This CVPR paper is the Open Access version, provided by the Computer Vision Foundation.
Except for this watermark, it is identical to the accepted version;
the final published version of the proceedings is available on IEEE Xplore.

Mudslide: A Universal Nuclear Instance Segmentation Method

Jun Wang
School of Electronic and Computer Engineering, Peking University, Shenzhen, China
junw@stu.pku.edu.cn

Abstract

Nuclear instance segmentation has played a critical role in pathology image analysis. The main challenges arise from the difficulty in accurately segmenting densely overlapping instances and the high cost of precise mask-level annotations. Existing fully-supervised nuclear instance segmentation methods, such as boundary-based methods, struggle to capture differences between overlapping instances and thus fail in densely distributed blurry regions. They also face challenges transitioning to point supervision, where annotations are simple and effective. Inspired by natural mudslides, we propose a universal method called Mudslide that uses simple representations to characterize differences between different instances and can easily be extended from fully-supervised to point-supervised. Concretely, we introduce a collapse field and leverage it to construct a force map and initial boundary, enabling a distinctive representation for each instance. Each pixel is assigned a collapse force, with distinct directions between adjacent instances. Starting from the initial boundary, Mudslide executes a pixel-by-pixel collapse along various force directions. Pixels that collapse into the same region are considered as one instance, concurrently accounting for both inter-instance distinctions and intra-instance coherence. Experiments on public datasets show superior performance in both fully-supervised and point-supervised tasks.

1. Introduction

Pathological slide analysis is widely regarded as the gold standard for cancer diagnosis, treatment, and prevention. Nuclear instance segmentation is a critical step in this process, as it enables the accurate identification and analysis of nuclei within pathology images. The morphology of nuclei, including their shape, appearance, and density, provides interpretable features that can be used as diagnostic and prognostic indicators of cancer [18].

Nuclear instance segmentation presents unique challenges compared to natural image instance segmentation tasks. Firstly, the densely packed distribution of nuclei often results in clustered overlapping instances. Secondly, the blurry boundaries between touching or overlapping nuclei make it difficult to distinguish individual instances. Third, obtaining precise pixel-level annotations is time-consuming and expensive for professional pathologists.

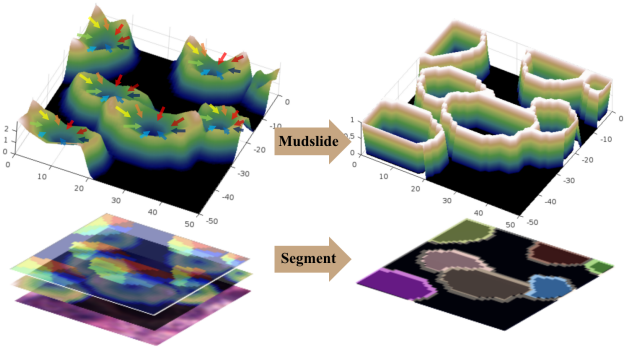

Figure 1. **An analogy between mudslide and nuclear instance segmentation.** (Top) Topographic map before and after the mudslide. The colored arrows indicate the collapse forces in different directions. (Bottom) Nuclear instance segmentation process in pathological images. The different layers on the left represent different feature representations of nuclear instances, and the right shows the result of instance segmentation.

To address the first two challenges, many fully-supervised methods based on deep learning have made significant progress. Several algorithms based on constructing innovative network architectures [8, 20, 23] have been applied to this task. Some contour-based techniques have been investigated to enhance the accurate prediction of nucleus boundaries [3, 9, 29] which aids in splitting the touched and overlapped nuclei. However, these methods lack the capability to characterize the differences among neighboring nuclear instances, making it difficult to accurately segment densely stacked overlapping nuclei. Moreover, their applicability is hindered in scenarios where fine pixel-level annotations are insufficient, as obtaining precise boundary labels becomes a formidable task. Addressing the challenge posed by the scarcity of fine annotations, certain point annotation-based methods have emerged [2, 19, 21], leveraging the accessibility of coarse point labels. How-

ever, these approaches frequently prioritize the acquisition of pseudo-mask labels through design optimization methods, neglecting the differentiated representation of overlapping instances.

Inspired by the mudslide in nature, we propose a simple universal method, called Mudslide. This approach excels in providing a differential representation of instances while remaining readily adaptable to the point-supervised domain.

As shown in Fig. 1, for a continuous geographical region, mudslides induced by external forces lead to the creation of distinct partitions within the area. The unaffected section sets the partition boundary, while the collapsing segment forms a contiguous region. Analogously, this parallels the process of nucleus instance segmentation, where a method partitions the foreground with diverse instances into distinct instance regions. Furthermore, the mudslide collapse process concurrently considers the distinctions between regions and the coherence within each region. Variations in force directions acting on different regions induce noticeable changes in adjacent areas. Internal forces within a region, specifically the collapse of one soil mass exerting force on neighboring soil masses, facilitate the clustering and rationalization of soil masses in the same area. This aligns with the objectives of nuclear instance segmentation.

Therefore, drawing inspiration from the characteristics of the muslide collapse process, we introduce Mudslide algorithm to induce the collapse of the foreground region of the nucleus and acquire instances. Specifically, we formulate a collapse field and leverage it to construct a force map to allocate forces directed to the center of the instance to individual pixels within the collapse field. This allocation ensures that the force direction in adjacent regions of distinct nuclei exhibits pronounced disparities, facilitating the distinctive representation of neighboring nuclei. Simultaneously, the forces within the same instance exhibit uniformity directed towards the center, enhancing a cohesive representation within instances. Subsequently, we devise a collapse algorithm commencing from the collapse boundary, proceeding pixel-by-pixel along the force direction. Ultimately, pixels collapsing into the same region are deemed components of the same instance. Due to the consideration of external differences between instances and the internal consistency within instances, Mudslide excels in accurately segmenting overlapping instances. The simplicity of collapse field design renders this approach easily extendable to the domain of point supervision.

The contributions can be summarized as follows:

- We propose Mudslide, a novel nucleus instance segmentation algorithm inspired by natural phenomenon of mudslide, which excels in achieving precise segmentation of adjacent or overlapping nuclei.
- We propose a collapse field and generate a force map along with an initial boundary to provide differentiated representations for neighboring instances and enhance the internal coherence within instances.
- Mudslide is a universal method that can be easily extended from full supervision tasks to point supervision domains.
- We justify the superiority of our Mudslide on public nuclear datasets with state-of-the-art performance.

2. Related Works

2.1. Fully-supervised nuclei segmentation

Fully-supervised nuclear instance segmentation has been widely studied and benefited from the application of deep learning techniques [3, 7, 22]. One of the most influential models in this field is the U-Net model proposed by Ronneberger et al. [23], which has become a benchmark model for medical image segmentation and has inspired many subsequent works. Qu et al. [20] proposed a full-resolution convolutional neural network (FullNet), a full-resolution convolutional neural network that avoids downsampling operations in the network structure to improve localization accuracy. Raza et al. [22] proposed Micro-Net, which utilizes multi-resolution and weighted loss functions to achieve robustness against large internal/external variances in nucleus size. He et al. [8] proposed Han-Net, a hybrid attention nested U-shaped network that extracts informative features from multiple layers to enhance segmentation performance.

Leveraging contour information for distinguishing interacting or overlapping nuclei, Chen et al. [3] introduced the integration of contour details into a multi-level Fully Convolutional Network (FCN), which resulted in the development of a profound contour-aware network designed specifically for nuclear instance segmentation. Following this, Zhou et al. [29] introduced the Contour-Aware Information Aggregation Network, which integrates spatial and textural features between nuclei and contours. Additionally, He et al. [9] proposed a Topology-Aware Loss to integrate meaningful topology constraints into the model optimization process for more precise boundary predictions.

Furthermore, certain studies [4, 6, 14, 16] have employed distance maps to address the segmentation challenges posed by contacting or overlapping nuclei. Naylor et al. [16] tackled the segmentation of touching nuclei by framing it as a regression task involving intra-nuclear distance maps. Graham et al. [6] ntroduced the Hover-Net, which utilizes vertical and horizontal distances between a nucleus pixel and its center of mass to delineate nuclei. Additionally, He et al. [7] proposed the Centripetal Directional Network (CDNet) for nuclear instance segmentation, integrating directional information into the network.

The above models have achieved performance improvements in nuclear segmentation. However, the challenge of segmenting overlapping cell nuclei persists. To further en-

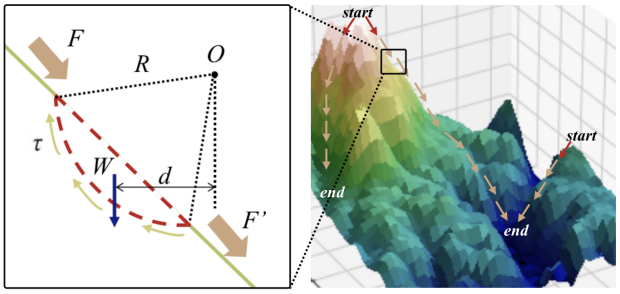

Figure 2. **The simple modeling of natural mudslide.** (Left) Schematic diagram of the force on a soil mass on the slope. The green line represents the slope and the red dashed line represents the soil mass. (Right) A mountain where the mudslide occurred. The mudslide starts at the position pointed by the force indicated by the red arrow and stops at the horizontal slope.

hance the model's ability to distinguish instances, there is a need for a distinct and differential representation of nucleus instances.

2.2. Pointly-supervised nuclei segmentation

The pixel-wise annotation of nucleus instances incurs high costs. Point annotation, as one of the most efficient labeling manners, has garnered increasing attention in the field of nuclear segmentation. Current approaches [2, 19, 21] primarily transform point annotations into coarse pixel-level labels, including Voronoi labels [25], cluster labels [19], and pseudo edge maps [27]. These methods commonly integrate prior knowledge about the shape and texture of nuclei into these coarse pixel-level labels. However, these techniques face challenges related to incomplete coarse labels and imprecise boundary information. Several strategies have been proposed to address the issues arising from these coarse labels. These include multi-stage optimization for refining segmentation in a bootstrapping manner [21, 25], and the incorporation of additional constraints based on the local contrast between nuclei and their surrounding cytoplasm. Additionally, some unique training strategies have been employed to optimize pseudo-labels. [12]utilizes the co-training method, optimizing mask labels through the simultaneous training of two models with mutual supervision. [28] employs a collaborative knowledge-sharing approach, sharing knowledge between a principal model and a very lightweight collaborator model.

Despite the progress in the field of making full use of point labels and generating accurate pseudo mask labels, however, the problem of nucleus overlapping instance segmentation in point-supervised prediction still needs further attention.

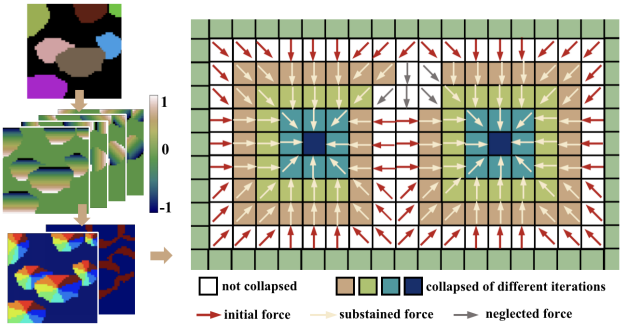

Figure 3. **The processing of Mudslide algorithm.** (Left) The instance map, the four-channel collapse field, the eight-directional force map, and the initial boundary. Each color in the force map represents a direction. (Right) Schematic diagram of the mudslide collapse process performed on two adjacent nuclei.

3. Method

3.1. Natural Mudslide

Mudslide refers to the natural phenomenon where the soil or rock mass on a slope, influenced by external forces such as river erosion, rainfall, and seismic events, undergoes downhill movement collectively along the slope [5]. To simulate the mudslide process, we simplify it as a continuous sliding problem of soil mass on a slope [24]. This assumes an arc-shaped sliding surface adhering to the Coulomb criterion during soil movement, and the soil undergoes rigid sliding. This allows the assessment of the stability of the soil slope based on the magnitudes of the sliding moment and anti-sliding moment.

As illustrated in Fig. 2, for the soil mass on the slope, the sliding moment is generated by the gravitational force W, while the resisting moment is provided by the frictional force on the sliding surface. The sliding moment M_s and resisting moment M_r are

$$M_s = W * d, \tag{1}$$

$$M_r = R * \int_0^l \tau dl, \tag{2}$$

where d represents the lever arm of gravity, R and l are respectively the radius and the length of the arc, and τ represents the shear force. Introducing the safety factor F_{os} [17]

$$F_{os} = \frac{M_r}{M_s}, \tag{3}$$

if $F_{os} > 1$, the soil mass remains stable and will not collapse. If $F_{os} = 1$, the soil mass is in a state of critical equilibrium, and if $F_{os} < 1$, the soil mass undergoes collapse.

Assuming each soil mass on the slope is in a critical state, i.e., $F_{os} = 1$, if an external force F with moment M_F is applied tangentially along the slope surface such that

$$F'_{os} = \frac{M_r}{M'_s} = \frac{M_r}{M_s + M_F} < F_{os}, \tag{4}$$

the corresponding soil mass undergoes collapse. This results in the application of a tangential force F' on the adjacent soil mass below, leading to a continued collapse of the lower soil mass, ultimately forming a mudslide. The collapse terminates when it reaches a horizontal slope surface. In the end, the mudslide induces the collapse of soil mass on the slope surface affected by external forces, while the unaffected soil mass remains intact. This delineates distinct regions on the slope.

3.2. Mudslide Segmentation

3.2.1 Mudslide Processing

Inspired by the natural mudslide mentioned in Sec. 3.1, we propose a Mudslide algorithm for nuclear instance segmentation. In natural landslide phenomena, the main issues include the characteristics of the slope that collapses, the forces acting on the soil mass on the slope, the location where the collapse begins, and how the collapse occurs continuously. Corresponding to this, the Mudslide algorithm primarily involves three steps: constructing a collapse field, obtaining the force map and initial collapse boundary, and executing the collapse.

Collapse field construction. In the first step, we initiate the process by constructing a collapse field, representing the initial slope surface of the mudslide. In image segmentation, we assume that each pixel represents a soil mass, located on a slope represented by the collapse field, and in a state of critical equilibrium. For simplicity, we utilize continuous normalized maps to characterize the collapse field, where pixel values close to 1 signify higher slope heights and values close to 0 indicate lower slope surfaces. Specifically, for ease of model learning and inspired by [6], we create a four-channel normalized map shown in Fig. 3. This map contains four distance maps generated in horizontal, vertical, and two diagonal directions, respectively, with each map indicating the distance from each pixel to the center of the nucleus in different directions. Pixel values are close to 0 when they are near the nucleus center and gradually increase outward until they reach 1 at the nucleus boundary. Thus, we construct a slope for the foreground nuclei in the image, where the slope is flat in the middle of each instance and steep around the edges, resembling the terrain that is prone to mudslide collapse.

Force map and initial boundary obtaining. For the second step, we generate a force map and initial boundary from the collapse field. The force map assigns a force F with different directions to each pixel, representing the directions in which pixels are expected to collapse during Mudslide. For two-dimensional images, we consider eight directions of force: up, down, left, right, upper-left, upper-right, lower-left, and lower-right (Fig. 3):

$$\begin{aligned}\mathbb{F} = \{&\begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\-1\end{pmatrix}, \begin{pmatrix}-1\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix},\\ &\begin{pmatrix}-1\\1\end{pmatrix}, \begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}-1\\-1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\}.\end{aligned} \tag{5}$$

Then the force map can be expressed as a set of forces:

$$FM = \{F(x, y) \mid x, y \in \mathbb{Z}, F \in \mathbb{F}\}, \tag{6}$$

where $\mathbb{Z}$ indicates the foreground of nuclei and x, y corresponds to the coordinate of each pixel. Here, we assign a force direction to each pixel that is approximately tangential to the slope. Since we assume that each pixel is in a state of critical equilibrium, we consider the force magnitude to be irrelevant to the collapse and therefore set the magnitude of each force to 1.

To obtain the starting position of mudslide collapse, we generate the initial boundary. We employ a straightforward edge detection filter, such as Sobel operator, on the collapse field to detect edges. The obtained edges are then utilized as the initial boundary for mudslide collapse. This approach is rational since, in our collapse field, there is a significant gradient between instance boundaries, allowing the Sobel operator to effectively identify them.

Mudslide collapse executing. After acquiring the force map and the initial boundary, we commence the mudslide collapse process. Originating from the initial boundary, each pixel p applies an initial force (red arrows in 3) to the directed pixel p', which can be calculated by

$$\begin{aligned}p^{'} &= p + F(p),\\ p &= (x, y),\\ p^{'} &= (x^{'}, y^{'}).\end{aligned} \tag{7}$$

The pixel p' then undergoes collapse. Subsequently, the collapsed pixel exerts a sustained force (yellow arrows in 3) on the pointed pixel, initiating the collapse of more adjacent pixels. This iterative process results in a sustained mudslide collapse. Eventually, when all pixels directed to have undergone collapse, those pixels collapsing to the same region are identified as a single instance. Note that those pixels that are never directed by forces from other pixels will not collapse and the force assigned to them will be neglected (gray arrows in 3), corresponding to the soil mass that is not affected by the mudslide.

3.2.2 Application to Nuclear Instance Segmentation

The framework of applying Mudslide to nuclear instance segmentation is illustrated in Fig. 4. Here, we unify the pipelines for both fully supervised and point-supervised scenarios, with the primary distinction lying in the labels used for training.

Specifically, we employ a dual-branch architecture, where one branch is dedicated to predicting a binary map

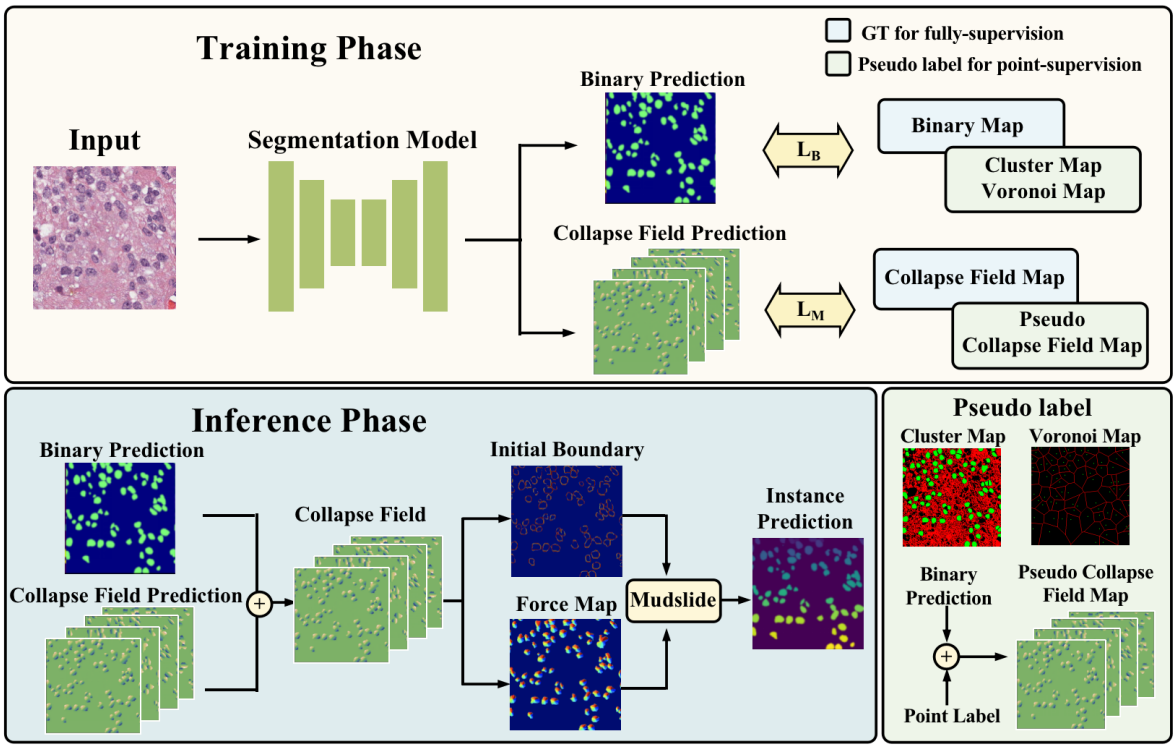

Figure 4. **The overall framework of the proposed Mudslide.** The network contains two basic prediction branches: the binary branch and the collapse field branch. During training, the difference between fully- and point-supervised is the type of ground truth. During testing, both branches' predictions are employed for the Mudslide algorithm.

for the nuclear foreground, and the other branch focuses on predicting the collapse field. The collapse field consists of four channels, corresponding to horizontal, vertical, and two diagonal directions. The segmentation model adopted is the Trans-UNet architecture from [10]. In fully supervised training, the binary map labels are the ground truth representing the nuclear foreground, while the collapse field labels are constructed from both the foreground and instance ground truth. For point-supervised training, we leverage point labels and the original input image to generate a cluster map [19] and Voronoi map [25] as pseudo-labels for foreground prediction, which is commonly observed in point-supervised tasks. Subsequently, we use the predicted foreground and point labels to generate collapse field labels.

We utilize cross-entropy loss as the binary classification loss and employ mean squared error loss and mean squared gradient error loss as the collapse field loss:

$$\mathcal{L}_B = \lambda_{B_{CE}} \mathcal{L}_{CE}, \tag{8}$$

$$\mathcal{L}_{CF} = \lambda_{CF_{MSE}} \mathcal{L}_{MSE} + \lambda_{CF_{MSGE}} \mathcal{L}_{MSGE}. \tag{9}$$

Then the overall loss function is

$$\mathcal{L} = \mathcal{L}_B + \mathcal{L}_{CF}. \tag{10}$$

It's worth noting that in point-supervised training, the cross-entropy loss comprises two terms, corresponding to cluster and Voronoi:

$$\mathcal{L}_{B_{point}} = \lambda_{B_{CE}} \mathcal{L}_{CE_{cluster}} + \lambda_{B_{CE}} \mathcal{L}_{CE_{voronoi}}. \tag{11}$$

During the inference process, we first combine the foreground and collapse field predictions to construct the collapse field, ensuring subsequent mudslide acts on the nuclear foreground regions. Subsequently, as detailed in Sec. 3.1, we utilize the collapse field to obtain the initial boundary and force map, initiating the mudslide collapse process. Ultimately, this process yields the nuclear instance segmentation results.

4. Experiment

4.1. Experimental Settings

Datasets. I) CPM17 [26] contains 64 H&E stained histopathology images with 7,570 annotated nuclear boundaries. It is from the MICCAI 2017 Digital Pathology Challenge [26] and images on two different scales: 500×500 and 600×600. This is a dataset for a single-class segmentation task. II) NuInsSeg [15] contains 665 image patches with more than 30,000 manually segmented nuclei from 31 human and mouse organs. It is one of the biggest fully manually annotated datasets of nuclei in Hematoxylin and

Eosin (H &E)stained histological images with the scale of 512×512. This is also a dataset for single-class segmentation task.

Evaluation Metrics. We use four instance-level evaluation metrics to measure the instance segmentation performance of the comparison models, which are: Detection Quality (DQ), Segmentation Quality (SQ), Panoptic Quality (PQ), and the Aggregated Jaccard Index (AJI).

Implementation details. The proposed Mudslide is trained and tested using the open-source software library Pytorch 1.8.0 on NVIDIA Tesla V100 with CUDA 10.1.

In order to build the dataset used for training, the CPM17 dataset is divided into multiple patches as it has an inconsistent original image size. We generate 512 training patches in the size of 256x256, thus allowing the network to be uniform and fully trained. In addition, we randomly divide NuInsSeg into training set and test set with the proportion of 80% and 20%, respectively. Moreover, for point-supervised tasks, we leverage the annotations of nucleus instances to generate a central point for each instance, serving as the point label. In the training phase, we apply a series of data augmentation methods, which include random flip, random rotation, random color, random crop, and Gaussian blur. The training epoch is set as 300, and the initial learning rate is set as 10^{-3}. The weights of each loss are $\lambda_{B_{CE}} = 1, \lambda_{CF_{MSE}} = 2.5, \lambda_{CF_{MSGE}} = 8$.

4.2. Ablation Study

We performed ablation studies on the CPM17 and NuInsSeg datasets to validate the efficacy of the proposed Mudslide method. All the following experiments were conducted under the setting of fully supervised tasks.

Instance separation methods. Table 1 shows the comparison of segmentation performance obtained using different instance separation methods. We apply four different methods to the predictions of the same model: the traditional watershed method [1], boundary partition method, HVmap segmentation method from Hover-Net [6], and our proposed method.

Specifically, the watershed method involves segmenting instances using the watershed algorithm on the binary probability map predicted by the model. The boundary separation method refers to predicting boundaries by adding an additional boundary prediction branch to the model and using the boundaries to directly separate different instances. The HVmap algorithm is an instance separation method in HoverNet that utilizes horizontal and vertical distance maps for segmentation.

The experimental results indicate significant variations in model performance based on different instance separation methods. Traditional methods such as watershed and boundary separation exhibit subpar performance. The HVmap approach, utilizing horizontal and vertical maps, enhances the differentiation between instances and achieves better results. Furthermore, our method, with more explicit instance differentiation representation and a more rational instance acquisition process, outperforms other methods. This underscores the superiority of our Mudslide approach.

Method	CPM				NuInsSeg			
	DQ ↑	SQ ↑	PQ ↑	AJI ↑	DQ ↑	SQ ↑	PQ ↑	AJI ↑
Watershed	0.797	0.730	0.582	0.694	0.601	0.735	0.461	0.502
Boundary	0.837	0.734	0.614	0.714	0.640	0.753	0.491	0.532
HVMap	0.879	**0.810**	0.713	0.737	0.786	0.754	0.596	0.638
Ours	**0.883**	0.808	**0.714**	**0.740**	**0.788**	**0.756**	**0.599**	**0.647**

Table 1. **Comparison of different adjacent instance separation methods**, including the watershed algorithm, boundary partition, HVmap and our method.

bound	collapse	CPM17				NuInsSeg			
		DQ ↑	SQ ↑	PQ ↑	AJI ↑	DQ ↑	SQ ↑	PQ ↑	AJI ↑
		0.772	0.710	0.548	0.610	0.587	0.703	0.414	0.497
✓		0.877	0.807	0.707	0.733	0.775	0.752	0.582	0.630
✓	✓	**0.883**	**0.808**	**0.714**	**0.740**	**0.788**	**0.756**	**0.599**	**0.647**

Table 2. **Ablation studies of components in Mudslide algorithm**, where bound indicates initial boundary, and collapse indicates mudslide collapse processing.

The ablation of components of Mudslide algorithm. We conducted ablation experiments on the key steps of the mudslide algorithm. Here, we considered two components: whether to use the initial boundary and whether to perform mudslide collapse. In Table 2, we explored three comparative experiments: a baseline approach with no additional components, relying solely on binary map partitioning based on connected regions; a scenario involving only the initial boundary extracted from the collapse field; and a comprehensive application of both the initial boundary and mudslide collapse, representing our full-fledged mudslide algorithm for instance segmentation. These experiments aim to scrutinize the individual contributions of each component to the overall segmentation performance.

The experimental results indicate that incorporating an initial boundary leads to performance gains of 16% PQ and 12% AJI for the CPM dataset and 17% PQ and 14% AJI for NuInsSeg. Furthermore, the additional mudslide collapse results in a performance boost of 7% PQ and 7% AJI for the CPM dataset and 16% PQ and 17% AJI for NuInsSeg. These findings underscore the substantial benefits of each component within the mudslide method for enhancing instance segmentation outcomes.

Force map construction methods. As a crucial manifestation of distinct instance differentiation, the construction of the force map is pivotal in the Mudslide algorithm. Therefore, we consider three methods for building the force map: direct prediction, collapse field segmentation, and collapse field extraction.

Method	CPM				NuInsSeg			
	DQ ↑	SQ ↑	PQ ↑	AJI ↑	DQ ↑	SQ ↑	PQ ↑	AJI ↑
predict	0.874	0.807	0.707	0.735	0.777	0.751	0.587	0.633
segment	0.864	**0.808**	0.701	0.728	0.772	0.753	0.582	0.629
extract	**0.880**	0.807	**0.711**	**0.739**	**0.784**	**0.755**	**0.592**	**0.643**

Table 3. **Comparison of different force map construction methods.**

In the first approach, we have the model predict the force map directly. Specifically, we augment the model with a branch that outputs nine channels, allowing it to learn the force map's eight directions and the background region. Subsequently, we obtain the force map by employing the argmax function to the predicted probability map. In the second approach, we perform segmentation on the collapse field. Here, we use a two-channel collapse field corresponding to the horizontal and vertical directions. By segmenting the collapse field based on angles, we categorize it into eight regions, representing the eight directions of the force map. In the third approach, we predict a four-channel collapse field for horizontal, vertical, and two diagonal directions. Then, we directly employ argmax to extract the direction of force that each pixel should possess.

The results presented in Table 3 indicate that the force map extracted through four-channel collapse fields performs the best. The direct prediction follows with slightly lower performance, while the force map obtained by splitting the two-channel collapse field performs the least favorably.

The four-channel collapsing field extraction method outperformed direct prediction by 4% PQ and 4% AJI on CPM17, and outperformed the two-channel segmentation method by 6% PQ, and 11% AJI. On the NuInsSeg dataset, it outperformed direct prediction by 5% PQ, and 1% AJI, and outperformed the two-channel segmentation method by 10% PQ, and 14% AJI. These results suggest that the four-channel collapsing field can better capture and characterize the differences between cell nucleus instances, resulting in greater differences between adjacent instances and better instance segmentation results.

Initial boundary generation methods. Regarding the issue of generating and selecting initial boundaries, we considered two methods: direct generation from the binary prediction map and generation from the collapsing field.

Specifically, for the first method, we first obtained the binary probability map predicted by the model's binary classification branch and then selected a threshold of 0.5 to segment the predicted foreground and background regions of the nucleus. Then, we used the part of the pixels that contacted the background, i.e., the edge of the foreground, as the initial boundary. For the second method, we first obtained the four-channel collapsing field predicted by the model and then used some edge extraction methods, such as using the Sobel operator to extract the regions with higher gradients in the collapsing field as the initial boundary.

Method	CPM				NuInsSeg			
	DQ ↑	SQ ↑	PQ ↑	AJI ↑	DQ ↑	SQ ↑	PQ ↑	AJI ↑
binary map	0.880	0.807	0.711	0.739	0.784	0.755	0.592	0.643
collapse field	**0.883**	**0.808**	**0.714**	**0.740**	**0.788**	**0.756**	**0.599**	**0.647**

Table 4. **Comparison of different initial boundary generation methods.**

As shown in Table 4, the method of generating initial boundaries using the collapsing field outperforms the method of using only the binary prediction map. Specifically, on the CPM dataset, the collapsing field method improved performance by 3% PQ, and 1% AJI compared to the binary map method. On the NuInsSeg dataset, it improved performance by 7% PQ, and 4% AJI. As anticipated, the construction of the initial boundary is crucial in our Mudslide collapse algorithm, as it determines the algorithm's starting position. Obtaining foreground edges directly from a binary image makes it challenging to capture the boundaries between adjacent instances. Although the Mudslide algorithm can capture this difference to some extent, its performance is not as effective as constructing a better initial boundary. The initial boundary extracted through the collapse field takes into account the differentiated representations between cell nuclei, effectively separating nuclei to a certain degree and proving to be a superior boundary.

4.3. Fully-Supervised Nuclei Segmentation

In Table 5, the proposed Mudslide is compared with the state-of-the-art fully-supervised nuclear instance segmentation models on CPM17 and NuInsSeg. Our Mudslide achieves 0.714 PQ and 0.740 AJI on the CPM17 dataset, and 0.599 PQ and 0.647 AJI on the NuInsSeg dataset, respectively. The Mudslide achieves competitive results for fully-supervised nuclei segmentation tasks.

We further carry out a qualitative visualization analysis of our method and different existing methods. Compared to all the other methods in Fig. 5, our Mudslide achieves superior instance segmentation results for nuclear instance segmentation, especially in some regions where nuclei are densely distributed with a large amount of overlap, where our instance segmentation is more accurate. The results, both quantitative and qualitative, unequivocally establish the competitiveness of our Mudslide algorithm in accurately segmenting densely packed and overlapping nuclear instances.

4.4. Pointly-Supervised Nuclei Segmentation

In Table 6, we extend the proposed Mudslide from fully-supervised to point-supervised nuclear segmentation task and compare it with the state-of-the-art point-supervised nuclear instance segmentation models on CPM17 and

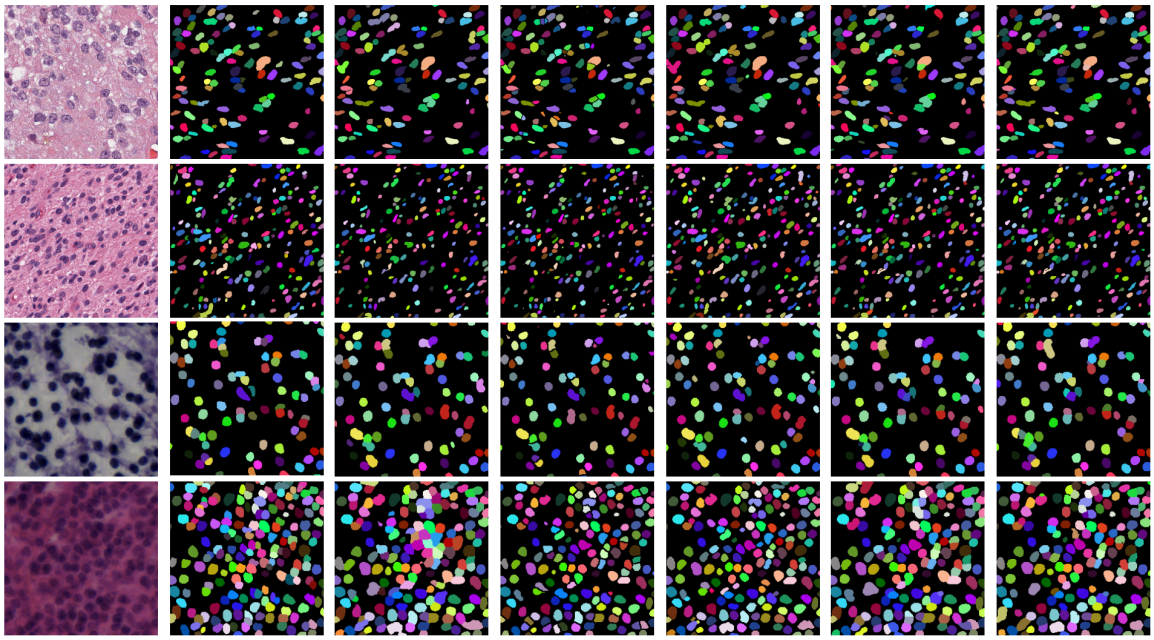

Figure 5. **Visualization of instance segmentation results with different methods.** (a) Input image; (b) Ground Truth; (c) U-Net [23]; (d) DIST [16]; (e) FullNet [20]; (f) Hover-Net [6]; (g) Our Mudslide.

Method	CPM				NuInsSeg			
	DQ ↑	SQ ↑	PQ ↑	AJI ↑	DQ ↑	SQ ↑	PQ ↑	AJI ↑
UNet [23]	0.778	0.734	0.578	0.643	0.635	0.688	0.436	0.508
DCAN [3]	0.732	0.740	0.545	0.561	0.563	0.644	0.363	0.435
DIST [16]	0.663	0.754	0.504	0.616	0.703	0.728	0.512	0.542
FullNet [20]	-	-	0.661	0.668	0.712	0.760	0.541	0.562
HoverNet [6]	0.854	**0.814**	0.697	0.705	0.743	0.764	0.568	0.611
CDNet [7]	-	-	-	0.733	-	-	-	0.640
CellViT [10]	0.879	0.810	0.713	0.737	0.786	0.754	0.596	0.638
Ours	**0.883**	0.808	**0.714**	**0.740**	**0.788**	**0.756**	**0.599**	**0.647**

Table 5. **Performance comparisons on fully-supervised nuclear segmentation.**

Method	CPM				NuInsSeg			
	DQ ↑	SQ ↑	PQ ↑	AJI ↑	DQ ↑	SQ ↑	PQ ↑	AJI ↑
PENet [27]	0.440	0.722	0.319	0.290	0.476	0.630	0.300	0.355
C2FNet [25]	0.429	0.630	0.271	0.421	0.405	0.580	0.235	0.400
S2Label [11]	0.654	0.698	0.459	0.503	0.613	0.667	0.409	0.496
SPN+IEN [13]	0.704	0.689	0.488	0.543	0.688	0.660	0.454	0.523
SCNet [12]	0.716	0.685	0.496	0.554	0.694	0.655	0.456	0.526
Ours	**0.747**	**0.703**	**0.529**	**0.581**	**0.707**	**0.689**	**0.487**	**0.558**

Table 6. **Performance comparisons on point supervised nuclear segmentation.**

NuInsSeg. Our Mudslide achieves 0.529 PQ and 0.581 AJI on the CPM17 dataset, and 0.456 PQ and 0.558 AJI on the NuInsSeg dataset, respectively. In the task of nuclear instance segmentation with only point annotations, our method can also achieve competitive results.

Although methods relying solely on point annotations still have ample room for improvement compared to fully supervised approaches, our approach has already surpassed other methods in the realm of point annotations.

5. Conclusion

This paper proposes a novel universal approach for nuclear instance segmentation, called Mudslide, which is inspired by the natural phenomenon of mudslides. To enhance the model's capability to differentiate between adjacent instances, we devise a collapsing field, and generate a force map along with an initial boundary, offering distinct representations for different instances and a unified representation for the same instance. We also propose Mudslide collapse algorithm that mimics the natural process by segmenting through the iterative collapse of individual pixels. Our collapse fields are simple and easily extendable to point-supervised tasks. The experimental results on public nuclei datasets demonstrate that Mudslide outperforms state-of-the-art methods.

Acknowledgments This work was supported in part by the National Key R&D Program of China (No. 2022ZD0118201), Natural Science Foundation of China (No.61972217, 32071459, 62176249, 62006133, 62271465), and Shenzhen Medical Research Funds in China (No. B2302037).

References

[1] Serge Beucher. Use of watersheds in contour detection. In *Proc. Int. Workshop on Image Processing, Sept. 1979*, pages 17–21, 1979. 6

[2] Alireza Chamanzar and Yao Nie. Weakly supervised multi-task learning for cell detection and segmentation. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 513–516. IEEE, 2020. 1, 3

[3] Hao Chen, Xiaojuan Qi, Lequan Yu, Qi Dou, Jing Qin, and Pheng-Ann Heng. Dcan: Deep contour-aware networks for object instance segmentation from histology images. *Medical Image Analysis*, 36:135–146, 2017. 1, 2, 8

[4] Shengcong Chen, Changxing Ding, Minfeng Liu, Jun Cheng, and Dacheng Tao. Cpp-net: Context-aware polygon proposal network for nucleus segmentation. *IEEE Transactions on Image Processing*, 32:980–994, 2023. 2

[5] FC Dai, Chin Fei Lee, and Y Yip Ngai. Landslide risk assessment and management: an overview. *Engineering Geology*, 64(1):65–87, 2002. 3

[6] Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis*, 58:1–15, 2019. 2, 4, 6, 8

[7] Hongliang He, Zhongyi Huang, Yao Ding, Guoli Song, Lin Wang, Qian Ren, Pengxu Wei, Zhiqiang Gao, and Jie Chen. Cdnet: Centripetal direction network for nuclear instance segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 4026–4035, 2021. 2, 8

[8] Hongliang He, Chi Zhang, Jie Chen, Ruizhe Geng, Luyang Chen, Yongsheng Liang, Yanchang Lu, Jihua Wu, and Yongjie Xu. A hybrid-attention nested unet for nuclear segmentation in histopathological images. *Frontiers in Molecular Biosciences*, 8:1–9, 2021. 1, 2

[9] Hongliang He, Jun Wang, Pengxu Wei, Fan Xu, Xiangyang Ji, Chang Liu, and Jie Chen. Toposeg: Topology-aware nuclear instance segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 21307–21316, 2023. 1, 2

[10] Fabian Hörst, Moritz Rempe, Lukas Heine, Constantin Seibold, Julius Keyl, Giulia Baldini, Selma Ugurel, Jens Siveke, Barbara Grünwald, Jan Egger, et al. Cellvit: Vision transformers for precise cell segmentation and classification. *arXiv preprint arXiv:2306.15350*, 2023. 5, 8

[11] Hyeonsoo Lee and Won-Ki Jeong. Scribble2label: Scribble-supervised cell segmentation via self-generating pseudo-labels with consistency. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23*, pages 14–23. Springer, 2020. 8

[12] Yi Lin, Zhiyong Qu, Hao Chen, Zhongke Gao, Yuexiang Li, Lili Xia, Kai Ma, Yefeng Zheng, and Kwang-Ting Cheng. Nuclei segmentation with point annotations from pathology images via self-supervised learning and co-training. *Medical Image Analysis*, 89:102933, 2023. 3, 8

[13] Weizhen Liu, Qian He, and Xuming He. Weakly supervised nuclei segmentation via instance learning. In *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)*, pages 1–5. IEEE, 2022. 8

[14] Xiaoming Liu, Zhengsheng Guo, Jun Cao, and Jinshan Tang. Mdc-net: a new convolutional neural network for nucleus segmentation in histopathology images with distance maps and contour information. *Computers in Biology and Medicine*, 135:1–12, 2021. 2

[15] Amirreza Mahbod, Christine Polak, Katharina Feldmann, Rumsha Khan, Katharina Gelles, Georg Dorffner, Ramona Woitek, Sepideh Hatamikia, and Isabella Ellinger. Nuinsseg: A fully annotated dataset for nuclei instance segmentation in h&e-stained histological images. *arXiv preprint arXiv:2308.01760*, 2023. 5

[16] Peter Naylor, Marick Laé, Fabien Reyal, and Thomas Walter. Segmentation of nuclei in histopathology images by deep regression of the distance map. *IEEE Transactions on Medical Imaging*, 38(2):448–459, 2018. 2, 8

[17] Hasmukh Pranshanker Oza and HP Oza. *Soil mechanics and foundation engineering*. Charotar Book Stall, 1969. 3

[18] H. Pinckaers, W. Bulten, J. van der Laak, and G. Litjens. Detection of prostate cancer in whole-slide images through end-to-end training with image-level labels. *IEEE Transactions on Medical Imaging*, 40(7):1817–1826, 2021. 1

[19] Hui Qu, Pengxiang Wu, Qiaoying Huang, Jingru Yi, Gregory M Riedlinger, Subhajyoti De, and Dimitris N Metaxas. Weakly supervised deep nuclei segmentation using points annotation in histopathology images. In *International Conference on Medical Imaging with Deep Learning*, pages 390–400. PMLR, 2019. 1, 3, 5

[20] Hui Qu, Zhennan Yan, Gregory M Riedlinger, Subhajyoti De, and Dimitris N Metaxas. Improving nuclei/gland instance segmentation in histopathology images by full resolution neural network and spatial constrained loss. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 378–386, 2019. 1, 2, 8

[21] Hui Qu, Pengxiang Wu, Qiaoying Huang, Jingru Yi, Zhennan Yan, Kang Li, Gregory M Riedlinger, Subhajyoti De, Shaoting Zhang, and Dimitris N Metaxas. Weakly supervised deep nuclei segmentation using partial points annotation in histopathology images. *IEEE Transactions on Medical Imaging*, 39(11):3655–3666, 2020. 1, 3

[22] Shan E Ahmed Raza, Linda Cheung, Muhammad Shaban, Simon Graham, David Epstein, Stella Pelengaris, Michael Khan, and Nasir M Rajpoot. Micro-net: A unified model for segmentation of various objects in microscopy images. *Medical Image Analysis*, 52:160–173, 2019. 2

[23] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241, 2015. 1, 2, 8

[24] Digvijay P Salunkhe, Rupa N Bartakke, Guruprasd Chvan, and Pooja R Kothavale. An overview on methods for slope stability analysis. *International Journal of Engineering Research & Technology (IJERT)*, 6(03):2278–0181, 2017. 3

[25] Kuan Tian, Jun Zhang, Haocheng Shen, Kezhou Yan, Pei Dong, Jianhua Yao, Shannon Che, Pifu Luo, and Xiao Han. Weakly-supervised nucleus segmentation based on point annotations: A coarse-to-fine self-stimulated learning strategy. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part V 23*, pages 299–308. Springer, 2020. 3, 5, 8

[26] Quoc Dang Vu, Simon Graham, Tahsin Kurc, Minh Nguyen Nhat To, Muhammad Shaban, Talha Qaiser, Navid Alemi Koohbanani, Syed Ali Khurram, Jayashree Kalpathy-Cramer, Tianhao Zhao, et al. Methods for segmentation and classification of digital microscopy tissue images. *Frontiers in Bioengineering and Biotechnology*, 53: 1–15, 2019. 5

[27] Inwan Yoo, Donggeun Yoo, and Kyunghyun Paeng. Pseudoedgenet: Nuclei segmentation only with point annotations. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22*, pages 731–739. Springer, 2019. 3, 8

[28] Ji Yu. Point-supervised single-cell segmentation via collaborative knowledge sharing. *arXiv preprint arXiv:2304.10671*, 2023. 3

[29] Yanning Zhou, Omer Fahri Onder, Qi Dou, Efstratios Tsougenis, Hao Chen, and Pheng-Ann Heng. Cia-net: Robust nuclei instance segmentation with contour-aware information aggregation. In *International Conference on Information Processing in Medical Imaging*, pages 682–693, 2019. 1, 2