Information Sciences 661 (2024) 120183

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Boundary-wise loss for medical image segmentation based on fuzzy rough sets

Qiao Lin *, Xin Chen, Chao Chen, Jonathan M. Garibaldi

School of Computer Science, University of Nottingham, Nottingham, UK

ARTICLE INFO

Keywords:
Fuzzy rough sets
Semantic segmentation
Medical image
Loss function

ABSTRACT

The loss function plays an important role in deep learning models as it determines the model convergence behavior and performance. In semantic segmentation, many methods utilize pixel-wise (e.g. cross-entropy) and region-wise (e.g. dice) losses while boundary-wise loss is underexplored. It is known that one of the key aims of semantic segmentation is to precisely delineate objects' boundaries. Hence, it is essential to design a loss function that measures the errors around objects' boundaries. Fuzzy rough sets are constituted by the fuzzy equivalence relation, which is commonly used to measure the difference between two sets. In this paper, the lower approximation of fuzzy rough sets is proposed to construct the boundary-wise loss in deep learning models for the first time. The experiments with various segmentation models and datasets have verified that the proposed fuzzy rough sets loss is superior to other boundary-wise losses in terms of segmentation accuracy and time complexity. Compared with the commonly used pixel-wise and region-wise losses, the proposed boundary-wise loss performs similarly in dice coefficient, pixel-wise accuracy, but has a better performance in Hausdorff distance and symmetric surface distance. It indicates that the proposed loss provides a better guidance for segmentation models in producing more accurate shapes of the target objects. Code is available online at Github: https://github.com/qiaolin1992/Boundary-Loss.

1. Introduction

With the rapid development of computer technology and the growth of data sources, deep learning is deemed as one of the cornerstones of artificial intelligence (AI) and successfully applied to various fields such as automatic speech recognition [1], image recognition [2], natural language processing [3], bioinformatics [4], and medical image analysis [5]. In some cases, the performance of deep learning is on a par with that of human experts [6]. One of the main components in deep learning methods that determines model performance is loss function.

Loss function (also called cost function or error function) is a function that measures the difference between the predicted values and the actual values. For deep learning optimization problems, the network parameters are estimated by minimizing the given loss function iteratively in a training process. The selection of loss functions is generally specified based on various tasks: classification, segmentation, object detection, etc. The loss function plays a considerably important role in the training process of deep learning networks as it instigates the convergence process and affects the performance of neural networks.

* Corresponding author.
E-mail addresses: qiao.lin@nottingham.ac.uk (Q. Lin), xin.chen@nottingham.ac.uk (X. Chen), chao.chen@nottingham.ac.uk (C. Chen), jon.garibaldi@nottingham.ac.uk (J.M. Garibaldi).

https://doi.org/10.1016/j.ins.2024.120183
Received 26 February 2022; Received in revised form 17 January 2024; Accepted 18 January 2024

Available online 24 January 2024

Researchers have designed various types of loss functions to address specific problems. In this paper, we primarily focus on the task of image segmentation. Cross entropy loss [7] and dice coefficient loss [8] are the commonly used image segmentation losses. Cross entropy loss is a type of pixel-wise loss, which is calculated by the negative average of the log of corrected predicted probabilities. This loss focuses on the predicted value for each pixel and performs less robust for unbalanced data. Thus many cross entropy variation losses are proposed to handle unbalanced data issues, including balanced cross entropy loss [9], focal loss [10]. Dice coefficient loss is a region-wise loss that quantifies the intersection regions of the predicted segmentation and the ground truth segmentation. This loss performs well on unbalanced datasets but its training error curve is unstable and gives no information for the convergence procedure. To take advantages of both dice and cross entropy loss, Taghanaki et al. [11] introduced a hybrid loss that combines both the dice loss and cross entropy loss.

All the aforementioned losses are pixel-wise and region-wise losses. In image segmentation tasks, uncertainty and misclassification normally happen at the boundaries [12]. Therefore, if a loss function is designed to concentrate on the boundary, the image segmentation performance can be potentially improved. The research conducted by Karimi et al. [13] discussed the boundary difference between the predicted segmentation and the ground truth segmentation. In their study, distance transforms and morphological operations were applied to construct the semantic segmentation boundary-wise loss, which calculated the Hausdorff distance between the predicted image boundary and the ground truth image boundary. However, as the range of Karimi's boundary-wise loss [13] is from 0 to infinity, the convergence efficiency is potentially low and the gradient descent process is unstable leading to unsatisfying segmentation performance. Moreover, Karimi et al. [13] only studied the performance of a combined loss based on region-wise and boundary-wise losses without in-depth investigation of the boundary-wise loss itself.

On the other hand, in 1990, Dubois et al. [14] introduced a novel theory named fuzzy rough sets which combined fuzzy sets with rough sets. The key idea of the fuzzy rough sets is to supersede the equivalence relation of rough sets by a fuzzy equivalence relation [15]. Thus, the fuzzy rough sets have the ability to manage data with fuzziness and vagueness based on the similarity of different attributes. As the generalizations of classical rough sets, fuzzy rough sets are commonly used in image segmentation [16], dimensionality reduction [17], feature selection [18], etc. Although fuzzy rough sets are applied widely in a multitude of AI tasks, no research proposed the use of fuzzy rough sets as the loss function in machine learning models. In this paper, as the lower approximation of fuzzy rough sets has the ability to evaluate the difference between two sets, we propose a novel boundary-wise loss based on fuzzy rough sets to pay more attention to the boundary with more satisfying segmentation accuracy and lower time complexity in comparison to other methods mentioned above.

The main contributions are summarized as follows.

- This is the first time that fuzzy rough sets are incorporated in deep learning models as a loss function for image segmentation.
- A Gaussian kernel is applied to normalize the boundary of the predicted segmentation and the ground truth segmentation, which plays a significant role in stabilizing the training process and saving computational time.
- Based on two public datasets with various of object shape and size, our proposed boundary loss has shown to outperform existing boundary-wise losses in terms of segmentation accuracy and time complexity based on three widely used deep learning image segmentation models. Moreover, the proposed boundary-wise loss was found to produce distinct and smoother borders compared to the pixel-wise loss (cross-entropy loss) and the region-wise loss (dice loss).

The structure of the remainder of this paper is as follows. Background information about semantic segmentation loss functions and fuzzy rough sets are introduced in Section 2. The mathematical derivation of fuzzy rough sets loss in semantic segmentation models is given in Section 3. The evaluation metrics, implementation details and experimental results are presented in Section 4. Finally, the discussion and conclusions are in Section 5 and Section 6 respectively.

2. Background

In this section, the background information about semantic segmentation loss functions and mathematical definitions of fuzzy rough sets are given.

2.1. Semantic segmentation loss functions

2.1.1. Semantic segmentation

Image segmentation is one of the most important tasks in image comprehension and computer vision. It aims to divide a given image into several disjoint areas based on distinct features such as shape, image intensity, and texture so that all the features share a high level of similarity in the same area. Classical image segmentation technologies are primarily unsupervised learning methods: edge detection [19], threshold [20], region growing [21], and clustering [22]. These methods are not robust and have low-quality segmentation results when the boundary of the original image is complicated and overlapped. To mitigate this problem, end-to-end convolutional neural networks (CNNs)-based semantic segmentation techniques have gradually become the mainstream image segmentation algorithms and achieved outstanding performance in numerous computer vision tasks.

Semantic segmentation can be treated as pixel-level classification. Every individual pixel in the input image is assigned to one of multiple classes by a trained segmentation model using a set of training images. The general structure of semantic segmentation models is constituted by convolutional, pooling, and de-convolutional layers. Convolutional and pooling layers are applied to extract the spatial and channel information of the input image, while de-convolutional layers are leveraged to resize the feature maps so

that the sizes of the output image and the ground truth image are consistent. The most widely used semantic models include Fully Convolutional Networks (FCN) [23], UNet [24], SegNet [25], Deconvnet [26] and DeepLab series [27–30].

2.1.2. Loss functions

The above mentioned CNN models aim to estimate a set of intrinsic model parameters by minimizing a loss function that measures the difference between the currently predicted segmentation and the ground truth segmentation. This minimization process is performed iteratively by error back propagation using gradient descent algorithm. As the loss function instigates the learning process of the semantic segmentation models, the selection of loss function is of great importance. The loss functions for semantic segmentation are divided into three categories: pixel-wise loss, region-wise loss and boundary-wise loss [31]. The pixel-wise losses include Cross Entropy Loss [7], Weighted Cross Entropy Loss [32], Balanced Cross Entropy Loss [9], and Focal Loss [10]. Cross Entropy Loss is the widely used pixel-wise loss. The others are the cross entropy variation losses, which is proposed to deal with unbalanced data issues. The Cross Entropy Loss (CELoss) is defined as:

$$L_{CE} = \frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{K} -y_{it}\log\left(\hat{y}_{it}\right), \tag{1}$$

where y is the label pixel value and $\hat{y}$ is the predicted pixel value, K is the number of classes, N is the number of pixels.

The region-wise losses contain Dice Loss [8], Sensitivity-Specificity Loss [33], Tversky Loss [34], Focal Tversky Loss [35], Log-Cosh Dice Loss [31]. Among them, Dice Loss is the most representative and commonly used region-wise loss. DiceLoss (DLoss) is represented as:

$$L_D = \frac{1}{K}\sum_{t=1}^{K}\left(1 - \frac{2\sum_{i=1}^{N} y_{it}\hat{y}_{it}}{\sum_{i=1}^{N} y_{it}^2 + \sum_{i=1}^{N}\hat{y}_{it}^2}\right), \tag{2}$$

where y, $\hat{y}$, K, N have the same meaning as L_{CE}.

The research about boundary-wise loss is limited. There are two popular boundary-wise losses namely Hausdorff loss (HDLoss) [13] and Dual HausdorffLoss (DHDLoss) [13] proposed to focus on the segmentation boundary. HDLoss is yielded from Hausdorff distance $d(X,Y) = \max_{x\in X}\min_{y\in Y}\|x-y\|_2$, where x and y are the elements of X and Y respectively. Since the formula of Hausdorff distance is non-convex, it cannot be directly applied to calculate the image segmentation loss. The variation of Hausdorff distance has the ability to make the HDLoss tractable, which is represented as:

$$L_{HD} = \frac{1}{|\Omega|}\sum_{\Omega}\left((p-q)^2 \otimes \left(\mathcal{D}_p^{\alpha}\right)\right), \tag{3}$$

where p and q are the predicted binary image and the ground truth image respectively, $\mathcal{D}_p$ refers to the distance map of p, α is equal to 2, $\otimes$ represents the pixel-wise multiplication operator, Ω is the whole area of the given image. Since $\mathcal{D}_p$ and $\mathcal{D}_q$ are not equal, a dual direction Hausdorff loss (DHDLoss) is proposed:

$$L_{DHD} = \frac{1}{|\Omega|}\sum_{\Omega}\left((p-q)^2 \otimes \left(\mathcal{D}_p^{\alpha} + \mathcal{D}_q^{\alpha}\right)\right), \tag{4}$$

From the equation of HDLoss and DHDLoss, the distance value in distance map has a large range which may cause instability in the training process of semantic segmentation models and lead the segmentation model to fail to converge to the optimal value. In this paper, to manage the limitations of existing boundary-wise losses and capture the more accurate segmentation boundary, a novel boundary-wise loss function based on fuzzy rough sets is proposed.

2.2. Fuzzy rough sets

Rough sets [36] use the lower and the upper approximation of the original set to approximate the original set, which are particularly useful in dealing with ambiguity, vagueness and general uncertainty problems. However, the classical rough sets are only effective when the data is symbolic-valued [37] and cannot manage the relationship between two sets. To address the above restrictions and broaden the application range of rough sets, Dubois et al. [14] proposed a novel theory named fuzzy rough sets which combined fuzzy sets and rough sets. The key idea of the fuzzy rough sets is to supersede the equivalence relation of rough sets by a fuzzy equivalence relation [38,39]. The introduction of a fuzzy equivalence relationship in fuzzy rough sets makes it possible to evaluate the relation between two sets.

Definition 1. Given a universe U, R is a fuzzy equivalence relation on U. For $\forall x, y \in U$, the fuzzy rough sets are defined as [14]

$$\begin{cases} \underline{R_{\max}}X(x) = \underset{y\in U}{inf}\max\left(1 - R(x,y), X(y)\right) \\ \overline{R_{\min}}X(x) = \underset{y\in U}{sup}\min\left(R(x,y), X(y)\right) \end{cases} \tag{5}$$

where X is a subset of U.

Based on [40][41], if the R is a fuzzy T-equivalence relation on U, the more general fuzzy operator t-norm, s-norm defined above are applied to substitute 'max' and 'min', that is

$$\begin{cases} \underline{R_s} X(x) = \inf\limits_{y \in U} s(N(R(x,y)), X(y)) \\ \overline{R_t} X(x) = \sup\limits_{y \in U} t(R(x,y), X(y)) \end{cases} \tag{6}$$

where X is the subset of U and N is the fuzzy complement. $\underline{R_{\max}} X(x)$ and $\underline{R_s} X(x)$ are the lower approximation and represent the degrees the x certainly belongs to the set X. $\overline{R_{\min}} X(x)$ and $\overline{R_t} X(x)$ are the upper approximation and denote the degrees the x possibly belongs to set X.

3. A new boundary-wise loss for semantic segmentation

In this section, a new boundary-wise loss for semantic segmentation is proposed. Based on the theory of fuzzy rough sets, the lower approximation $\underline{R_s} D_i(x)$ of fuzzy rough sets means the degrees the x certainly belongs to the set D_i. Therefore, the sum $\left(\sum_{x \in X} \underline{R_s} D_i(x)\right)$ is applied to evaluate the similarity between two sets X and D_i. The key point of the boundary-wise loss is to calculate the difference between the set of predicted image boundary pixels and the set of ground truth image boundary pixels. It is a natural idea to use the lower approximation of fuzzy rough sets to calculate the boundary-wise loss. The detailed mathematical derivation is as follows.

3.1. The lower approximation of fuzzy rough sets

Given a finite and nonempty set of samples U, and decision $\mathbb{D}$ which partitions the samples into subsets $\{D_1, D_2, \cdots, D_M\}$. For $\forall x \in U$, if $x \notin D_i(x)$, $D_i = 0$, otherwise $D_i(x) = 1$. Based on the definition of fuzzy rough sets, the membership degree of x certainty belonging to the given class D_i is calculated by the lower approximation of fuzzy rough sets $\underline{R_s} D_i(x) = \inf\limits_{y \in U} s\left(N(R(x,y)), D_i(y)\right)$. Herein, inspired by [42], the s-norm $s_{\cos}(a,b) = \min\left(a + b - ab + \sqrt{2a - a^2}\sqrt{2b - b^2}, 1\right)$ is chosen to compute the lower approximation:

$$\begin{aligned} \underline{R_s} D_i(x) &= \inf_{y \in U} s_{\cos}\left(1 - R(x,y), D_i(y)\right) \\ &= \inf_{y \in U} \min(1 - R(x,y) + D_i(y) - (1 - R(x,y) \\ &\times D_i(y) + \sqrt{1 - R^2(x,y)}\sqrt{2D_i(y) - D_i^2(y)}, 1) \\ &= \inf_{y \in U} \begin{cases} 1, y \in D_i \\ 1 - R(x,y), y \notin D_i \end{cases} \\ &= \inf_{y \notin D_i} (1 - R(x,y)) \end{aligned} \tag{7}$$

where $R(x,y)$ is the fuzzy equivalence relation between samples and is computed by some kernel functions [43].

$\underline{R_s} D_i(x) = \inf\limits_{y \notin D_i} (1 - R(x,y))$ means that the degree of x certainty belonging to D_i relies on the closest sample of another category. Consider a special situation in which there are only two classes D_1 and D_2. The formula $\underline{R_s} D_1(x) = \inf\limits_{y \in D_2} (1 - R(x,y))$ denotes the likelihood of sample x belonging to class D_1 increases with the distance between x and class D_2. That is if x is far apart from class D_2, then it is more likely to belong to class D_1.

3.2. Fuzzy rough sets loss

For a semantic segmentation model, the training process is that given the input image array $X \in \Re^{h \times w \times c}$ (h, w and c refer to the height, weight and channel of X), the output predicted image is obtained by

$$\hat{Y} = f(W \otimes X), \tag{8}$$

where $\hat{Y} \in \Re^{h' \times w' \times \mathscr{P}}$ (h' and w' are the height and weight of $\hat{Y}$, $\mathscr{P}$ is the number of pixel categories), W represents the model parameters, and f means the softmax function. Then the loss value is back-propagated to update the parameters. The abstract loss function is written as $loss = F\left(\hat{Y}, Y\right)$ where $\hat{Y}$ and Y are the predicted image and the corresponding ground truth image, and F means the selected loss function. The loss value increases with the distance between $\hat{Y}$ and Y.

Fig. 1 shows the predicted image that is the output of a given semantic segmentation model and the corresponding ground truth image. For the ground truth image, the boundary pixel sets belong to the class D_1, while the rest parts of pixels are assigned to class D_2. In the predicted image, the boundary pixel sets are $\{x_i : x_i \in X\}$. For the binary categories semantic segmentation, the ground

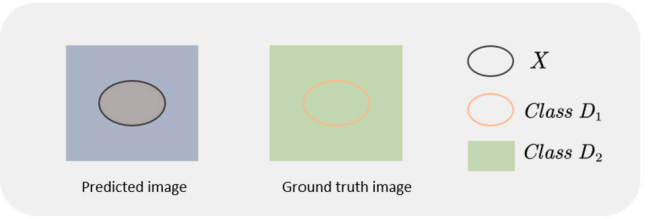

Fig. 1. The left part refers to the predicted image and the right part is the ground truth image. Black circle is the predicted image boundary pixel sets; orange circle is the ground truth image boundary pixel sets and belongs to the class D_1; the rest parts of the ground truth image pixels belong to the class D_2; gray part is the area of segmentation object.

truth image only has two categories D_1 and D_2 ($D_1 \cap D_1 = \oslash$, $D_1 \cup D_1 = \mho$, where $\mho$ refers to the whole ground truth image) so that the similarity between X and D_2 can be applied to assess the distance of X and D_1. The greater the similarity value of X and D_2 is, the farther X is from D_1, which means the loss difference between X and D_1 is greater. Therefore, the boundary loss is represented as

$$loss_P = \mathscr{L}\left(\hat{Y}_X, Y_{D_2}\right). \tag{9}$$

where $\mathscr{L}$ is the similarity of X and D_2.

From the section 3.1, the lower approximation $\underline{R_s} D_i(x) = \underset{y \notin D_i}{inf} (1 - R(x,y))$ of fuzzy rough sets means the degrees the x certainly belongs to the set D_i. Therefore, given the $\{x_i : x_i \in X\}$ and $D_i = D_2$, the similarity of X and D_2 is assessed by the average of the lower approximations for all x: $\frac{1}{m}\left(\sum_{x \in X} \underline{R_s} D_2(x)\right)$ where m is the number of X. The boundary loss is evaluated by

$$\begin{aligned} loss_P &= \frac{1}{m}\left(\sum_{i=1}^{m} \underline{R_s} D_2(x_i)\right) \\ &= \frac{1}{m}\sum_{i=1}^{m}\left(\underset{\substack{j \in \{1,2,\ldots,k\} \\ y_j \notin D_2}}{inf} \left(1 - R(x_i, y_j)\right)\right) \end{aligned} \tag{10}$$

where $x_i \in X$, $y_j \notin D_2$, R is the fuzzy equivalence relation of x_i, y_j. Furthermore, based on Fig. 1, the ground truth image only has two categories D_1, D_2 and $D_1 \cap D_2 = \oslash$, so $y \notin D_2 \Leftrightarrow y \in D_1$. It means that the boundary loss can be calculated by the boundary sets of the predicted image and the ground truth image. The widely used Gaussian kernel $R(x,y) = \exp\left(-\frac{\|x-y\|^2}{\sigma}\right)$ [37] is applied to calculate the fuzzy equivalence relation. The final boundary-wise loss formula is obtained by

$$\begin{aligned} loss_P &= \frac{1}{m}\sum_{i=1}^{m}\left(\underset{\substack{j \in \{1,2,\ldots,k\} \\ y_j \in D_1}}{inf} \left(1 - R(x_i, y_j)\right)\right) \\ &= \frac{1}{m}\sum_{i=1}^{m}\left(\underset{\substack{j \in \{1,2,\ldots,k\} \\ y_i \in D_1}}{inf} \left(1 - \exp\left(-\frac{\|x_i - y_j\|^2}{\sigma}\right)\right)\right) \end{aligned} \tag{11}$$

and named as Fuzzy Rough Sets Loss (FRSLoss).

To make the proposed loss more robust and satisfy the symmetric condition, the predicted image is separated into two categories and the boundary pixels sets of the ground truth image are the samples to be classified. Then the average of the lower approximation for all boundary pixels of the ground truth is adopted to calculate the boundary-wise loss:

$$loss_G = \frac{1}{k}\left(\sum_{j=1}^{k} \underline{R_s} D_2(y_j)\right)$$

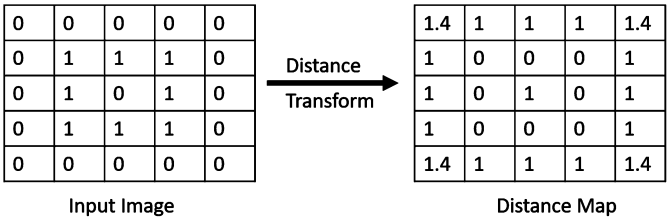

Fig. 2. Euclidean Distance Transform for a simple binary image: after applying the distance transform operator, the distance from each pixel to the nearest boundary pixel is calculated.

$$= \frac{1}{k}\sum_{j=1}^{k}\left(\inf_{\substack{i\in\{1,2,\ldots,m\}\\ x_i\in D_1}}\left(1-R\left(y_j,x_i\right)\right)\right) \tag{12}$$

It is noted that $Loss_P \neq Loss_G$. Therefore, the Dual Fuzzy Sets Loss (DFRSLoss) is introduced as $Loss_D = \left(Loss_P + Loss_G\right)$.

The aforementioned loss is only suitable for binary classification: one category is the object pixels and the other is the background pixels. Nevertheless, numerous practical segmentation tasks are required to address multi-class issues. Therefore, to expand the application scope of our proposed loss, the binary formula is extended to calculate the multi-class segmentation loss. Given the number of the categories is $\mathcal{P}$, the multi-class segmentation can be divided into $\mathcal{P}$ binary-class segmentation tasks. Thus, the multi-class FRSLoss is:

$$\begin{aligned} Loss_M &= \frac{1}{\mathcal{P}}\sum_{o=1}^{\mathcal{P}}\frac{1}{m}\left(\sum_{i=1}^{m}\underline{R_s}D_2^o\left(x_i^o\right)\right) \\ &= \frac{1}{\mathcal{P}}\sum_{o=1}^{\mathcal{P}}\frac{1}{m}\left(\sum_{i=1}^{m}\left(\inf_{\substack{j\in\{1,2,\cdots,k\}\\ y_j\in D_1^o}}\left(1-R\left(x_i^o,y_j^o\right)\right)\right)\right) \end{aligned} \tag{13}$$

3.3. Distance transform

Based on equation (11), the FRSloss are one non-convex function. In order to make the loss applicable in the segmentation model, the distance transform algorithm is utilized to calculate the FRSLoss in practical applications.

Distance transform is a type of image operator only applied to the binary image, which specifies the distance from each pixel to the nearest boundary pixel. Given a binary image I, the distance transform is defined as

$$\mathcal{D}(I) = \left\{\mathcal{D}(p) = \min_{q\in\mathfrak{B}}\left(d\left(p,q\right)\right)|p\in I\right\}, \tag{14}$$

where $\mathfrak{B}$ refers to the boundary pixel sets, d means one given distance metric that is utilized to determine the distance between the two pixels p and q. Herein, Euclidean Distance is chosen to evaluate the distance: $d\left(p,q\right) = \sqrt{\left(x_p - x_q\right)^2 + \left(y_p - y_q\right)^2}$, where $\left(x_p, y_p\right), \left(x_q, y_q\right)$ represent the coordinate values of pixel p and q. Fig. 2 shows the Euclidean Distance Transform for a simple binary image.

Due to the practicability and effectiveness of Distance Transform in computer vision [44] [45] [46] [47], many researchers devote themselves to find the optimal algorithm to calculate the distance map. In this paper, a linear-time algorithm based on min-convolutions and squared Euclidean distance is adopted to calculate the Distance Transform [48]. This algorithm is less time-consuming (O(N)) and suitable for the application in arbitrary dimension images. The primary computation procedure is as below:

- define squared Euclidean distance transform formula of p as $\mathcal{D}_f(p) = \min\left((p-q)^2 + f(q)\right)$, where q is the argument of the parabola function, f is a sampled function [48]
- iterate through all the candidate pixels and obtain the distance transform based on the intersection point value s between the parabolas $s = \frac{(f(r)+r^2)-(f(q)+q^2)}{2r-2q}$ and lower envelope of each parabola, where r and q represent different parabolas.

3.4. Computational details of FRSLoss

Based on the FRSLoss formula (11) and Euclidean Distance Transform (Section 3.3), the computational procedure of the proposed FRSLoss is described in Algorithm 1.

Algorithm 1 Fuzzy Rough Sets Loss.

Input: semantic segmentation model M, raw input image I, corresponding ground truth image Y
Output: FRSLoss
1: input image I into the model M and generate the predicted image P
2: obtain the predicted binary image $\hat{Y}$ using the binaryzation operator $\hat{Y} = \begin{cases} 1, p \geq 0.5 \\ 0, p < 0.5 \end{cases}$, where p is the pixel value of P
3: obtain the boundary pixel set $\mathfrak{B}$ of Y
4: calculate the difference map $F = |\hat{Y} - Y|$
5: use Euclidean Distance Transform equation (14) to get the distance map $D(Y)$
6: calculate the minimum distance between predicted image boundary and ground truth image boundary $\Lambda = F \otimes D$, where $\otimes$ means pixel-wise multiplication
7: apply Gaussian kernel $k_G = \exp\left(-\frac{|\Lambda|^2}{\sigma}\right)$
8: $FRSLoss = Average(1 - Gaussian(\Lambda))$.
9: **return** FRSLoss

From the Algorithm 1, to obtain the minimum distance between predicted image boundary and ground truth image boundary, Euclidean Distance Transform algorithm has played the pivotal role. As each value in the product map Λ refers to the minimum distance from the boundary pixels in $\hat{Y}$ to the boundary pixels in Y, the product map Λ can be treated as the equivalence results of the inf operators. In this way, the non-convex issue of the FRSLoss is handled successfully. Next section, several experiments are implemented to evaluate the effectiveness of FRSLoss.

4. Evaluation

In this section, two public datasets with various object shape and size and three widely-used semantic segmentation models including UNet, FCN and SegNet are utilized to assess the performance of the proposed boundary-wise loss. Evaluation metrics, implementation details and experimental results are described in the following subsections.

4.1. Datasets

In the next experiments, nuclei and kidney cell datasets were adopted. Both of the datasets have multiple segments in each image. However, the nuclei dataset has more complicated boundaries than the cell datasets (seen in Fig. 4). The primary merit of our proposed boundary-loss is that it pays more attention to the boundary compared with other types of losses. Therefore, to verify the effectiveness and applicability of the proposed new boundary-based loss, it is reasonable and meaningful to choose datasets with diverse boundaries.

- Nuclei dataset comes from the 2018 Data Science Bowl, the goal of which is to segment nuclei boundary from given divergent images (https://www.kaggle.com/c/data-science-bowl-2018/data). It includes 670 raw images and the corresponding ground truth images are also available. As the nuclei dataset has multiple segments in each image, its boundaries are relatively complicated and diverse.
- The kidney renal clear cell dataset [49] contains 462 raw images with corresponding ground truth images annotated by experts. The pixel size of each image is 400×400. For sake of avoiding the overfitting issue, data augmentation including flipping and rotation was utilized to increase the quantity of this dataset. Compared with the nuclei dataset, this cell dataset has less complicated boundaries.

4.2. Metrics

To evaluate the segmentation performance of our proposed boundary-wise loss (FRSLoss and DFRSLoss) and compare it against other pixel-wise loss (CELoss), region-wise loss (DLoss), and boundary-wise losses (HDLoss and DHDLoss), pixel-wise metric (pixel accuracy (PA)), region-wise metric (Dice coefficient (DC)), and boundary-wise metrics (95th-percentile of Hausdorff distance (HD) and average symmetric surface distance (ASD)) are used and expressed as below.

$$DC = \frac{2|P \cap R|}{|P| + |R|} \tag{15}$$

where P and R present the foreground areas of the ground truth image and the predicted image respectively.

$$PA = \frac{TP + FP}{TP + FP + TN + FN} \tag{16}$$

where TN, FN, FP and TP refer to the true negative, false negative, false positive and true positive rates respectively. Note that DC belongs to the region-wise metric and PA is the pixel-wise metric [23].

$$HD(X, Y) = \max_{x \in X} \min_{y \in Y} \|x - y\|_2 \tag{17}$$

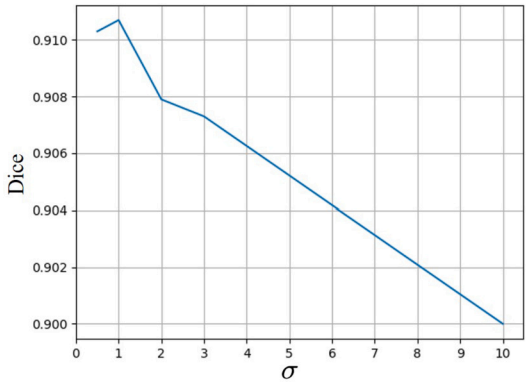

Fig. 3. Segmentation performance with different σ, when $\sigma = 1$, the semantic segmentation model achieves the best performance.

where X and Y are the boundaries of the predicted image and ground truth image, respectively. The standard HD (equation (17)) is sensitive to outliers and not robust for the performance measurement since it exclusively takes the maximum value into consideration. To avoid the effects of outliers, 95th-percentile HD is utilized to measure segmentation performance.

$$ASD = \frac{1}{\left|S\left(\hat{Y}\right)\right| + |S(Y)|}\left(\sum_{\alpha \in S(Y)} \min_{\beta \in S\left(\hat{Y}\right)} \|\alpha - \beta\| + \sum_{\beta \in S\left(\hat{Y}\right)} \min_{\alpha \in S(Y)} \|\beta - \alpha\|\right) \tag{18}$$

where $S(Y)$ refers to the boundary pixels set of the ground truth image and $S(\hat{Y})$ is the boundary pixels set of the predicted image, α and β are the elements of boundary sets $S(Y)$ and $S(\hat{Y})$ respectively. ASD is a symmetric measure of the average distance between two segmented boundaries.

4.3. Implementation details

In this paper, without loss of generality, three commonly used semantic segmentation models, FCN [23], UNet [24] and SegNet [25], were applied as the backbone and evaluated on two public datasets. FCN is the most classical semantic segmentation model and firstly proposed the encoder-decoder framework for semantic segmentation. In the encoding stage, the semantic segmentation model constituted by convolutional layers, maxpooling layers and relu activation function is applied to capture the deep features. In the decoding stage, deconvolutional operators are adopted to resize the feature maps in order to make the output image and the ground truth image have the same size. In the FCN model, the size of convolutional filter is 3×3 with the corresponding stride is 1, while the size of deconvolutional filter is 3×3 with the corresponding stride is 2. UNet and SegNet models are the improvement of FCN. They share the same encoder-decoder framework. The difference is that UNet uses skip connection to transmit the features extracted from the encoding stage to the decoding process. Whereas, the SegNet introduces an index function to record the maximum value in each sliding window of pooling layers and during the decoding stage uses the index function to recover the feature maps by up-maxpooling operators instead of the deconvolutional operator. In the Unet mode, the size of convolutional filter is 3×3 with the corresponding stride is 1, while the size of deconvolutional filter is 2×2 with the corresponding stride is 2. In the SegNet mode, the size of convolutional filter is 3×3 with the corresponding stride is 1, while the size of up-maxpooling filter is 2×2 with the corresponding stride is 2.

Numerous semantic segmentation losses have achieved exceptional performance in various sorts of datasets [7,8]. To verify the effectiveness of our proposed loss, comparison experiments were conducted. Herein, we choose pixel-wise loss (CELoss), region-wise loss (DLoss), and boundary-wise loss (HDLoss and DHDLoss) for comparison. CELoss, DLoss, HDLoss and DHDLoss are calculated by the equation (1)—(4). Each dataset was divided into five groups for five-fold cross-validation. In each fold, four groups were used for training (80%) and validation (20%), and the remaining group was used for testing.

The detailed algorithm procedure of our proposed FRSLoss is given in Section 3.4. Moreover, based on equation (11), the Gaussian kernel in the FRSLoss, which is applied to normalize the distance between the predicted image and the ground truth image, has a hyperparameter σ. From equation (11), the value of σ controls the shape of the Gaussian that helps to map the distance between the predicted image and the ground truth image to a proper loss value range. The higher the precision in delineating the relationship between loss and distance, the more effectively the segmentation model will be trained. It is necessary to determine an optimal value for σ. Thus, we conducted related experiments with the different values of $\sigma = \{0.5, 1, 2, 3, 5, 10\}$ to find the optimal value that yields the best performance on a validation set. Fig. 3 shows variation curves of the performance evaluation indices with various values of σ. The experimental results show that when σ was equal to 1, the semantic segmentation model performed best. Therefore, in all following experiments, the value of σ was set to 1, which achieved the best performance.

Table 1
Experimental results for boundary-wise losses, pixel-wise loss and region-wise loss in Unet, FCN and SegNet on the nuclei dataset. Mean ± standard deviation values are reported for all the evaluation measures. ∗ represents FRLoss and CELoss are significantly different with $p < 0.05$; ♯ represents FRLoss and DLoss are significantly different with $p < 0.05$.

Model	Loss	DC (%) ↑	PA (%) ↑	ASD (*mm*) ↓	95th-percentile of HD (*mm*) ↓
UNet	FRSLoss	91.02 ± 0.04	**97.64 ± 0.03**	**0.82 ± 0.02** ∗	**4.74 ± 0.20** ∗ ♯
	HDLoss	74.75 ± 4.05	94.90 ± 0.41	4.43 ± 0.94	35.78 ± 5.70
	DFRSLoss	**91.07 ± 0.04**	97.60 ± 0.02	0.83 ± 0.02	5.09 ± 0.31
	DHDLoss	75.39 ± 2.17	94.49 ± 0.34	2.88 ± 0.22	17.84 ± 1.62
	CELoss	90.97 ± 0.07	97.58 ± 0.01	0.92 ± 0.01	7.83 ± 0.13
	DLoss	90.73 ± 0.13	97.51 ± 0.08	0.83 ± 0.04	5.92 ± 1.01
FCN	FRSLoss	90.39 ± 0.11	97.58 ± 0.04	**0.81 ± 0.01** ♯	**4.10 ± 0.03** ∗ ♯
	HDLoss	82.77 ± 0.60	96.26 ± 0.17	1.81 ± 0.38	8.95 ± 1.73
	DFRSLoss	90.57 ± 0.08	97.62 ± 0.06	0.84 ± 0.01	4.15 ± 0.12
	DHDLoss	83.41 ± 0.45	96.37 ± 0.05	1.35 ± 0.05	5.26 ± 0.28
	CELoss	90.49 ± 0.07	97.55 ± 0.03	0.84 ± 0.01	4.97 ± 0.08
	DLoss	**90.74 ± 0.10**	**97.63 ± 0.02**	0.86 ± 0.02	5.62 ± 1.35
SegNet	FRSLoss	89.56 ± 0.53	**97.22 ± 0.12**	**0.98 ± 0.06** ∗ ♯	6.23 ± 1.28 ∗ ♯
	HDLoss	77.66 ± 2.76	95.11 ± 0.72	3.07 ± 0.59	22.15 ± 2.94
	DFRSLoss	**89.98 ± 0.15**	97.19 ± 0.13	0.98 ± 0.12	**5.81 ± 2.18**
	DHDLoss	81.93 ± 1.29	95.75 ± 0.31	2.16 ± 0.50	11.18 ± 2.17
	CELoss	89.46 ± 0.26	97.17 ± 0.04	1.14 ± 0.20	8.89 ± 1.74
	DLoss	88.97 ± 0.47	96.98 ± 0.16	1.11 ± 0.08	7.86 ± 0.51

All the experiments were implemented based on Pytorch. Adam optimization algorithm was applied to update the weights with an initial learning rate of 0.0001. The training time was about 1.5 hours for one semantic segmentation model on a workstation with i7-3820 CPU and NVIDIA GeForce GTX1080Ti.

4.4. Experimental results

4.4.1. Nuclei

Table 1 depicts the experimental results for boundary-wise losses in different semantic segmentation models. DC is generally used to assess the segmentation performance and belongs to the region-wise metric, pixel accuracy (PA) is one type of pixel-wise evaluation method, ASD and 95th-percentile of HD are adopted to evaluate the boundary distance of the predicted image and the ground truth image. It should be noted that ASD and 95th-percentile of HD have the inverse trend compared to DC and PA: when ASD and 95th-percentile of HD have lower values, segmentation performance is better; while when DC and PA have lower values, segmentation performance is worse.

FRSLoss, DFRSLoss are our proposed boundary-wise losses. HDLoss and DHDLoss are the other boundary-wise losses proposed in [13]. To make a fair comparison, FRSLoss and HDLoss only take the distance transform of the ground truth image into account, while DFRSLoss and DHDLoss consider both directions using the distance maps of the predicted image and the ground truth image. It can be seen in Table 1 that the performance of UNet, FCN, and SegNet with FRSLoss and DRSLoss are improved in DC, PA, ASD and 95th-percentile of HD compared with HDLoss and DHDLoss, which means that our proposed losses are superior to the Hausdroff distance based boundary-wise losses in all given segmentation performance evaluation. Moreover, in comparison of the experimental results of UNet, FCN and SegNet with FRSLoss, HDLoss, DFRSLoss and DHDLoss, the performance variation of our proposed loss function is much smaller across different segmentation models than Hausdroff distance based losses. In addition, the standard deviations of the experimental results indicate that FRSLoss and DFRSLoss are more stable than HDLoss and DHDLoss.

Table 1 also shows the experimental results for pixel-wise loss and region-wise loss in different semantic segmentation models. Herein, we only discuss widely used pixel-wise (CELoss) and region-wise (DLoss) losses. As reported in Table 1, CELoss and DLoss have similar segmentation accuracy to FRSLoss in DC and PA but FRSLoss performes better in boundary-wise metrics (measured by ASD and 95th-percentile of HD). This means that our proposed novel boundary-wise loss method pays more attention to the boundaries than the region-wise and pixel-wise losses. Furthermore, to further verify if there is a statistically significant difference between FRSLoss, CELoss and DLoss, wilcoxon sign rank test with $p < 0.05$ is adopted. In Table 1, ∗ represents FRLoss and CELoss are significantly different; ♯ represents FRLoss and DLoss are significantly different. Hence, for DC and PA, no statistical difference between our method and DLoss/CELoss. However, FRLoss has a statistically significant difference with CELoss and DLoss in 95th-percentile of HD for Unet, FCN and SegNet models, while for the ASD boundary metric, FRSLoss has a statistically significant difference with DLoss on the FCN and SegNet models, with CELoss on the Unet and SegNet models. On the other hand, for different segmentation models, the proposed boundary-wise loss, the widely-used pixel-wise loss and the region-wise loss have consistent performance: they perform best in the UNet model and perform worst in the SegNet model.

Table 2 shows the training time to achieve model convergence of the UNet, FCN and SegNet models using FRSLoss, DFRSloss, HDLoss and DHDloss as the loss function for the nuclei dataset. The segmentation models with FRSloss and DFRSloss require less

Table 2
Convergence time of the UNet, FCN and SegNet models with different boundary-wise losses.

Model	FRSLoss	HDLoss	DFRSLoss	DHDLoss
UNet	**40.71 min**	67.17 min	**65.10 min**	90.43 min
FCN	**57.20 min**	93.53 min	**91.18 min**	101.55 min
SegNet	**68.57 min**	115.12 min	**87.45 min**	137.95 min

Table 3
Experimental results for boundary-wise losses, pixel-wise loss and region-wise loss in Unet, FCN and SegNet on the cell dataset. Mean ± standard deviation values are reported for all the evaluation measures. $*$ represents FRLoss and CELoss are significantly different with $p < 0.05$; ♯ represents FRLoss and DLoss are significantly different with $p < 0.05$.

Model	Loss	DC (%) ↑	PA (%) ↑	ASD (*mm*) ↓	95th-percentile of HD (*mm*) ↓
UNet	FRSLoss	75.11 ± 0.21	**92.57 ± 0.08**	2.42 ± 0.03 ♯	**16.59 ± 0.17** $*$ ♯
	HDLoss	72.84 ± 0.13	89.97 ± 0.58	2.93 ± 0.08	21.06 ± 1.46
	DFRSLoss	75.16 ± 0.15	92.56 ± 0.19	**2.42 ± 0.01**	16.90 ± 0.09
	DHDLoss	72.79 ± 0.21	91.29 ± 0.98	2.79 ± 0.04	18.34 ± 0.88
	CELoss	75.55 ± 0.12	92.31 ± 0.15	2.43 ± 0.02	17.20 ± 0.09
	DLoss	**75.68 ± 0.05**	92.05 ± 0.04	2.47 ± 0.01	18.19 ± 0.22
FCN	FRSLoss	74.65 ± 0.13	**92.64 ± 0.05**	2.42 ± 0.00 $*$	**16.15 ± 0.11** $*$ ♯
	HDLoss	72.90 ± 0.81	91.88 ± 0.37	3.04 ± 0.21	22.43 ± 0.93
	DFRSLoss	75.59 ± 0.41	92.55 ± 0.06	**2.39 ± 0.05**	16.74 ± 0.27
	DHDLoss	73.53 ± 0.83	91.82 ± 0.30	2.74 ± 0.13	18.74 ± 0.96
	CELoss	75.52 ± 0.17	91.91 ± 0.19	2.50 ± 0.03	17.83 ± 0.40
	DLoss	**75.70 ± 0.04**	92.04 ± 0.11	2.44 ± 0.02	17.63 ± 0.34
SegNet	FRSLoss	74.21 ± 0.43	**92.20 ± 0.13**	2.55 ± 0.03 ♯	17.40 ± 0.25 $*$ ♯
	HDLoss	71.07 ± 0.23	91.90 ± 0.15	3.43 ± 0.07	25.73 ± 0.56
	DFRSLoss	74.41 ± 0.17	91.97 ± 0.18	**2.54 ± 0.03**	**17.18 ± 0.25**
	DHDLoss	72.20 ± 0.77	91.91 ± 0.09	2.82 ± 0.07	18.88 ± 0.76
	CELoss	74.37 ± 0.20	92.10 ± 0.16	2.56 ± 0.04	17.94 ± 0.32
	DLoss	**74.86 ± 0.19**	91.84 ± 0.10	2.60 ± 0.04	18.61 ± 0.34

time to reach the optimal state than that with HDLoss and DHDloss. It means that our proposed boundary-wise losses are capable of boosting the convergence speed compared with other boundary-wise losses. Note that the memory consumption of all methods is the same as they use the same backbone models for comparison.

4.4.2. Cell

Table 3 shows the performance on the cell dataset using different losses. FRSLoss and DFRSloss performed better than HDLoss and DHDLoss in DC, PA, ASD and 95th-percentile of HD. It further verifies that our proposed boundary-wise losses are superior to Hausdroff distance based losses and achieved higher stability in training segmentation models.

Table 3 also depicts the experimental results for the pixel-wise loss and region-wise loss in the UNet, FCN and SegNet models. It can be seen that the FRSLoss achieves better performance than CELoss and DLoss in the boundary-wise measures (measured by ASD and 95th-percentile of HD), which indicates that our proposed FRSLoss pays more attention to the boundaries than the pixel-wise and region-wise losses. As the same with the Nuclei dataset, wilcoxon sign rank test with $p < 0.05$ also is adopted to explore if there is a statistically significant difference between FRSLoss, CELoss and DLoss. Experimental results in Table 3 show that FRLoss has a statistically significant difference with CELoss and DLoss in 95th-percentile of HD for all three semantic segmentation models, while for the ASD boundary metric, FRSLoss has a statistically significant difference with DLoss on the Unet and SegNet models, with CELoss on the FCN model. For DC and PA, no statistical difference between our method and DLoss/CELoss

4.4.3. Qualitative analysis

Table 1 and Table 3 show the results of quantitative analysis in comparison of different loss functions. In this section, the corresponding qualitative analysis is provided. Fig. 4 shows some segmentation results for the three semantic segmentation models and two datasets. Fig. 4 (a)(b) are the original image and the corresponding label image, respectively. From Fig. 4 (b), the nuclei dataset has more complicated segmentation boundaries and the boundaries appear as various shapes compared with the cell dataset. Fig. 4 (c)(e) are the predicted segmentation image of our proposed FRSLoss and DFRSloss. Fig. 4 (d)(f)(g)(h) are the predicted segmentation images based on HDLoss, DHDLoss, CELoss and DLoss, respectively. As shown in Fig. 4, the semantic segmentation models with FRSLoss yield more precise boundary. Although the evaluation indexes of segmentation performance are nearly the same for FRSLoss, CELoss and DLoss, the predicted segmentation images in Fig. 4 show that the boundary of CELoss is relatively blurry and the boundary of DLoss is not as smooth as that of FRSLoss (The red circles and arrows in Fig. 4 represent the boundary difference

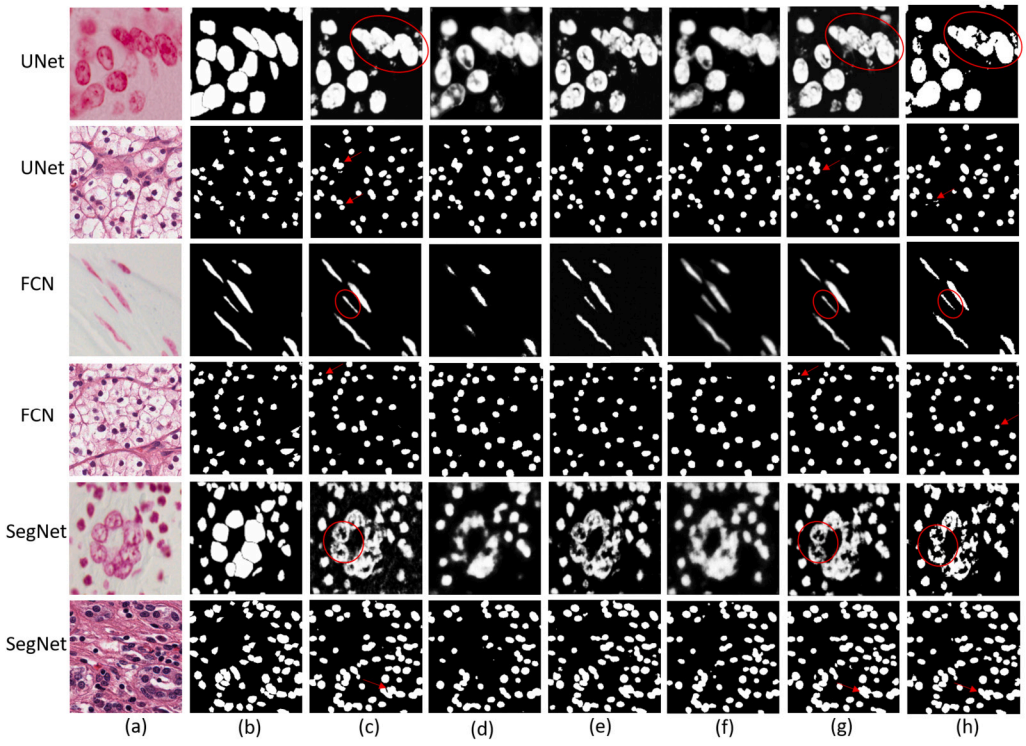

Fig. 4. Segmentation results for two datasets and three models: (a) original image; (b) benchmark image; (c) predicted segmentation image of FRSLoss; (d) predicted segmentation image of HDLoss; (e) predicted segmentation image of DFRSLoss; (f) predicted segmentation image of DHDLoss;(g) predicted segmentation image of CELoss; (h) predicted segmentation image of DLoss. The red circle and arrow represent the boundary difference between FRSLoss, CELoss and DLoss.

between FRSLoss, CELoss and DLoss). It suggests that the boundary-wise loss pays much attention to the boundary, resulting in FRSLosss achieving more accuracy and distinct boundaries than CELoss and DLoss.

5. Discussion

Table 1 and Table 3 show that our proposed novel boundary-wise losses, FRSLoss and DFRSLoss, have the ability to enhance the segmentation performance considerably compared with the other boundary-based losses, HDLoss and DHDLoss. To further explore the differences between FRSLoss and HDLoss, the training process and testing process for the nuclei dataset are visualized. Fig. 5 and Fig. 6 depict the variation curves of the training and testing processes for FRSLoss, HDLoss, DFRSLoss and DHDLoss. Blue curves mean the UNet model, orange curves refer to the FCN model and green curves are the SegNet model. As shown in Fig. 5 and Fig. 6, the training process and testing process for HDLoss and DHDLoss fluctuate violently, while the training process and testing process for FRSLoss and DFRSLoss are more stable. Moreover, the range of FRSLoss and DFRLoss value is from 0 to 1, while the range of HDLoss and DHDLoss value is from 0 to infinity. The aforementioned differences in Fig. 5 and Fig. 6 are due to the fact that there is a Gaussian function in the formula of FRSLoss and DFRLoss, which plays an important role in normalizing the distance. There are two merits for the normalization operator: 1) it excludes the influence of outliers and makes the gradient descent process more stable and robust (shown in Fig. 5 and Fig. 6); 2) it accelerates the convergence rate and reduces the training time (shown in Table 2). Thus, our proposed novel boundary-wise losses, FRSLoss and DFRSLoss are more stable and efficient than the other boundary-based losses, HDLoss and DHDLoss.

Compared Table 1 with Table 3, our proposed FRSLoss and DFRSLoss perform better on the nuclei dataset than that on the cell dataset. The segmentation performance difference between our proposed boundary-wise losses and other boundary-wise losses on the nuclei dataset is more than four times than that on the cell dataset. The reason is that the nuclei dataset has more complicated boundaries than the cell dataset. From Fig. 4, the boundaries of the nuclei dataset have various shapes while the boundaries of the cell dataset are relatively simple and uniform. The primary merit of our proposed boundary-loss is that it pays more attention to the boundary compared with other types of losses. Therefore, FRSLoss and DFRSloss are more beneficial to be applied to objects with complicated boundaries.

Furthermore, Table 1 and Table 3 show that our proposed loss concentrates more on the boundaries than the pixel-wise and region-wise losses. In many practical applications, for instance 3D multi-class image segmentation problems, one single loss is generally unable to obtain a satisfying segmentation result. The popular solution is to integrate the pixel-wise loss, region-wise loss and boundary-wise loss. In this way, the semantic segmentation models are capable of focusing the pixel, region and boundary simultaneously, which helps to improve the segmentation performance. Therefore, a stable and robust boundary-wise loss is of

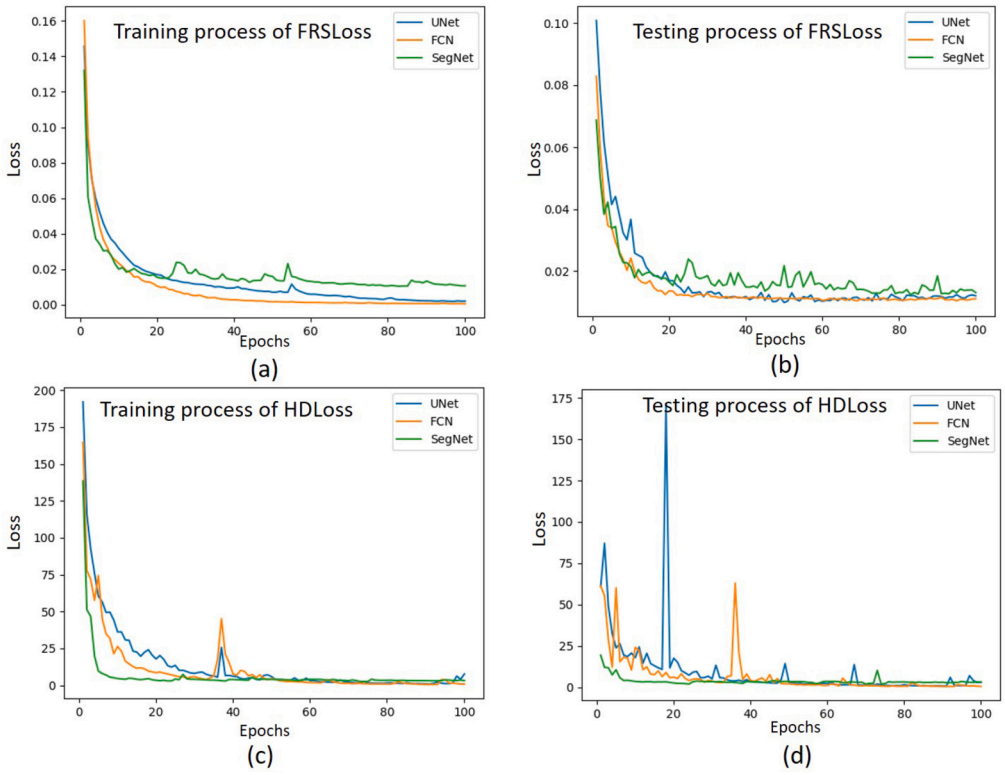

Fig. 5. Loss variation curves for UNet, FCN and SegNet: (a) training loss for FRSLoss; (b) testing loss for FRSLoss; (c) training loss for HDLoss; (d) testing loss for HDLoss.

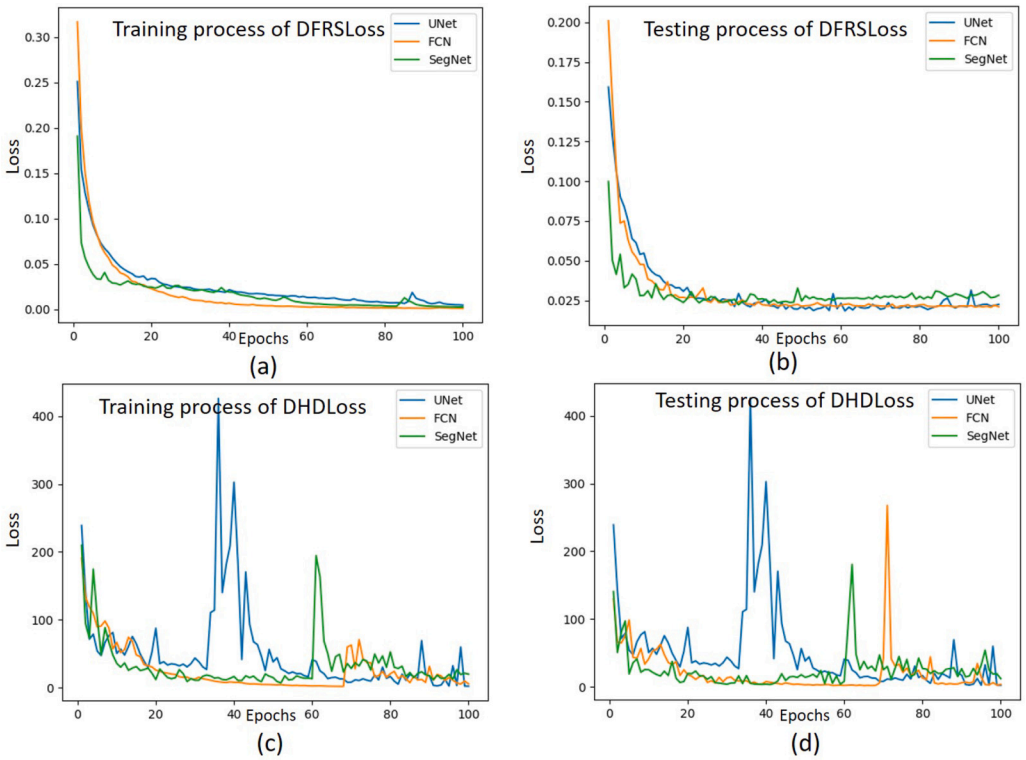

Fig. 6. Loss variation curves for UNet, FCN and SegNet: (a) training loss for DFRSLoss; (b) testing loss for DFRSLoss; (c) training loss for DHDLoss; (d) testing loss for DHDLoss.

considerable importance. According to the experimental results, our proposed boundary-wide losses, FRSLoss and DFRSLoss, are superior to the other boundary-wide losses, HDLoss and DHDLoss, and have the potential to compete with commonly used pixel-wise and region-wise losses, which indicates that our proposed boundary-wise loss is more suitable to combine with the pixel-wise loss, region-wise loss than other boundary-wise losses.

6. Conclusions

In this paper, we have proposed a novel boundary-wise loss namely FRSL-oss that can be used in various semantic segmentation models. The FRSLoss derives from the lower approximation of fuzzy rough sets. In our proposed FRSLoss formula, the introduction of the Gaussian kernel is able to normalize the boundary difference between the predicted segmentation and the ground truth segmentation, which stabilizes the convergence procedure and consumes less time. Considering the non-convex nature of the lower approximation of fuzzy rough sets, the distance transform algorithm is applied to calculate the FRSLoss in semantic segmentation tasks. Moreover, the extension of our proposed FRSloss to multi-class semantic segmentation is also investigated and discussed in this paper, which broadens the application range further. It should be noted that it is the first time the fuzzy rough sets theory is directly implemented in semantic segmentation models. The experiments with various segmentation models and datasets have verified that the proposed fuzzy rough sets loss is superior to other boundary-wise losses in terms of the segmentation accuracy and time complexity. Compared with the commonly used pixel-wise and region-wise losses, the proposed boundary-wise loss has a similar performance but pays more attention to the boundaries.

The selection of s-norm and fuzzy relation in the FRSLoss formula is fixed in this paper. Whether the segmentation performance will be enhanced by other s-norms and fuzzy relations is still an open question and requires further exploration. Furthermore, in this paper, we only take the signal loss into consideration. However, for many complicated datasets, the combination of the pixel-wise loss, region-wise loss and boundary-wise loss will be utilized in semantic segmentation models to deal with the pixel, region and boundary simultaneously, which is beneficial for the improvement of segmentation performance. Therefore, in the future, we will associate the proposed FRSLoss with CELoss and DLoss to handle 3D multi-class image segmentation problems.

CRediT authorship contribution statement

Qiao Lin: Writing – review & editing, Writing – original draft, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Xin Chen:** Writing – review & editing, Supervision. **Chao Chen:** Supervision. **Jonathan M. Garibaldi:** Visualization.

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

Data will be made available on request.

Acknowledgements

This work was supported by the University of Nottingham.

References

[1] A. Hannun, C. Case, J. Casper, B. Catanzaro, G. Diamos, E. Elsen, R. Prenger, S. Satheesh, S. Sengupta, A. Coates, et al., Deep speech: scaling up end-to-end speech recognition, arXiv preprint arXiv:1412.5567, 2014.
[2] X. Chen, R. Girshick, K. He, P. Dollár Tensormask, A foundation for dense object segmentation, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2019, pp. 2061–2069.
[3] A. Vaswani, S. Bengio, E. Brevdo, F. Chollet, A.N. Gomez, S. Gouws, L. Jones, Ł. Kaiser, N. Kalchbrenner, N. Parmar, et al., Tensor2tensor for neural machine translation, arXiv preprint arXiv:1803.07416, 2018.
[4] E. Choi, A. Schuetz, W.F. Stewart, J. Sun, Using recurrent neural network models for early detection of heart failure onset, Journal of the American Medical Informatics Association 24 (2017) 361–370.
[5] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.
[6] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (2017) 115–118.
[7] M. Yi-de, L. Qing, Q. Zhi-Bai, Automated image segmentation using improved pcnn model based on cross-entropy, in: Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004, IEEE, 2004, pp. 743–746.
[8] C.H. Sudre, W. Li, T. Vercauteren, S. Ourselin, M.J. Cardoso, Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, Springer, 2017, pp. 240–248.
[9] S. Xie, Z. Tu, Holistically-nested edge detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1395–1403.
[10] T.-Y. Lin, P. Goyal, R. Girshick, K. He, P. Dollár, Focal loss for dense object detection, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2980–2988.

[11] S.A. Taghanaki, Y. Zheng, S.K. Zhou, B. Georgescu, P. Sharma, D. Xu, D. Comaniciu, G. Hamarneh, Combo loss: handling input and output imbalance in multi-organ segmentation, Comput. Med. Imaging Graph. 75 (2019) 24–33.
[12] A. Mehrtash, W.M. Wells, C.M. Tempany, P. Abolmaesumi, T. Kapur, Confidence calibration and predictive uncertainty estimation for deep medical image segmentation, IEEE Trans. Med. Imaging 39 (2020) 3868–3878.
[13] D. Karimi, S.E. Salcudean, Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks, IEEE Trans. Med. Imaging 39 (2019) 499–513.
[14] D. Dubois, H. Prade, Rough fuzzy sets and fuzzy rough sets, Int. J. Gen. Syst. 17 (1990) 191–209.
[15] V. Murali, Fuzzy equivalence relations, Fuzzy Sets Syst. 30 (1989) 155–163.
[16] Q. Chen, M. Huang, H. Wang, G. Xu, A feature discretization method based on fuzzy rough sets for high-resolution remote sensing big data under linear spectral model, IEEE Trans. Fuzzy Syst. (2021).
[17] N. Mac Parthaláin, R. Jensen, R. Diao, Fuzzy-rough set bireducts for data reduction, IEEE Trans. Fuzzy Syst. 28 (2019) 1840–1850.
[18] Q. Hu, L. Zhang, Y. Zhou, W. Pedrycz, Large-scale multimodality attribute reduction with multi-kernel fuzzy rough sets, IEEE Trans. Fuzzy Syst. 26 (2017) 226–238.
[19] T. Lindeberg, M.-X. Li, Segmentation and classification of edges using minimum description length approximation and complementary junction cues, Comput. Vis. Image Underst. 67 (1997) 88–98.
[20] K.J. Batenburg, J. Sijbers, Adaptive thresholding of tomograms by projection distance minimization, Pattern Recognit. 42 (2009) 2297–2305.
[21] R. Nock, F. Nielsen, Statistical region merging, IEEE Trans. Pattern Anal. Mach. Intell. 26 (2004) 1452–1458.
[22] T. Lei, P. Liu, X. Jia, X. Zhang, H. Meng, A.K. Nandi, Automatic fuzzy clustering framework for image segmentation, IEEE Trans. Fuzzy Syst. 28 (2019) 2078–2092.
[23] J. Long, E. Shelhamer, T. Darrell, Fully convolutional networks for semantic segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 3431–3440.
[24] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.
[25] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: a deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 2481–2495.
[26] H. Noh, S. Hong, B. Han, Learning deconvolution network for semantic segmentation, in: Proceedings of the IEEE International Conference on Computer Vision, 2015, pp. 1520–1528.
[27] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Semantic image segmentation with deep convolutional nets and fully connected crfs, arXiv preprint arXiv:1412.7062, 2014.
[28] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (2017) 834–848.
[29] L.-C. Chen, G. Papandreou, F. Schroff, H. Adam, Rethinking atrous convolution for semantic image segmentation, arXiv preprint arXiv:1706.05587, 2017.
[30] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 801–818.
[31] S. Jadon, A survey of loss functions for semantic segmentation, in: 2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), IEEE, 2020, pp. 1–7.
[32] V. Pihur, S. Datta, S. Datta, Weighted rank aggregation of cluster validation measures: a Monte Carlo cross-entropy approach, Bioinformatics 23 (2007) 1607–1615.
[33] S.R. Hashemi, S.S.M. Salehi, D. Erdogmus, S.P. Prabhu, S.K. Warfield, A. Gholipour, Asymmetric loss functions and deep densely-connected networks for highly-imbalanced medical image segmentation: application to multiple sclerosis lesion detection, IEEE Access 7 (2018) 1721–1735.
[34] S.S.M. Salehi, D. Erdogmus, A. Gholipour, Tversky loss function for image segmentation using 3d fully convolutional deep networks, in: International Workshop on Machine Learning in Medical Imaging, Springer, 2017, pp. 379–387.
[35] N. Abraham, N.M. Khan, A novel focal Tversky loss function with improved attention u-net for lesion segmentation, in: 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), IEEE, 2019, pp. 683–687.
[36] Z. Pawlak, Rough sets, Int. J. Comput. Inf. Sci. 11 (1982) 341–356.
[37] Q. Hu, L. Zhang, D. Chen, W. Pedrycz, D. Yu, Gaussian kernel based fuzzy rough sets: model, uncertainty measures and applications, Int. J. Approx. Reason. 51 (2010) 453–471.
[38] L.A. Zadeh, Fuzzy sets, Inf. Control 8 (1965) 338–353.
[39] L.A. Zadeh, Fuzzy sets, in: Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems: Selected Papers by Lotfi a Zadeh, World Scientific, 1996, pp. 394–432.
[40] J.-S. Mi, W.-X. Zhang, An axiomatic characterization of a fuzzy generalization of rough sets, Inf. Sci. 160 (2004) 235–249.
[41] D.S. Yeung, D. Chen, E.C. Tsang, J.W. Lee, W. Xizhao, On the generalization of fuzzy rough sets, IEEE Trans. Fuzzy Syst. 13 (2005) 343–361.
[42] Y. Wang, Q. Hu, P. Zhu, L. Li, B. Lu, J.M. Garibaldi, X. Li, Deep fuzzy tree for large-scale hierarchical visual classification, IEEE Trans. Fuzzy Syst. 28 (2019) 1395–1406.
[43] M.G. Genton, Classes of kernels for machine learning: a statistics perspective, J. Mach. Learn. Res. 2 (2001) 299–312.
[44] J. Goutsias, S. Batman, Morphological methods for biomedical image analysis, in: Handbook of Medical Imaging, vol. 2, 2000, pp. 175–272.
[45] G.J. Grevera, J.K. Udupa, Shape-based interpolation of multidimensional grey-level images, IEEE Trans. Med. Imaging 15 (1996) 881–892.
[46] G. Penney, J. Little, J. Weese, D. Hill, D. Hawkes, Deforming a preoperative volume to represent the intraoperative scene, Comput. Aided Surg. 7 (2002) 63–73.
[47] M. Herk, Image registration using chamfer matching, in: Handbook of Medical Imaging Processing and Analysis, 2000.
[48] P.F. Felzenszwalb, D.P. Huttenlocher, Distance transforms of sampled functions, Theory Comput. 8 (2012) 415–428.
[49] H. Irshad, L. Montaser-Kouhsari, G. Waltz, O. Bucur, J. Nowak, F. Dong, N.W. Knoblauch, A.H. Beck, Crowdsourcing image annotation for nucleus detection and segmentation in computational pathology: evaluating experts, automated methods, and the crowd, in: Pacific Symposium on Biocomputing Co-Chairs, World Scientific, 2014, pp. 294–305.