Improving the Boundary Recognition Performance of Visual Elements on Product Packaging by Integrating the Image Segmentation Mechanism of SAM-Adapter

Cuihua Meng[1], Xiaoming Huang[2], Yuting Li[1]
[1]College of Art and Design, Guangdong University of Finance & Economics, Guangzhou 510320, Guangdong, China
[2]Guangzhou Academy of Fine Arts, Guangzhou 510006, Guangdong, China
Corresponding author: Yuting Li (e-mail: x20250608@126.com)

ABSTRACT Accurate boundary recognition under high-noise, complex-texture, and reflective conditions remains a fundamental challenge for intelligent visual perception systems and is equally important for electromagnetic imaging and radar-based target analysis that require reliable feature localization. This study presents a boundary-aware segmentation framework by integrating a gradient-guided Adapter mechanism into the Segment Anything Model (SAM). Pixel-level boundary response maps generated by the Laplacian of Gaussian operator are combined with gradient magnitude and directional sensitivity weights to guide adaptive feature calibration, while boundary-aware cross-entropy loss and Hausdorff distance constraints jointly optimize parameter learning and geometric consistency. The proposed strategy enhances multi-scale boundary representation without sacrificing the lightweight characteristics of Adapter-based fine-tuning. Experimental results demonstrate that the method achieves a Boundary Intersection over Union (BIoU) of 0.923 under 20% noise interference and maintains a contour F1 score of 0.951 under 20% reflectivity, significantly outperforming existing comparison methods in boundary continuity and localization accuracy. The framework provides an effective solution for high-precision image segmentation and offers potential value for electromagnetic imaging, microwave sensing, and intelligent perception systems requiring robust boundary extraction under complex interference conditions.

INDEX TERMS image segmentation, boundary-aware adapter, segment anything model, electromagnetic imaging, boundary recognition

I. INTRODUCTION

PRODUCT packaging, serving as the core carrier of commodity information, plays a crucial role in e-commerce platforms and automated quality inspection systems [1,2]. The accuracy of visual element boundary recognition on product packaging directly determines the reliability of packaging design optimization and anti-counterfeiting analysis [3,4]. The application demand of image segmentation technology in the field of industrial vision continues to grow, but research on dedicated methods for high-complexity scenarios of product packaging is still insufficient, and it is urgent to break through the bottleneck of boundary ambiguity to support the actual deployment of intelligent packaging inspection systems [5-7].

Product packaging images commonly suffer from high noise interference, overlapping printing textures, and material reflectivity [8,9]. These characteristics cause existing segmentation models to produce blurred edges and discontinuous contours when locating boundaries [10,11]. Visual elements such as text labels and barcodes often exhibit positioning deviations in sub-pixel segmentation due to low contrast and local occlusion [12,13]; while changes in ambient lighting further exacerbate the boundary continuity problem, affecting the accuracy of automated analysis [14,15].

Full model fine-tuning of SAM requires a large amount of labeled data and has a high computational cost in the packaging scenario [16]. Although the lightweight Adapter method reduces the amount of parameter updates, its weight allocation mechanism treats boundary and non-boundary regions equally and cannot specifically strengthen the expression of boundary features, resulting in insufficient boundary sensitivity [17]. Improved schemes based on attention mechanisms are prone to over-segmentation in complex texture regions and cannot effectively suppress false detections caused by reflective interference [18-20]. Existing methods all face the core challenge of maintaining boundary continuity and positioning accuracy under high noise conditions.

To address the shortcomings of the general SAM model in adapting to the high noise and reflective characteristics of visual elements on product packaging, this paper innovatively integrates boundary prior knowledge and the Adapter mechanism to propose a gradient-guided adaptive boundary segmentation framework. After extracting multi-scale feature maps from the output layer of the SAM's ViT (Visual Transformer) encoder, the LoG (Laplacian of Gaussian) operator is applied to generate pixel-level boundary response maps. Gradient magnitude and orientation sensitivity weight matrices are dynamically calculated using a differentiable function, and these weights are embedded into the Adapter linear transformation layer to achieve feature calibration. Finally, the parameter update and optimization fine-tuning process is achieved by combining the boundary-aware cross-entropy loss function and HD constraints. This method retains the model's generality while accurately enhancing the feature representation capability of the boundary regions of packaging visual elements. This research achieves the organic integration of boundary prior knowledge and deep learning models, significantly improving the sub-pixel-level boundary recognition accuracy of packaging visual elements, a methodology highly relevant for the textile industry in achieving precise pattern segmentation and flaw localization in fabrics at a comparable resolution.

II. BOUNDARY-AWARE ADAPTER MECHANISM

A. BOUNDARY RESPONSE MAP GENERATION AND GRADIENT WEIGHT EMBEDDING MECHANISM

Figure 1 illustrates the boundary response map generation and gradient weight calculation process. The multi-scale feature map output by the SAM encoder is processed by the LoG operator to generate the boundary response map. This response map is then input into the gradient magnitude weight calculation and orientation sensitivity weight calculation modules, respectively. The two types of weights are fused to form a gradient sensitivity weight matrix. This matrix participates in the generation of the boundary adaptive transformation matrix and is finally embedded in the Adapter linear transformation layer to complete feature calibration.

After extracting multi-scale feature maps from the output layer of the ViT encoder, the LoG operator is applied [21]. Through depthwise convolution with channel-wise processing followed by weighted aggregation of boundary responses. The same 2D LoG kernel is independently applied to each channel to generate channel-specific boundary response maps, which are aggregated using learnable weights. This implementation preserves spatial boundary information while leveraging multi-scale semantic features from the ViT encoder containing richer spatial details at multiple resolutions critical for accurate boundary detection under high noise conditions. The feature tensor output by the SAM encoder is represented as $F \in \mathbb{R}^{H\times W\times C}$, where H, W, and C correspond to the height, width, and number of channels of the feature map, respectively [22]. The boundary response map $B \in \mathbb{R}^{H\times W}$ is obtained through convolution operation, and its mathematical expression is:

$$B(x, y) = \nabla^2 G_\sigma(x, y) * F(x, y) \tag{1}$$

∇^2 represents the LoG operator; $G_\sigma(x, y)$ is a Gaussian kernel function with standard deviation σ; $*$ represents the convolution operation; x and y are the pixel coordinates. The boundary response map is normalized to obtain the boundary weight map $W_b = (B - \min(B))/(\max(B) - \min(B))$, where B is the boundary response map and $W_b \in [0, 1]^{H\times W}$. The gradient magnitude weight W_m and direction sensitivity weight W_d are calculated using a differentiable boundary response function as input to the boundary weight map. The gradient magnitude weight is calculated based on the local gradient magnitude of the boundary response map.

$$W_m(x, y) = \frac{1}{1 + e^{-\lambda\cdot\sqrt{\left(\frac{\partial B}{\partial x}\right)^2 + \left(\frac{\partial B}{\partial y}\right)^2}}} \tag{2}$$

λ is a hyperparameter controlling gradient sensitivity, and $\partial B/\partial x$ and $\partial B/\partial y$ represent the partial derivatives of the boundary response map in the horizontal and vertical directions, respectively. Gradient direction sensitivity weights are obtained by calculating the cosine similarity of the gradient directions of adjacent pixels. The two types of weights are fused to form the gradient sensitivity weight matrix $W_g = W_m \odot W_d$, where $\odot$ represents element-wise multiplication. The gradient sensitivity weight matrix participates in the parameter update process of the Adapter linear transformation layer. The Adapter layer parameters are represented as $A \in \mathbb{R}^{C\times C}$. The weight matrix is fused with the Adapter parameters through a linear mapping along the channel dimension to form a boundary adaptive transformation matrix $A' = A \odot (W_g \times I_C)$. I_C is a C-order identity matrix, and $\times$ represents scalar multiplication of the matrix. This boundary adaptive transformation matrix is used in the subsequent feature calibration process.

B. FEATURE CALIBRATION AND ADAPTER OPTIMIZATION BASED ON BOUNDARY WEIGHTS

The boundary weight map W_b and the feature tensor F are multiplied element-wise to generate a boundary-enhanced feature representation $F_b = F \odot W_b$ [23-25]. The boundary weight map serves distinct functional purposes across different processing stages: in feature calibration, it modulates spatial feature responses to enhance boundary contrast; in adapter parameter transformation, it guides selective parameter updates through gradient sensitivity; in loss formulation, it establishes differential optimization priorities for boundary regions. These applications operate

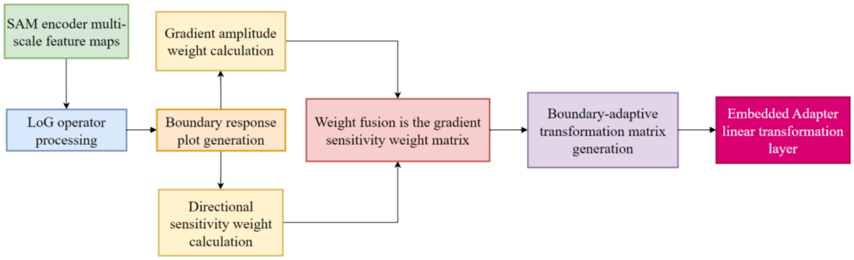

FIGURE 1. Boundary response diagram generation and gradient weight calculation process

at different levels of model processing - spatial feature representation, parameter update mechanics, and optimization objective definition - with complementary effects that address the multi-faceted challenge of boundary recognition under noise interference. This operation strengthens the feature response of the boundary region in the spatial dimension and suppresses noise interference in the non-boundary region. The boundary-enhanced feature representation F_b is used as input to the feature calibration module and enters the processing flow consisting of convolutional layers and layer normalization layers [26,27].

A standard convolutional structure with 3×3 kernels is used. A normalization layer follows the convolution operation to standardize the feature channels, and its mathematical expression is:

$$LN(x_i) = \gamma \cdot \frac{x_i - \mu}{\sqrt{\sigma^2 + \varepsilon}} + \beta \tag{3}$$

x_i represents the i-th element of the eigenvector; μ is the mean of the eigenvector; σ^2 is the variance; ε is a small constant to avoid division by zero error; γ and β are learnable scaling and offset parameters. The boundary adaptive transformation matrix A' participates in the model parameter update process, implementing a gradient-sensitive parameter selection mechanism. This mechanism strictly limits the number of updatable parameters, updating gradients only for Adapter layer parameters related to key boundary features, significantly reducing the number of training parameters. Through gradient sensitivity threshold control, the model maintains the main parameters of the original SAM in a frozen state, effectively reducing training computational overhead and storage requirements, achieving efficient parameter fine-tuning. The selection criteria are defined as $\|W_g\|_F > \tau$, where $\|\cdot\|_F$ represents the Frobenius norm, and τ is a preset threshold parameter. The set of parameters satisfying the selection criteria is updated via gradient propagation, while the remaining parameters remain frozen. The parameter update expression is:

$$\Delta A = \eta \cdot \nabla_A \mathcal{L} \tag{4}$$

η is the learning rate, and $\nabla_A \mathcal{L}$ is the gradient of the loss function $\mathcal{L}$ with respect to the adapter parameter A. The gradient sensitivity weight matrix W_g achieves sparse optimization of the model by controlling the parameter update range, which enhances the ability to recognize the boundaries of packaging visual elements while maintaining the model's general expressive power. The boundary adaptive transformation matrix and the feature calibration module work together to complete the boundary-sensitive feature enhancement process of multi-scale feature maps [28,29].

C. DESIGN OF JOINT LOSS FUNCTION CONSTRAINTS

The boundary-aware cross-entropy loss function applies a boundary weight map to differentiate boundary regions based on the standard cross-entropy loss function. This weight map is generated by the LoG operator [30-32], giving higher attention to boundary regions. The boundary-aware cross-entropy loss function imposes a stronger penalty on pixel classification errors in boundary regions, ensuring that the optimization direction of the model in boundary regions remains consistent with the overall segmentation objective. The loss function is defined as follows:

$$\mathcal{L}_{bce} = -\frac{1}{N}\sum_{i=1}^{N} W_b^{(i)} \left[y_i \log(\hat{y}_i) + (1-y_i)\log(1-\hat{y}_i)\right] \tag{5}$$

N represents the total number of pixels in the image; y_i represents the true label value of the i-th pixel; $\hat{y}_i$ represents the probability value predicted by the model; $W_b^{(i)}$ represents the weight value of W_b generated in section "Boundary Response Map Generation and Gradient Weight Embedding Mechanism" at the corresponding position, ranging from [0, 1], with the weight value of the boundary region being higher than that of the non-boundary region.

HD constraints are used to measure the geometric distance between the predicted boundary and the true boundary [33,34], forcing the boundary contour generated by the model to closely fit the true boundary in spatial location. HD constraints quantify boundary continuity by calculating the maximum and minimum distance between two sets of points, and this metric is sensitive to boundary discontinuity and edge blurring. HD constraints are defined as follows:

$$\mathcal{L}_{hd} = \max\left(\sup_{a\in\partial G}\inf_{b\in\partial P}\|a-b\|, \sup_{b\in\partial P}\inf_{a\in\partial G}\|a-b\|\right) \tag{6}$$

∂G represents the true boundary point set; ∂P represents the predicted boundary point set; $\|a-b\|$ represents the Euclidean distance; sup and inf represent the supremum and infimum operations, respectively. This constraint ensures that the model focuses not only on pixel-level classification accuracy during optimization but also on maintaining the geometric integrity of the boundary structure. Joint loss function combines boundary-aware cross-entropy loss and HD constraint to optimize boundary recognition [35]. Cross-entropy loss guides pixel classification while HD constraint preserves geometric accuracy. During backpropagation, these losses provide complementary gradients that the boundary adaptive transformation matrix directs to enhance boundary features specifically. This ensures the model prioritizes boundary details during optimization while maintaining segmentation performance.

III. EXPERIMENTAL SETUP AND RESULTS

A. DATASET AND EVALUATION PROTOCOL

In the development of the dataset, the primary focus is given to the diversity and complexity of visual elements in product packaging. The image acquisition process follows standardized methods in order to allow for sample representativeness and include a variety of visual elements across different packaging products. The matter of real-world challenges like the possibility of noise interference, overlapping textures, and glare are also given consideration through respected variables during the image acquisition process. The dataset contains 40% high-noise samples, 30% reflective samples, and 30% complex texture samples. To guarantee it is in line with actual packaging inspection situations, industrial vision experts conduct triple verification for every sample. The images adhere to the industry standard of image quality and resolution, so they clearly capture the details of the packaging designs.

The image dataset is then split into a training, validation, and test set to ensure all sets have wrapper images with different categories, and that amount of imagines in each category is balanced to ensure that data leakage does not impact the testing phase. The test set is used solely for evaluating the model and is independent from the training and validation sets.

The evaluation scheme utilizes BIoU and contour F1 score regarding the measures of evaluation. The BIoU evaluates how much overlap exists between the predicted boundary and ground truth boundary, whereas the F1 score provides a measure for the model's continuity and accuracy of the contoured boundary under high noise and complex conditions. Any ontological metrics used for evaluation guarantees the model can handle complex scenes in packaged images accurately.

Table 1 shows that the image resolution for each dataset is 1024×1024 with BIoU and contour F1-score as the major evaluation metrics. Dataset splitting ensures balanced representation and integrity of each category to support model training and evaluation.

Hyperparameter λ controlling gradient sensitivity was set to 0.8 determined through systematic grid search over the range [0.5, 1.2] with 0.1 increments on the validation set where the value yielding the highest BIoU score was selected demonstrating optimal balance between boundary continuity preservation and noise suppression across diverse packaging scenarios with varying noise levels and reflective properties.

B. BOUNDARY POSITIONING ACCURACY VERIFICATION

In order to address the challenges encountered in recognizing visual element boundaries within product packaging contexts that include overlapping printed textures, high levels of noise across the entire image, and the reflection characteristics of the material, this research makes a comprehensive comparison of the accuracy of multiple segmentation methods for boundary localization. The evaluation implements BIoU as the primary accuracy metric, as it provides a useful assessment of the extent of overlap between the predicted boundary and its associated true boundary labels, in particular for product packaging applications that may have sub-pixel level boundaries. This comparison specifically evaluates original SAM, Adapter-SAM, a variant of Attention-SAM that integrates the standard multi-head self-attention mechanism after the Adapter layer, and the method proposed in this paper. Figure 2 (a) shows the variation trends of BIoU values for each method across noise levels, and (b) displays the associated BIoU decrease rate. The key result that our method achieves a BIoU value of 0.923 at a 20% noise level is highlighted with an asterisk.

In Figure 2, at a noise level of 20%, the BIoU value of the proposed method remains at 0.923, which is superior to the original SAM's 0.812, Adapter-SAM's 0.889, and Attention-SAM's 0.896. The BIoU reduction rate is only 6.7%, while the other methods all exceed 9%. The gradient guidance mechanism effectively enhances the feature representation of sub-pixel-level boundaries through the boundary weight map generated by the Gaussian-Laplacian operator and the gradient sensitivity weight matrix, enabling the model to maintain sensitivity to the boundaries of packaging visual elements even under high noise conditions. The synergistic effect of the boundary adaptive transformation matrix and the feature calibration module suppresses noise interference in non-boundary regions while preserving the geometric integrity of the boundary structure, thus achieving a BIoU improvement superior to the comparative methods at a noise level of 20%. Experimental results verify the effective improvement of the proposed method in the performance of product packaging visual element boundary recognition.

Figure 3 presents a qualitative comparison of the segmentation results of various methods under 20% noise interference. The figure displays three distinct groups of packaging visual elements: Group 1 represents high-density text labels, Group 2 displays barcode identification areas, and Group 3 illustrates complex texture and logo boundaries. Outputs of the proposed method and baseline models. As observed, the proposed method maintains a clear boundary structure and sub-pixel-level localization accuracy across all diverse

TABLE 1. Dataset composition and evaluation protocol

Dataset category	Number of images	Boundary marking accuracy	Average image resolution	Evaluation metrics
Training set	1500			
Validation set	400	High-precision marking		
Test set	600		1024×1024	BIoU, F1
Total	2500	–	–	–
Number of categories	15	–	–	–

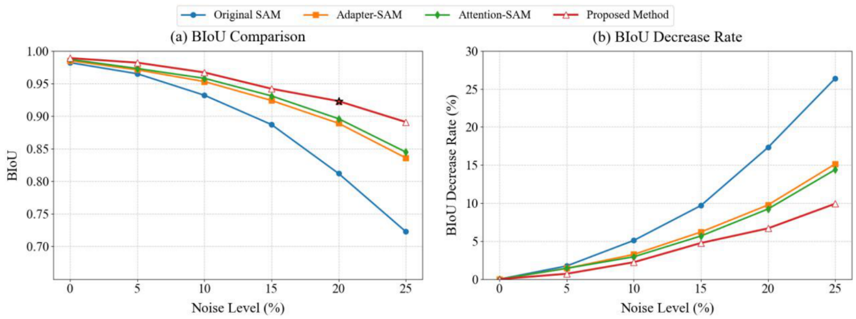

FIGURE 2. Comparison of boundary positioning accuracy methods under noise interference conditions: (a) BIoU comparison; (b) BIoU decrease rate

samples. In contrast, other methods exhibit varying degrees of boundary blurring, breakage, or false detections under the same noise conditions. This visual evidence verifies the consistent performance of the proposed method in recognizing boundaries of visual elements on packaging in high-noise environments.

C. BOUNDARY CONTINUITY ASSESSMENT

In the visual element boundary recognition of product packaging, boundary continuity plays a decisive role in automated quality inspection and anti-counterfeiting analysis. The contour F1 score, as a key indicator for evaluating boundary continuity, can accurately reflect the structural integrity of the model in sub-pixel-level boundary recognition, and is particularly suitable for high-precision boundary recognition scenarios such as text labels and barcodes in product packaging. Figure 4 compares the trend of contour F1 scores of various methods under different reflectivity and texture complexity conditions: Figure 4(a) shows the impact of five levels of reflectivity on boundary continuity, and Figure 4(b) presents the performance changes under five levels of texture complexity.

In Figure 4, the proposed method maintains a contour F1 score of 0.951 under 20% reflectivity, which is superior to the original SAM's 0.842, Adapter-SAM's 0.884, and Attention-SAM's 0.891, confirming that the gradient sensitivity weight matrix effectively suppresses boundary breaks caused by reflectivity interference. With texture complexity level 5, the proposed method produces an F1 score of 0.948 that is much better than the original SAM's score of 0.853, Adapter-SAM's score of 0.881, and Attention-SAM's score of 0.892, indicating the feature calibration module can effectively enhance the boundary recognition capacity of the complex texture region. The differentiation grows as the interference intensity grows. The additional boundary adaptive transformation matrix with the feature calibration module works together to optimize the gradient response of the boundary region, allowing for the model to successfully enhance the geometric integrity recognition ability of the boundary structure of the packaging visuals while the semantic segmentation capacity maintains similar performance.

D. ABLATION STUDY OF KEY COMPONENTS

To quantify the individual contributions of the three core components, systematic ablation experiments were conducted using Adapter-SAM as the baseline model. The gradient-guided adapter weighting mechanism was incrementally integrated first, followed by the boundary-aware cross-entropy loss function, and finally the HD constraint. All experiments were performed under the same 20% noise and reflectivity conditions as the main experiments to ensure consistent evaluation.

The results presented in Table 2 demonstrate the progressive performance improvement when adding each component. The baseline Adapter-SAM achieved a BIoU of 0.889 under 20% noise conditions. The integration of gradient-guided adapter weighting increased the BIoU to 0.898, representing a 0.009 improvement. Adding the boundary-aware cross-entropy loss further improved the BIoU to 0.912, showing an additional 0.014 gain. The final addition of the HD constraint elevated the BIoU to 0.923, consistent with the main experimental results. The contour F1 score followed a similar progression, starting from 0.884 for the baseline model, increasing to 0.928 with gradient-guided adapter weighting, then to 0.938 with the boundary-aware cross-entropy loss, and finally reaching 0.951 with the complete model. The gradient-guided adapter weighting provided the most substantial individual contribution to boundary recognition accuracy, while the HD constraint

FIGURE 3. Qualitative comparison of segmentation results of various methods under 20% noise interference conditions

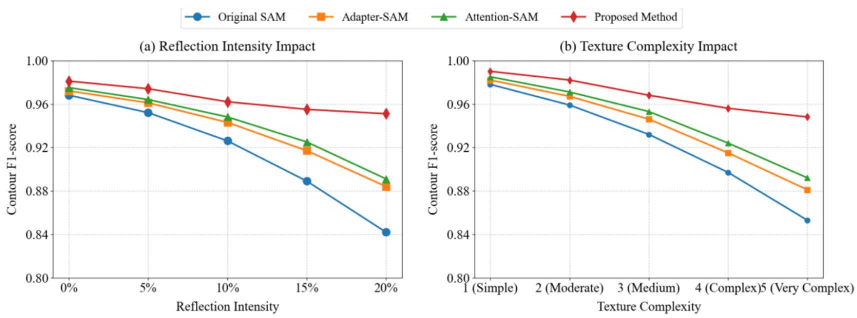

FIGURE 4. Performance comparison of the continuity of packaging visual element boundaries under different reflectivity and texture complexity conditions: (a) Reflection intensity impact; (b) Texture complexity impact

demonstrated critical importance for maintaining geometric integrity under reflective interference.

The ablation study confirms that each component contributes distinct improvements to boundary precision and structural continuity. The gradient-guided adapter weighting significantly enhances the model's sensitivity to boundary features, the boundary-aware cross-entropy loss improves classification accuracy in critical boundary regions, and the HD constraint ensures geometric fidelity of the predicted boundaries. The synergistic integration of all three components produces the highest performance metrics, validating the design choices made in this research.

E. PARAMETER EFFICIENCY AND INFERENCE SPEED

This study evaluates the computational efficiency of the boundary-aware Adapter for the recognition of elements of visual information on product packages by considering the effect of paramater update ratios during training. Training time is an indication of the algorithm's iterative speed and reflects the model's need for computational resources when operationalized; the GPU (Graphics Processing Unit) memory tracks on which hardware the model can feasibly be used, as well as the practicability of deploying the model for edge computing use; the model increment size represents the number of parameters that need to be added during fine-tuning, directly effecting the cost to store and transmit the fine-tuned models. These three measures represent a multi-dimensional standard of measuring the practicality outcomes of applying lightweight fine-tuning methods and are especially relevant for industrial inspections that are often resource constrained. The relationship between the parameters of the model update ratio and measures of computational efficiency is illustrated in Figure 5.

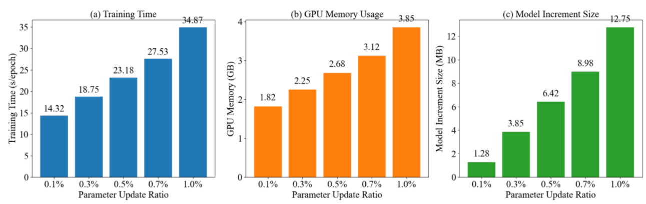

FIGURE 5. Analysis of the impact of parameter update ratio on computational efficiency in the boundary-aware Adapter: (a) Training time; (b) GPU memory usage; (c) Model increment size

As illustrated in Figure 5, at a parameter update ratio of 0.5%, the training time per epoch is 23.18 seconds; the GPU memory consumption is 2.68 GB; the model increment is 6.42 MB. Furthermore, when the parameter update ratio increases from 0.1% to 1.0%, the training time increases from 14.32 seconds per epoch to 34.87 seconds per epoch; the GPU memory consumption increases from 1.82 GB to 3.85 GB; the model increment increases from 1.28 MB to 12.75 MB. This observation indicates a non-linear growth pattern that reflects the more complicated relationship between the parameter update ratio and their computational resource consumption. The boundary adaptive transformation matrix utilizes the gradient sensitivity weight matrix to ascertain which parameters are most relevant in updating the model, avoiding the additional computational expense that performing full model fine-tuning would incur, while allowing the model to maintain an accurate recognition of the boundary features of the packaging visual elements. This synergistic parameter selection mechanization, alongside the feature calibration trimming module, enables the model to reduce

TABLE 2. Ablation study of key components under challenging conditions

Component combination	BIoU (20% noise)	Contour F1 (20% reflectivity)	Improvement over baseline
Adapter-SAM (baseline)	0.889	0.884	–
+ Gradient-guided adapter weighting	0.898	0.928	+0.009
+ Boundary-aware cross-entropy loss	0.912	0.938	+0.014
+ HD constraint	0.923	0.951	+0.011
Full model (all components)	0.923	0.951	+0.034

the computational costs significantly while preserving a general expressive power, thus validating the operational potential of lightweight adapter design in industrial applications.

IV. CONCLUSION

This paper proposes a gradient-guided boundary adaptive Adapter method, which addresses the high noise, complex texture, and reflective interference characteristics of product packaging by extracting multi-scale feature maps from the ViT encoder output in SAM. Then, a LoG operator is applied to generate pixel-level boundary weight maps. Gradient magnitude and direction sensitivity weights are calculated using a differentiable boundary response function and embedded into the Adapter linear transformation matrix. The weight map and feature map are element-wise multiplied and then input into a lightweight convolution-normalization module to calibrate the features. Parameter optimization is achieved by combining a boundary-aware cross-entropy loss function and HD constraints. Experimental results show that the BIoU reaches 0.923 under 20% noise conditions, and the contour F1 score remains at 0.951 under 20% reflectivity. This method effectively improves the problems of blurred edges, poor continuity, and high false detection rates of visual elements on packaging, achieving sub-pixel-level precise boundary localization of visual elements such as text labels and barcodes on product packaging. This meets the stringent requirements of e-commerce platforms and automated quality inspection systems for boundary recognition accuracy.

AUTHOR CONTRIBUTIONS

Conceptualization – Cuihua Meng, Xiaoming Huang and Yuting Li; methodology – Cuihua Meng, Xiaoming Huang and Yuting Li; investigation – Cuihua Meng and Yuting Li; writing-original draft preparation – Cuihua Meng, Xiaoming Huang and Yuting Li. All authors have read and agreed to the published version of the manuscript.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] T. Abekoon, B. Buthpitiya, H. Sajindra, E. R. J. Samarakoon, J. A. D. C. A. Jayakody, K. Kantamaneni, et al., “A comprehensive review to evaluate the synergy of intelligent food packaging with modern food technology and artificial intelligence field,” Discover Sustainability, vol. 5, no. 1, pp. 160-180, 2024, doi: 10.1007/s43621-024-00371-7.

[2] S. Liu and Z. Yu, “Modeling and efficiency analysis of blockchain agriculture products E-commerce cold chain traceability system based on Petri net,” Heliyon, vol. 9, no. 11, Art. no. e21302-e21317, 2023, doi: 10.1016/j.heliyon.2023.e21302.

[3] T. Guan and L. Wang, “Computer-Aided Optimization Design of Intelligent Commodity Packaging Based on Generative Adversarial Network,” Computer-Aided Design and Applications, vol. 21, no. S3, pp. 107-120, 2023, doi: 10.14733/cadaps.2024.S3.107-120.

[4] D. Zhou and M. Miao, “Research on cloud anti-counterfeiting product packaging design based on internet of things,” International Journal of Internet Manufacturing and Services, vol. 10, no. 1, pp. 60-76, 2024, doi: 10.1504/IJIMS.2024.136718.

[5] C. Zhang, M. Han, J. Jia, and C. Kim, “Packaging Design Image Segmentation Based on Improved Full Convolutional Networks,” Applied Sciences, vol. 14, no. 22, pp. 10742-10759, 2024, doi: 10.3390/app142210742.

[6] M. M. Rahman, S. Rahman, S. Bhatt, and M. Faezipour, “A Systematic Review on Advancement of Image Segmentation Techniques for Fault Detection Opportunities and Challenges,” Electronics, vol. 14, no. 5, pp. 974-996, 2025, doi: 10.3390/electronics14050974.

[7] G. Wang, Z. Li, G. Weng, and Y. Chen, “An overview of industrial image segmentation using deep learning models,” Intelligence & Robotics, vol. 5, no. 1, pp. 143-180, 2025, doi: 10.20517/ir.2025.09.

[8] U. Ulusoy, “A review of particle shape effects on material properties for various engineering applications: from macro to nanoscale,” Minerals, vol. 13, no. 1, pp. 91-171, 2023, doi: 10.3390/min13010091.

[9] J. Liu, S. Liu, W. Zhou, H. Ren, W. Zhao, and Z. Li, “Zero-Reference Depth Curve Estimation-Based Low-Light Image Enhancement Method for Coating Workshop Inspection,” Coatings, vol. 15, no. 4, pp. 478-495, 2025, doi: 10.3390/coatings15040478.

[10] A. Sonavale, M. Chakkaravarthy, S. Srinivasa Rao, H. B. M. Salleh, and J. Jadhav, “Automatic Semantic Segmentation and Classification of Remote Sensing Image Data for Flood Detection Using Novel LSTM Neural Network,” SN Computer Science, vol. 5, no. 8, pp. 992, 2024, doi: 10.1007/s42979-024-03336-9.

[11] Y. Gong, M. Liu, X. Wang, C. Liu, and J. Hu, “Research on surface defects detection method and system in manufacturing processes based on the fusion of multi-scale features and semantic segmentation for intelligent manufacturing,” Journal of Intelligent & Fuzzy Systems, vol. 44, no. 4, pp. 6463-6481, 2023, doi: 10.3233/JIFS-223041.

[12] V. Guimaraes, J. Nascimento, P. Viana, and P. Carvalho, “A review of recent advances and challenges in grocery label detection and recognition,” Applied Sciences, vol. 13, no. 5, pp. 2871-2908, 2023, doi: 10.3390/app13052871.

[13] G. Duan, S. Zhang, Y. Shang, Y. Shao, and Y. Han, “Research on a Multi-Type Barcode Defect Detection Model Based on Machine Vision,” Applied Sciences, vol. 15, no. 15, pp. 8176-8196, 2025, doi: 10.3390/app15158176.

[14] M. Abbasi, P. Vaz, J. Silva, and P. Martins, “Enhancing visual perception in immersive VR and AR environments: AI-driven color and clarity adjustments under dynamic lighting conditions,” Technologies, vol. 12, no. 11, pp. 216-230, 2024, doi: 10.3390/technologies12110216.

[15] T. Kar, P. Kanungo, S. N. Mohanty, S. Groppe, and J. Groppe, “Video shot-boundary detection: Issues, challenges and solutions,” Artificial Intelligence Review, vol. 57, no. 4, pp. 104-141, 2024, doi: 10.1007/s10462-024-10742-1.

[16] Y. Jiang, J. Chen, and J. Lu, "Leveraging Vision Foundation Model via PConv-Based Fine-Tuning with Automated Prompter for Defect Segmentation," Sensors, vol. 25, no. 8, pp. 2417-2438, 2025, doi: 10.3390/s25082417.
[17] X. Lin, X. Huang, and L. Wang, "Underwater object detection method based on learnable query recall mechanism and lightweight adapter," PLOS One, vol. 19, no. 2, Art. no. e0298739-e0298755, 2024, doi: 10.1371/journal.pone.0298739.
[18] Y. Lin, T. Fan, and C. Fang, "A Classification and Segmentation Model for Diamond Abrasive Grains Based on Improved Swin-Unet-SAM," Electronics, vol. 13, no. 21, pp. 4213-4231, 2024, doi: 10.3390/electronics13214213.
[19] Z. Zhou, W. Li, Y. Wang, H. Liu, and N. Wu, "A Feasible Domain Segmentation Algorithm for Unmanned Vessels Based on Coordinate-Aware Multi-Scale Features," Journal of Marine Science and Engineering, vol. 13, no. 8, pp. 1387-1407, 2025, doi: 10.3390/jmse13081387.
[20] A. A. Almazroey, S. kammoun Jarraya, and R. Alnanih, "SAM for Road Object Segmentation: Promising but Challenging," Journal of Imaging, vol. 11, no. 6, pp. 189-219, 2025, doi: 10.3390/jimaging11060189.
[21] A. Xie, Q. Lin, Y. He, X. Zeng, Y. Cao, Z. Man, et al., "Metastasis lesion segmentation from bone scintigrams using encoder-decoder architecture model with multi-attention and multi-scale learning," Quantitative Imaging in Medicine and Surgery, vol. 15, no. 1, pp. 689-708, 2024, doi: 10.21037/qims-24-1246.
[22] H. Chen, S. Zhou, K. Li, J. Yin, and J. Huang, "A Hybrid Framework for Referring Image Segmentation: Dual-Decoder Model with SAM Complementation," Mathematics, vol. 12, no. 19, pp. 3061-3081, 2024, doi: 10.3390/math12193061.
[23] K. Yuan, X. Li, Y. Ren, L. Zhang, W. Liu, and E. Li, "MSFE-Net: Multi-Scale Feature Enhancement Network for Remote Sensing Object Detection," Applied Artificial Intelligence, vol. 39, no. 1, pp. 2514324-2514360, 2025, doi: 10.1080/08839514.2025.2514324.
[24] X. Zhang, C. Zhao, W. Lu, and K. Liang, "Load Equipment Segmentation and Assessment Method Based on Multi-Source Tensor Feature Fusion," Electronics, vol. 14, no. 5, pp. 1040-1063, 2025, doi: 10.3390/electronics14051040.
[25] J. Xiao, T. Chen, X. Hu, G. Zhang, and S. Wang, "Boundary-guided context-aware network for camouflaged object detection," Neural Computing and Applications, vol. 35, no. 20, pp. 15075-15093, 2023, doi: 10.1007/s00521-023-08502-3.
[26] J. Shi, G. Chen, and Y. Chen, "Enhanced boundary perception and streamlined instance segmentation," Scientific Reports, vol. 15, no. 1, pp. 23612-23634, 2025, doi: 10.1038/s41598-025-09139-z.
[27] M. Tang, M. Ling, J. Tang, and J. Hu, "A micro-expression recognition algorithm based on feature enhancement and attention mechanisms," Virtual Reality, vol. 27, no. 3, pp. 2405-2416, 2023, doi: 10.1007/s10055-023-00808-w.
[28] L. Yang, Y. Gu, and H. Feng, "Multi-scale feature fusion and feature calibration with edge information enhancement for remote sensing object detection," Scientific Reports, vol. 15, no. 1, pp. 15371-15392, 2025, doi: 10.1038/s41598-025-99835-7.
[29] J. Q. Wang, T. Chen, L. Zheng, J. Tie, Y. Zhang, P. Chen, et al., "A multi-scale remote sensing semantic segmentation model with boundary enhancement based on UNetFormer," Scientific Reports, vol. 15, no. 1, pp. 1-19, 2025, doi: 10.1038/s41598-025-99663-9.
[30] J. Wang, N. Ding, and G. He, "A boundary enhancement loss function for semantic segmentation of land cover," International Journal of Remote Sensing, vol. 44, no. 12, pp. 3637-3659, 2023, doi: 10.1080/01431161.2023.2224101.
[31] I. Ahmad, F. Shang, M. S. Pathan, A. Wajahat, and Y. S. Kim, "Dual-stream hybrid architecture with adaptive multiscale boundary-aware mechanisms for robust urban change detection in smart cities," Scientific Reports, vol. 15, no. 1, pp. 30729-30749, 2025, doi: 10.1038/s41598-025-16148-5.
[32] M. Li, Y. Liu, D. Chen, L. Chen, and S. Liu, "Transformer with difference convolutional network for lightweight universal boundary detection," PLOS One, vol. 19, no. 4, Art. no. e0302275-e0302301, 2024, doi: 10.1371/journal.pone.0302275.
[33] K. Lange, "Computation of the Hausdorff distance between two compact convex sets," Algorithms, vol. 16, no. 10, pp. 471-488, 2023, doi: 10.3390/a16100471.
[34] Y. Wang, X. Jing, Y. Xu, L. Cui, Q. Zhang, and H. Li, "Geometry-guided semantic segmentation for post-earthquake buildings using optical remote sensing images," Earthquake Engineering & Structural Dynamics, vol. 52, no. 11, pp. 3392-3413, 2023, doi: 10.1002/eqe.3966.
[35] M. E. Omati and F. T. Mahmoudi, "Efficiency analysis of ITN loss function for deep semantic building segmentation," Earth Science Informatics, vol. 17, no. 3, pp. 2011-2025, 2024, doi: 10.1007/s12145-024-01267-w.