

RESEARCH

Open Access



B EGL: boundary enhancement with Gaussian Loss for rock-art image segmentation

Chuanping Bai^{1,2}, Yangyang Liu^{1,3}, Pengbo Zhou⁴, Xiaofeng Wang^{1,3} and Mingquan Zhou^{1,3,4*}

Abstract

Rock-art has been scratched, carved, and pecked into rock panels all over the world resulting in a huge number of engraved figures on natural rock surfaces that record ancient human life and culture. To preserve and recognize these valuable artifacts of human history, 2D digitization of rock surfaces has become a suitable approach due to the development of powerful 2D image processing techniques in recent years. In this article, we present a novel systematical framework for the segmentation of different petroglyph figures from 2D high-resolution images. The novel boundary enhancement with Gaussian loss (B EGL) function is proposed aiming at refining and smoothing the rock-arts boundaries in the basic UNet architecture. Several experiments on the 3D-pitoti dataset demonstrate that our proposed approach can achieve more accurate boundaries and superior results compared with other loss functions. The comprehensive framework of petroglyph segmentation from 2D high-resolution images provides the foundation for recognizing multiple petroglyph marks. The framework can then be extended to other cultural heritage digital protection domain easily.

Keywords Petroglyph segmentation, Boundary enhancement, Cultural heritage, Rock-art

Introduction

Petroglyphs are the most widespread, ancient and long-lasting rock-art in the world which have been incised, pecked, scratched or carved into rock surfaces [1]. Many figures and significant marks are present on rock surfaces. Rock-art is an important way of recording and exhibiting ancient human life and culture. Since rock paintings have a long history, natural weathering and

man-made destruction have been threatening the life of petroglyphs [1]. There is an urgent need to protect and identify petroglyphs.

Traditionally, rock-arts around the world have been recorded and preserved using a broad variety of approaches, including manual contact tracing, casting with plaster and frottage [2]. Due to the large quantity of petroglyphs which have been found out so far and some of rock-arts are in the cliffs, many manual documenting methods become infeasible [2–4]. Furthermore, this is an extremely time-consuming and repetitive work for documenting these pre-historic resources [4]. With the advances of digital photography and automatic image processing techniques, the number of digital images of complete petroglyphs will grow steadily [2, 5]. The automatic segmentation of petroglyph shapes is a basic and upstream task for recognizing rock-art and distinguishing rock painting artistic styles [6]. Segmentation of rock-art is to, firstly, classify the image in pecked and

*Correspondence:

Mingquan Zhou
mqzhou@nwnu.edu.cn

¹ School of Information Science and Technology, Northwest University, 1 Xuefu Avenue, Guodu Education Technology Industrial Park, Chang'an District, Xi'an 710127, China

² School of Mathematics and Computer science, Ningxia Normal University, Guyuan, China

³ National and Local Joint Engineering Research Center for Cultural Heritage Digitization, Northwest University, Xi'an, China

⁴ Virtual Reality Research Center of Ministry of Education, Beijing Normal University, Beijing, China



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

unpecked regions, and secondly, segment the different figures as well as different symbols in details. Related research mainly pays attention to the interactive segmentation with the appropriate combination of different visual features and automated classification of rock surfaces in terms of feature descriptors [2, 7]. Existing works also consider petroglyph shape similarity measure approaches for data mining and shape retrieval [6, 8, 9]. Recent work has mainly focused on the surface segmentation utilizing native 3D attributes of rock surfaces and discriminating pecking styles in a hybrid 2D/3D method [10–12]. Also, a publicly available dataset has been published for 2D or 3D rock-art surfaces segmentation [4]. Besides, valuable information acquired from automated tracing can be added to a rock-art inventory that can improve interpretation on rock-art artistic styles [13]. Although those methods achieved promising performance on petroglyph segmentation, the complexity of petroglyph makes it a very challenging problem.

The automated segmentation for rock-art shapes is still unsolved and even considered infeasible which is an significant pre-processing step in this field [6]. Just a little works for the pixel-wise classification of petroglyph shapes have been done. Zhu et al. [9] proposed a collaborative manual segmentation approach that utilizes completely automated public turing test to tell computers and humans apart (CAPTCHA) for rock art image segmentation. Seidl and Breiteneder [2] developed a method for the pixel-wise classification of petroglyphs directly from images of natural rock panels. An integration of support vector machine (SVM) classifiers was trained for the appropriate combination of lots of visual features, then they devised a fusion of the classified results that allowed the interactive refining of the segmentation by the user. Vincenzo and Paolino [14] proposed a novel method for the segmentation of rock-art figures and recognition of carving symbols. A shape descriptor derived by 2D Fourier transform is applied to identify petroglyph figures, which is insensitive to shape deformations and robust to scale and rotation. Recently, the work [4] presented a 3D-pitoti dataset of high-resolution surface reconstructions which consists of the whole geometric information as well as color information. The 3D scanner acquired both the tactile and visual appearance of the rock panels at a millimetre scale. Of course, the intelligent segmentation methods benefit strongly from full 3D geometric information in contrast to only 2D textures [15]. Furthermore, they tested and verified various tasks on this dataset [4] that should serve as first public baseline in rock-art field. They evaluate the performance of semantic segmentation for petroglyph with common approaches based on random forests(RF) and fully convolutional networks (FCN). In contrast to these previous approaches

for rock-art image segmentation, we focus on fully automatic segmentation framework based on convolutional neural networks (CNNs).

Objective or loss function is especially significant while devising complicated image segmentation models based on deep learning architectures as it advances the learning effects step by step [16]. Binary cross entropy loss [17] is the most universal objective function in the domain of image semantic segmentation. Cross entropy loss function achieves the better results on balanced dataset, but not on imbalanced dataset, so some variants of cross entropy are devised, such as weighted cross entropy (WCE) [18], balanced cross entropy (BCE) [19]. Focal loss (FL) [20] assigns different weights to foreground pixels and background pixels, in order to change the case that some foreground pixels are overlapped or surrounded by many background pixels. In addition, it draws into hyper parameters and expects to update parameters. Dice loss (DL) [21] is designed to solve the phenomenon that many pixels are overlapped each other. When predicting, each category is calculated separately, and then the final result is obtained by averaging. The novel loss function called boundary enhancement (BE) loss [22] is introduced to concentrate on the boundary regions while training, so as to further improve the segmentation performance for the samples owning many blurred boundaries.

Deep learning technology has tremendously advanced the performance of image segmentation models, usually attaining the highest precise rates on popular benchmarks in recent years [23]. The milestone of image segmentation model based on deep learning inevitably is FCN, proposed by Long et al. [24] in 2015. Subsequently, the variants of FCN have created a boom in the field of image segmentation. The FCN model consists of only the convolution layers instead of the fully connected layers, which enforce it to achieve a segmentation map whose size is the same as the input image. Badrinarayanan et al. [25] proposed SegNet that contains an encoder network and a symmetrical decoder network, utilizing pooling indices calculated in the max-pooling step of the corresponding encoder to perform unpooling in the decoder network. UNet is one of the most distinguished architectures for medical image segmentation, initially introduced by Ronneberger et al. [26] using the principle of deconvolution. The UNet architecture consists of two components, a shrinking branch to extract features, and a symmetric enlarging branch that focuses on precise localization. The most important property of UNet is the skipping connections between layers of the same resolution in encoding path to decoding path. These shortcut connections contain local detailed data that providing crucial high-resolution features to the deconvolution layers. Moreover, the UNet training tactic depends on the

applying of data augmentation to learn effectively from very little labeled data. Finally, UNet is also great rapider to learn than the most other segmentation architectures due to its global based learning strategies [27]. Our rock-art segmentation network is based on UNet [26] that has won the first places in many international segmenting and tracking contests.

Owing to the vast diversity of different signs and symbols, many kinds of carving styles, lots of pecking tools and pecking styles, as well as various forms of rock surfaces, diverse degrees of deterioration and scribble noises, the situation of rock drawings segmentation is especially difficult [4]. One of the main challenges for rock drawings segmentation is that component of the rock-art lacks of sharp boundaries with surrounding degraded regions. Without adequate training data is another major challenge, which makes it difficult to get complicated networks completely trained as enough labeled data is a critical pole of the success of convolutional neural networks (CNNs). This work makes the effort to solve the aforementioned difficulties and challenges, a comprehensive petroglyph segmentation framework is proposed for pixel-wise classification of extremely deteriorated training data, especially, blurred and superimposed figures in petroglyph data. Moreover, to accelerate the rock drawings segmentation network rapidly converge to segmentation boundaries, we propose a novel boundary enhancement with Gaussian loss (BEGL) as the supervised loss of segmenting network for petroglyphs. The segmentation effects show that our framework can achieve better and precise masks while segmenting blurred boundaries. For evaluation, we demonstrate our method on the 3D-pitoti dataset benchmark [4]. We also compare BEGL to other state-of-the-art loss function utilized in the proposed framework performed on the benchmark dataset [4].

The innovative contributions of the proposed method can be summarized as follows:

- 1 We propose a systematic petroglyph segmentation framework for accurate surface segmentation of complex rock-art.
- 2 We propose a novel loss function named BEGL aiming at refining and smoothing the rock-art boundaries, which could be easily implemented and plugged into any backbone networks.
- 3 The new framework desired for rock-art segmentation is an exploration in the cultural heritage digital protection domain.

The remainder of this paper is organized as follows: "Methods" section describes the methods in detail. "Overview on framework of petroglyph segmentation"

section lays out the experimental setup, objective, design and evaluation metrics. We introduce the results and discussion in "BEGL loss function" section. Finally, several concluding remarks are drawn in "Segmentation network" section.

Methods

In this section, we first introduce the framework of our ancient rock-art segmentation, which heavily augments the training dataset and employs a novel BEGL loss function for emphasizing rock-art boundaries in UNet. Moreover, a novel BEGL loss function aiming at enhancing and refining the rock-art boundaries is described. Finally, We describe the segmented network architecture in detail.

Overview on framework of petroglyph segmentation

Segmentation of rock-art is an incredibly challenging task due to different levels of degradation of petroglyph boundary and much scribble noises on rock panels. For a more efficient rock-art segmenting, we concentrate on the systematic framework of petroglyph segmentation. The proposed boundary enhancement based rock-art image segmentation framework is presented in Fig. 1. It comprises two phases, namely the image preprocessing and segmenting phases.

Due to the petroglyph orthophotos are tilted in general, it is necessary to apply image rotation correction based on Fourier transform. The principle of 2D discrete Fourier transform (2D-DFT) can be defined as Eq. (1):

$$y(k, l) = \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} x(i, j) e^{-i2\pi \left(\frac{ki}{M} + \frac{lj}{N} \right)}. \quad (1)$$

$$e^{iz} = \cos z + i \sin z. \quad (2)$$

where $x(i, j)$ is the value of image spatial domain, i and j are the indices of image position, $y(k, l)$ is the value of image frequency domain, k and l are the discrete spatial frequencies, M and N are the number of pixels in the 2D image space. Also, Eq. (2) is Euler's formula, which establishes a connection between the complex exponential functions and the trigonometric functions. In essence, the application of 2D-DFT enables to convert signals in the spatial domain into the frequency domain conveniently. The Fourier spectrum is comprised of the sizes of the 2D-DFT complex coefficients, which are proportional to the strength of the spatial frequencies. Next, the corrected petroglyph images are sliced into small patches which can be taken into ResNet classifier as input. Because the large background often exists on the ancient rock-art panels which draws into great class imbalance, ResNet is selected as the classifier of the framework,

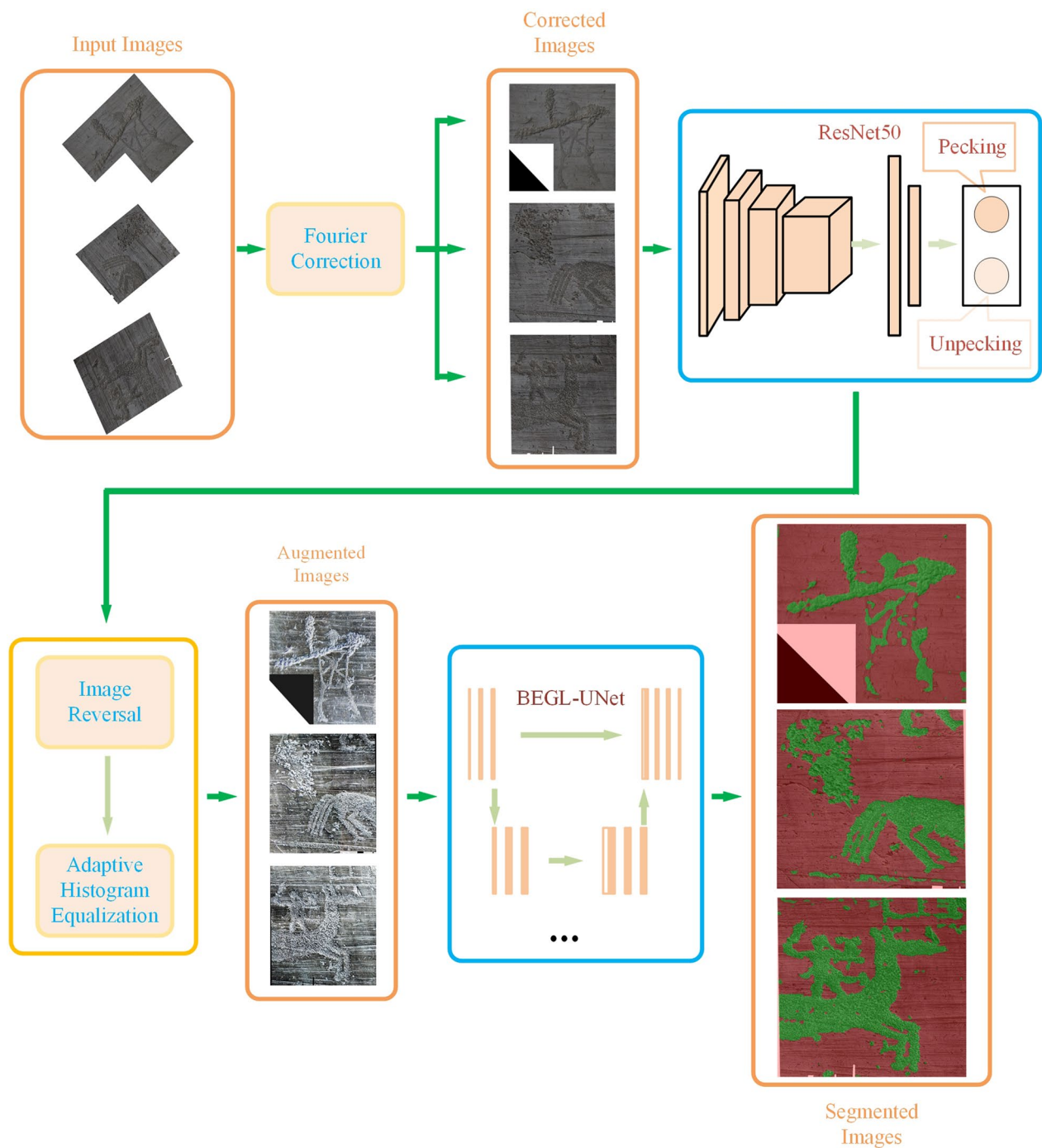


Fig. 1 Overview of the rock-art segmentation framework

which filters rock-art patches with no pecking marks. Figure 2 shows a class activation map (CAM) which selects pecked regions in red and drops unpecked regions in blue obtained from ResNet. Finally, in order to extract and emphasize the geometric patterns and boundaries related to the pecked-marks in the map that make up

petroglyph shapes, image reversal and image adaptive histogram equalization are applied in the framework.

The second phase is based on an UNet [26], which is an auto-encoder network with skip connections between layers of the same shape. We modify the

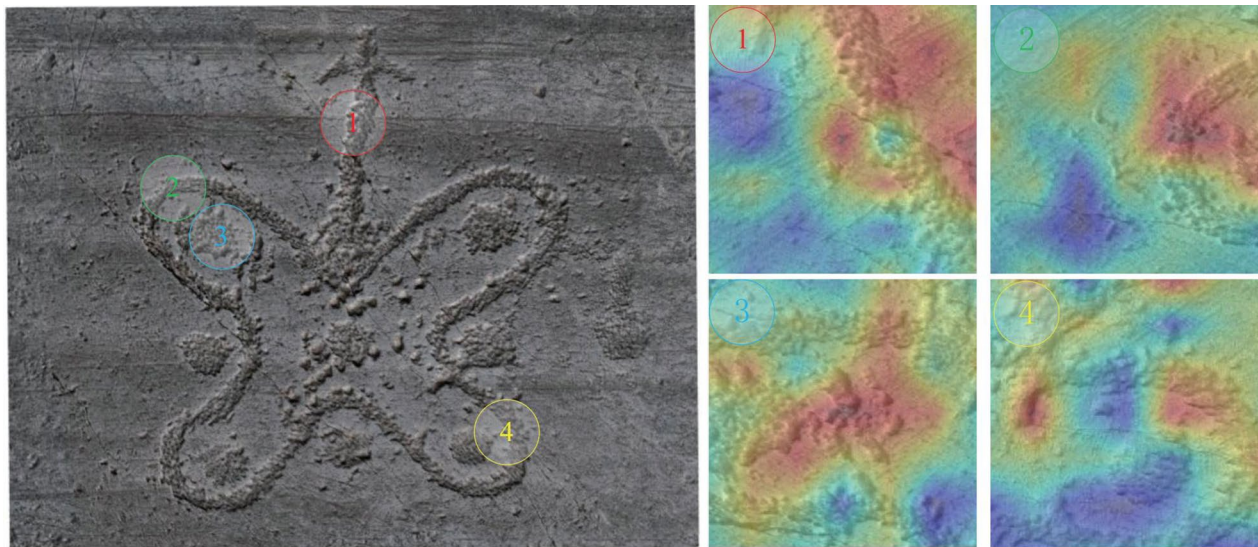


Fig. 2 This is a CAM which illustrates the ResNet selects pecked regions in red

-1	0	1	1	2	1
-2	0	2	0	0	0
-1	0	1	-1	-2	-1

(a) h-component (b) v-component

Fig. 3 Sobel operator (templates of filtering)

network by introducing a novel loss function named BEGL, allowing it to better learn rock-art boundary features.

BEGL loss function

In order to emphasize the boundary regions, we apply the Sobel operator, which generates strong responses around the boundary areas and little response elsewhere, to each point in a 2D image x in Eq. (3) and Eq. (4).

$$S_h = T_h * x \quad (3)$$

$$S_v = T_v * x \quad (4)$$

It is useful to express this as weighted density summations using the following weighting functions for h and v components. The two T_h and T_v templates used by Sobel are showed as Fig. 3a, b. The filters enable to be utilized individually to the input image, to generate individual

measures of the gradient components in each orientation. Then, These can be added to obtain the absolute magnitude of the gradient at every point. The orientation of the spatial gradient is given by Eq. (5):

$$\theta = \arctan\left(\frac{S_h}{S_v}\right) \quad (5)$$

The gradient magnitude S is given by Eq. (6):

$$|S| = \sqrt{S_h^2 + S_v^2} \quad (6)$$

Gaussian kernels are the most broadly applied in smoothing filters. These filters have been proved to play an important role in edge detection in human vision system, and to be very useful as detectors for edge and boundary detection [28]. The 2D Gaussian filter is also the only rotationally symmetric filter that is separable in Cartesian coordinates. Separability is important for computational efficiency when implementing the smoothing operation by convolutions in the spatial domain. The Gaussian filter in two dimensions can be defined as Eq. (7):

$$G(i, j) = \frac{1}{2\pi\sigma^2} e^{-\left(\frac{i^2+j^2}{2\sigma^2}\right)} \quad (7)$$

where ($\sigma = 0.8$) is the standard deviation of the Gaussian function and (i, j) are the Cartesian coordinates of the image. Standard 2D convolution operation can be used to calculate the discrete Gaussian filter. Hence, we can easily achieve the difference between filtered output of predictions of the CNNs and filtered output of the ground truth labels. Minimizing the divergence between two filtered

outputs enables to close the gap between the results of CNNs and ground truth labels. Following the analyses above, the boundary enhancement with Gaussian loss is defined as a L_2 -norm shown in Eq. (8):

$$L_G = \|G(S(y)) - G(S(\hat{y}))\|_2 \quad (8)$$

where y are the ground truth labels, \hat{y} are the prediction labels, $S(\cdot)$ is Sobel operator, and $G(\cdot)$ is Gaussian filter. Meanwhile, L_{BCE} effectively suppresses false positives and remote outliers, which are far away from the boundary regions. The formula of L_{BCE} is defined as Eq. (9).

$$L_{BCE}(y, \hat{y}) = -(\beta * y \log(\hat{y}) + (1 - \beta) * (1 - y) \log(1 - \hat{y})) \quad (9)$$

Here, y are the ground truth labels, \hat{y} are the prediction labels, β is defined as $1 - \frac{y}{H * W}$, H and W are height as well as width of the image. The overall BEGL loss function is defined as Eq. (10) that is derived from Eqs. (8, 9):

$$L_{BEGL} = \lambda_1 L_G + \lambda_2 L_{BCE} \quad (10)$$

where λ_1 is 0.001 and λ_2 is 1 respectively. The BEGL loss function is the combination of BCE loss [19] and Gaussian loss.

Segmentation network

The details of the segmentation network used in our work are provided in this section. In order to fully leverage the spatial contextual and boundary information of pecking rock-art data to accurately segment petroglyph images, a new BEGL loss function is designed for rock-art image segmentation network (BEGL-UNet) with inspiration from the work [22]. The BEGL-UNet architecture is showed in Fig. 4. It consists of an encoder-decoder structure resulting in an U-shape. The encoder applies max-pooling and a double convolution which halves the image size and doubles the number of feature maps, respectively. The decoder is comprised of three parts: a bilinear upsampling operation to double the feature map size, the feature maps of the encoder path are directly concatenated onto the corresponding layers in the decoder path, and lastly a double convolution to half the number of feature maps. The skip-connections enable the model to use multiple scales of the input to generate the output. This aids the network by propagating more semantic information between the two paths, thereby enabling it to segment images more accurately.

Experiments

Experimental setup

The proposed methods are implemented using the open source deep learning library TensorFlow1.10 [29] and python3.5. Each model is trained end-to-end with Adam

optimization method. In the training phase, the learning rate is initially set to 0.0001 and decreased by a weight decay of 1.0×10^{-6} after each epoch. The experiments were carried out on a NVIDIA GTX 2080ti GPU with 12GB memory. Due to the limitation of the GPU memory, we chose 2 as the batch size. In the testing phase, the segmented maps were stitched together once again.

Experimental objective

First of all, the aim of the current experiments is to test the availability of the systematical rock-art segmentation framework. Then, the purpose of the various experiments is to examine the effectiveness of BEGL loss function in image segmentation for ancient petroglyphs, and the performance of the BEGL loss function is tested by comparing those of other loss functions.

Experimental design

The public 3D-pitoti dataset [22] consists of 26 high-resolution surface reconstructions of natural rock surfaces with a large number of petroglyphs. The petroglyph dataset provides orthophotos of all surface reconstructions with a pixel-accurate ground truth. To alleviate the problem of extremely little training data, we use a sliding window to crop original high-resolution images to 512×512 small images without overlapping which also are processed with ease for BEGL-UNet. Then, we achieve an augmented dataset containing 548 images for training and evaluation. Experiments are conducted with two kinds of data splits that set aside 10% of the total images for the test set and other 90% of the total images for training. The normalization strategy with standard mean and deviation is employed to further boost the image data. As the rock-art orthophotos usually aren't aligned, image rotation correction based on Fourier transform is applied to original images. Furthermore, ResNet classifier is used to eliminate the unpecked small rock-art patches. Finally, we use data augmentation, in which images are reversed and equalized with adaptive histogram.

Evaluation metrics

Evaluation metric plays an important role in assessing the outcomes of segmentation models. In this work, we have analyzed our results using pixel accuracy, average precision, recall, F1-score, mean intersection over union (MIoU) and dice similarity coefficient (DSC) metrics. The pixel accuracy is the ratio between correctly classified pixels and the overall number of pixels. The average precision is measuring the average percentage of correct positive predictions among all predictions made. The recall rate refers to the proportion of pixels marked correctly in the mark of the result of artificial marking. The F1 score is a "harmonious"

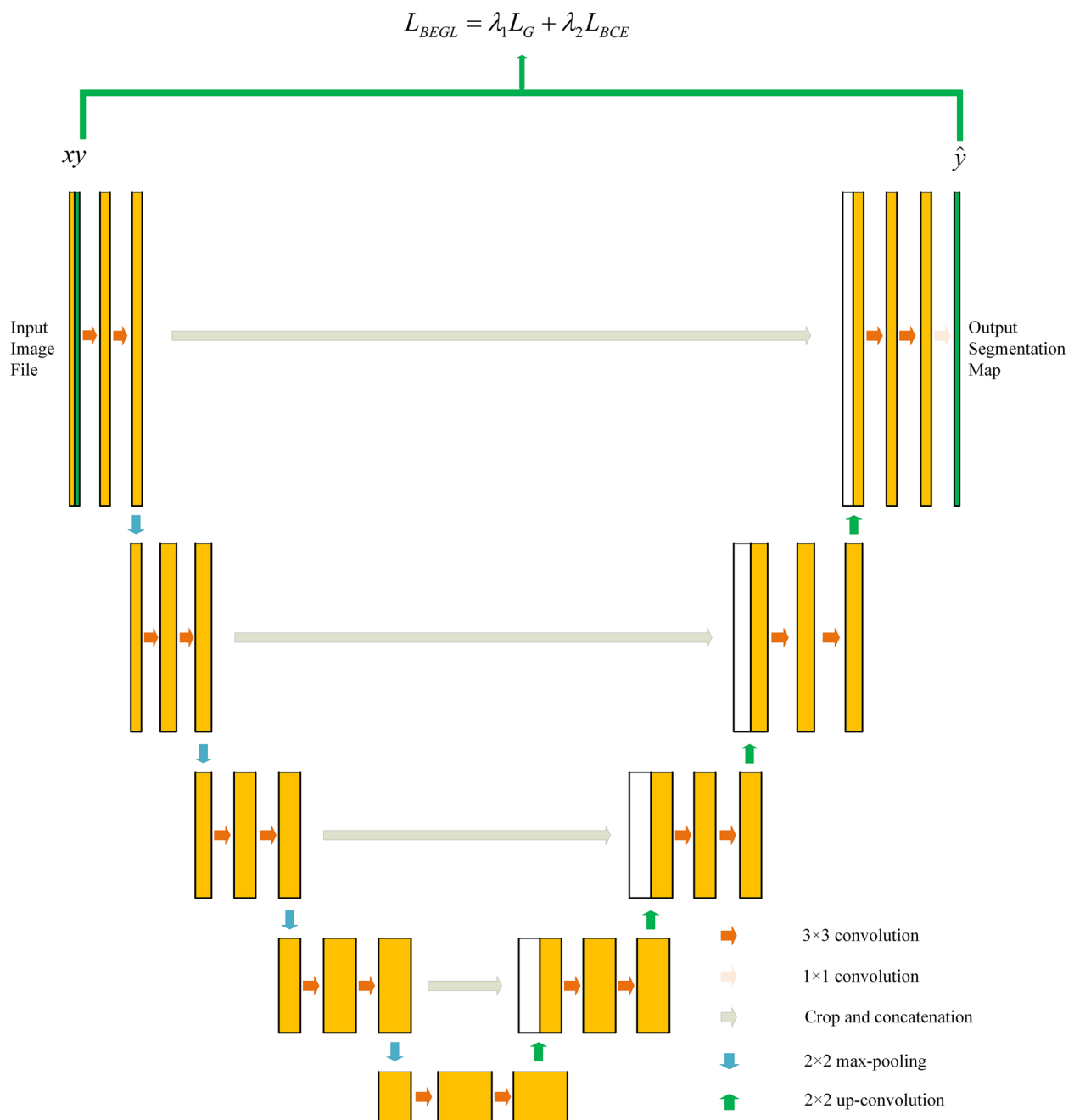


Fig. 4 This is an overview of the BEGL-UNet architecture based on the basic UNet

balance between precision and recall. MIOU is defined as the mean intersection of the predicted segmentation mask and the ground truth mask over their union. DSC, also known as overlapping index measures the overlapping between ground truth and predicted output.

Results and discussion

Comparison with other loss functions

The results in Table 1 describe the quantitative comparisons on the test set without overlap with the training set. It shows the rock-art segmentation performance

Table 1 The comparisons of the rock-art segmentation performance of BEGL-UNet with various loss functions on the test set

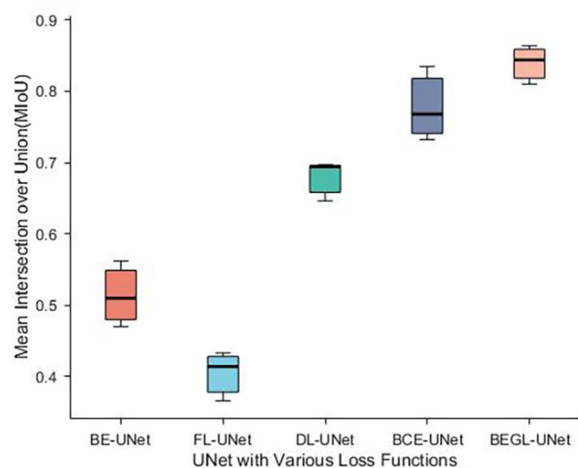
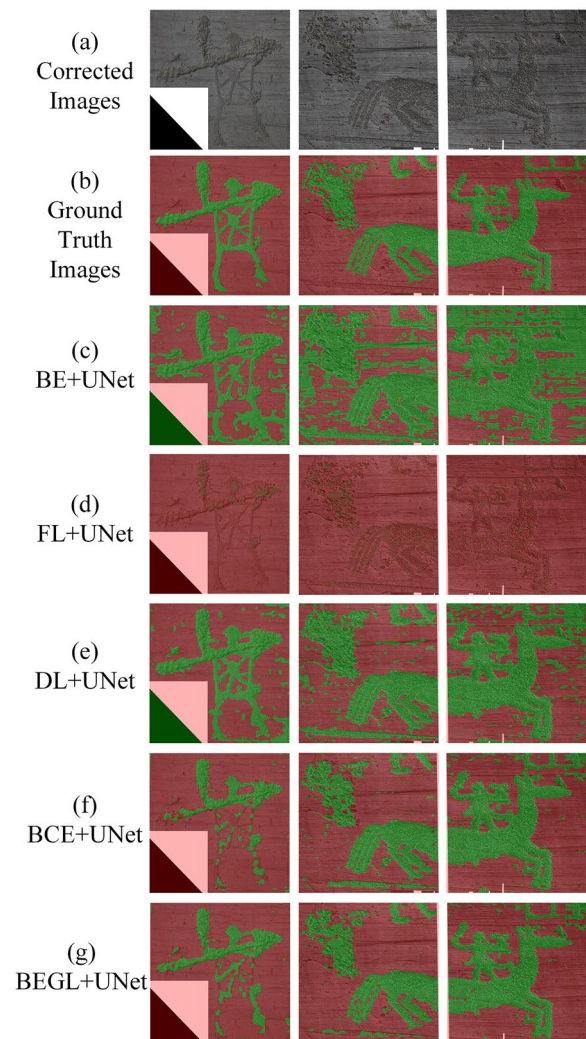
Methods	Accuracy	Precision	Recall	F1-score	MIoU	DSC
BE-UNet	0.695	0.450	0.964	0.613	0.514	0.613
FL-UNet	0.757	0.899	0.056	0.104	0.404	0.104
DL-UNet	0.829	0.591	0.992	0.739	0.679	0.739
BCE-UNet	0.907	0.804	0.832	0.804	0.778	0.804
BEGL-UNet	0.935	0.863	0.873	0.865	0.840	0.865

BEGL-UNet achieves the best results on Accuracy (0.935), F1 (0.865), MIoU (0.840) and DSC (0.865), only the worse results on precision and recall which are competitive with the best results

of BEGL-UNet with various loss functions that use the basic UNet architecture. From Table 1, we see that our approach achieves the best results on Accuracy (0.935), F1 (0.865), MIoU (0.840) and DSC (0.865), only the worse results on precision and recall which are competitive with the best results. The results in Table 1 clearly show the necessity for BEGL loss function to obtain refining and precise results on average. In addition, Fig. 5 shows visualization on the MIoU metric which makes great advance compared with other loss functions. The segmentation results of our proposed BEGL loss function have much smaller variance and less outliers compared to others.

Figure 6 demonstrates the visualization of segmented maps with various loss functions. From the results we observe that the BE-UNet, DL-UNet and BCE-UNet are insensitive to noise, whereas the BEGL-UNet yields more consistent as well as refining segmented results. In particular, BEGL loss function help enhance the performance of petroglyph segmentation network. The FL-UNet correctly detects small and thin pecked regions but misses larger pecked regions. Fig. 7 shows that BEGL-UNet achieves more smooth and refined segmented

maps than other loss functions in the zooming in maps. Furthermore, the zooming in maps in Fig. 7 illustrate rock-art boundary is the vital element for petroglyph segmentation.

**Fig. 5** Mean Intersection over Union (MIoU) across UNet architecture with various loss functions**Fig. 6** This is the visualization of segmented maps with UNet based on various loss functions

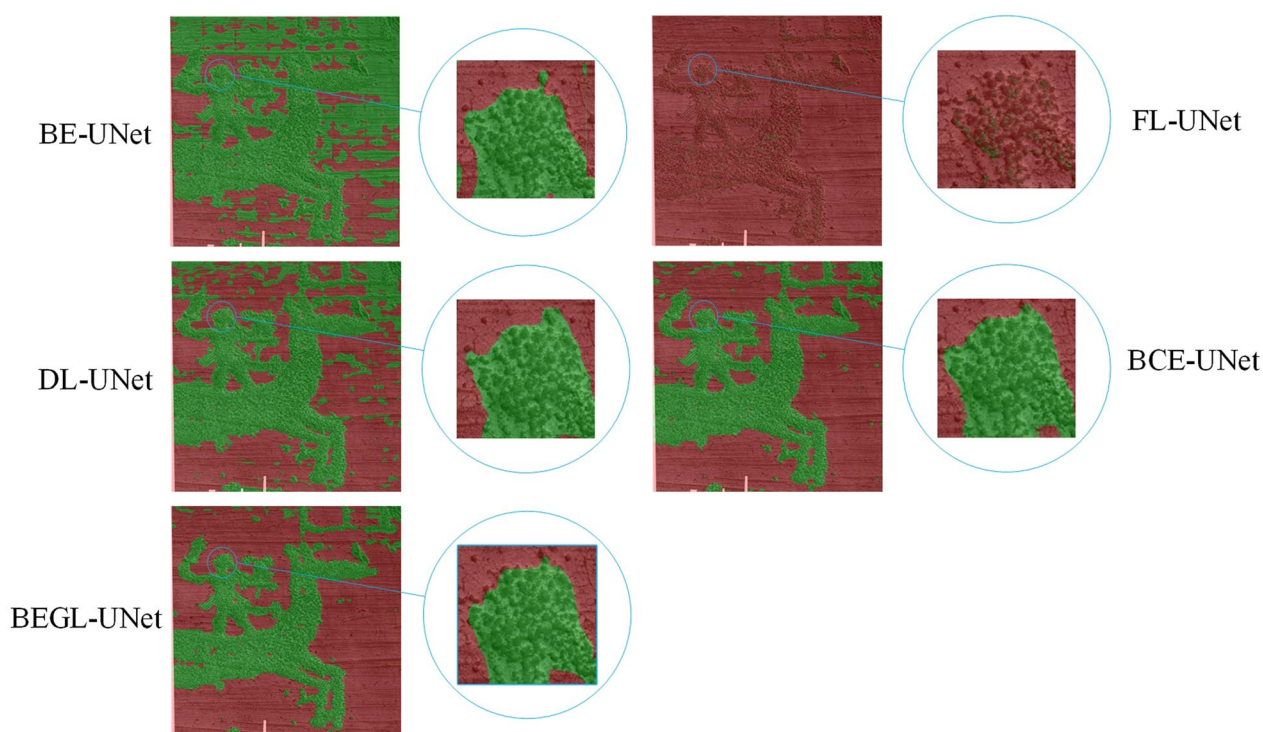


Fig. 7 Qualitative comparisons on the different loss functions in the zooming in views

Conclusions

In this paper, we have presented a novel framework for the segmentation of petroglyph shapes from 2D high-resolution images. The novel BEGL loss function is deployed in the basic UNet architecture. It addresses two challenges in rock-art image segmentation, which are the lack of clear boundary and the lack of enough annotated data for training CNNs. Several experiments on the 3D-pitoti dataset demonstrate that our proposed method can get more accurate boundaries and achieve superior results compared with other loss functions. In our future work, we will extend the proposed method to segment petroglyphs from other imaging modalities.

Abbreviations

BEGL	Boundary enhancement with Gaussian Loss
CNNs	Convolutional neural networks
CAPTCHA	Completely Automated Public Turing test to tell Computers and Humans Apart
SVM	Support vector machine
RF	Random forests
FCN	Fully convolutional networks
WCE	Weighted cross entropy
BCE	Balanced cross entropy
FL	Focal loss
DL	Dice loss
BE	Boundary enhancement
2D-DFT	2D discrete fourier transform
CAM	Class activation map
MIoU	Mean intersection over union
DSC	Dice similarity coefficient

Acknowledgements

We acknowledge the National Key Research and Development Program of China (No. 2019YFC1521103, No. 2020YFC1523301 and No. 2020YFC1523303), the Key Research and Development Project of Qinghai Province (No. 2020-SF-142) National Natural Science Foundation of China under grant (No. 62262054) and the Key Research and Development Program of Shaanxi Province (No. 2021GY-171) for supporting our work.

Author contributions

CB mainly contributed to the design and implementation of the research, to the analysis of the experiments and results. YL, XW and PZ contributed to some experiments and data curation. CB wrote the main manuscript in consultation with MZ. All authors read and approved the final manuscript. All authors commented on previous versions of the manuscript. Besides, all authors read and approved the final manuscript.

Funding

This work is mainly supported by the National Key Research and Development Program of China under grant (No. 2019YFC1521103, No. 2020YFC1523301 and No. 2020YFC1523303). Besides, this study is also partly supported by the Key Research and Development Project of Qinghai Province under grant (No. 2020-SF-142), National Natural Science Foundation of China under grant (No. 62262054) and the Key Research and Development Program of Shaanxi Province under grant (No. 2021GY-171).

Availability of data and materials

Data used in this research is publicly available at <https://www.tugraz.at/institute/icg/research/team-bischof/learning-recognition-surveillance/downloads/3dpitotidataset/>.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Received: 29 September 2022 Accepted: 31 December 2022

Published online: 25 January 2023

References

- Bendicho VML-M, Gutiérrez MF. Holistic approaches to the comprehensive management of rock art in the digital age. *Quantitative Methods in the Humanities and Social Sciences*. In: Vincent ML, López-Mencheró Bendicho VM, Ioannides M, Levy TE, editors. *Heritage and Archaeology in the Digital Age*. Cham: Springer; 2017. p. 27–47.
- Seidl M, Breiteneder C. Automated petroglyph image segmentation with interactive classifier fusion. In: *Proceedings of the Eighth Indian Conference on Computer Vision, Graphics and Image Processing*, 2012; pp. 1–8. Association for Computing Machinery, New York, NY, USA. <https://doi.org/10.1145/2425333.2425399>
- Zeppelzauer M, Poier G, Seidl M, Reinbacher C, Schuster S, Breiteneder C, Bischof H. Interactive 3D segmentation of rock-art by enhanced depth maps and gradient preserving regularization. *JOCCH*. 2016;9(4):1–30.
- Poier G, Seidl M, Zeppelzauer M, Reinbacher C, Schaich M, Bellandi G, Marretta A, Bischof H. The 3d-pitoti dataset: a dataset for high-resolution 3D surface segmentation. In: *Proceedings of the 15th International Workshop on Content-Based Multimedia Indexing*, 2017; pp. 1–7.
- Fiorucci M, Khoroshiltseva M, Pontil M, Travaglia A, Del Bue A, James S. Machine learning for cultural heritage: a survey. *Pattern Recognit Lett*. 2020;133:102–8.
- Zhu Q, Wang X, Keogh E, Lee S-H. Augmenting the generalized hough transform to enable the mining of petroglyphs. In: *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. Association for Computing Machinery, New York, NY, USA. 2009; pp. 1057–1066. <https://doi.org/10.1145/1557019.1557133>
- Seidl M, Wieser E, Alexander C. Automated classification of petroglyphs. *DAACH*. 2015;2(2–3):196–212.
- Seidl M, Wieser E, Zeppelzauer M, Pinz A, Breiteneder C. Graph-based shape similarity of petroglyphs. In: Agapito L, Bronstein MM, Rother C, editors. *Computer Vision - ECCV 2014 Workshops*. Cham: Springer; 2015. p. 133–48.
- Qiang Z, Wang X, Keogh E, Lee SH. An efficient and effective similarity measure to enable data mining of petroglyphs. *Data Min Knowl Discov*. 2011;23(1):91–127.
- Zeppelzauer M, Poier G, Seidl M, Reinbacher C, Breiteneder C, Bischof H, Schuster S. Interactive segmentation of rock-art in high-resolution 3D reconstructions. In: *2015 Digital Heritage*, vol. 2, 2015; pp. 37–44. <https://doi.org/10.1109/DigitalHeritage.2015.7419450>
- Seidl M, Zeppelzauer M. Towards distinction of rock art pecking styles with a hybrid 2D/3D approach. In: *2019 International Conference on Content-Based Multimedia Indexing (CBMI)*, 2019; pp. 1–4. <https://doi.org/10.1109/CBBI.2019.8877469>
- Horn C, Ivarsson O, Lindhé C, Potter R, Green A, Ling J. Artificial intelligence, 3D documentation, and rock art-approaching and reflecting on the automation of identification and classification of rock art images. *J Archaeol Method Theory*. 2022;29(1):188–213.
- Jalandoni A, Shuker J. Automated tracing of petroglyphs using spatial algorithms. *DAACH*. 2021;22:00191. <https://doi.org/10.1016/j.daach.2021.e00191>.
- Deufemia V, Paolino L. Segmentation and recognition of petroglyphs using generic fourier descriptors. *Lect Notes Comput Sci*. 2014;8509:487–94.
- Poier G, Seidl M, Zeppelzauer M, Reinbacher C, Bischof H. PetroSurf3D - a high-resolution 3D dataset of rock art for surface segmentation. 2016. <https://arxiv.org/pdf/1610.01944>
- Jadon S. A survey of loss functions for semantic segmentation. In: *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*, 2020; pp. 1–7. <https://doi.org/10.1109/CIBCB48159.2020.9277638>
- Yi-de M, Qing L, Zhi-Bai Q. Automated image segmentation using improved pcnn model based on cross-entropy. In: *Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing*, IEEE, 2004, pp. 743–746.
- Pihur V, Datta S, Datta S. Weighted rank aggregation of cluster validation measures: a monte carlo cross-entropy approach. *Bioinformatics*. 2007;23(13):1607.
- Xie S, Tu Z. Holistically-nested edge detection. In: *2015 IEEE International Conference on Computer Vision (ICCV)*, 2016.
- Lin TY, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* PP(99), 2017; pp. 2999–3007
- Milletari F, Navab N, Ahmadi SA. V-net: Fully convolutional neural networks for volumetric medical image segmentation. In: *2016 Fourth International Conference on 3D Vision (3DV)*, 2016.
- Yang D, Roth H, Wang X, Xu Z, Xu D. Enhancing foreground boundaries for medical image segmentation. 2020. 10.48550/arXiv.2005.14355
- Minaee S, Boykov Y, Porikli F, Plaza A, Kehtarnavaz N, Terzopoulos D. Image segmentation using deep learning: a survey. *IEEE Trans Pattern Anal Mach Intell*. 2022;44(7):3523–42. <https://doi.org/10.1109/TPAMI.2021.3059968>.
- Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015; pp. 3431–3440.
- Badrinarayanan V, Kendall A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans Pattern Anal Mach Intell*. 2017;39(12):2481–95.
- Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-assisted Intervention*. Springer, 2015; pp. 234–241
- Siddique N, Paheding S, Elkin CP, Devabhaktuni V. U-net and its variants for medical image segmentation: a review of theory and applications. *IEEE Access*. 2021;9:82031–57.
- Basu M. Gaussian-based edge-detection methods-a survey. *IEEE Trans Syst Man Cybern C*. 2002;32(3):252–60.
- Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M, et al. {TensorFlow}: a system for {Large-Scale} machine learning. In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016; pp. 265–283

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)