DOI: 10.19884/j.1672-5220.202501002

# Segmentation of Line Laser Stripe on Weld Image under Strong Interferences

HE Yiyan, LI Yan*, XU Yang
College of Mechanical Engineering, Donghua University, Shanghai 201620, China

**Abstract: To accurately and quickly segment the line laser stripes under interferences of strong noise and strong reflection, a lightweight weld line laser stripe segmentation network based on the DeepLabv3+ network, named WLS-Net, is proposed. To improve the segmentation speed of the network, the shallow residual network ResNet-18 is selected as the backbone network. Combining the multi-Dconv head transposed attention (MDTA) and the convolutional block attention module (CBAM), the multi-dimensional CBAM (MD-CBAM) is proposed, and the dynamic upsampling method (DySample) is chosen to replace the traditional bilinear interpolation to improve the segmentation accuracy. To address the foreground-background class imbalance in the weld line laser stripe image, the sum function of the Dice loss function (Dice Loss) and the pixel-wise cross-entropy loss function is chosen as the loss function of the model. The experimental results show that compared with the original DeepLabv3+ network, the WLS-Net achieves absolute improvements of 6.00% in IoU, 5.67% in precision, and 3.76% in F1 score on the weld line laser dataset, and the inference time of a single image is reduced by 18 ms. Compared with other semantic segmentation networks, the WLS-Net also achieves better segmentation effect and higher segmentation rate.**

**Keywords**: weld seam image; laser stripe segmentation; DeepLabv3+ network; attention mechanism
**CLC number**: TP391 **Document code**: A
**Article ID**: 1672-5220(2026)02-0120-08

Open Science Identity (OSID)

## 0 Introduction

Intelligent welding that integrates sensor technology and image processing is emerging as a future development trend[1]. The welding method based on linear laser vision sensors involves projecting a line laser onto the weld surface and identifying the line laser stripe features in the camera images. According to the triangulation principle, these line laser stripe features are converted into weld features, achieving the localization of weld features in three-dimensional space[2-3]. However, during actual welding processes, the collected line laser stripe images are severely disturbed by intense interferences such as arc light, spatter, smoke and metallic reflections. The strong reflection produced by the line laser on the weld surface also significantly reduces the imaging quality of the structured light images. These interferences greatly affect the recognition of line laser stripe features.

Traditional image processing methods struggle to accurately and quickly identify line laser stripe features from noise with significant randomness in both grayscale values and spatial distribution. Consequently, the application of deep learning-based semantic segmentation for weld line laser stripe segmentation has become a research focus. In pursuit of a lightweight line laser stripe segmentation network, Huang et al.[4] trimmed some short-range dense connection modules in the BiSeNet to make the network more lightweight. Zhang et al.[5] introduced parallel downsampling modules and reduced the number of convolutional kernels in the encoder to lighten the U-Net network. Chen et al.[6] utilized a lightweight network as the backbone for the DeepLabv3+ network to enhance the inference speed. Noise filtering capability is a crucial performance metric for the network. Yang et al.[7] adopted the convolutional long short-term memory (ConvLSTM) module and dense attention module to enhance the U-Net network's noise filtering ability. Addressing the challenge of increased segmentation difficulty due to the imbalance of positive and negative samples in weld line laser stripe images, Lu et al.[8] employed online hard example mining (OHEM) and the Dice loss function (Dice Loss)[9] as the network's loss function to achieve better segmentation accuracy.

Aiming at the problem that the repeated stacked pooling layer and subsampling layer in the neural network would lead to the loss of a large number of location information and difficult to recover, the Deeplab series semantic segmentation network uses the method of atrous convolution to expand the receptive field of the feature map in the coding phase, so that each convolution output contains a large range of information. Within the DeepLabv3+ network, the atrous spatial pyramid pooling (ASPP) module sets dilated convolutions with different dilation rates in parallel, utilizing multi-scale contextual information to achieve more accurate segmentation results.

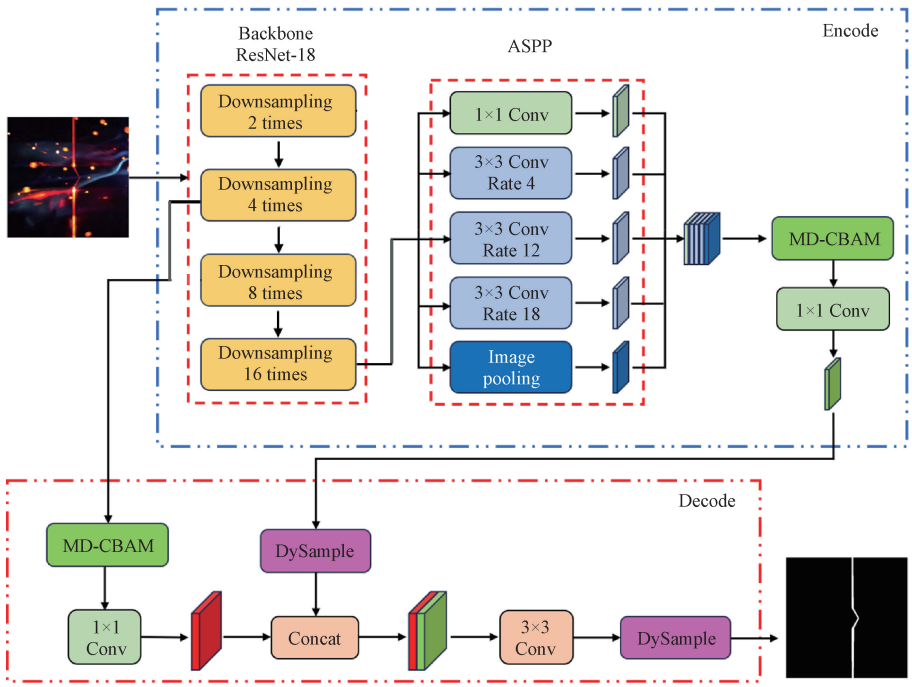
Received date: 2025-01-10
Foundation item: National Natural Science Foundation of China (No. U1831123)
* Correspondence should be addressed to LI Yan, email: liyanly@dhu.edu.cn
Citation: HE Y Y, LI Y, XU Y. Segmentation of line laser stripe on weld image under strong interferences[J]. *Journal of Donghua University (English Edition)*, 2026, 43(2): 120-127.

This paper addresses the challenge of extracting line laser stripe features under strong interferences such as arc light, spatter, smoke, and metallic reflections during the welding process. A lightweight weld line laser stripe segmentation network based on the DeepLabv3 + architecture (named WLS-Net) is proposed. The network employs a shallow residual network, ResNet-18, as the backbone, and incorporates an attention mechanism that combines multi-Dconv head transposed attention (MDTA)[11] and convolutional block attention module (CBAM)[12]. It uses the dynamic upsampling method (DySample)[13] for upsampling. The loss function is a combination of the pixel-wise cross-entropy loss function and Dice Loss. The network is trained on a self-compiled weld line laser dataset. The results indicate that the WLS-Net enables rapid and accurate segmentation of line laser stripes in weld images under strong interferences, thereby providing support for subsequent extraction of the line laser stripe centerline and recognition of structural features.

## 1 WLS-Net

As illustrated in Fig. 1, the WLS-Net is constructed based on the DeepLabv3 + architecture to balance the precision and real-time performance of weld line laser stripe segmentation. The network consists of a classic encoder and decoder. In the encoder phase, the ASPP module of the DeepLabv3 + network is retained, with a shallow residual network ResNet-18 serving as the backbone. The line laser stripe image is processed through the backbone to obtain high-level and low-level feature maps. The high-level feature map, after undergoing atrous convolutions with various dilation rates in the ASPP module, is concatenated to create a feature map with multi-scale information. This map is then passed through the MD-CBAM, which combines MDTA and CBAM, before being fed into the decoder. The low-level feature map proceeds directly to the decoder phase.

In the decoder phase, the high-level feature map is resized to align with the low-level feature map after undergoing dynamic upsampling with DySample. At this point, the low-level feature map is also processed by MD-CBAM and concatenated with the upscaled high-level feature map for fusion. The fused feature map, after undergoing a 3×3 convolution and another upsampling step, is restored to the original resolution, yielding the final image segmentation result.



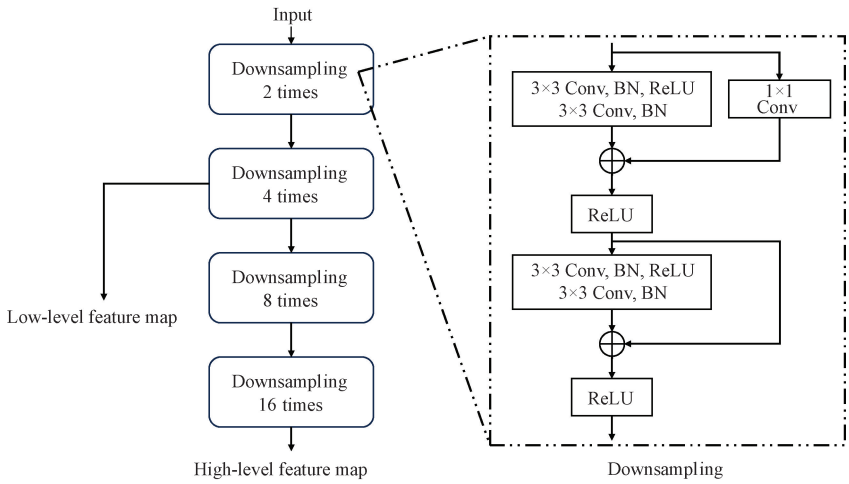

Conv—convolution; Concat—concatenate.

Fig. 1 Architecture of WLS-Net

### 1.1 Lightweight backbone network: ResNet-18

To meet the requirements of real-time weld seam tracking with both precision and speed, the WLS-Net employs a more lightweight yet performance-effective shallow residual network, ResNet-18, as the backbone network within its encoder. The layers with residual structures in ResNet-18 allow the network's input to bypass the layer directly, creating an identity mapping, which addresses issues such as network degradation, vanishing gradients, and gradient explosion during neural network training. Additionally, ResNet-18 has fewer network layers, parameters, and computational complexity, demonstrating superior performance in the task of weld line laser stripe segmentation. The structure of ResNet-18 within the WLS-Net is shown in Fig. 2. The input image undergoes four downsampling operations

within the backbone network to extract features, ultimately producing a low-level feature map with four times downsampling and a high-level feature map with 16 times downsampling for subsequent feature processing.

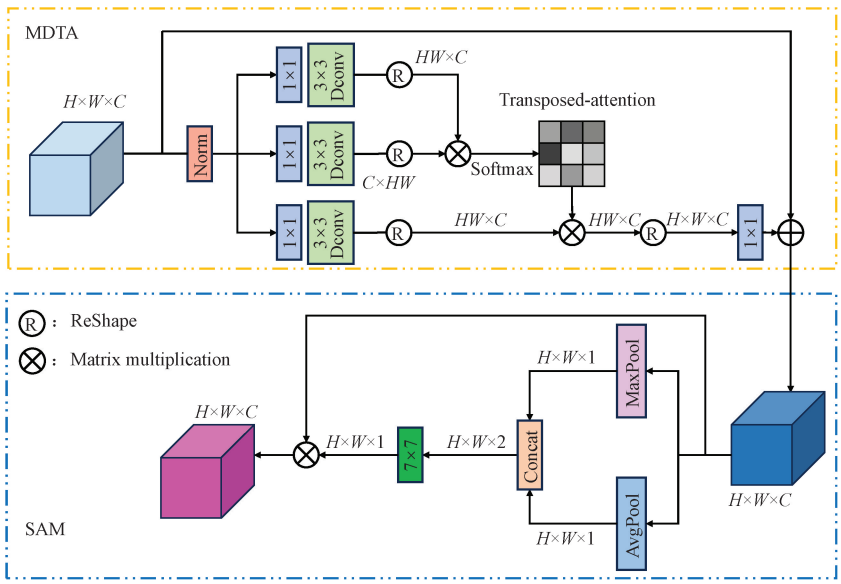



BN—batch normalization; ReLU—rectified linear unit; ⊕ —element-wise addition.

Fig. 2 ResNet-18 and downsampling structure

## 1.2 Attention mechanism: MD-CBAM

The CBAM is a dual attention module composed of the channel attention module (CAM) and the spatial attention module (SAM). In the task of weld line laser stripe image segmentation, the dual attention module can enhance the model's sensitivity to the features and positions of line laser stripes, thereby improving the accuracy of model segmentation. Since self-attention modules are sensitive to the correlations within features, they can utilize structured light stripes that are not interfered with by noise such as arc light, spatter, smoke, and metallic reflections to improve the segmentation accuracy of the noisy parts. Compared to ordinary self-attention modules, MDTA has better feature representation capabilities and higher parallelism. Its design, which performs attention calculations across channel dimensions, also makes it lightweight. To incorporate the advantages of MDTA into CBAM, the CAM in CBAM is replaced with MDTA to form the MD-CBAM attention module, as shown in Fig. 3. In the WLS-Net, MD-CBAM is inserted before the fusion of the high-level and low-level feature maps. The feature maps processed by the attention calculation have a more accurate feature representation, thereby enhancing the model's expressive power.

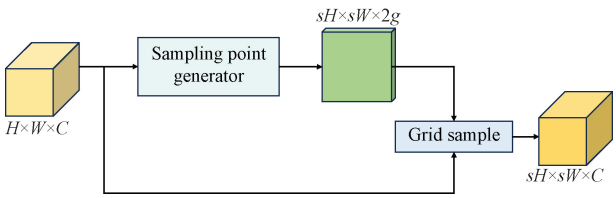



Norm—layer normalization; Dconv—depth-wise convolution; *H*—height; *W*—width; *C*—number of channels; MaxPool—max pooling; AvgPool—average pooling.

Fig. 3 MD-CBAM structure

### 1.3 Dynamic upsampling: DySample

Traditional bilinear interpolation upsampling often results in the loss of target edges and detailed information, while kernel-based dynamic upsampling introduces a significant workload. In the WLS-Net, DySample is chosen to restore the image resolution. As a lightweight and efficient upsampling operator, DySample addresses the issue of information loss during the upsampling process through dynamic point sampling. Additionally, as shown in Fig. 4, the DySample network structure adopts a grouped upsampling strategy, dividing the feature map into groups along the channel dimension before performing upsampling calculations, which reduces the computational load of the upsampling process. Therefore, DySample preserves more information detail when restoring low-resolution features to high resolution, achieving efficient and high-quality upsampling. This makes it a promising technique for real-time applications.




*s*—upsampling scale factor; *g*—number of sampling groups.

Fig. 4 DySample network structure

### 1.4 Loss function

The essence of semantic segmentation is the classification of image pixels, with the pixel-wise cross-entropy loss function being a common choice in semantic segmentation. The segmentation of weld line laser stripe images constitutes a binary classification task. The calculation formula for this loss function $L_{bce}$ in binary classification is

$$L_{bce} = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \lg p_i + (1-y_i)\lg(1-p_i)\right], \quad (1)$$

where $N$ represents the number of pixels; $y_i$ is the true label of the $i$th pixel; $p_i$ is the predicted probability of the $i$th pixel by the model.

However, in the weld line laser stripe images, the number of background pixels significantly exceeds that of the laser stripe pixels. Relying solely on a pixel-wise cross-entropy loss function can result in the model's training outcome favoring the background over the more critical laser stripes in the segmentation task. Incorporating Dice Loss effectively mitigates the adverse effects of the foreground-background imbalance on the model. The expression for Dice Loss $L_{dice}$ is

$$L_{dice} = 1 - \frac{2|X \cap Y|}{|X| + |Y|} = 1 - \frac{2\sum_{i=1}^{N} y_i p_i}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} p_i}, \quad (2)$$

where $X$ represents the true label of the line laser stripe; $Y$ represents the prediction result of the model; $|X \cap Y|$ denotes the number of intersecting pixels between the line laser pixels and the true label; $|X|$ and $|Y|$ represent the number of pixels in the line laser stripe for the true label and the prediction result, respectively. Different from the pixel-wise cross-entropy loss function, Dice Loss focuses only on the foreground pixels. Therefore, the more intersecting foreground pixels there are, the smaller the Dice Loss value, allowing the network to pay more attention to the laser stripe pixels as the foreground during training, thereby mitigating issues caused by the imbalance between foreground and background.

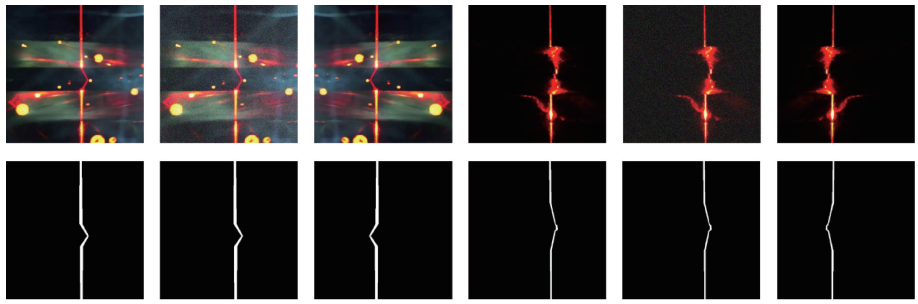
Due to the tendency of Dice Loss to cause severe oscillations, it is not used as a loss function in isolation. Therefore, in the WLS-Net, the sum function $L$ of Dice Loss and the pixel-wise cross-entropy loss function is chosen as the model's loss function to guide model training.

$$L = L_{dice} + L_{bce} = 1 - \frac{1}{N}\sum_{i=1}^{N}\left[y_i \lg p_i + (1-y_i)\times \lg(1-p_i)\right] - \frac{2\sum_{i=1}^{N} y_i p_i}{\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} p_i}. \quad (3)$$

## 2 Results and Discussion

### 2.1 Weld line laser dataset

The robot was taught to weld on the test welding plate. During the welding process, 240 line laser stripe images of the welding seam were collected. All of them were 1 920 × 1 200 pixels in size. They were then trimmed to 512 × 512 pixels by using region of interest (ROI) extraction. Considering the complexity of the welding site, the data were augmented through operations such as rotation, noise addition, and mirroring, resulting in a total of 1 200 weld line laser stripe images. Each weld line laser stripe image was labeled by using LabelMe to create the weld line laser dataset. Some of the images are shown in Fig. 5. Figure 5(a) shows the laser stripe images and label images of the weld with strong noise. The second and third column images are obtained from the first column images after adding Gaussian noise and flipping, respectively. Figure 5(b) shows the expanded weld line laser stripe images with strong reflection and their label images. The gray value of the welding noise randomly distributed in the strong noise image is equivalent to that of the line laser, and the line laser stripe at the groove in the strong reflective image is even seriously damaged, which makes it difficult for the traditional image processing to recognize the weld characteristics. The entire dataset was randomly divided into training and validation sets in a ratio of 8∶2 and fed into the network for training.



(a) (b)

Fig. 5 Laser stripe images (top) and their label images (bottom) of some welds: (a) expanded weld with strong noise; (b) expanded weld with strong reflections

### 2.2 Training parameters

The network design was experimentally trained by using the dataset produced above. The deep learning framework employed was PyTorch, version 1.10.0, with CUDA version 11.3. The development language was Python 3.8. The specific experimental environment parameters included the Ubuntu 20.04 operating system and an NVIDIA RTX 3080Ti graphics card with 12 GB of video memory. During the training process, the batch size was set to 8, the number of training epochs was 80, and the optimizer used was the adaptive moment estimation (Adam). The initial learning rate was set to 0.0001, and the learning rate decay method was cosine annealing.

### 2.3 Performance metrics

To evaluate the segmentation performance of the network, three common metrics in semantic segmentation are selected: intersection over union (IoU) $I_{ou}$, precision $P$ and F1 score $F_1$. The calculation formulas for these metrics are

$$I_{ou} = \frac{N_{TP}}{N_{TP} + N_{FP} + N_{FN}}, \tag{4}$$

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}, \tag{5}$$

$$F_1 = \frac{2N_{TP}}{N_{FP} + 2N_{TP} + N_{FP}}, \tag{6}$$

where $N_{TP}$ represents the count of correctly predicted the structured light stripe pixels as structured light stripes; $N_{FP}$ represents the count of incorrectly predicted background pixels as structured light stripes; $N_{FN}$ represents the count of incorrectly predicted structured light stripe pixels as background; $N_{TN}$ represents the count of correctly predicted background pixels as background. Additionally, the average time taken for the network to segment a single-line laser stripe image is used as a metric to evaluate the model's segmentation speed.

### 2.4 Ablation experiments

To validate the segmentation performance of each module selected for the WLS-Net on the weld line laser stripe with strong noises and intense metallic reflections, ablation experiments were conducted based on a weld line laser dataset. Method 1 employs the original Deeplabv3+ network model with the Xception backbone, bilinear interpolation for upsampling, and pixel-wise cross-entropy as the loss function. Method 2 opts for the shallow residual network ResNet-18 as the backbone. Method 3 uses the sum of pixel-wise cross-entropy and Dice Loss as the model's loss function. Methods 4 and 5 add CBAM and MD-CBAM, respectively. Method 6 selects DySample as the upsampling method. Method 7 represents the complete WLS-Net model. The experimental results are shown in Table 1.

**Table 1** Results of ablation experiments

| Method | Backbone | Attention | Upsampling | Loss | IoU/% | Precision/% | F1 score/% | Inference time/ms |
|---|---|---|---|---|---|---|---|---|
| 1 | Xception | — | — | $L_{bce}$ | 75.83 | 82.71 | 86.25 | 56 |
| 2 | ResNet-18 | — | — | $L_{bce}$ | 76.07 | 83.03 | 86.41 | 32 |
| 3 | ResNet-18 | — | — | $L_{bce}+L_{Dice}$ | 77.07 | 85.90 | 87.05 | 32 |
| 4 | ResNet-18 | CBAM | — | $L_{bce}+L_{Dice}$ | 79.90 | 87.99 | 88.82 | 35 |
| 5 | ResNet-18 | MD-CBAM | — | $L_{bce}+L_{Dice}$ | 80.71 | 86.33 | 89.33 | 36 |
| 6 | ResNet-18 | — | DySample | $L_{bce}+L_{Dice}$ | 77.68 | 86.41 | 91.87 | 34 |
| 7 | ResNet-18 | MD-CBAM | DySample | $L_{bce}+L_{Dice}$ | 81.83 | 88.38 | 90.01 | 38 |

Compared to Method 1, Method 2 achieves a more desirable segmentation accuracy and reduced inference time. Compared to Method 2, Method 3 employs the sum of pixel-wise cross-entropy and Dice Loss as the loss function, which improves the IoU of the line laser stripes in absolute terms by 1.00%. Methods 4 and 5 achieve higher IoU for line laser stripes at the cost of increased inference time compared to Method 3. Specifically,

Method 4 shows an absolute improvement of 2.83% (from 77.07% to 79.90%), while Method 5 shows an absolute improvement of 3.64% (from 77.07% to 80.71%). This result demonstrates that replacing CAM with MDTA in CBAM yields better segmentation accuracy. Method 6 slightly increases the inference time but delivers absolute improvements of 0.61% in IoU and 0.51% in precision for weld line laser stripe segmentation compared to Method 3. The WLS-Net model proposed in this paper achieves the best overall performance. Compared Method 7 with Method 1, the WLS-Net model yields absolute gains of 6.00% in IoU, 5.67% in precision, and 3.76% in F1 score, while also reducing the inference time per image by 18 ms. This result conclusively demonstrates the effectiveness of the proposed network.

### 2.5 Comparison of different models

In order to further verify the feasibility of the WLS-Net model in the task of weld line laser stripe image segmentation, the WLS-Net model was compared with the classic semantic segmentation models: U-Net, DeepLabv3+, PSP-Net, and FastSCNN. To ensure that the hyperparameter settings during the training process affect the results of comparative experiments between different models, the hyperparameter settings of all models are consistent, and the image segmentation results are shown in Fig. 6.

From Fig. 6, it can be observed that all models can accurately segment the line laser stripes to some extent despite interferences such as arc light, spatter, smoke, and metallic reflections during the welding process. DeepLabv3+ and PSP-Net show better segmentation effects, while the U-Net and FastSCNN models produce irregular sawtooth edges at the weld groove. When it comes to segmenting strong reflective weld line laser stripe images, both the U-Net and PSP-Net models exhibit interruptions at the groove to varying degrees, and the FastSCNN model misses the entire groove. The DeepLabv3+ model, however, demonstrates a better detection effect, although there are jagged edges at the groove. The segmented stripes do not show any interruption and are relatively complete. Compared to other network models, the WLS-Net model excels in both filtering out welding process noise and recognizing strong reflective weld line laser stripe images, performing the segmentation task more effectively and closely aligning with the label images.

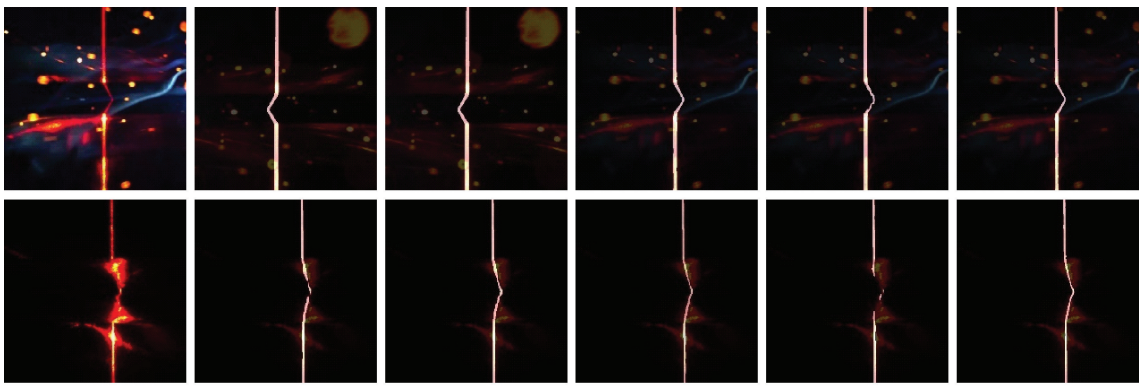



Fig. 6 Comparison of segmentation effect between WLS-Net model and other models: (a) laser stripe image; (b) U-Net model; (c) DeepLabv3+ model; (d) PSP-Net model; (e) FastSCNN model; (f) WLS-Net model

The performance metric comparison of different networks is shown in Table 2. The WLS-Net model outperforms all other comparative networks in the segmentation of line laser stripes. This is not only due to the ASPP module with dilated convolutions, which increases the receptive field of the convolution kernel, but also because of the addition of the MD-CBAM attention mechanism. It allows the network to utilize line laser stripes that are not interfered with by arc light, spatter, smoke, and metallic reflections, thereby enhancing the segmentation accuracy of the noisy parts. At the same time, the choice of DySample as the upsampling method in the WLS-Net enables the network to capture more information during the upsampling process. Additionally, the introduction of Diss Loss mitigates the issue of foreground-background imbalance in the segmentation of weld line laser stripe images. Therefore, the WLS-Net demonstrates superior segmentation performance, especially for strong reflective line laser stripe images. Although the WLS-Net model has a longer segmentation inference time than the FastSCNN model, its superior segmentation effectiveness compared to other classical networks results in an overall significant advantage.

**Table 2** Experimental results of different models

| Model | IoU/% | Precision/% | F1 score/% | Inference time/ms |
|---|---|---|---|---|
| U-Net | 70.38 | 75.49 | 82.61 | 52 |
| DeepLabv3+ | 75.83 | 82.71 | 86.25 | 56 |
| PSP-Net | 71.49 | 83.14 | 83.38 | 63 |
| FastSCNN | 66.50 | 80.65 | 79.88 | 11 |
| WLS-Net | 81.83 | 88.38 | 90.01 | 38 |

## 3 Conclusions

In order to recognize weld features from weld line laser stripe images with strong interferences such as arc light, spatter, smoke, and metallic reflections, a lightweight weld line laser stripe segmentation network named WLS-Net is proposed to filter image noise. To enhance the efficiency of image segmentation, the shallow residual network ResNet-18 is selected as the backbone network for feature extraction; an MD-CBAM attention mechanism combined with a multi-head self-attention mechanism is designed to improve segmentation accuracy; a more advanced dynamic upsampling method, DySample, is chosen to restore image resolution; and a sum function of pixel-wise cross-entropy loss and Dice Loss is used as the network's loss function. The experimental results demonstrate the effectiveness of the WLS-Net.

1) The shallow residual network ResNet-18, as the backbone of the WLS-Net, can reduce the inference time of each image by 24 ms without reducing the segmentation accuracy.

2) By adding the MD-CBAM attention mechanism, using DySample, and adding Dice Loss as a supplement to the pixel-wise cross-entropy loss function, the WLS-Net model achieves absolute IoU gains of 3.64%, 0.61%, and 1.00%, respectively, for weld line laser stripe segmentation.

3) Compared with different semantic segmentation networks, the WLS-Net has better segmentation performance and higher segmentation efficiency in the task of weld line laser stripe segmentation. Compared with the original model, it shows a 6.00%, 5.67%, and 3.76% absolute improvement in IoU, precision, and F1 score, respectively, which proves the effectiveness of the WLS-Net in the task of weld line laser stripe segmentation.

## References

[1] GUO J C, ZHU Z M, SUN B W. A multifunctional monocular visual sensor based on combined laser structured lights[J]. *Transactions of the China Welding Institution*, 2019, 40(10): 1-7, 161. (in Chinese)

[2] GILLIES D. Close range photogrammetry[J]. *The Photogrammetric Record*, 2015, 30(151): 318-322.

[3] BÁSACA-PRECIADO L C, SERGIYENKO O Y, RODRÍGUEZ-QUINONEZ J C, et al. Optical 3D laser measurement system for navigation of autonomous mobile robot[J]. *Optics and Lasers in Engineering*, 2014, 54: 159-169.

[4] HUANG Y G, WANG D Q, JIANG M, et al. Laser fringe segmentation and feature points location method of weld image based on multi-task learning[J]. *Chinese Journal of Lasers*, 2023, 50(16): 79-89. (in Chinese)

[5] ZHANG S K, WU Q X, LIN Z Y. Detection and segmentation of structured light stripe in weld image[J]. *Acta Optica Sinica*, 2021, 41(5): 88-96. (in Chinese)

[6] CHEN B, HE S, LIU J, et al. Weld structured light image segmentation based on lightweight DeepLab v3+ network[J]. *Chinese Journal of Lasers*, 2023, 50(8): 49-58. (in Chinese)

[7] YANG L, FAN J F, HUO B Y, et al. Image denoising of seam images with deep learning for laser vision seam tracking[J]. *IEEE Sensors Journal*, 2022, 22(6): 6098-6107.

[8] LU J, ZHANG J, LUO J, et al. Plate additive, seam-tracking technology based on feature segmentation[J]. *Optics & Laser Technology*, 2024, 168: 109848.

[9] MILLETARI F, NAVAB N, AHMADI S A. V-net: fully convolutional neural networks for volumetric medical image segmentation[C]// 2016 Fourth International Conference on 3D Vision (3DV). New York: IEEE, 2016: 565-571.

[10] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(4): 834-848.

[11] ZAMIR S W, ARORA A, KHAN S, et al. Restormer: efficient transformer for high-resolution image restoration[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2022: 5728-5739.

[12] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[C]// Computer Vision-ECCV 2018. Cham: Springer, 2018: 3-19.

[13] LIU W Z, LU H, FU H T, et al. Learning to upsample by learning to sample[C]//2023 IEEE/CVF International Conference on Computer Vision. New York: IEEE, 2023: 6004-6014.

# 强干扰下焊缝图像线激光条纹的分割

何艺彦，李　艳*，徐　洋
东华大学 机械工程学院，上海 201620

**摘　要：**为了在强噪声和强反光等干扰下精确快速地分割出线激光条纹，提出一种基于 DeepLabv3+网络的轻量级焊缝线激光条纹分割网络（weld laser stripe segmentation network，WLS-Net）。为了提高网络的分割速度，选择浅层残差网络 ResNet-18 为主干网络。结合多头转置注意力（multi-Dconv head transposed attention，MDTA）与卷积块注意力模块（convolutional block attention module，CBAM），提出 MD-CBAM 注意力模块。选择动态上采样法（dynamic upsampling method，DySample）取代传统的双线性插值法，提高分割精度。针对焊缝线激光图像前景−背景类别不平衡的问题，选择 Dice 损失函数与逐像素交叉熵损失函数的加和函数作为模型的损失函数。实验结果表明，相比原始 DeepLabv3+网络，WLS-Net 网络使焊缝线激光数据集的交并比、像素精确率、F1 分数分别提升了 6.00、5.67、3.76 个百分点，并且单张图像推理时间减少了 18 ms。与其他语义分割网络相比，WLS-Net 网络也能取得较好的分割效果和较高的分割速率。

**关键词：**焊缝图像；激光条纹分割；DeepLabv3+网络；注意力机制