



Low-contrast X-ray image defect segmentation via a novel core-profile decomposition network

Xiaoyuan Liu^a, Jinhai Liu^{a,*}, Huanqun Zhang^b, Huaguang Zhang^a

^a College of Information Science and Engineering, Northeastern University, Shenyang, 110819, China

^b Shenyang Paidelin Technology Company Ltd., Shenyang, 110081, China

ARTICLE INFO

Keywords:

Defect segmentation
Low-contrast X-ray image
Core feature
Profile refinement
Quality control

ABSTRACT

Accurate X-ray image defect segmentation is of paramount importance in industrial contexts, as it is the foundation for product quality control and production safety. Deep learning (DL) has demonstrated powerful image scene understanding capabilities and has achieved unprecedented performance in defect segmentation tasks. However, existing DL methods suffer from significant performance degradation when facing low-contrast X-ray images, as the core information of defects is often obscured and the profile details are ambiguous. To address this issue, this paper explicitly decomposes the X-ray defect segmentation task into two subtasks: core feature learning and elasticity profile refinement, allowing task “serial” decomposition and performance “parallel” improvement. On this basis, a novel core-profile decomposition network (CPDNet) is developed to achieve accurate defect segmentation of X-ray images. Specifically, the core feature learning module is designed to construct the effective feature space from two views, discriminative and structural, to extract defect-related core features from X-ray images. Subsequently, the elasticity profile refinement module is developed to further improve the defect segmentation performance, which makes the first attempt to define the profile refinement as an out-of-distribution detection and leverage the elasticity score to refine the profile details at the pixel level. To fully evaluate the presented method, we conduct a series of experiments using two real-world X-ray defect datasets, and the results demonstrate that the CPDNet outperforms state-of-the-art methods.

1. Introduction

Welding technology plays a pivotal role in driving the development of the manufacturing industry (Le et al., 2024; Li et al., 2024b). During the welding process, affected by factors such as improper heat control, poor material quality, and inadequate joint preparation, some manufacturing defects will inevitably occur (Liu et al., 2023a; Li et al., 2024a). Welding defects can significantly compromise the strength and durability of structures (Wang et al., 2024), increase safety risks, and potentially lead to equipment failure or catastrophic accidents (Geng et al., 2024; Liu et al., 2024). Specific defects include porosity, crack, slag, and lack of penetration, as shown in Fig. 1(a). To effectively assess defects thereby eliminating safety hazards, the X-ray nondestructive testing (NDT) technology is commonly used to inspect internal defects within products without compromising their performance (Mashrafi et al., 2020; Liu et al., 2022), while providing X-ray data shown in Fig. 1(b) for detailed defect analysis (Zhang et al., 2024). As a core task in defect analysis, defect segmentation can precisely determine the position of defects at the pixel level, which plays a significant role in ensuring product quality in the manufacturing process (Sime et al.,

2023; Yu et al., 2020). However, X-ray images typically exhibit low-contrast, making it difficult to distinguish defects from the background. Additionally, fuzzy boundaries make texture, shape, and color features between the defect and the background unclear, which also causes shifted or incorrect defect locations. Therefore, defect segmentation based on X-ray images is an essential but challenging task. All along, studies on X-ray defect segmentation proliferated (Du et al., 2021; Yang et al., 2023a), and these studies have profound theoretical and engineering significance.

The traditional defect segmentation approaches focus on extracting features related to grayscale, color, texture, shape and spectral of images, thereby segmenting defects from the background. Representative methods include threshold segmentation (de Queiroz et al., 2000), edge detection (Kim and Lee, 2015), and feature statistics (Diniz et al., 2021). Threshold segmentation algorithms are only suitable for cases with a large grayscale difference between the defect and the background. Edge detection algorithms (e.g., Sobel and Canny) are sensitive to small pixel variations and susceptible to noise interference, and thus are unsuitable for images with complex backgrounds as well

* Correspondence to: No. 11, Lane 3, Culture Road, Heping District, Shenyang, Liaoning Province, China.

E-mail address: liujinhai@mail.neu.edu.cn (J. Liu).

<https://doi.org/10.1016/j.compind.2024.104123>

Received 21 March 2024; Received in revised form 12 June 2024; Accepted 4 July 2024

Available online 17 July 2024

0166-3615/© 2024 Elsevier B.V. All rights are reserved, including those for text and data mining, AI training, and similar technologies.

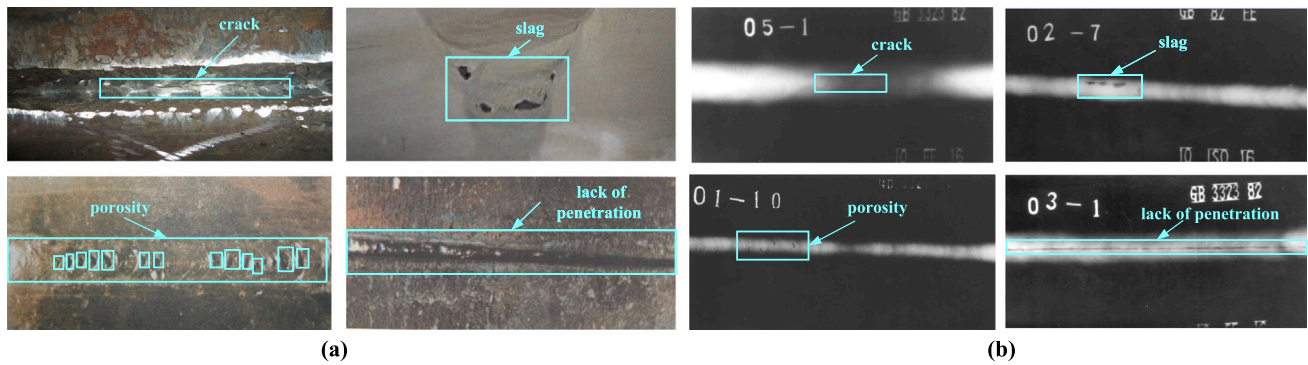


Fig. 1. Illustration of weld defects. (a) Physical images of weld defects. (b) X-ray images of weld defects.

as low-contrast. Further, whether the texture or color feature statistics, the segmentation result is merely a response to the underlying feature, lacking more high-level semantic relationship understanding. Therefore, the traditional methods are difficult to capture complex and high-level features, leading to limited performance in handling diverse and intricate image segmentation tasks.

Recently, deep learning (DL) approaches have made breakthroughs in various defect segmentation fields (Zhou et al., 2024; Yang et al., 2023b), such as solar cell defects (Wang et al., 2023), rail surface defects (Ma et al., 2023), material surface defects (Shan et al., 2023), etc. Based on the above excellent performance, DL approaches are considered to be a promising and reasonable solution for defect segmentation (Bao et al., 2021; Niu et al., 2023). Research on DL approaches includes two main directions (Yeung and Lam, 2023; Shen et al., 2023; Du et al., 2022): (i) region-level algorithms, and (ii) pixel-level algorithms. Region-level algorithms typically model and segment defects with simple bounding boxes, which can only provide a rough area of defect and cannot accurately describe defect boundaries, especially for low-contrast and complex defects (Liu et al., 2023b). In contrast, the pixel-level algorithms utilize the neural network to automatically learn and extract intricate features from X-ray images, which can provide more accurate segmentation results at the pixel level, and thus are more suitable for our X-ray defect segmentation task. In Shelhamer et al. (2017), by replacing the fully-connected layers in convolutional neural networks (CNN) with an up-sampling layer, the fully convolutional neural network (FCN) was proposed, which was the cornerstone of deep learning applied to semantic segmentation. Due to the poor performance of FCN in segmenting multi-scale objects, PSPNet, Deeplabv3, and Deeplabv3+ were proposed to improve the segmentation performance (Chen et al., 2018a). Moreover, U-Net (Ronneberger et al., 2015) was proposed which was a breakthrough progress of deep learning in medical image segmentation. Following U-Net, UNet++, Attention U-Net and MultiRseUNet were proposed to further improve the robustness in defect segmentation. Despite the remarkable performance that has been achieved with the above methods, due to the low-contrast characteristics of X-ray images in defect segmentation, they still suffer from two issues: (1) They are difficult to extract core features accurately as they may be hidden or missed (Zhu et al., 2020; Lin et al., 2022). (2) They are only suitable for datasets with clear profiles and struggle to capture the ambiguous profile details in X-ray datasets (Mei et al., 2021; Ni et al., 2022).

To solve the first issue, the key question is reconsidered as: “How to create a good core feature extraction space for X-ray images?” Ideally, it can not only extract discriminative features, but is also well structured for intra-class compactness and inter-class dispersion. Therefore, the feature space is constructed from two views: discriminative and structural. To achieve discriminative, we expect to convert “expert knowledge” into detailed natural language information, thus providing more flexible and comprehensive descriptions of defects. It is different from previous studies where semantic information was limited to

abstract instance relations or labeling information (Rong et al., 2023), leaving the feature extraction process divorced from the guidance of X-ray domain knowledge. To overcome this issue, a semantic-driven mining module is designed that tightly integrates the feature extraction process with “expert knowledge” in the X-ray domain, thereby learning discriminative features from the low-contrast X-ray images. On this basis, contrastive learning is introduced to achieve structural. Unlike previous studies (He et al., 2020; Khosla et al., 2020), we mine the similarity between different pixels, and develop soft confidence for similarity measurement between anchor and positives, for which a soft-contrast gathering method is designed to pull similar pixels closer and push dissimilar pixels farther, so that the feature space can be structured.

To solve the second issue, we tried to devise an accurate profile refinement method to capture the X-ray defect profile details. Previous works (Yu et al., 2012; Chaibou et al., 2016) have focused on searching for similar pixels in neighboring regions, but this approach has its limitations as neighboring pixels could also be mislabeled simultaneously. For this, we consider the overall pixel distribution and define profile refinement as an out-of-distribution detection problem. Inspired by Joseph et al. (2021), the Energy-Based Models (EBMs) can assign low energy values to in-distribution data and vice versa, which motivates us to use the energy metric to determine whether a pixel is from the background. Therefore, an elasticity profile refinement module is designed, which leverages the elasticity score to refine the profile details at the pixel level and introduces EBMs to achieve the coarse-to-fine retrieval of the profile, further accomplishing accurate defect segmentation. This is the first attempt to define defect profile refinement as an out-of-distribution detection problem, and the proposed method excels at defect segmentation of low-contrast X-ray images.

Based on the above analysis, a core-profile decomposition network (CPDNet) is proposed to achieve accurate defect segmentation of X-ray images. Specifically, the CPDNet is decomposed into two stages: core feature learning and elasticity profile refinement, the two stages achieve defect segmentation in a “serial” manner, but achieve performance improvement in a “parallel” manner. To the best of our knowledge, this paper is the first to explicitly decompose core feature and profile feature, and is driven by X-ray practical problems to achieve accurate defect segmentation. The main contributions include:

1. A CPDNet decomposed into core feature learning and elasticity profile refinement is proposed, which can achieve accurate defect segmentation of X-ray images.
2. The core feature learning module is designed to create an effective feature space from two views, discriminative and structural, so as to extract distinctive features from X-ray images by semantic-driven mining and soft-contrast gathering, respectively, thus deriving powerful feature learning ability for core defect segmentation.

3. The elasticity profile refinement module is developed, which leverages the elasticity score to refine the profile features of defects at the pixel level, thus the accuracy of defect segmentation can be effectively improved.

2. Data acquisition and typical X-ray defects

During the welding process①, multiple defects are formed by various factors such as poor material quality, incorrect welding techniques or adverse environmental conditions, etc. X-ray radiographic inspection② utilizes X-rays to penetrate the workpiece and assess the internal defects, structure or integrity by receiving and recording defect information on film. Subsequently, the X-ray films are converted into X-ray digital images using film scanning technology③ for more detailed processing and analysis. The sample labeling process in the early stages necessitates significant manual involvement and expertise, typically provided by professionals, to enhance label quality and set the groundwork for intelligent image assessment. In the practical scenario, different types of defects have different damage to the weld, to better assess the quality and reliability of the weld, several typical defect types④ are analyzed: *crack defect (CR)* is slender in shape and variable in scale, as the severest type of defects can lead to weld breakage; *round defect (RD)* with an aspect ratio of less than 3:1, which reduce the strength of welds and increase the risk of cracking; *bar defect (LD)* increase the brittleness of the weld, thus reducing its durability, and its aspect ratio is typically more than 3:1; *lack of penetration (LP)* occurs in the middle of the weld and appears as a continuous or intermittent linear or band-like dark line. Fig. 2 shows the above typical defect X-ray images, from which it can be seen that the X-ray images are low-contrast where defects are difficult to recognize, which creates great challenges for X-ray defect segmentation: (1) The low-contrast causes small grayscale differences between the defect and the background, causing some core features of defects to be hidden or missed, and therefore difficult to extract accurately. (2) The low-contrast leads to blurring of defect edges, making it tough to accurately capture the profile information, thus reducing the defect segmentation accuracy.

3. Method

In this section, we present an accurate X-ray defect segmentation architecture, the CPDNet, which is decomposed into the core feature learning and elasticity profile refinement, as shown in Fig. 3. Within the core feature learning, distinctive defect-related features are initially extracted via a semantic-driven module, subsequently improved for intra-class compactness and inter-class dispersion by a soft-contrast gathering module, and ultimately input into the decoder to generate the core segmentation masks. During the elasticity profile refinement, the elasticity score of each pixel is first calculated from the overall pixel energy value distribution, which is then used for detailed profile refinement of the defects, leading to an updated defect segmentation mask with improved accuracy. Further details are provided in the following subsections.

3.1. Core feature learning

Core feature extraction is essential for understanding defect characteristics involving grayscale, gradient, texture, shape, etc. Considering the small blackness difference between defect and background may cause core features to be hidden or lost during extraction, we introduce a semantic-driven mining module for the discriminative view of the feature space and a soft-contrast gathering module for its structural view, to effectively extract core features for defect segmentation.

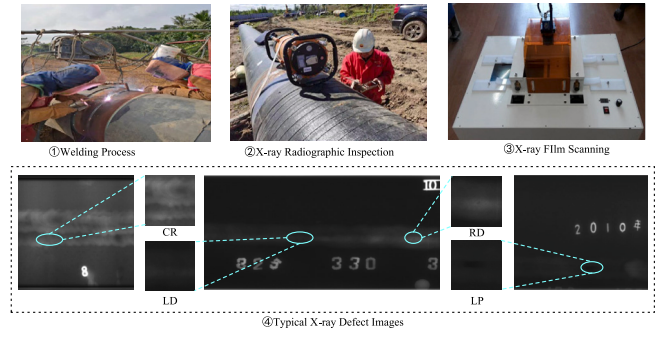


Fig. 2. Typical X-ray defect image acquisition.

3.1.1. Semantic-driven mining

Visual features can provide more abundant and distinctive descriptions for specific scenes compared to semantic features, but for low-contrast X-ray images, visual features are often plagued by spurious features such as background clutter and noise interference. Consequently, relying solely on visual features is insufficient for effective X-ray image defect segmentation. In contrast, semantic features contain relationship information that provides contextual cues, aiding the model in understanding the shape and position of defects within an image. However, it cannot accurately segment defects due to the lack of interaction with visual features, so individual semantic information has limited benefit for defect segmentation. Facing this dilemma, we are inspired by the “cognitive penetration” of humans (Maier and Abdel Rahman, 2019), and expect to use semantic features as priors to adjust the ongoing visual perception, promoting the accurate X-ray defect segmentation.

(a) *Semantic Construction.* The semantic descriptions are derived from “expert knowledge”, which are natural language descriptions of defect shape, scale and blackness information. To transfer “expert knowledge” into our model training, it is essential to collect the accurate and comprehensive descriptions associated with various defect types. In this work, domain rules established by experts and perpetually employed serve as important reference, and the semantic description of defect X-ray images is summarized in Fig. 4.

(b) *Semantic Driving.* Given dataset $D = \{(I_i, S_i, \mathcal{Y}_i)\}_{i=1}^N$, where (I_i, S_i) is an image-semantic pair, \mathcal{Y}_i denotes the true label, and N denotes the number of samples, the image encoder $\psi_\theta(\cdot)$ and the semantic encoder $\gamma_\alpha(\cdot)$ are designed to extract features. First, the defect semantics are encoded as $g^m = \gamma_\alpha(S)$, $g^m \in \mathbb{R}^m$ based on a pre-trained text encoder, simultaneously, the X-ray images are fed into the image encoder to obtain the image feature map $\psi_\theta(I) \in \mathbb{R}^{c \times w \times h}$, where w , h and c denote the width, height and feature channel, respectively. Then, to fully fuse the semantic prior to facilitate the image feature extraction, image feature map $\psi_\theta(I)$ is compressed into a feature $f^m = \hat{h}_\delta(\psi_\theta(I))$, $f^m \in \mathbb{R}^m$ to guarantee the same dimension as the semantic feature, where compressing network $\hat{h}_\delta(\cdot) = \text{Averagepooling} + \text{Dropout} + \text{FC} + \text{BN}$. Next, the prior semantic features are fused into the image features by the convex combination, the updated image features can be computed as:

$$\tilde{f}^m = \kappa f^m + (1 - \kappa) g^m, \quad (1)$$

where κ is the coefficient generated by the adaptive mixing network $X(\cdot)$ with parameter ϵ , and computed as follows:

$$\kappa = \frac{1}{1 + \exp(-X(g^m, \epsilon))}, \quad (2)$$

Finally, to better capture defect features at different scales and effectively segment defects of different sizes, the atrous spatial pyramid pooling (ASPP) network (Chen et al., 2018b) is followed by the image encoder since it can acquire feature information under different

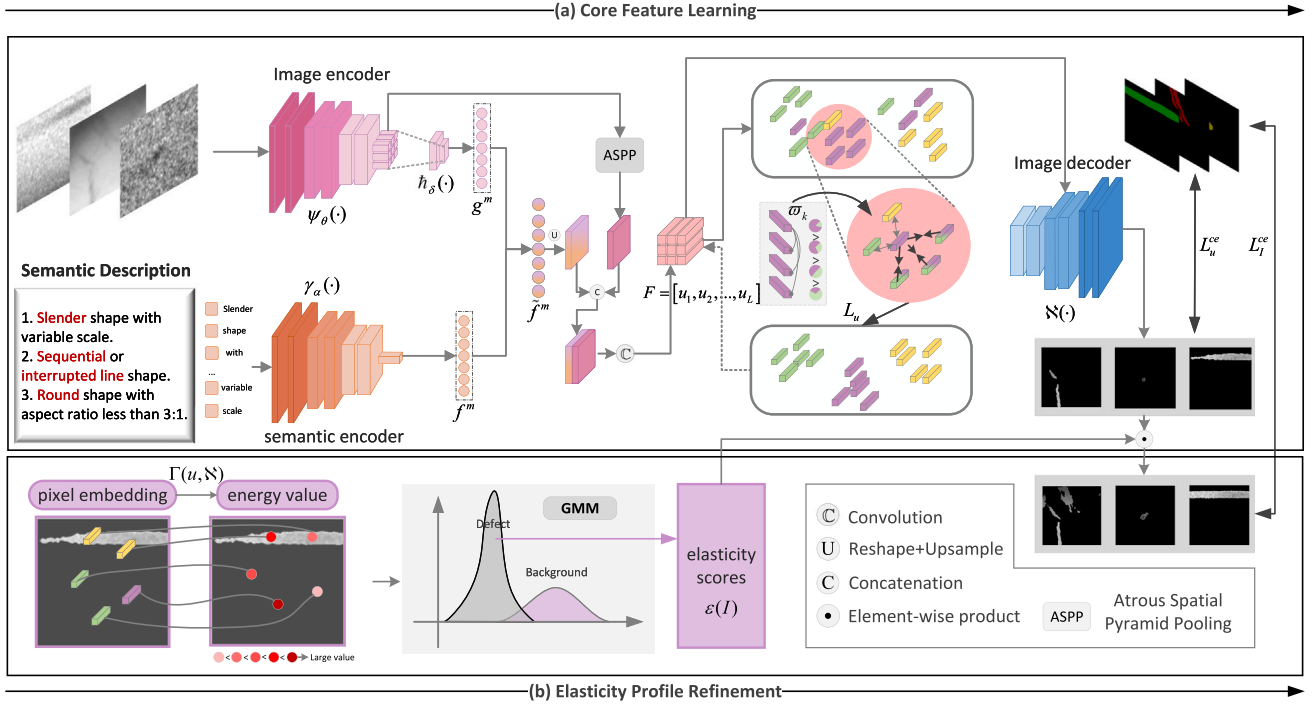


Fig. 3. Architecture of proposed CPDNet. In the core feature learning process, the features are first extracted using a semantic-driven module, then the intra-class compactness and inter-class dispersion of these features are enhanced by a soft-contrast gathering module, and finally fed into a decoder to generate a core segmentation mask. In the elasticity profile refinement process, the elasticity score of each pixel is first calculated based on the overall pixel energy distribution, which is then used to refine defect profiles in detail, resulting in an updated defect segmentation mask with higher accuracy.

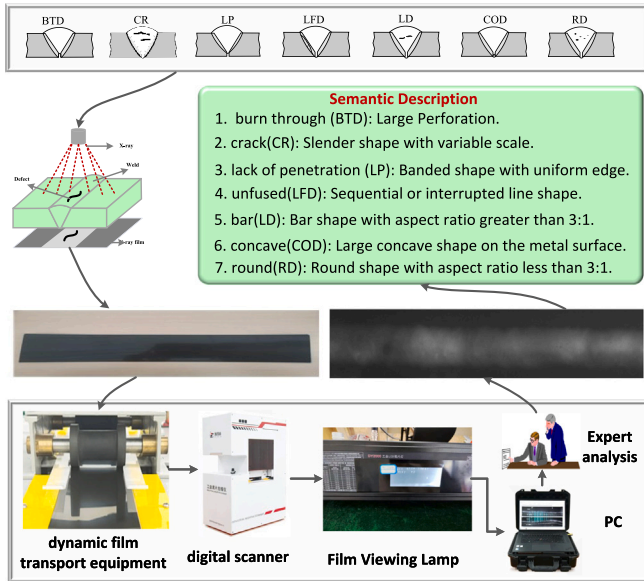


Fig. 4. Semantic description construction of defect X-ray images.

receptive fields by using multi-scale cavity convolution, and its output feature map is concatenated and convolved with the reshaped and upsampled \tilde{f}^m to generate the feature map:

$$F = \mathbb{C} [\text{cat}(\mathbb{U}(\tilde{f}^m), \mathbb{A}(\psi_\theta(I)))], \quad (3)$$

where $F \in \mathbb{R}^{c \times w \times h}$, \mathbb{U} denotes the reshape layer followed by an upsampling layer, \mathbb{A} denotes the ASPP network, $\text{cat}(\cdot)$ represents the concatenation operation, and \mathbb{C} represents the convolutional layer.

3.1.2. Soft-contrast gathering

With semantic-driven mining, effective image features related to defect areas are embedded into the feature space, while they still suffer from perturbation, which arises from the intra-class defects being different while inter-class defects are similar, causing some pixels to be incorrectly predicted. To eliminate this impact, we develop a soft-contrast gathering module to enhance intra-class compactness and inter-class separability, so that extracted core features can be structured for accurate defect segmentation.

First, the pixel embedding $u \in \mathbb{R}^c$ is derived from the dense feature $F \in \mathbb{R}^{c \times w \times h}$ (i.e., $u \in F$). Given the pixel $\{u_j, y_j\}_{j=1}^L$, where $L = w \times h$, y_j denotes each pixel label, from which the anchor u , positives \mathfrak{P}_u from the same class and negatives \mathfrak{Y}_u from different classes are selected. Existing studies based on contrastive learning directly utilize labels as supervision for feature optimization (He et al., 2020; Khosla et al., 2020), without mining the similarity relationships between samples, which greatly increases the risk of misclassification. To address this limitation, we devise a soft confidence to assess the similarity between an anchor and positives, which is formulated as follows:

$$\varpi_k = \frac{1}{\chi_k} \left(e^{u \cdot u_{(k)}^+} / \sum_{k=1}^{|\mathfrak{P}_u|} e^{u \cdot u_{(k)}^+} \right) \quad (4)$$

where $u_{(k)}^+ \in \mathfrak{P}_u$, and χ_k is a scaling factor to adjust ϖ_k . By utilizing the derived soft confidence, we can integrate similarity rankings into the loss function so that the soft-contrast gathering module can be well-trained, thus enhancing the intra-class compactness and inter-class separability at the pixel level:

$$\mathcal{L}_u = -\frac{1}{|\mathfrak{P}_u|} \log \frac{\sum_{u_{(k)}^+ \in \mathfrak{P}_u} \varpi_k \exp(\frac{u \cdot u_{(k)}^+}{\tau})}{\sum_{u_{(k)}^+ \in \mathfrak{P}_u} \exp(\frac{u \cdot u_{(k)}^+}{\tau}) + \sum_{u^- \in \mathfrak{Y}_u} \exp(\frac{u \cdot u^-}{\tau})} \quad (5)$$

where τ is a temperature parameter. During the training of the soft-contrast gathering module, similarity mining has the following significance. First, high similarity pixels will be prioritized, thus accelerating

the model convergence. Second, pixels with little similarity may be noisy or anomalous, removing them from positives can eliminate the impact on model training. Third, the similarity order can be memorized by the model, thus guiding the model to select more helpful samples in the subsequent training.

With extracted core features, each pixel u would be categorized into a specific class $y \in \mathcal{Y}$ with the image decoder $\aleph(\cdot)$, which maps dense feature F into a categorical score map $\hat{\mathcal{Y}} = \aleph(F) \in \mathbb{R}^{q \times w \times h}$, and q denotes the number of categories. Further let $\hat{y} \in [\hat{y}_1, \hat{y}_2, \dots, \hat{y}_q] \in \mathbb{R}^q$ be the score vector for pixel u , derived from \mathcal{Y} , i.e., $\hat{y} \in \hat{\mathcal{Y}}$, the pixel-level cross-entropy loss can be optimized with softmax:

$$\mathcal{L}_u^{ce} = -\mathbf{1}_y^T \log(\text{softmax}(\hat{y})). \quad (6)$$

3.2. Elasticity profile refinement

With core feature learning, the core part of the defect can be segmented from X-ray images, yet the profile of the defect is still not finely segmented, which impedes the improvement of segmentation accuracy. Inspired by Joseph et al. (2021), we propose to formulate the background area (except defects) recognition as an out-of-distribution detection problem, and it is also the first attempt in the defect segmentation field. To achieve this, the elasticity profile refinement module is designed to refine the defect profile with elasticity scores, whereby the defect segmentation accuracy is further enhanced.

Compared to the traditional and intuitive softmax score, existing studies suggest that Energy-based models (EBMs) can more reliably distinguish ambiguous profile pixels, as EBMs tend to assign low energy values to defective areas and vice versa. Consequently, we opt to introduce the EBMs to determine whether a pixel is from the background area. Given the learned core pixel embedding $u \in \mathbb{R}^c$ and its corresponding pixel label y , we expect to learn a function $\Gamma(u) : \mathbb{R}^c \rightarrow \mathbb{R}$ to estimate the compatibility between u and y . Firstly, the Helmholtz free energy $\Gamma(u)$ of u is formulated as:

$$\Gamma(u) = -\zeta \log \int_{\mathcal{Y}} \exp\left(-\frac{\Gamma(u, y')}{\zeta}\right). \quad (7)$$

where ζ is the temperature parameter. According to the connection relation between the softmax score and the energy score (Joseph et al., 2021), the category probability $p(y|u)$ of label y is obtained:

$$p(y|u) = \frac{\exp(\frac{\aleph_y(u)}{\zeta})}{\sum_{e=1}^q \frac{\aleph_e(u)}{\zeta}} = \frac{\exp(-\frac{\Gamma(u, y)}{\zeta})}{\exp(-\frac{\Gamma(u)}{\zeta})}. \quad (8)$$

where $\aleph_y(u)$ indicates the y th index of image decoder $\aleph(u)$, $\Gamma(u, y)$ represents the energy score of y th category. Accordingly, it can be derived that $\Gamma(u, y) = -\aleph_y(u)$, substituting it into Eq. (7), the energy value $\Gamma(u, \aleph)$ of pixel u can be obtained:

$$\Gamma(u, \aleph) = -\zeta \log \sum_{y=1}^q \exp\left(\frac{\aleph_y(u)}{\zeta}\right). \quad (9)$$

Secondly, we can gather the energy values from all the pixels within the image and suggest modeling the distribution of these energy values using a spherical Gaussian mixture model (GMM). Given one particular class $y = k$, to determine whether a pixel belonged to the background or the defect, all pixels within this class are divided into two types: defect-area pixels and background-area pixels. Therefore, the mixed Gaussian distribution of class k is defined as:

$$\mathcal{G}(u_{y=k}) = \sum_{t=0}^1 \rho(t) \mathcal{N}(\Gamma(u, \aleph) | t). \quad (10)$$

where $t = 0$ and $t = 1$ refer to the Gaussian distribution of pixel energy values of the defect-area and background-area, respectively. $\rho(\cdot)$ represents the weight of the Gaussian distribution $\mathcal{N}(\cdot)$, and $\mathcal{G}(u_{y=k})$ denotes the generated mixed Gaussian distribution when the pixel u is labeled k . After modeling the mixed Gaussian distribution $\mathcal{G}(u_{y=k})$, the

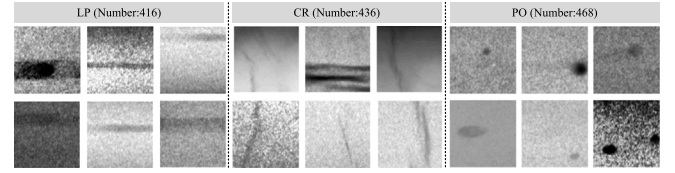


Fig. 5. Examples of RIAWELC dataset.

elasticity scores for pixels within the defect area \mathbb{D} can be determined:

$$\varepsilon(\mathbb{D} | u)^k = \Lambda(\mathcal{G}(u_{y=k})) \quad (11)$$

where $\Lambda(\cdot)$ represents the probability prediction function for the defect area. For the q pixel categories in the X-ray image, the whole elasticity scores $\varepsilon(I)$ can be collected:

$$\varepsilon(I) = \varepsilon(\mathbb{D} | u)^1 \cup \varepsilon(\mathbb{D} | u)^2 \dots \cup \varepsilon(\mathbb{D} | u)^k \cup \varepsilon(\mathbb{D} | u)^q. \quad (12)$$

where q represents the number of categories. With the elasticity score $\varepsilon(I) \in \mathbb{R}^{q \times w \times h}$ of X-ray image I , to refine the defect profile to achieve accurate defect segmentation, the categorical score map is predicted and updated:

$$\hat{\mathcal{Y}}' = \text{argmax}(\hat{\mathcal{Y}}) \times \varepsilon(I). \quad (13)$$

where $\hat{\mathcal{Y}} \in \mathbb{R}^{q \times w \times h}$ represents the categorical score map, $\text{argmax}(\cdot)$ is used to predict the category of each pixel, and $\hat{\mathcal{Y}}'$ is the updated prediction for the X-ray image mask. Finally, the image-level cross-entropy loss can be adopted to optimize defect segmentation:

$$\mathcal{L}_I^{ce} = ce(\hat{\mathcal{Y}}', \mathcal{Y}). \quad (14)$$

Based on the above calculation, the total loss can be obtained to train the whole CPDNet, so that defects can be accurately segmented from X-ray images:

$$\mathcal{L}_{total} = \mathcal{L}_u + \zeta \mathcal{L}_u^{ce} + \mu \mathcal{L}_I^{ce}. \quad (15)$$

where ζ and μ represent the trade-off parameters. So far, defects have been accurately segmented through core feature learning followed by elasticity profile refinement, while segmentation performance has been improved simultaneously.

4. Experiment and discussion

4.1. Experimental setting

4.1.1. Dataset

To assess the effectiveness of our proposed CPDNet, we utilize the RIAWELC dataset (Perri et al., 2023) and a pipeline dataset sourced from real industrial manufacturing settings. Both datasets contain low-contrast X-ray images for experimental evaluation.

The RIAWELC dataset is an open-access collection of X-ray weld images, obtained using Lincoln Electric's Power Wave[®] advanced welding platform in combination with submerged arc welding technology and a COMET EVO X-ray machine. A total of 1320 X-ray images have been collected and annotated into four categories: lack of penetration (LP), cracks (CR), porosity (PO) and no defect (ND), some defect examples along with sample sizes are presented in Fig. 5. For the defect segmentation task, the ratio of the training set, validation set, and testing set for RIAWELC is 8:1:1. It can be seen from Fig. 5 that defects are variable and complex in shape and size, and the complex background of many materials makes defects difficult to be segmented. Furthermore, the low-contrast characteristics presented in X-ray images further increase the difficulty of defect segmentation.

The pipeline dataset was collected from X-ray images derived from actual pipeline welds in northern China, and the collection scenarios

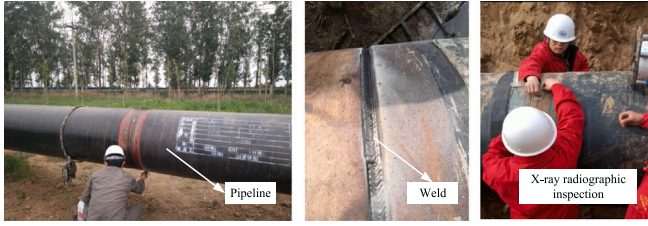


Fig. 6. Pipeline dataset collection scenarios.

Table 1
Detailed Structures of Encoders.

Encoder/Input	Layer	Output
$\psi_\theta(\cdot)$ $3 \times 227 \times 227$	Conv2d, kernel size=(3, 3), stride=(2, 2), padding=(1, 1)	$32 \times 114 \times 114$
	Inverted residual block	$16 \times 114 \times 114$
	Inverted residual block	$24 \times 57 \times 57$
	Inverted residual block	$24 \times 57 \times 57$
$\psi_\theta(\cdot)$ $3 \times 180 \times 180$	Conv2d, kernel size=(3, 3), stride=(2, 2), padding=(1, 1)	$32 \times 90 \times 90$
	Inverted residual block	$16 \times 90 \times 90$
	Inverted residual block	$24 \times 45 \times 45$
	Inverted residual block	$24 \times 45 \times 45$
$\gamma_\alpha(\cdot)$	Embedding layer	77×512
	Residual Attention Block+Reshape	512

are shown in Fig. 6. These images were captured using an X-ray radiography inspection system, which is adept at translating potential defect information within the welds into X-ray images, and the system details are demonstrated in Section 2. Considering that cracks are the most dangerous defect increasing the pipeline rupture risk, and that bar and round defects as the commonest defects offering substantial information about pipeline health, the crack defect (CR), round defect (RD) and bar defect (LD) are adopted for experiments. The pipeline dataset consists of 609 X-ray images, with a distribution of 157 RD, 212 LD, and 240 CR images, allocated in an 8:1:1 ratio for the training, validation, and testing sets respectively.

4.1.2. Implementation details

All experiments are performed using the PyTorch toolbox on a PC equipped with Intel Core i7-10700 CPU @3.80 GHz processor and NVIDIA GeForce RTX 2070 SUPER. The image encoder of the CPDNet is based on pre-trained MobileNetV2 due to its light weight and proven performance. The semantic encoder is initialized with the CLIP model pre-trained on large-scale corpora and is available for public use, and only the text encoder of CLIP is introduced for semantic feature extraction. The detailed structure of the two encoders is shown in Table 1. We use the stochastic gradient descent (SGD) optimizer with an initial learning rate of 0.007, where the momentum and weight decay are set as 0.9 and 0.0001, respectively. We run 200 epochs on two datasets and set the batch size to 8. The image sizes of RIAWELC and pipeline datasets are 227×227 and 180×180 , respectively.

4.1.3. Evaluation metrics

Following the segmentation work in Shen et al. (2023), Niu et al. (2022), we adopt the Precision (Pre), Recall (Rec), F1-measure (F1), and intersection over union (IoU) to evaluate segmentation performance, and they are defined as follows:

$$\text{Pre} = \frac{TP}{TP + FP}, \text{Rec} = \frac{TP}{TP + FN}. \quad (16)$$

$$\text{F1} = \frac{2 \times \text{Pre} \times \text{Rec}}{\text{Pre} + \text{Rec}}, \text{IoU} = \frac{TP}{TP + FP + FN}. \quad (17)$$

where TP, FP, TN, and FN represent the number of true positive, false positive, true negative, and false negative. Based on this, mIoU

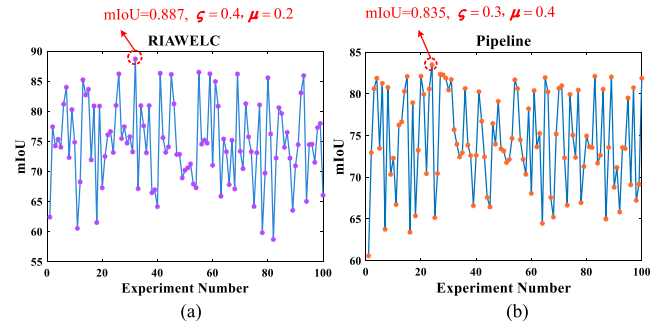


Fig. 7. Results of trade-off parameters analysis, and the red circles indicate the highest mIoU achieved on the two datasets. Due to limited space, only the results with the highest mIoU are circled. (a) results on RIAWELC dataset. (b) results on pipeline dataset. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

is calculated as the average of IoU. A higher mIoU value indicates that the model can accurately segment each class and has a high overlap with the ground truth.

4.1.4. Trade-off parameters analysis

To investigate the effect of parameters ζ and μ on the performance of the proposed CPDNet, we search the ζ and μ in $[0.1, 0.2, 0.3, \dots, 1.0]$, and performing 100 groups of experiments on each dataset to find the optimal value. Based on the experimental results shown in Fig. 7, it can be found that the 32nd experiment on the RIAWELC dataset achieved the optimal mIoU value, where mIoU was 0.887, ζ was 0.4, and μ was 0.2. For pipeline dataset, the 24th experiment achieved the highest mIoU of 0.835, with the values of ζ and μ set to 0.3 and 0.4, respectively.

4.2. Performance comparison

To evaluate the effectiveness and performance of the CPDNet, comparative studies between the proposed method and eight excellent methods are performed with weld X-ray images from the RIAWELC dataset and pipeline dataset. Concretely, the comparison methods include: (1) attention-guided segmentation network (AGSN) (Yang et al., 2022); (2) two-stream CNN segmentation architecture (TSCNN) (Du et al., 2021); (3) semantic information decomposition network (SIDN) (Yang et al., 2023a); (4) shape-consistent one-shot unsupervised domain adaptation framework (SC-OSDA) (Ma et al., 2023); (5) optimal bilateral feature transport network (OBFTNet) (Shan et al., 2023); (6) triplet-graph reasoning network (TGRNet) (Bao et al., 2021); (7) fully convolutional network (FCN) (Shelhamer et al., 2017) and (8) DeepLabv3+ network (DeepLabv3+) (Chen et al., 2018a).

4.2.1. Segmentation results of RIAWELC dataset

The experiment results are shown in Table 2, from which it can be noted that the proposed method achieves optimal experimental results on the average of four evaluation metrics compared to the other eight comparative methods. Specifically, for the eight comparison methods, the average values of four evaluation metrics are as follows: Pre ranges from 87.6% to 91.4%, Rec ranges from 86.6% to 91.8%, F1 ranges from 87.1% to 91.6%, and IoU ranges from 78.6% to 85.1%. On this basis, the mean value of our method in Pre, Rec, F1, and IoU has reached 93.5%, 94.2%, 93.9%, and 88.7%, which is an improvement of 3.51%, 3.65%, 3.66%, and 5.60%, respectively. Furthermore, it can be noted that the TSCNN method achieves suboptimal segmentation results, to compare the segmentation performance more intuitively, the segmentation results of three defective X-ray images from the RIAWELC dataset are visualized, which is shown in Fig. 8. Despite facing low-contrast X-ray defects that even make labeling difficult, the proposed

Table 2

X-ray defect segmentation results on RIAWELC dataset.

Method	Pre					Rec					F1					IoU				
	Normal	LP	CR	PO	Mean.	Normal	LP	CR	PO	Mean.	Normal	LP	CR	PO	Mean.	Normal	LP	CR	PO	Mean.
AGSN (Yang et al., 2022)	98.7%	90.7%	72.9%	88.2%	87.6%	98.8%	92.1%	67.5%	88.1%	86.6%	98.8%	91.4%	70.1%	88.2%	87.1%	97.5%	84.1%	54.0%	78.8%	78.6%
TSCNN (Du et al., 2021)	99.2%	92.7%	80.9%	92.8%	91.4%	99.1%	92.9%	80.8%	94.2%	91.8%	99.1%	92.8%	80.8%	93.5%	91.6%	98.3%	86.6%	67.9%	87.7%	85.1%
SIDN (Yang et al., 2023a)	99.1%	91.4%	79.4%	90.7%	90.2%	99.0%	93.3%	79.7%	93.1%	91.3%	99.0%	92.3%	79.6%	91.9%	90.7%	98.1%	85.8%	66.1%	85.0%	83.8%
SC-OSDA (Ma et al., 2023)	99.1%	93.4%	81.3%	91.9%	91.4%	99.1%	92.3%	79.8%	95.2%	91.6%	99.1%	92.9%	80.6%	93.5%	91.5%	98.3%	86.7%	67.4%	87.7%	85.0%
OBFTNet (Shan et al., 2023)	99.1%	92.9%	81.1%	92.1%	91.3%	99.1%	93.2%	79.7%	92.5%	91.1%	99.1%	93.1%	80.4%	92.3%	91.2%	98.2%	87.1%	67.2%	85.6%	84.5%
TGRNet (Bao et al., 2021)	99.1%	92.4%	81.9%	91.2%	91.2%	99.1%	93.3%	77.3%	95.5%	91.3%	99.1%	92.8%	79.5%	93.3%	91.2%	98.2%	86.6%	66.0%	87.4%	84.6%
FCN (Shelhamer et al., 2017)	99.0%	93.0%	74.6%	84.4%	87.8%	98.7%	90.0%	78.4%	94.2%	90.3%	98.9%	91.5%	76.5%	89.0%	89.0%	97.7%	84.3%	61.9%	80.2%	81.0%
DeepLabv3+ (Chen et al., 2018a)	99.1%	91.0%	77.8%	88.0%	89.0%	98.8%	94.1%	74.2%	94.3%	90.4%	99.0%	92.5%	75.9%	91.1%	89.6%	98.0%	86.1%	61.2%	83.6%	82.2%
Proposed	99.4%	93.3%	87.1%	94.1%	93.5%	99.3%	94.2%	88.3%	94.8%	94.2%	99.4%	93.8%	87.7%	94.5%	93.9%	98.8%	88.2%	78.1%	89.5%	88.7%

The bold values represent the optimal experimental results, and the underlined values represent suboptimal experimental results.

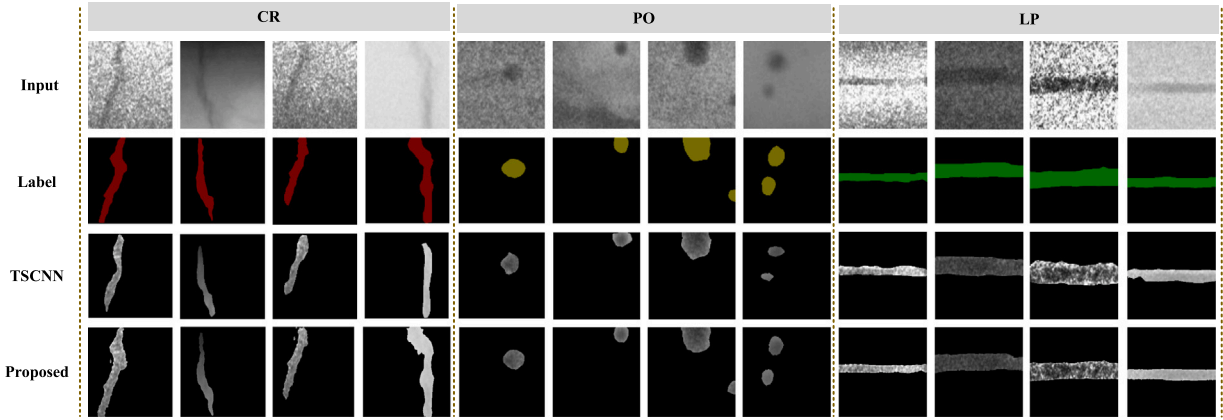


Fig. 8. Segmentation results of RIAWELC dataset.

Table 3

X-ray defect segmentation results on pipeline dataset.

Method	Pre					Rec					F1					IoU				
	Normal	RD	LD	CR	Mean.	Normal	RD	LD	CR	Mean.	Normal	RD	LD	CR	Mean.	Normal	RD	LD	CR	Mean.
AGSN (Yang et al., 2022)	99.8%	83.5%	75.4%	68.1%	81.7%	99.6%	90.2%	86.0%	86.0%	90.5%	99.7%	86.7%	80.4%	76.0%	85.7%	99.4%	76.5%	67.2%	61.3%	76.1%
TSCNN (Du et al., 2021)	99.9%	84.8%	82.1%	68.9%	83.9%	99.6%	90.8%	91.1%	89.8%	92.8%	99.7%	87.7%	86.3%	78.0%	87.9%	99.5%	78.1%	75.9%	63.9%	79.4%
SIDN (Yang et al., 2023a)	99.8%	84.5%	78.7%	72.1%	83.8%	99.6%	87.0%	92.7%	85.8%	91.3%	99.7%	85.7%	85.1%	78.4%	87.2%	99.4%	75.0%	74.1%	64.4%	78.2%
SC-OSDA (Ma et al., 2023)	99.9%	86.0%	82.4%	68.6%	84.2%	99.5%	88.2%	93.1%	89.6%	92.6%	99.7%	87.1%	87.4%	77.7%	88.0%	99.4%	77.1%	77.7%	63.6%	79.5%
OBFTNet (Shan et al., 2023)	99.9%	86.6%	82.4%	66.9%	84.0%	99.6%	90.1%	91.5%	92.9%	93.5%	99.7%	88.3%	86.7%	77.8%	88.1%	99.5%	79.1%	76.6%	63.6%	79.7%
TGRNet (Bao et al., 2021)	99.7%	85.9%	81.2%	73.5%	85.1%	99.7%	82.4%	86.5%	81.2%	87.5%	99.7%	84.1%	83.8%	77.2%	86.2%	99.4%	72.6%	72.0%	62.8%	76.7%
FCN (Shelhamer et al., 2017)	99.8%	81.3%	78.8%	67.4%	81.8%	99.6%	86.6%	87.6%	80.4%	88.6%	99.7%	83.8%	83.0%	73.3%	85.0%	99.3%	72.2%	70.9%	57.9%	75.1%
DeepLabv3+ (Chen et al., 2018a)	99.7%	81.3%	82.1%	67.8%	82.7%	99.6%	86.6%	84.1%	76.3%	86.7%	99.7%	83.9%	83.1%	71.8%	84.6%	99.3%	72.2%	71.1%	56.0%	74.7%
Proposed	99.9%	87.4%	88.9%	76.6%	88.2%	99.7%	92.2%	93.9%	87.4%	93.3%	99.8%	89.7%	91.3%	81.7%	90.6%	99.6%	81.3%	84.0%	69.0%	83.5%

The bold values represent the optimal experimental results, and the underlined values represent suboptimal experimental results.

method can still accurately segment defects from complex backgrounds, which is attributed to our designed CPDNet that achieves accurate segmentation in a decomposed way through core feature learning and elasticity profile refinement.

4.2.2. Segmentation results of pipeline dataset

Table 3 indicates that the means of our method outperforms comparative approaches in terms of Pre, F1, and IoU, and are higher by 4.80%, 4.01%, and 6.07% on average for each evaluation metric. However, it is slightly less effective in terms of Rec, trailing by 0.2% behind the optimal OBFTNet. This suboptimal performance can be attributed to the inherent characteristics of pipeline crack defects, which often have irregular, elongated shapes and diverse orientations, complicating accurate detection. To further analyze the segmentation results on mIoU, we visualize the defect segmentation results to compare our method with the suboptimal OBFTNet method, which are shown in Fig. 9. It can be found that our CPDNet can accurately segment various realistic defective images, such as round, bar, and crack. The performance advantage benefits from the CPDNet, which can derive powerful core feature learning ability and implement profile refinement with elasticity score, thus the accuracy of defect segmentation can be significantly improved.

4.3. Ablation studies

4.3.1. Ablation study of CPDNet

This experiment is conducted on the RIAWELC dataset to evaluate the effectiveness of each component in CPDNet. Three ablation methods are used to make comparisons: CPDNet without semantic driving (CPDNet without SD), CPDNet without soft-contrast gathering (CPDNet without SG), and CPDNet without elasticity profile refinement (CPDNet without EPR). All variants of CPDNet use the same parameter setting to guarantee a fair comparison. The experimental results are shown in Table 4. Observations reveal that the removal of SD or SG from the core feature learning, the mIoU decreases by 2.4% and 4.2%, respectively, and the decrease in IoU is most pronounced for CR defects compared to LP and PO, dropping by 6.9% and 10.9%. Furthermore, as a key step in profile refinement, the results in Table 4 reflect that without EPR, the mIoU decreases by 1.6%, which implies that solely focusing on core feature learning without refining the profile cannot lead to accurate defect segmentation.

4.3.2. Ablation study of CFL

This study is performed to explore the importance of SD and SG in core feature learning (CFL), the training and experiment performance are shown in Fig. 10, which demonstrates that the CFL achieves a better

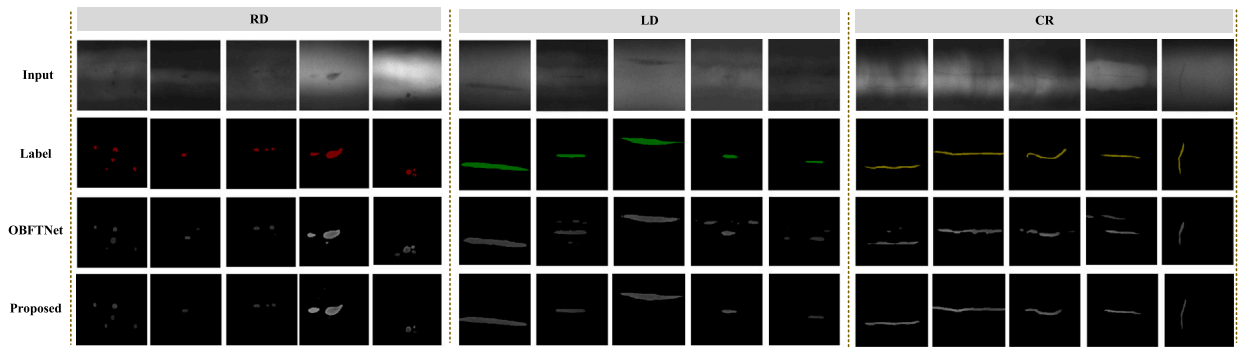


Fig. 9. Segmentation results of pipeline dataset.

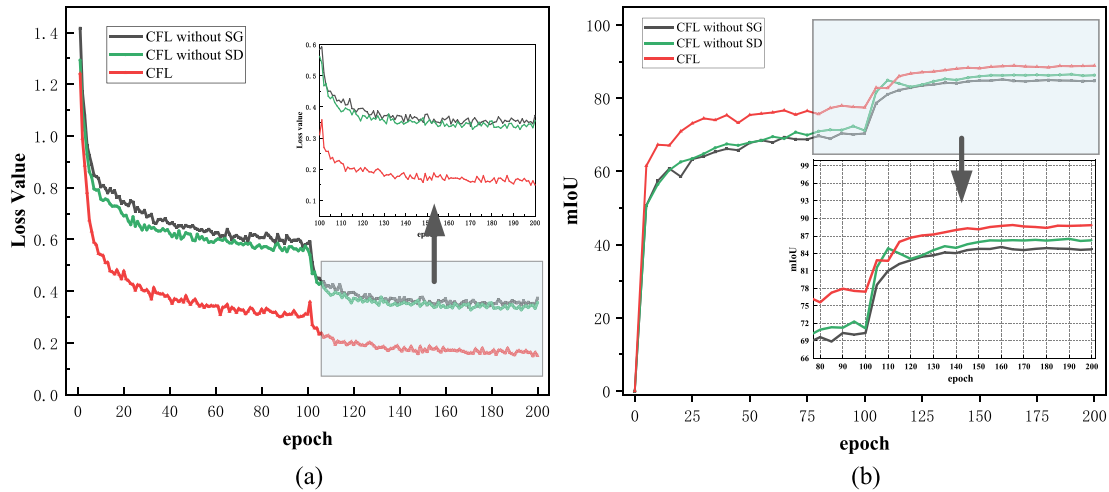


Fig. 10. Ablation experiment results of CFL. (a) Loss variation curves of three cases. (b) mIoU curves for CFL ablation study.

Table 4
Ablation experiment results of CPDNet.

Method	Normal	LP	CR	PO	Mean.
CPDNet without SD	98.4%	87.7%	71.2%	87.9%	86.3%
CPDNet without SG	98.2%	87.1%	67.2%	85.6%	84.5%
CPDNet without EPR	98.5%	88.2%	72.6%	89.2%	87.1%
CPDNet (proposed)	98.8%	88.2%	78.1%	89.5%	88.7%

mIoU score compared to other two cases, either without SG or SD. The loss and mIoU curves show significant performance fluctuations at 100th epoch, which is because the encoder is frozen before 100th epoch to maintain the feature extraction capability while reducing the number of parameters, and then is unfrozen to adjust the parameters to the core feature learning gradually. By proposed core feature learning, the hidden feature of low-contrast X-ray images can be mined for achieving core defect segmentation.

5. Conclusion

To accurately segment the low-contrast X-ray defect images, a CPDNet decomposed into core feature learning and elasticity profile refinement has been proposed in this paper. On the one hand, considering the small blackness difference between defect and background may cause core features to be hidden or lost during extraction, we explore two views, discriminant and structure of the feature space, and extract the core features by semantic-driven mining and soft-contrast gathering, thus deriving powerful feature learning ability to achieve core defect segmentation. On the other hand, to further improve the defect

segmentation performance upon the core feature learning, we consider the overall pixel distribution and make the first attempt to define the defect profile refinement as an out-of-distribution detection problem. In this case, the elasticity profile refinement module is developed, which leverages the elasticity score to refine the profile features of defects at the pixel level, thus the accuracy of defect segmentation can be significantly improved. The extensive experimental results on RIAWELC dataset and pipeline dataset demonstrated that the CPDNet method achieved the state-of-the-art segmentation performance. Consequently, the proposed CPDNet enables accurate and efficient localization of weld defects, thereby pinpointing the areas in equipment that require repair and maintenance. This approach significantly reduces risks and enhances the safety of both the equipment and the overall system.

However, for the task of defect segmentation in low-contrast X-ray images, annotating training and testing samples is challenging and costly, especially for large-scale datasets. Therefore, future research may focus on developing semi-supervised or unsupervised learning algorithms to reduce reliance on annotated data, enhancing the model's scalability and generalization. Additionally, exploring data augmentation techniques to improve the model's adaptability to various data distributions is a crucial research direction.

CRedit authorship contribution statement

Xiaoyuan Liu: Writing – original draft, Validation, Software, Methodology. **Jinhai Liu:** Methodology, Funding acquisition. **Huanqun Zhang:** Data curation. **Huaguang Zhang:** Methodology, Conceptualization.

Declaration of competing interest

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

Data availability

The authors do not have permission to share data.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (62373085, U21A20481, 61973071)

References

- Bao, Yanqi, Song, Kechen, Liu, Jie, Wang, Yanyan, Yan, Yunhui, Yu, Han, Li, Xingjie, 2021. Triplet-graph reasoning network for few-shot metal generic surface defect segmentation. *IEEE Trans. Instrum. Meas.* 70, 1–11.
- Chaibou, Mahaman Sani, Kalti, Karim, Solaiman, Basel, Mahjoub, Mohamed Ali, 2016. A combined approach based on fuzzy classification and contextual region growing to image segmentation. In: 2016 13th International Conference on Computer Graphics, Imaging and Visualization. CGIV, pp. 172–177.
- Chen, Liang-Chieh, Zhu, Yukun, Papandreou, George, Schroff, Florian, Adam, Hartwig, 2018a. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Proceedings of the European Conference on Computer Vision. ECCV*, pp. 801–818.
- Chen, Liang-Chieh, Zhu, Yukun, Papandreou, George, Schroff, Florian, Adam, Hartwig, 2018b. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: Ferrari, Vittorio, Hebert, Martial, Sminchisescu, Cristian, Weiss, Yair (Eds.), *Computer Vision – ECCV 2018*. Springer International Publishing, Cham, pp. 833–851.
- de Queiroz, R.L., Fan, Zhigang, Tran, T.D., 2000. Optimizing block-thresholding segmentation for multilayer compression of compound images. *IEEE Trans. Image Process.* 9 (9), 1461–1471.
- Diniz, Rafael, Freitas, Pedro Garcia, Farias, Mylène C.Q., 2021. Color and geometry texture descriptors for point-cloud quality assessment. *IEEE Signal Process. Lett.* 28, 1150–1154.
- Du, Wangzhe, Shen, Hongyao, Fu, Jianzhong, 2021. Automatic defect segmentation in X-Ray images based on deep learning. *IEEE Trans. Ind. Electron.* 68 (12), 12912–12920.
- Du, Wangzhe, Shen, Hongyao, Zhang, Ge, Yao, Xinhua, Fu, Jianzhong, 2022. Interactive defect segmentation in X-Ray images based on deep learning. *Expert Syst. Appl.* 198, 116692.
- Geng, Chen, Buyun, Sheng, Gaocai, Fu, Xiangxiang, Chen, Guangde, Zhao, 2024. A GAN-based method for diagnosing bodywork spot welding defects in response to small sample condition. *Appl. Soft Comput.* 157, 111544.
- He, Kaiming, Fan, Haoqi, Wu, Yuxin, Xie, Saining, Girshick, Ross, 2020. Momentum contrast for unsupervised visual representation learning. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 9726–9735.
- Joseph, K J, Khan, Salman, Khan, Fahad Shahbaz, Balasubramanian, Vineeth N, 2021. Towards open world object detection. In: 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition. CVPR, pp. 5826–5836.
- Khosla, Prannay, Teterwak, Piotr, Wang, Chen, Sarna, Aaron, Tian, Yonglong, Isola, Phillip, Maschinot, Aaron, Liu, Ce, Krishnan, Dilip, 2020. Supervised contrastive learning. *Adv. Neural Inf. Process. Syst.* 33, 18661–18673.
- Kim, Jaewoong, Lee, Sukhan, 2015. Extracting major lines by recruiting zero-threshold Canny edge links along Sobel highlights. *IEEE Signal Process. Lett.* 22 (10), 1689–1692.
- Le, Jian, Li, Fayuan, Zeng, Mingru, Zhang, Hua, 2024. Research on the online detection method of fillet welding arc stability based on the adaptive model of rotating arc stability sensor. *Mech. Syst. Signal Process.* 217, 111500.
- Li, Chen, Pan, Xiakai, Zhu, Peiyuan, Zhu, Shidong, Liao, Chengwei, Tian, Haoyang, Qian, Xiang, Li, Xiu, Wang, Xiaohao, Li, Xinghui, 2024a. Style adaptation module: Enhancing detector robustness to inter-manufacturer variability in surface defect detection. *Comput. Ind.* 157–158, 104084.
- Li, Liangliang, Wang, Peng, Ren, Jia, Lü, Zhigang, Li, Xiaoyan, Gao, Hui, Di, RuoHai, 2024b. Synthetic data augmentation for high-resolution X-ray welding defect detection and classification based on a small number of real samples. *Eng. Appl. Artif. Intell.* 133, 108379.
- Lin, Qiangqiang, Zhou, Jinzhu, Ma, Qiurui, Ma, Yongji, Kang, Le, Wang, Jianjun, 2022. EMRA-net: A pixel-wise network fusing local and global features for tiny and low-contrast surface defect detection. *IEEE Trans. Instrum. Meas.* 71, 1–14.
- Liu, Molin, Lv, Jun, Du, Shichang, Deng, Yafei, Shen, Xiaoxiao, Zhou, Yulu, 2024. Multi-resource constrained flexible job shop scheduling problem with fixture-pallet combinatorial optimisation. *Comput. Ind. Eng.* 188, 109903.
- Liu, Moyun, Xie, Jingming, Hao, Jing, Zhang, Yang, Chen, Xuzhan, Chen, Youping, 2022. A lightweight and accurate recognition framework for signs of X-ray weld images. *Comput. Ind.* 135, 103559.
- Liu, Boheng, Zhang, Tianrui, Yu, Yao, Miao, Ligang, 2023a. A data generation method with dual discriminators and regularization for surface defect detection under limited data. *Comput. Ind.* 151, 103963.
- Liu, Jinhai, Zhao, He, Chen, Zhaolin, Wang, Qiannan, Shen, Xiangkai, Zhang, Huaguang, 2023b. A dynamic weights-based wavelet attention neural network for defect detection. *IEEE Trans. Neural Netw. Learn. Syst.* 1–11.
- Ma, Shuai, Song, Kechen, Niu, Menghui, Tian, Hongkun, Wang, Yanyan, Yan, Yunhui, 2023. Shape-consistent one-shot unsupervised domain adaptation for rail surface defect segmentation. *IEEE Trans. Ind. Inform.* 19 (9), 9667–9679.
- Maier, Martin, Abdel Rahman, Rasha, 2019. No matter how: Top-down effects of verbal and semantic category knowledge on early visual perception. *Cogn. Affect. Behav. Neurosci.* 19, 859–876.
- Mashrafi, Sheikh T., Deng, Junjing, Preissner, Curt, Salapaka, Srinivasa M., 2020. Optimal control for X-Ray microscopes. *IEEE/ASME Trans. Mechatronics* 25 (2), 627–637.
- Mei, Hongwei, Jiang, Huaiyuan, Yin, Fanghui, Wang, Liming, Farzaneh, Masoud, 2021. Terahertz imaging method for composite insulator defects based on edge detection algorithm. *IEEE Trans. Instrum. Meas.* 70, 1–10.
- Ni, Xuefeng, Liu, Hongli, Ma, Ziji, Wang, Chao, Liu, Jianwei, 2022. Detection for rail surface defects via partitioned edge feature. *IEEE Trans. Intell. Transp. Syst.* 23 (6), 5806–5822.
- Niu, Shuanlong, Li, Bin, Wang, Xinggang, Peng, Yaru, 2022. Region- and strength-controllable GAN for defect generation and segmentation in industrial images. *IEEE Trans. Ind. Inform.* 18 (7), 4531–4541.
- Niu, Shuanlong, Peng, Yaru, Li, Bin, Wang, Xinggang, 2023. A transformed-feature-space data augmentation method for defect segmentation. *Comput. Ind.* 147, 103860.
- Perri, Stefania, Spagnolo, Fanny, Frustaci, Fabio, Corsonello, Pasquale, 2023. Welding defects classification through a convolutional neural network. *Manuf. Lett.* 35, 29–32.
- Rong, Xuee, Wang, Peijin, Diao, Wenhui, Yang, Yiran, Yin, Wenxin, Zeng, Xuan, Wang, Hongqi, Sun, Xian, 2023. Micro: Modeling cross-image semantic relationship dependencies for class-incremental semantic segmentation in remote sensing images. *IEEE Trans. Geosci. Remote Sens.* 61, 1–18.
- Ronneberger, Olaf, Fischer, Philipp, Brox, Thomas, 2015. U-net: Convolutional networks for biomedical image segmentation. In: Navab, Nassir, Hornegger, Joachim, Wells, William M., Frangi, Alejandro F. (Eds.), *Medical Image Computing and Computer-Assisted Intervention. MICCAI 2015*, Springer International Publishing, Cham, pp. 234–241.
- Shan, Dexing, Zhang, Yunzhou, Coleman, Sonya, Kerr, Dermot, Liu, Shitong, Hu, Ziqiang, 2023. Unseen-material few-shot defect segmentation with optimal bilateral feature transport network. *IEEE Trans. Ind. Inform.* 19 (7), 8072–8082.
- Shelhamer, Evan, Long, Jonathan, Darrell, Trevor, 2017. Fully convolutional networks for semantic segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* 39 (4), 640–651.
- Shen, Xiangkai, Liu, Jinhai, Jiang, Lin, Liu, Xiaoyuan, Zhang, Huaguang, 2023. A novel weld defect detection method for intelligent magnetic flux leakage detection system via contextual relation network. *IEEE Trans. Ind. Electron.* 1–11.
- Sime, Dejene M., Wang, Guotai, Zeng, Zhi, Peng, Bei, 2023. Uncertainty-aware and dynamically-mixed pseudo-labels for semi-supervised defect segmentation. *Comput. Ind.* 152, 103995.
- Wang, Chuhan, Chen, Haiyong, Zhao, Shenshen, 2023. RERN: Rich edge features refinement detection network for polycrystalline solar cell defect segmentation. *IEEE Trans. Ind. Inform.* 1–12.
- Wang, Xiaopeng, Zhang, Baixin, Yu, Xinghua, 2024. Zoom in on the target network for the prediction of defective images and welding defects location. *NDT E Int.* 143, 103059.
- Yang, Hua, Hu, Jiale, Yin, Zhouping, Wang, Zhengjia, 2023a. A semantic information decomposition network for accurate segmentation of texture defects. *IEEE Trans. Ind. Inform.* 19 (7), 8319–8327.
- Yang, Lei, Song, Shouan, Fan, Junfeng, Huo, Benyan, Li, En, Liu, Yanhong, 2022. An automatic deep segmentation network for Pixel-level welding defect detection. *IEEE Trans. Instrum. Meas.* 71, 1–10.
- Yang, Wei, Xiao, Yancai, Shen, Haikuo, Wang, Zhipeng, 2023b. An effective data enhancement method of deep learning for small weld data defect identification. *Measurement* 206, 112245.
- Yeung, Ching-Chi, Lam, Kin-Man, 2023. Attentive boundary-aware fusion for defect semantic segmentation using transformer. *IEEE Trans. Instrum. Meas.* 72, 1–13.
- Yu, Han, Li, Xingjie, Song, Kechen, Shang, Erfeng, Liu, Huiyu, Yan, Yunhui, 2020. Adaptive depth and receptive field selection network for defect semantic segmentation on castings X-rays. *NDT E Int.* 116, 102345.
- Yu, Peter, Qin, A.K., Clausi, David A., 2012. Unsupervised polarimetric SAR image segmentation and classification using region growing with edge penalty. *IEEE Trans. Geosci. Remote Sens.* 50 (4), 1302–1317.

- Zhang, Rui, Liu, Donghao, Bai, Qiaofeng, Fu, Liuhu, Hu, Jing, Song, Jinlong, 2024. Research on X-ray weld seam defect detection and size measurement method based on neural network self-optimization. *Eng. Appl. Artif. Intell.* 133, 108045.
- Zhou, Yulu, Du, Shichang, Liu, Molin, Shen, Xiaoxiao, 2024. Machine-fixture-pallet resources constrained flexible job shop scheduling considering loading and unloading times under pallet automation system. *J. Manuf. Syst.* 73, 143–158.
- Zhu, Zongwei, Han, Guangjie, Jia, Gangyong, Shu, Lei, 2020. Modified DenseNet for automatic fabric defect detection with edge computing for minimizing latency. *IEEE Internet Things J.* 7 (10), 9623–9636.