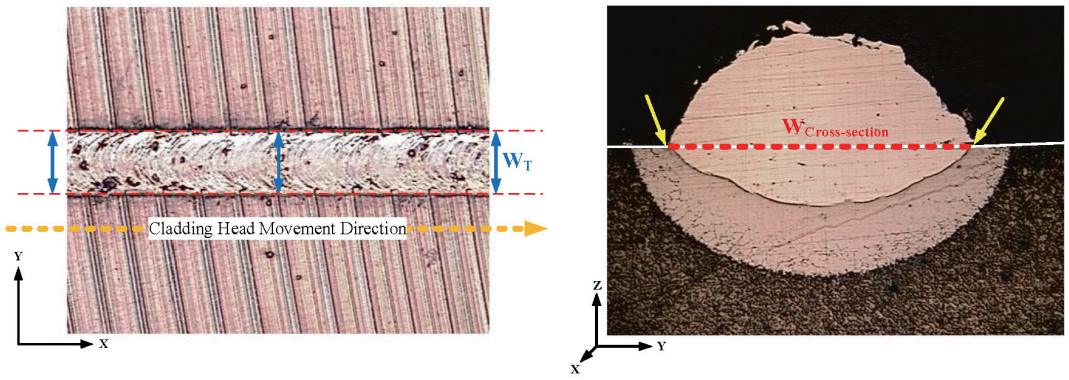
The weld bead features investigated in this study include the appearance bead width, cross-sectional width, deposition region (DR), build-up region (BR), and penetration region (PR). Figure 2 shows the definition of the appearance and cross-sectional bead widths. As shown in Fig. 2(a), the laser head moves along the X-axis (yellow dashed arrow), forming a continuous bead. The bead expands laterally in the Y-axis direction, perpendicular to the processing path, and its maximum lateral extent is defined as the appearance bead width. During manual annotation, the bead boundaries are first identified (red dashed lines), and the appearance width (W_T) is calculated as the distance between the midpoints of the two boundary lines (blue double-headed arrow).

In Fig. 2(b), the cross-sectional width is determined from the bead cross-sectional image. Specifically, the deposition region (brighter central area) is identified, and its intersection with the substrate baseline (white solid line) is located (yellow arrows). The distance between the left and right intersection points defines the cross-sectional width (Wcross-section, red dotted line), which also provides the basis for calculating the deposition area.

Figure 3 further shows the subdivision of the deposition region. In Fig. 3(a), the red area corresponds to the DR, bounded by the previously defined cross-sectional width. This region can be further divided into the BR (green) and the PR (blue), as shown in Fig. 3(b).

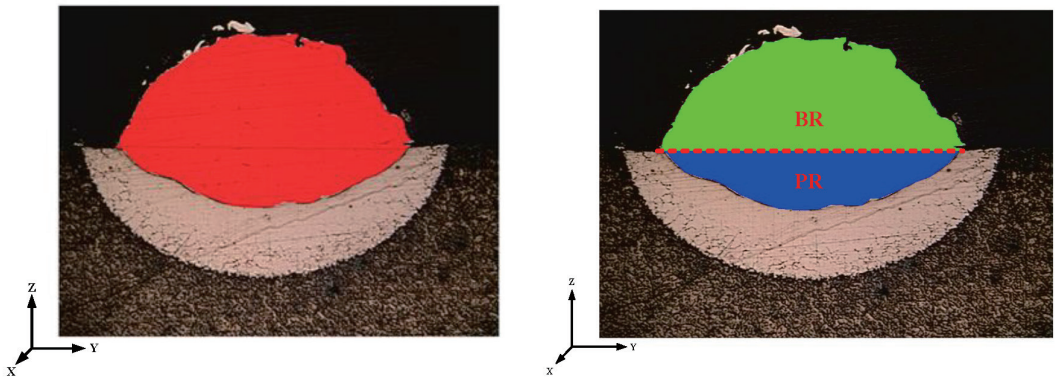
By computing the respective areas of DR, BR, and PR, the dilution (D_C) can be derived to evaluate powder accumulation and penetration behavior. The dilution is calculated according to the definition adopted by Li *et al.*[16] and Pasarkar and Balaguru.[17] Specifically, it is defined as the ratio of the penetration region area (A_{PR}) to the total deposition area ($A_{BR} + A_{PR}$). The dilution is expressed as

$$D_C\ (\%) = \frac{A_{PR}}{A_{BR} + A_{PR}}\,. \tag{1}$$



(a) (b)

Fig. 2. (Color online) Geometric features of weld bead images: (a) appearance bead width and (b) cross-sectional bead width.



(a) (b)

Fig. 3. (Color online) Cross-sectional area feature calculation: (a) deposition region (DR) and (b) build-up region (BR) and penetration region (PR).

2.2 Imaging equipment for weld bead acquisition

Weld bead images were acquired from metallic specimens produced by the LDED process. The imaging system employed was a high-magnification digital microscope (MAX-4530, Max-Plus Optical Measurement System), shown in Fig. 4 with its specifications listed in Table 1. Both weld bead appearance images and cross-sectional images were captured at a native resolution of 800 × 600 pixels and stored in BMP format to preserve full color and detail information. For appearance imaging, each single deposited track was divided into five equal segments, and the second to fourth central segments were selected for imaging. This strategy was adopted to avoid instability at the start and end regions, which may occur owing to the laser head entering and exiting the deposition path, thereby ensuring the geometric stability of the acquired images.

For cross-sectional imaging, corresponding positions were selected, and the specimens were subsequently subjected to cutting, mounting, grinding, and etching processes. These procedures enabled the visualization of the metallurgical bonding between the deposited layers and the substrate, as well as the detailed cross-sectional profiles.

2.3 Preprocessing of weld bead images

To ensure consistent input quality and minimize irrelevant artifacts, all raw images underwent preprocessing before analysis. The procedure consisted of three steps.

1. Removal of background artifacts, including white borders and automatic calibration marks from the microscope system.
2. Cropping and resizing to a standardized resolution of 640 × 480 pixels.
3. Conversion from BMP to PNG format, enabling efficient mask generation and batch processing for deep learning models.

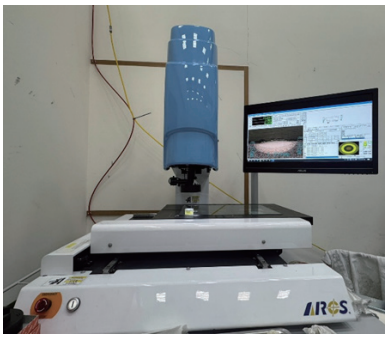


Fig. 4. (Color online) High-magnification digital microscope (MAX-4530) used for weld bead imaging.

Table 1
Specifications of MAX-4530 optical measurement system.

Item	Specification
Magnification	17.5–112.5 × coaxial optical lens
Imaging sensor	High-resolution color CCD
Linear accuracy	*X*, *Y* axes: (2 + *L*/200) μm
Repeatability	0.002 mm
Travel-X,Y axis	450 × 300 mm^2

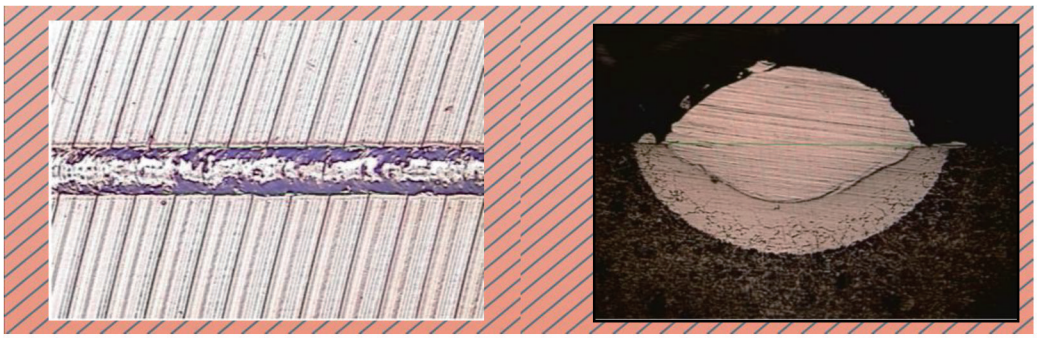
Representative examples of preprocessed images are shown in Fig. 5, where the removal of background elements allowed the model to focus exclusively on weld bead width and deposition geometry during training. After annotation, the labeled information was stored in a JavaScript Object Notation (JSON) file, which was then converted via scripting into semantic segmentation mask images in PNG format.

2.4 Dataset construction workflow

The overall workflow for dataset construction is presented in Fig. 6, which illustrates the sequential process from image acquisition to annotation and classification. After preprocessing and annotation, the dataset comprised 556 weld bead appearance images and 398 cross-sectional images, which were subsequently used for independent model training. For dataset partitioning, all images were randomly assigned to training, validation, and testing sets at a ratio of 8:1:1. No distinction was made between material types or processing groups during allocation. Original file names were retained as sample identifiers to ensure traceability, allowing direct comparison between model outputs and the corresponding geometric features of the raw images.

3. Development of Feature Measurement System

To address the challenges of weld bead geometric feature measurement—such as time-consuming procedures, high costs, significant errors, and the difficulty of defining standardized feature boundaries for recognition—in this study, we aimed to develop an automated geometric measurement system that integrates deep learning and image processing techniques. The



(a) (b)

Fig. 5. (Color online) Weld bead images after preprocessing: (a) weld bead appearance image and (b) cross-sectional image.

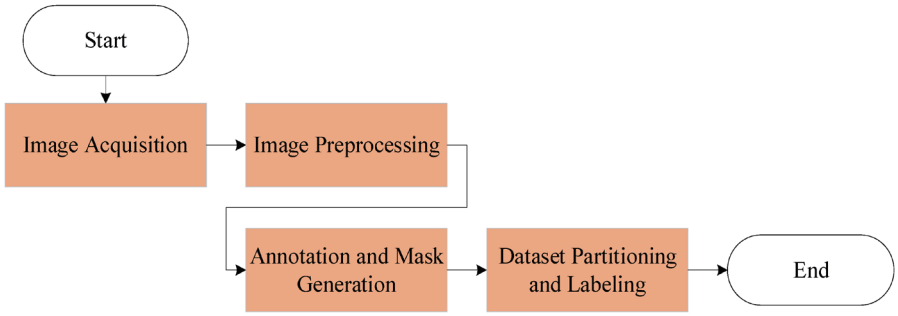


Fig. 6. (Color online) Workflow of weld bead image dataset construction.

proposed framework encompasses the following stages: raw image acquisition, semantic segmentation processing, feature mask generation, geometric measurement computation, and indicator output. The system is expected not only to reduce measurement time but also to improve measurement accuracy and consistency through the incorporation of deep learning and image processing.

3.1 System architecture

The system is composed of five major modules: (1) image acquisition and preprocessing, (2) training and inference of the semantic segmentation model (U-Net), (3) postprocessing of segmentation results, (4) geometric feature measurement, and (5) evaluation and output of measurement results. Data are exchanged between modules in a unified format to ensure system scalability and operational flexibility. As illustrated in Fig. 7, two types of weld bead images are acquired: weld bead appearance images and cross-sectional images. Corresponding segmentation masks are annotated manually prior to training. The preprocessing module then performs size normalization (resize) and format conversion (one-hot encoding or binary mask) to generate standardized input data for the model.

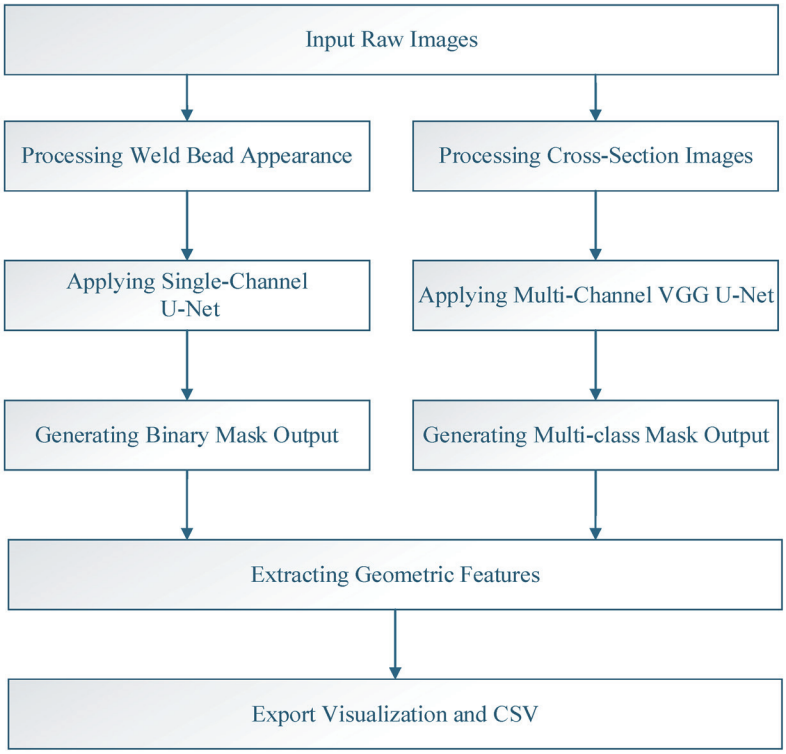


Fig. 7. (Color online) System architecture of proposed weld bead feature measurement framework.

Since image recognition and model design constitute the core of this study, different architectures are adopted in accordance with image type to enhance recognition accuracy. Table 2 summarizes the input data types, task characteristics, and model designs. The differences are described below.

1. Weld Bead Appearance: A U-Net architecture[18] is employed for binary semantic segmentation. The model outputs a single-channel mask image, where each pixel is classified as either belonging to the weld bead region (WBR) or not, using a sigmoid activation function. Because weld bead appearance images generally exhibit clear boundaries and strong contrast, a standard U-Net architecture is sufficient for effective segmentation and boundary learning.
2. Cross-sectional image recognition: For cross-sectional weld bead images, a U-Net architecture incorporating a VGG-style encoder module[19] is adopted to support multi-class semantic segmentation. The model produces a multi-channel mask, with each channel corresponding to a specific feature region, including the DR, the heat-affected zone and base material (HAZ + BM), and the background. A softmax activation function is applied to assign each pixel to the class with the highest probability. Cross-sectional weld images often present blurred boundaries, subtle texture transitions, and class imbalance, which pose challenges for accurate segmentation. To address these issues, a deeper VGG-style encoder is incorporated to enhance feature extraction capability and improve spatial representation. Moreover, the uniform 3 × 3 convolutional structure of the VGG network facilitates stable

Table 2
Comparative analysis of model tasks.

Items	Weld bead appearance image recognition	Cross-sectional image
Task type	Binary semantic segmentation	Multi-class semantic segmentation
Region annotation	WBR (single class)	DR / HAZ + BM / Background (3 classes)
Model architecture	U-Net	VGG-style U-Net
Activation	Sigmoid	Softmax
Output channels	1	3

and robust feature learning, enabling the more accurate discrimination of adjacent material regions and the effective reconstruction of fine geometric details required for precise segmentation. The predicted segmentation maps are subsequently processed by the postprocessing module, which performs geometric operations, connected-component extraction, and boundary analysis to quantify geometric features such as bead width, penetration depth, cross-sectional area, and dilution coefficient. In addition, the system supports visualization through overlay images and standardized data export formats (CSV files and mask images), enabling further statistical analysis and process parameter modeling. Overall, the proposed architecture exhibits high reusability and scalability, allowing model replacement and module retraining to accommodate different weld materials, imaging conditions, or recognition tasks. This flexibility extends the applicability of the system to a wide range of manufacturing scenarios. Within the geometric feature measurement module, the system automatically quantifies the following indicators.

(1) Weld bead appearance width, derived from the short edge of the maximum inscribed rectangle
(2) Total and subdivided areas of the cross-sectional deposition region and their ratios
(3) Other extensible geometric features such as bead height and thickness

All measurement results are consolidated and recorded through the output module, which supports both visualization and structured data export for subsequent analysis and statistical evaluation.

3.2 Loss function and evaluation metrics

In the weld bead appearance recognition task, the target feature region (e.g., the weld bead region, WBR) occupies only a very small proportion of the overall image. Consequently, foreground and background distributions are highly imbalanced in weld bead image segmentation. If the conventional binary cross-entropy loss function is employed, the model tends to underperform in recognizing these small regions. To enhance learning effectiveness for such minority areas, we adopted the Tversky loss,[20] an extension of the Dice loss, as the primary loss function. The parameters of the Tversky loss were further adjusted in accordance with the characteristics of the task. The formulation of the Tversky loss is expressed as

$$\mathcal{L}_{Tversky} = 1 - \frac{TP + \varepsilon}{TP + \alpha \cdot FP + \beta \cdot FN + \varepsilon}. \tag{2}$$

Equation (2) defines the Tversky loss, where TP denotes the number of true positives, FP and FN represent false positives and false negatives, respectively, and ε is a numerical stability constant. The parameters α and β are adjustable to control the penalty weights, thereby balancing the effect of different error types on the model. For the binary semantic segmentation model of weld bead appearance images, we set $\alpha = 0.3$ and $\beta = 0.7$, which increases the penalty on false negatives and improves the detection of small feature regions. For the cross-sectional image model, the segmentation task involves three classes with clearly imbalanced distributions, where the primary target region (Class 1) accounts for the smallest proportion. To strengthen the model's ability to detect such small regions, the same configuration ($\alpha = 0.3$, $\beta = 0.7$) is adopted under the multi-class setting, following the same design rationale as in the binary case. For comparison, the Dice loss function is also considered, and its formulation is given as

$$D_{loss} = 1 - \frac{2 \times TP}{2 \times TP + FP + FN}. \tag{3}$$

For the weld bead appearance model, we adopted a weighted combination of loss functions, integrating Dice loss and Tversky loss. The weighted scheme is defined such that Tversky loss serves as the primary component (weight = 0.9), providing the flexible penalization of false positives and false negatives, while Dice loss is incorporated as a secondary component (weight = 0.1) to preserve its ability to monitor overall overlap regions and prevent excessive bias toward class imbalance compensation. This weighted combination enhances the model's robustness and generalization capability in small-object detection tasks.

In terms of evaluation, the weld bead appearance model employs the Mean Intersection over Union (Mean IoU) metric implemented in Keras,[21] with the number of classes set to 2, to quantify the spatial overlap between predicted and ground-truth masks. For the cross-sectional model, the same evaluation function is used with the number of classes set to 3. The IoU is computed separately for each of the three classes, and the mean value is reported as the primary evaluation criterion. This design enables class-wise performance assessment and facilitates cross-validation with confusion matrices and classification reports.

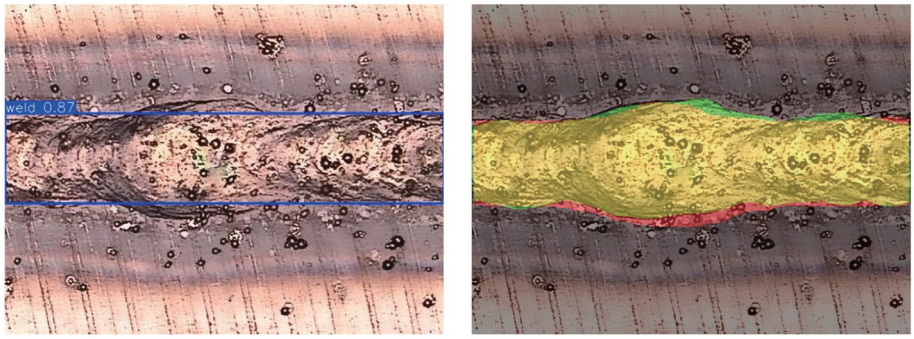
3.3 Development of feature model

3.3.1 Weld bead appearance model

For the semantic segmentation of weld bead regions in weld bead appearance images, we adopted the U-Net architecture as the primary model design. To validate the feasibility of semantic segmentation using U-Net, a comparative analysis is conducted with the YOLOv8 model,[22] which also aligns with the task orientation but employs a different detection strategy. Table 3 summarizes the key differences between YOLOv8 and U-Net across multiple aspects, while Fig. 8 illustrates the distinct inference outputs of the two models. As shown in Fig. 8(a), the YOLOv8 model provides the rapid localization of weld bead regions through bounding boxes. However, its outputs are limited to coarse rectangular boundaries, which fail to capture the fine

Table 3
Characteristics of YOLOv8 and U-Net models.

Items	YOLOv8	U-Net
Primary task orientation	Object Detection: predicting bounding boxes and class labels	Pixel-wise Semantic Segmentation: classifying each pixel
Output type	Bounding boxes + class scores	Full-image pixel labels (each pixel assigned to a class)
Pixel-level localization ability	No: provides only object bounding box regions	Yes: precise localization at the pixel level
Boundary precision	Coarse: limited to rectangular boxes, cannot capture fine contours	High precision: capable of reconstructing complex and irregular boundaries
Applicable image types	RGB / grayscale (commonly applied to RGB inputs)	RGB / grayscale, with channel number adjustable depending on the task



(a) (b)

Fig. 8. (Color online) Model outputs: (a) YOLOv8 and (b) U-Net.

details of the weld bead edges. In contrast, Fig. 8(b) presents the pixel-level mask output from the U-Net model, demonstrating its ability to precisely delineate the weld bead geometry, including irregular boundaries. While YOLOv8 provides faster inference and coarse object localization through bounding-box-based detection, it is limited in capturing fine-grained geometric features of the weld bead. Relative to YOLOv8, the U-Net model enables dense, pixel-level localization, resulting in higher boundary accuracy and more detailed shape representation. On the basis of the results of this comparison, U-Net is selected as the primary model architecture for weld bead segmentation tasks in this study.

Within the U-Net framework, different architectural variations are further explored to identify the most suitable configuration for weld bead detection. Table 4 lists the structural differences among the models examined, including variations in input channel numbers, encoder/decoder layers and configurations, and dropout settings. Models U01 and U02 adopt grayscale inputs, whereas models U03 and U04 use RGB color inputs. Compared with U01 and U03, models U02 and U04 employ deeper encoder–decoder structures with larger feature map channel capacities, theoretically offering stronger feature extraction capabilities but at the cost of increased model size and computational complexity. Additionally, the choice of dropout ratio affects the generalization ability and training stability of the models. These architectural

Table 4
Differences in model architectures for weld bead appearance.

Model ID	U01	U02	U03	U04
Input channels	Grayscale (1)	Grayscale (1)	RGB (3)	RGB (3)
Number of encoder layers	4	5	4	5
Encoder configuration	16 → 128	32 → 512	16 → 128	32 → 512
Dropout	0.3	0.5	0.3	0.5
Number of decoder layers	4	5	4	5
Decoder configuration	128 → 32	512 → 32	128 → 32	512 → 32

differences are expected to be reflected in performance variations, as measured by evaluation metrics such as IoU and Dice scores, which will be discussed in the results section.

3.3.2 Cross-sectional weld bead model

For the design of the cross-sectional weld bead model, in addition to employing the baseline U-Net architecture as Model V01, we further investigated a U-Net variant (Model V02) that integrates a VGG-based encoder.[23–25] Specifically, the encoder of Model V02 is constructed on the basis of the VGG16 architecture, as illustrated in Fig. 9.

Figure 9 shows the operational flow of the VGG16 architecture: the input image of size 224 × 224 × 3 passes through five downsampling stages before reaching the bottleneck, followed by three fully connected layers at the classification head. VGG16 consists of 13 convolutional layers (3 × 3 filters) and five max-pooling layers (2 × 2), arranged in an alternating stack. The channel dimensions progress sequentially as 64 → 128 → 256 → 512 → 512. At each pooling operation, the spatial resolution is halved while the feature dimensionality is doubled. The architecture concludes with two fully connected layers of 4,096 units each, followed by a 1,000-unit softmax layer for classification across 1,000 ImageNet categories.

Table 5 summarizes the comparative differences between the constructed models. The primary distinction lies in the depth of the encoder and bottleneck. Model V01 employs a single convolutional layer per stage (1 × Conv) and a single convolutional layer in the bottleneck [1 × Conv(512)]. In contrast, Model V02 applies three convolutional layers per stage (3 × Conv) and two convolutional layers in the bottleneck [2 × Conv(512)]. As a result, Model V02 provides higher representational capacity and a larger number of parameters but requires longer training, stronger regularization, and a larger dataset to achieve optimal performance. The decoder and skip-connection designs are identical in both models, consisting of an upsampling layer followed by a single convolutional layer (1 × Conv) concatenated with the corresponding encoder feature maps. Therefore, the performance differences between V01 and V02 can primarily be attributed to the increased depth of the encoder and bottleneck in V02.

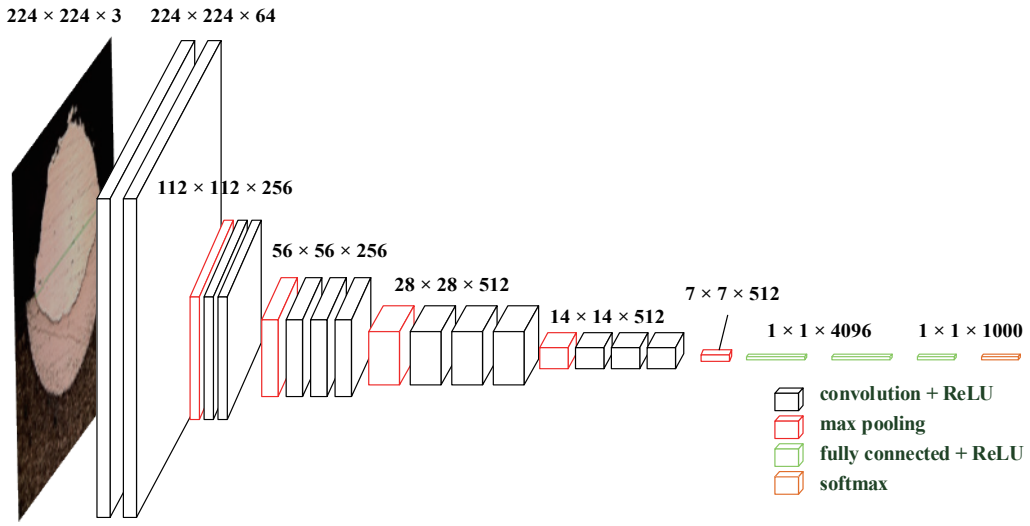


Fig. 9. (Color online) Schematic illustration of VGG16 architecture, employed as the encoder in Model V02 (adapted from Ref. 25).

Table 5
Differences between cross-sectional weld bead models V01 and V02.

Model ID	V01 (Baseline U-net model)	V02 (VGG-style model)
Encoder	1 × Conv (+ BN + ReLU): one layer for each of c1–c4	3 × Conv (+ BN + ReLU): three layers for each of c1–c4
Bottleneck	1 × Conv (512) + BN + ReLU	2 × Conv (512) + BN + ReLU
Pixel-level localization ability	Upsampling (2 × 2); after each upsampling, 1 × Conv (3 × 3) → BN → ReLU;	Upsampling (2 × 2); after each upsampling, 1 × Conv (3 × 3) → BN → ReLU;
	Channels 256 → 128 → 64 → 32	Channels 256 → 128 → 64 → 32
Boundary precision	After each upsampling, concatenated with the corresponding encoder feature maps (from encoder stage 4 down to stage 1).	After each upsampling, concatenated with the corresponding encoder feature maps (from encoder stage 4 down to stage 1)

4. Experimental Analysis and Results

4.1 Training and analysis of weld bead appearance models

The dataset for weld bead appearance model training consisted of 522 images captured under various experimental conditions. After preprocessing, four U-Net-based architectures (U01–U04) were trained to evaluate the effects of model depth, channel width, and input format. Table 6 summarizes the F1-score-based evaluation metrics of the trained models. Among the evaluated models, U02 achieved the highest Macro F1 (0.9719) and Weighted F1 (0.9795) scores, indicating superior classification performance and robustness under class-imbalanced conditions. Figure 10 further illustrates the per-image distribution of Mean IoU across the four models. U01 and U02 form the leading group, with U02 exhibiting a slightly higher median IoU, a narrower interquartile range, and a higher lower whisker (i.e., a higher minimum non-outlier value). These characteristics indicate a more concentrated performance distribution and improved control over

Table 6
Evaluation metrics for weld bead appearance models.

Model ID	U01	U02	U03	U04
Macro avg. F1	0.9703	0.9719	0.9666	0.9622
Weighted avg. F1	0.9785	0.9795	0.9759	0.9724

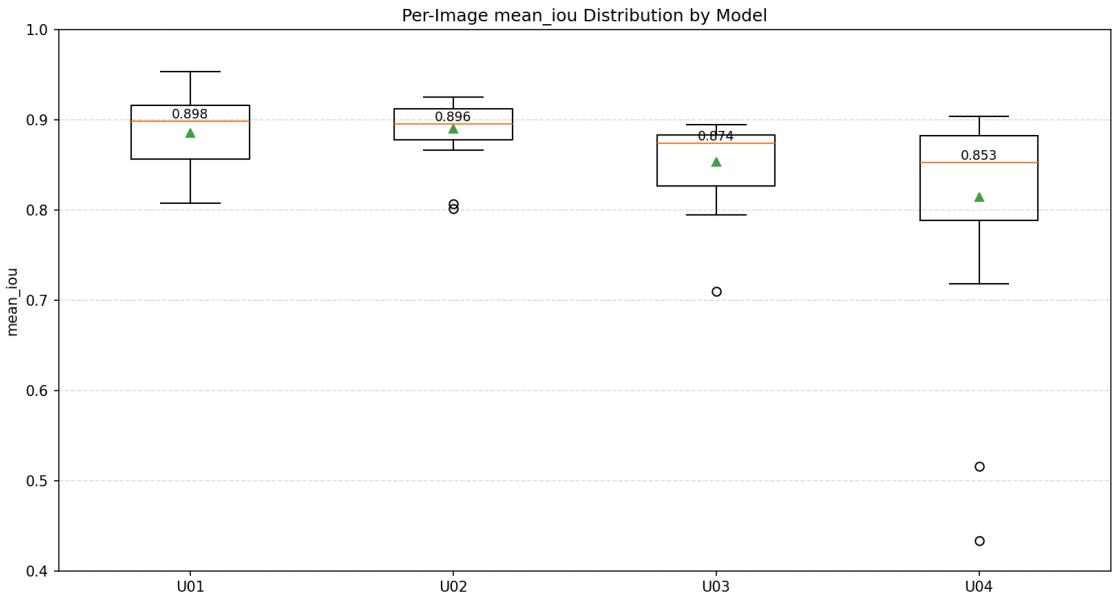


Fig. 10. (Color online) Boxplot comparison of validation metrics for weld bead appearance models (U01–U04).

low-score cases. In contrast, U03 shows wider score variation, while U04 presents noticeable outliers, suggesting less stable segmentation performance.

From an architectural perspective, both U02 and U04 use deeper encoder–decoder networks (five layers) with larger channel widths (32 → 512) and higher dropout (0.5), theoretically promoting richer feature extraction. However, U02 uses single-channel grayscale input, reducing redundancy and enabling more efficient learning with the deeper structure. U04 relies on RGB input, which may introduce feature redundancy given the limited dataset size, thereby affecting generalization. U01 and U03, based on shallower four-layer architectures, differ mainly in input format; U01 (grayscale) slightly outperforms U03 (RGB), suggesting that network depth and input characteristics must be aligned for optimal performance. Overall, U02 is the preferred model for appearance segmentation because of its superior accuracy and stable distribution.

4.2 Training and analysis of weld cross-sectional models

The cross-sectional dataset consisted of 398 images collected under different experimental conditions. After preprocessing, the dataset was used to train two models: V01, a baseline U-Net with a conventional encoder–decoder architecture, and V02, a U-Net variant incorporating a

VGG16 encoder. Both models were trained using identical inputs, loss functions, and output configurations. RGB images were retained during training to preserve class-discriminative color and intensity information between the deposition region and the surrounding heat-affected and base material regions, as such information contributes to accurate region separation in cross-sectional images. Consequently, conversion to grayscale was avoided to prevent the loss of color-related features relevant to the segmentation task. The segmentation task involved three classes: background (Class 0), deposition region (DR, Class 1), and heat-affected zone combined with base material (HAZ + BM, Class 2). The primary focus of this study was the accurate identification of the DR region, with HAZ and BM grouped together to simplify the classification problem.

Table 7 presents the evaluation results. Overall, V02 outperformed V01 in both macro- and weighted-average F1 scores, confirming its superior ability to balance minority and majority classes while maintaining high overall accuracy. Although V01 yielded slightly lower scores, the difference was minor, and its performance remained stable, making it a conservative yet reliable choice. The architectural distinction between the models explains the performance gap. V02, based on VGG16, applies three convolutional layers per encoder stage (3 × Conv + BN + ReLU) and two convolutional layers at the bottleneck [2 × Conv(512)]. This deeper design enables more comprehensive feature extraction and captures high-level feature relationships. In contrast, V01 adopts the original U-Net encoder with only one convolutional layer per stage and a single bottleneck convolution [1 × Conv(512)], resulting in a shallower structure. While less expressive, V01 offers lower computational cost and reduced data requirements. Both models share identical decoder designs, using 2× upsampling followed by 1 × Conv(3 × 3) + BN + ReLU, with channel reductions of 256→128→64→32 and standard skip connections.

In summary, V02 is the most accurate and stable model for cross-sectional segmentation owing to its deeper encoder and enhanced feature representation. V01 remains a computationally efficient option for resource-constrained deployment scenarios.

4.3 Independent dataset-based evaluation

To evaluate the generalization capability of the proposed models, additional datasets independent of the training and testing sets were employed for inference and evaluation. The test data consisted of weld bead appearance and cross-sectional images obtained under different experimental conditions, none of which had appeared in the training process, thereby ensuring a rigorous assessment of model robustness and generalizability. In total, 25 weld bead appearance images and 34 cross-sectional images were collected as the benchmark dataset. Only basic cropping and resizing were applied to meet the input requirements of the models. The test data

Table 7
Evaluation metrics for weld cross-sectional models.

Model ID	V01	V02
Macro avg. F1	0.996	0.997
Weighted avg. F1	0.995	0.997

were manually annotated in accordance with the defined regions to enable the calculation of accuracy metrics.

4.3.1 Validation of weld bead appearance models

For the validation of the weld bead appearance models, four U-Net variants (U01–U04) were evaluated using an independent test dataset. For each image, pixel-level prediction accuracy in the semantic segmentation task was assessed using the IoU and F1-score, which were further summarized as Mean IoU, Macro F1, and Weighted F1 metrics. As shown in Table 8, the averaged evaluation on the external test set indicates that Model U02 achieved Mean IoU, Macro F1, and Weighted F1 scores of 0.8903, 0.9407, and 0.9580, respectively, reflecting consistent segmentation performance in terms of region overlap and pixel-level classification. In addition, Fig. 10 illustrates the per-image distribution of Mean IoU across models. In the box plots, the orange line denotes the median, the green triangle indicates the mean, the box represents the interquartile range (IQR, 25%–75%), and the whiskers and black dots correspond to the non-outlier range and outliers, respectively. The results show that Models U01 and U02 form the leading group, with U02 exhibiting the most favorable overall performance. Specifically, U02 achieved the highest mean and median scores, exhibited a narrower IQR, and showed a higher lower whisker, indicating a more concentrated distribution and better control of low-score cases, thereby confirming its stability. By contrast, U03 demonstrated a wider performance distribution with occasional low-score samples, while U04 exhibited noticeable outliers, indicating less stable segmentation behavior. Taken together, these results establish U02 as the most accurate and stable model in external testing, making it the preferred choice for subsequent inference and deployment.

4.3.2 Validation of weld cross-sectional models

For the validation of cross-sectional models, the independent test dataset was again employed to evaluate the two developed models (V01 and V02). The results are summarized in Table 9 and illustrated in Fig. 11.

As shown in Table 9, Model V02 achieved scores of 0.9435, 0.9709, and 0.9715 for Mean IoU, Macro F1, and Weighted F1, respectively, indicating robust and high-level accuracy in both overlap precision and pixel-level classification. In contrast, Model V01 exhibited significantly weaker overall performance, underscoring the superiority of V02 in both accuracy and consistency.

Table 8
Validation results for weld bead appearance models.

Model ID	U01	U02	U03	U04
Mean IoU	0.8859	0.8903	0.8536	0.8147
F1-score (macro)	0.9379	0.9407	0.9187	0.8912
F1-score (weighted)	0.9564	0.9580	0.9426	0.9146

Table 9
Validation results for weld cross-sectional models.

Model ID	V01	V02
Mean IoU	0.5154	0.9435
F1-score (macro)	0.6705	0.9709
F1-score (weighted)	0.6283	0.9715

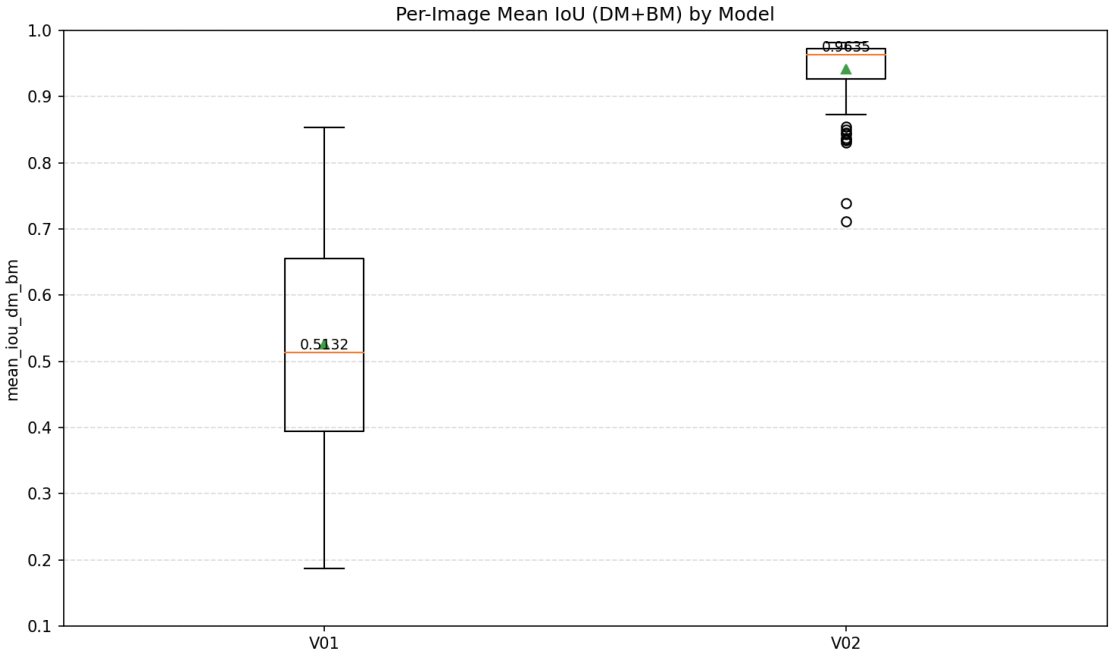


Fig. 11. (Color online) Boxplot of validation metrics for cross-sectional weld bead models (V01–V02).

Figure 11 presents the per-image Mean IoU distribution on the independent test set (considering only the DR and BM classes). The orange line represents the median, the green triangle indicates the mean, the box body corresponds to the interquartile range (IQR, 25th–75th percentiles), and the whiskers and black dots denote the non-outlier range and outliers, respectively. The results demonstrate that Model V02 was consistently clustered in the high-score range, with a median of approximately 0.9635 and a mean close to the median. The narrow IQR and the higher lower whisker further confirm that most samples achieved stable segmentation overlaps close to the ground truth annotations. By comparison, Model V01 exhibited a median of only about 0.5132, a much wider distribution, and numerous low-scoring samples, reflecting substantial variability and insufficient stability across different cases. Overall, the combined analysis of both central tendency and distribution confirms that Model V02 achieves higher accuracy and greater stability in Mean IoU, thereby making it the preferred choice for subsequent inference and deployment.

4.4 Evaluation of inspection efficiency

In the preceding sections, we demonstrated the generalization ability and geometric measurement accuracy of the proposed models, emphasizing that appropriate model selection can effectively reduce both measurement errors and processing time. To further compare the efficiency of automated inference and manual measurement, a small-scale experiment was conducted to evaluate differences in time and cost. The analysis considered the total processing time from image acquisition to measurement completion, as well as the uncertainty associated with manual operations. Table 10 summarizes the differences between the conventional manual inspection workflow and the proposed model-driven automated inspection.

The results show that the traditional method requires an average of approximately 15 s per image for storage, feature annotation, and data integration, with manual annotation being the most time-consuming step. In addition, manual measurement introduces significant variability due to operator subjectivity. In contrast, the automated inspection process requires only 1.263 s per image (0.677 s for feature annotation and 0.586 s for measurement computation), reducing the overall processing time by about 91.6% and saving an average of 13.737 s per image.

In terms of measurement quality, the manual workflow is prone to operator-dependent variations, which tend to accumulate as dataset size increases. By comparison, the automated approach ensures consistent measurement conditions through model-based computation and remains unaffected by dataset scale. Moreover, the automated system not only reproduces traditional measurements of bead width and cross-sectional height/depth, but also enables the automatic calculation of cross-sectional area and dilution coefficient, thereby providing more comprehensive measurement indicators.

In summary, the proposed automated inspection framework not only reduces processing time and improves overall efficiency but also minimizes human-induced variability, thereby

Table 10
Summary of inspection methods.

Category	Traditional inspection	Proposed model-driven automated inspection
Workflow	Multiple manual steps; hand measurement and logging	Model-driven inference with automated steps; minimal human intervention
Image storage	~ 10 s per image	~ 10 s per image
Feature processing/ annotation	~1 s per image (manual or semi-automatic)	0.677 s per image (automatic)
Quantity/count statistics	Handled by ad-hoc operations; difficult to quantify consistently.	0.586 s per image (automatic)
Integration & recordkeeping	~5 s per image (manual entry and merging)	Exported by system; not timed
Measurement quality	Influenced by human variations; lower stability	Computed by model under fixed conditions; high consistency
Measurement items	Bead width; cross-sectional height/depth	Bead width; cross-sectional height/depth; cross-sectional area; dilution coefficient
Overall	Time-consuming; results vary by operator and dataset size.	Faster and more stable; minimally affected by dataset size

demonstrating strong potential to replace manual workflows in quality assessment for laser-based additive manufacturing.

5. Conclusions

In this study, we developed a deep-learning-based automated inspection framework for weld bead analysis in LDED processes. Using U-Net and its variants, the proposed system accurately segmented and quantified both weld bead appearance and cross-sectional geometric features. Experimental results demonstrated that the U-Net model incorporating a VGG16 encoder achieved the best performance, while the automated workflow improved efficiency by 91.6% compared with manual inspection. In addition to substantially reducing measurement time, the proposed system also overcomes several inherent limitations of traditional measurement methods, such as strong operator dependence, inconsistent repeatability, and high sensitivity to illumination and surface variations. By providing stable and high-precision geometric measurements while reducing the subjectivity associated with manual interpretation, the framework effectively minimizes measurement errors and operator-induced variability. Consequently, it delivers consistent and reliable evaluation results and establishes a practical and extensible solution suitable for deployment in real-world additive manufacturing environments.

Acknowledgments

The authors would like to thank the National Science and Technology Council of the Republic of China, Taiwan, for supporting this research under Grant Nos. NSTC 113-2622-E-992-010 and NSTC 114-2622-E-992-004. The authors also gratefully acknowledge the support of Mr. Jui-Hsiung Yen, Chairman, and Ms. Lu Yen, Vice President, of Tongtai Machine & Tool Co., Ltd.

References

1 J. Lee, H. Choi, and J. Kim: Electronics **12** (2023) 2523. https://doi.org/10.3390/electronics12112523
2 T. Lei, S. Gong, and C. Wu: Mater. **17** (2024) 4683. https://doi.org/10.3390/ma17194683
3 P. P. Rebouças Filho, T. S. Cavalcante, V. H. C. de Albuquerque, J. M. R. S. Tavares, and P. C. Cortez: Proc. 2nd South-East European Conf. Computational Mechanics: An IACM–ECCOMAS Special Interest Conf., Rhodes, Greece (2009) 94–100.
4 J. Lu, S. Yamane, W. Wang, N. Li, S. Wang, and Y. Xia: Appl. Sci. **15** (2025) 1814. https://doi.org/10.3390/app15041814
5 H. I. Lukács, B. Z. Beregi, B. Porteleki, T. Fischl, and J. Botzheim: Sci. Rep. **15** (2025) 2156. https://doi.org/10.1038/s41598-025-90355-3
6 Y. He, R. Cai, F. Dai, Z. Yu, Y. Deng, J. Deng, Z. Wang, G. Ma, and W. Zhong: J. Manuf. Process. **131** (2024) 2275. https://doi.org/10.1016/j.jmapro.2024.11.030
7 W. He, A. Zhang, and P. Wang: Machines **11** (2023) 335. https://doi.org/10.3390/machines11030335
8 J. Long, E. Shelhamer, and T. Darrell: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR) (IEEE, 2015) 3431–3440
9 C. Knaak, G. Kolter, F. Schulze, M. Kröger, and P. Abels: Appl. Mach. Learn. **11139** (2019) 10. https://doi.org/10.1117/12.2527890
10 W. Cai, P. Jiang, L. Shu, S. Geng, and Q. Zhou: J. Manuf. Process. **76** (2022) 695. https://doi.org/10.1016/j.jmapro.2022.06.042.
11 G. S. Jin, S. J. Oh, Y. S. Lee, and S. C. Shin: Appl. Sci. **11** (2021) 12051. https://doi.org/10.3390/app112412051

12 F. Shang, H. Chen, B. Sun, Y. Han, D. Cai, S. Wang, T. Gao, C. Ji, and H. Zhang: Eng. Res. Express **1** (2025) 015555.
13 L. Yang, H. Wang, W. Meng, and H. Pan: Heliyon **10** (2024) e34738.
14 J. Yang, G. Liu, W. Zhu, Y. Zhang, W. Zhou, D. Liu, and Y. Lin: Addit. Manuf. Front. **4** (2025) 200199. https://doi.org/10.1016/j.amf.2025.200199.
15 A. Shah, R. Weissbach, D. A. Griggs, A. J. Hart, E. Ertekin, and S. Tawfick: arXiv:2409.18326 (2024).
16 Z. Li, S. Sui, X. Ma, H. Tan, C. Zhong, G. Bi, A. T. Clare, A. Gasser, and J. Chen: Int. J. Mach. Tools Manuf. **181** (2022) 103942.
17 A. Pasarkar and S. Balaguru: Nucl. Eng. Technol. **57** (2025) 103185.
18 N. Siddique, S. Paheding, C. P. Elkin, and V. Devabhaktuni: IEEE Access **9** (2021) 82031.
19 K. Simonyan and A. Zisserman: Proc. Int. Conf. Learn. Representations (ICLR) (2015).
20 S. S. M. Salehi, D. Erdogmus, and A. Gholipour: Proc. Machine Learning in Medical Imaging (MLMI), MICCAI 2017, Lecture Notes in Computer Science, Vol. 10541 (Springer, 2017) 379–387. https://doi.org/10.1007/978-3-319-67389-9_44
21 Keras - Deep Learning for humans: https://keras.io (accessed July 9, 2025).
22 M. Yaseen: arXiv:2408.15857 (2024).
23 K. Simonyan and A. Zisserman: arXiv:1409.1556 (2015).
24 S. Mascarenhas and M. Agarwal: Proc. 2021 Int. Conf. Disruptive Technol. Multi-Disciplinary Res. Appl. (CENTCON) (2021) 96–99.
25 G. Learning, Everything you need to know about VGG16 https://medium.com/@mygreatlearning/everything-you-need-to-know-about-vgg16-7315defb5918 (accessed July 9, 2025).

About the Authors

Ke-Han Su received his Ph.D. degree in electrical engineering from National Cheng Kung University, Tainan, Taiwan, in 2008. He is currently an assistant professor with the Department of Electrical Engineering, National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan. His research interests are in motion control, fuzzy control, metal additive manufacturing, and real-time CNC interpolation. (kehansu@nkust.edu.tw)

Sung-Chu Wang received his B.S. degree in electrical engineering from Cheng Shiu University, Kaohsiung, Taiwan, in 2020, and his M.S. degree in electrical engineering from National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan, in 2025. His research interests include motion control, fuzzy control, and metal additive manufacturing. (f112111114@nkust.edu.tw)

Yi-Jun Chen received his B.S. degree in electrical engineering from Southern Taiwan University of Science and Technology, Tainan, Taiwan, in 2023, and his M.S. degree in electrical engineering from National Kaohsiung University of Science and Technology, Kaohsiung, Taiwan, in 2025. His research interests include motion control, fuzzy control, and metal additive manufacturing. (f112111121@nkust.edu.tw)

Keng-Pin Chang received his B.S. degree in mechanical engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2009, and his M.S. degree in mechanical engineering from National Cheng Kung University, Tainan, Taiwan, in 2011. He is currently an engineer with Tongtai Machine & Tool Co., Ltd., Taiwan. His research interests include CNC applications, mechanism design, and metal additive manufacturing.
(brentchang@tongtai.com.tw)

Haw-Ching Yang received his B.S. degree in engineering science from National Cheng Kung University, Tainan, Taiwan, in 1989, his M.S. degree in control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1991, and his Ph.D. degree from the Institute of Manufacturing Engineering, National Cheng Kung University, in 2003. Since September 2004, he has been an assistant professor with the Institute of System and Control, National Kaohsiung First University Science and Technology, Kaohsiung, Taiwan. In 2013, he became an associate professor with the Department of Electrical Engineering, National Kaohsiung University Science and Technology, Kaohsiung. He is currently a full professor. His research interests include intelligent manufacturing and intelligent methods for industrial, additive manufacturing, and cloud-based applications. (hao@nkust.edu.tw)