



Learning Reliability of Multi-modality Medical Images for Tumor Segmentation via Evidence-Identified Denoising Diffusion Probabilistic Models

Jianfeng Zhao¹ and Shuo Li²(✉)

¹ School of Biomedical Engineering, Western University, London, ON, Canada

² School of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, USA

slishuo@gmail.com

Abstract. Denoising diffusion probabilistic models (DDPM) for medical image segmentation are still a challenging task due to the lack of the ability to parse the reliability of multi-modality medical images. In this paper, we propose a novel evidence-identified DDPM (EI-DDPM) with contextual discounting for tumor segmentation by integrating multi-modality medical images. Advanced compared to previous work, the EI-DDPM deploys the DDPM-based framework for segmentation tasks under the condition of multi-modality medical images and parses the reliability of multi-modality medical images through contextual discounted evidence theory. We apply EI-DDPM on a BraTS 2021 dataset with 1251 subjects and a liver MRI dataset with 238 subjects. The extensive experiment proved the superiority of EI-DDPM, which outperforms the state-of-the-art methods.

Keywords: Denoising diffusion probabilistic models · Evidence theory · Multi-modality · Tumor segmentation

1 Introduction

Integrating multi-modality medical images for tumor segmentation is crucial for comprehensive diagnosis and surgical planning. In the clinic, the consistent information and complementary information in multi-modality medical images provide the basis for tumor diagnosis. For instance, the consistent anatomical structure information offers the location feature for tumor tracking [22], while the complementary information such as differences in lesion area among multi-modality medical images provides the texture feature for tumor characterization. Multi-modality machine learning aims to process and relate information from multiple modalities [4]. But it is still tricky to integrate multi-modality medical images due to the complexity of medical images.

Existing methods for multi-modality medical image integration can be categorized into three groups: (1) input-based integration methods that concatenate multi-modality images at the beginning of the framework to fuse them directly

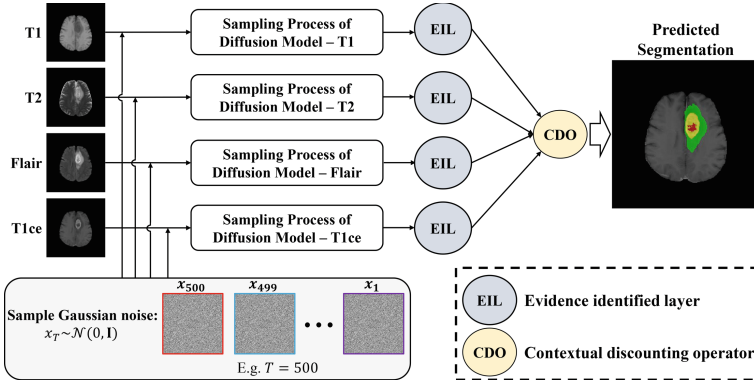


Fig. 1. The workflow of EI-DDPM for brain tumor segmentation of BraTS 2021 dataset. It mainly contains three parts: parallel DDPM path for multi-modality medical images feature extraction, EIL for preliminary multi-modality medical images integration, and CDO for parsing the reliability of multi-modality medical images.

[19, 21], (2) feature-based fusion methods that incorporate a fusion module to merge feature maps [16, 23], and (3) decision-based fusion methods that use weighted averaging to balance the weights of different modalities [11, 15]. Essentially, these methods differ in their approach to modifying the number of channels, adding additional convolutional layers with a softmax layer for attention, or incorporating fixed modality-specific weights. However, there is no mechanism to evaluate the reliability of information from multi-modality medical images. Since the anatomical information of different modality medical images varies, the reliability provided by different modalities may also differ. Therefore, it remains challenging to consider the reliability of different modality medical images when combining multi-modality medical image information.

Dempster-Shafer theory (DST) [18], also known as evidence theory, is a powerful tool for modeling information, combining evidence, and making decisions by integrating uncertain information from various sources or knowledge [10]. Some studies have attempted to apply DST to medical image processing [8]. However, using evidence theory alone does not enable us to weigh the different anatomical information from multi-modality medical images. To explore the reliability of different sources when using evidence theory, the work by Mercier et al. [13] proposed a contextual discounting mechanism to assign weights to different sources. Furthermore, for medical image segmentation, denoising diffusion probabilistic models (DDPM [7]) have shown remarkable performance [9, 20]. Inspired by these studies, if DDPM can parse the reliability of multi-modality medical images to weigh the different anatomy information from them, it will provide a significant approach for tumor segmentation.

In this paper, we propose an evidence-identified DDPM (EI-DDPM) with contextual discounting for tumor segmentation via integrating multi-modality medical images. Our basic assumption is that we can learn the segmentation

feature on single modality medical images using DDPM and parse the reliability of different modalities medical images by evidence theory with a contextual discounting mechanism. Specifically, the EI-DDPM first utilizes parallel conditional DDPM to learn the segmentation feature from a single modality image. Next, the evidence-identified layer (EIL) preliminarily integrates multi-modality images by comprehensively using the multi-modality uncertain information. Lastly, the contextual discounting operator (CDO) performs the final integration of multi-modality images by parsing the reliability of information from multi-modality medical images. The contributions of this work are:

- Our **EI-DDPM** achieves tumor segmentation by using DDPM under the guidance of evidence theory. It provides a solution to integrate multi-modality medical images when deploying the DDPM algorithm.
- The proposed EIL and CDO apply contextual discounting guided DST to parse the reliability of information from different modalities of medical images. This allows for the integration of multi-modality medical images with learned weights corresponding to their reliability.
- We conducted extensive experiments using the BraTS 2021 [12] dataset for brain tumor segmentation and a liver MRI dataset for liver tumor segmentation. Experimental results demonstrate the superiority of EI-DDPM over other State-of-The-Art (SoTA) methods.

2 Method

The EI-DDPM achieves tumor segmentation by parsing the reliability of multi-modality medical images. Specifically, as shown in Fig. 1, the EI-DDPM is fed with multi-modality medical images into the parallel DDPM path and performs the conditional sampling process to learn the segmentation feature from the single modality image (Sect. 2.1). Next, the EIL preliminarily integrates multi-modality images by embedding the segmentation features from multi-modality images into the combination rule of DST (Sect. 2.2). Lastly, the CDO integrates multi-modality medical images for tumor segmentation by contextual discounting mechanism (Sect. 2.3).

2.1 Parallel DDPM Path for Segmentation Feature Learning

Background of DDPM: As an unconditional generative method, DDPM [7] has the form of $p_\theta(\mathbf{x}_0) := \int p_\theta(\mathbf{x}_{0:T})d\mathbf{x}_{1:T}$, where $\mathbf{x}_1, \dots, \mathbf{x}_T$ represent latents with the same dimensionality as the data $\mathbf{x}_0 \sim q(\mathbf{x}_0)$. It contains the forward process of diffusion and the reverse process of denoising. The forward process of diffusion that the approximate posterior $q(\mathbf{x}_{1:T}|\mathbf{x}_0)$, it is a Markov Chain by gradually adding Gaussian noise for converting the noise distribution to the data distribution according to the variance schedule β_1, \dots, β_T :

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) := \mathcal{N}(\mathbf{x}_t; \sqrt{1 - \beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}), \quad q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1 - \bar{\alpha}_t)\mathbf{I}) \quad (1)$$

The reverse process of denoising that the joint distribution $p_\theta(\mathbf{x}_{0:T})$, it can be defined as a Markov chain with learnt Gaussian transitions starting from $p(\mathbf{x}_T) = \mathcal{N}(\mathbf{x}_T; \mathbf{0}, \mathbf{I})$:

$$p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) := \mathcal{N}(\mathbf{x}_{t-1}; \mu_\theta(\mathbf{x}_t, t), \sigma_t^2 \mathbf{I}), \quad p_\theta(\mathbf{x}_{0:T}) := p(\mathbf{x}_T) \prod_{t=1}^T p_\theta(\mathbf{x}_{t-1}|\mathbf{x}_t) \quad (2)$$

where $\alpha_t := 1 - \beta_t$, $\bar{\alpha}_t := \prod_{s=1}^t \alpha_s$, and $\sigma_t^2 = \frac{1 - \bar{\alpha}_{t-1}}{1 - \bar{\alpha}_t} \beta_t$.

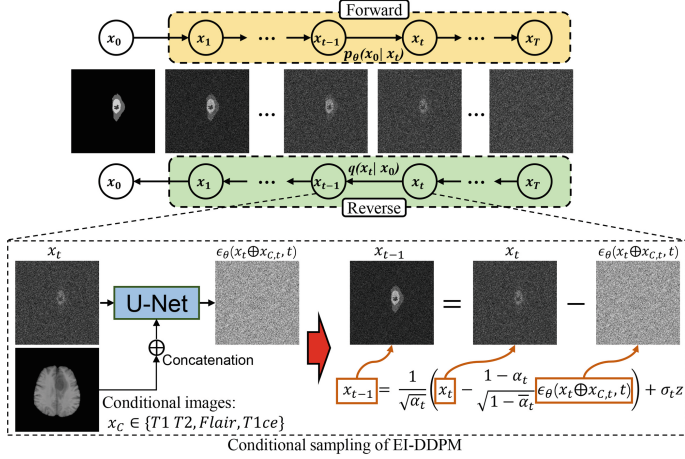


Fig. 2. The algorithm of multi-modality medical images conditioned DDPM.

Multi-modality Medical Images Conditioned DDPM: In Eq. 2 from DDPM [7], the unconditional prediction \mathbf{x}_{t-1} at each step is obtained by subtracting the predicted noise from the previous \mathbf{x}_t , which can be defined as:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(\mathbf{x}_t, t) \right) + \sigma_t \mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, \mathbf{I}) \quad (3)$$

As shown in Fig. 2, to perform the conditional sampling in EI-DDPM, the prediction \mathbf{x}_{t-1} at each step t is on the basis of the concatenation " \oplus " of previous \mathbf{x}_t and the conditional image $\mathbf{x}_{C,t}$. Thus, the \mathbf{x}_{t-1} here can be defined as:

$$\mathbf{x}_{t-1} = \frac{1}{\sqrt{\alpha_t}} \left(\mathbf{x}_t - \frac{1 - \alpha_t}{\sqrt{1 - \bar{\alpha}_t}} \epsilon_\theta(\mathbf{x}_t \oplus \mathbf{x}_{C,t}, t) \right) + \sigma_t \mathbf{z}, \quad \mathbf{z} \sim \mathcal{N}(0, \mathbf{I}) \quad (4)$$

where the conditional image $\mathbf{x}_C \in \{T1, T2, Flair, T1ce\}$ corresponding to the four parallel conditional DDPM path. And in each step t , the $\mathbf{x}_{C,t}$ was also performed the operation of adding Gaussian noise to convert the distribution:

$$\mathbf{x}_{C,t} = \mathbf{x}_{C,0} + \mathcal{N}(0, (1 - \bar{\alpha}_t) \mathbf{I}) \quad (5)$$

where $\mathbf{x}_{C,0}$ presents the multi-modality medical images.

2.2 EIL Integrates Multi-modality Images

As shown in Fig. 1, the EIL is followed at the end of the parallel DDPM path. The hypothesis is to regard multi-modality images as independent and different sources knowledge. And then embedding multi-phase features into the combination rule of DST for evidence identification, which comprehensively parses the multi-phase uncertainty information for confident decision-making. The basic concepts of evidence identification come from DST [18]. It is assumed that $\Theta = \{\theta_1, \theta_2, \dots, \theta_n\}$ is a finite domain called discriminant frame. *Mass function* is defined as \mathcal{M} . So, the evidence about Θ can be represented by \mathcal{M} as:

$$\mathcal{M} : 2^\Theta \rightarrow [0, 1] \quad \text{where} \quad \mathcal{M}(\emptyset) = 0 \quad \text{and} \quad \sum_{A \subseteq \Theta} \mathcal{M}(A) = 1 \quad (6)$$

where the $\mathcal{M}(A)$ denotes the whole belief and evidence allocated to A . The associated *belief* (Bel) and *plausibility* (Pls) functions are defined as:

$$Bel(A) = \sum_{B \subseteq A} \mathcal{M}(B) \quad \text{and} \quad Pls(A) = \sum_{B \cap A \neq \emptyset} \mathcal{M}(B) \quad (7)$$

And using the contour function pls to restrict the *plausibility* function Pls of singletons (i.e. $pls(\theta) = Pls(\{\theta\})$ for all $\theta \in \Theta$) [18]. According to the DST, the mass function \mathcal{M} of a subset always has lower and upper bound as $Bel(A) \leq \mathcal{M}(A) \leq Pls(A)$. From the evidence combination rule of DST, for the mass functions of two independent items \mathcal{M}_1 and \mathcal{M}_2 (i.e. two different modality images), it can be calculated by a new mass function $\mathcal{M}_1 \oplus \mathcal{M}_2(A)$, which is the orthogonal sum of \mathcal{M}_1 and \mathcal{M}_2 as:

$$\mathcal{M}_1 \oplus \mathcal{M}_2(A) = \frac{1}{1 - \gamma} \sum_{B \cap C = A} \mathcal{M}_1(B) \mathcal{M}_2(C) \quad (8)$$

where $\gamma = \sum_{B \cap C = \emptyset} \mathcal{M}_1(B) \mathcal{M}_2(C)$ is *conflict degree* between \mathcal{M}_1 and \mathcal{M}_2 . And the combined contour function Pls_{12} corresponding to $\mathcal{M}_1 \oplus \mathcal{M}_2$ is:

$$Pls_{12} = \frac{Pls_1 Pls_2}{1 - \gamma} \quad (9)$$

The specific evidence identification layer mainly contains three sub-layers (i.e. activation layer, mass function layer, and belief function layer). For activation layer, the activation of i unit can be defined as: $y_i = a_i \exp(-\lambda_i \|x - w_i\|^2)$. The w_i is the weight of i unit. $\lambda_i > 0$ and a_i are parameters. The mass function layer calculates the mass of each K classes using $\mathcal{M}_i(\{\theta_k\}) = u_{ik} y_i$, where $\sum_{k=1}^K u_{ik} = 1$. u_{ik} means the degree of i unit to class θ_k . Lastly, the third layer yields the final belief function about the class of each pixel using the combination rule of DST (Eq. 8).

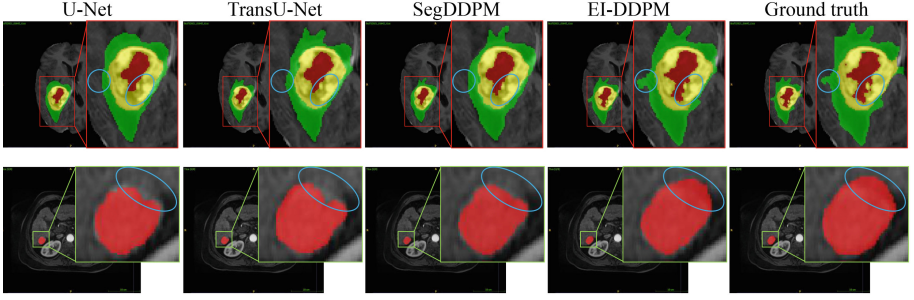


Fig. 3. Visualized segmentation results on BraTS 2021 dataset and liver MRI dataset. The first row presents tumor segmentation on BraTS 2021, where the label of green, yellow, and red represent the ED, ET, and NCR, respectively. The second row is HCC segmentation on liver MRI. The major segmentation differences between different methods are marked with blue circles. (Color figure online)

2.3 CDO Parses the Reliability of Multi-modality Medical Images

CDO performs the contextual discounting operation to EIL for parsing the reliability of multi-modality medical images on tumor segmentation. Firstly, define η as a coefficient in $[0,1]$. Then, transforming \mathcal{M} using the discounting operation with discount rate $1 - \eta$ into a new mass function ${}^\eta\mathcal{M}$ [6]:

$${}^\eta\mathcal{M} = \eta\mathcal{M} + (1 - \eta)\mathcal{M}_? \quad (10)$$

where $\mathcal{M}_?$ is a vacuous mass function defined by $\mathcal{M}(\Theta) = 1$, ${}^\eta\mathcal{M}$ is a mixture of \mathcal{M} and $\mathcal{M}_?$. The coefficient η plays the role of weighting mass function ${}^\eta\mathcal{M}$. The corresponding contour function of ${}^\eta\mathcal{M}$ is:

$${}^\eta pls(\{\theta_k\}) = 1 - \eta_k + \eta_k pls(\{\theta_k\}), \quad k = 1, \dots, K \quad (11)$$

Advantage: *EIL and CDO parse the reliability of different modality medical images in different contexts.* For example, if we feed two modality medical images like T_1 and T_2 into the EI-DDPM, with the discount rate $1 - \eta_{T_1}$ and $1 - \eta_{T_2}$, we will have two contextual discounted contour functions ${}^{\eta_{T_1}} pls_{T_1}$ and ${}^{\eta_{T_2}} pls_{T_2}$. The combined contour function in Eq. 9 is proportional to the ${}^{\eta_{T_1}} pls_{T_1} {}^{\eta_{T_2}} pls_{T_2}$. In this situation, the η_{T_1} and η_{T_2} can be trained to weight the two modality medical images by parsing the reliability.

3 Experiment and Results

Dataset. We used two MRI datasets that BraTS 2021 [1, 2, 12] and a liver MRI dataset. BraTS 2021¹ contains 1251 subject with 4 aligned MRI modalities: T1, T2, Flair, and contrast-enhanced T1 (T1ce). The segmentation labels consist of

¹ <http://braintumorsegmentation.org/>.

Table 1. The quantitative evaluation of the comparison between EI-DDPM and other methods. The criteria of Dice evaluated the performance. Avg presents the average value in the whole dataset.

	Dice for BraTS 2021				Dice for Liver MRI		
	WT	ET	TC	Avg	Hemangioma	HCC	Avg
U-Net [17]	91.30	83.23	86.53	86.49	91.89	84.21	87.12
TransU-Net [5]	91.56	84.72	87.26	87.43	92.44	85.35	88.42
SegDDPM [20]	92.04	85.74	88.37	88.65	93.10	86.27	89.07
EI-DDPM	93.52	87.31	90.23	90.17	94.57	88.03	90.47

Table 2. Ablation studies for EIL and CDO. The criteria of Dice evaluated the performance. Avg presents the average value in the whole dataset.

	Dice for BraTS 2021				Dice for Liver MRI		
	WT	ET	TC	Avg	Hemangioma	HCC	Avg
No CDO	92.10	86.42	89.27	89.57	94.01	87.26	89.89
No EIL&CDO	92.33	85.83	88.50	89.03	93.24	86.33	89.25
EI-DDPM	93.52	87.31	90.23	90.17	94.57	88.03	90.47

GD-enhancing tumor (ET), necrotic tumor core (NCR), and peritumoral edematous (ED), which are combined into 3 nested subregions: Enhancing Tumor (ET) region, Tumor Core (TC) region (i.e. ET+NCR), and Whole Tumor (WT) region (i.e. ET+ED+NCR). The liver MRI dataset contains 238 subjects with 110 hemangioma subjects and 128 hepatocellular carcinoma (HCC) subjects. Each subject has corresponding 3 aligned MRI modalities: T1, T2, and T1ce (protocols were gadobutrol 0.1 mmol/kg on a 3T MRI scanner). The segmentation labels of hemangioma and HCC were performed by two radiologists with more than 5-year-experience. The resolution of the BraTS 2021 image and liver MRI image are 240×240 and 256×256 , respectively. Both BraTS 2021 and liver MRI datasets were randomly divided into 3 groups following the ratio of training/validation/test as 7:1:2.

Implementation Details. For the number of DDPM paths, BraTS 2021 dataset is equal to 4 corresponding to the input 4 MRI modalities and the liver MRI dataset is equal to 3 corresponding to the input 3 MRI modalities. In the parallel DDPM path, the noise schedule followed the improved-DDPM [14], and the U-Net [17] was utilized as the denoising model with 300 sampling steps. In EIL, the initial values of a_i equal 0.5 and λ_i is equal to 0.01. For CDO, the initial of parameter η_k is equal to 0.5. With the Adam optimization algorithm, the denoising process was optimized using \mathcal{L}_1 , and the EIL and CDO were optimized using Dice loss. The learning rate of EI-DDPM was set to 0.0001. The

Table 3. The learnt reliability coefficients η of BraTS 2021

η	ED	ET	NCR
T1	0.2725	0.3751	0.3318
T2	0.8567	0.9458	0.6912
Flair	0.9581	0.7324	0.5710
T1ce	0.7105	0.9867	0.9879

Table 4. The learnt reliability coefficients η of Liver MRI

η	Hemangioma	HCC
T1	0.8346	0.5941
T2	0.8107	0.5463
T1ce	0.9886	0.9893

framework was trained on Ubuntu 20.04 platform by using Pytorch and CUDA library, and it ran on an RTX 3090Ti graphics card with 24 GB memory.

Quantitative and Visual Evaluation. The performance of EI-DDPM is evaluated by comparing with three methods: a classical CNN-based method (U-Net [17]), a Transformer-based method (TransU-Net [5]), and a DDPM-based method for multi-modality medical image segmentation (SegDDPM [20]). The Dice score is used for evaluation criteria. Figure 3 shows the visualized segmentation results of EI-DDPM and compared methods. It shows some ambiguous area lost segmentation in three compared methods but can be segmented by our EI-DDPM. Table 1 reports the quantitative results of EI-DDPM and compared methods. Our EI-DDPM achieves highest Dice value on both BraTS dataset and liver MRI dataset. All these results proved EI-DDPM outperforms the three other methods.

Ablation Study. To prove the contribution from EIL and CDO, we performed 2 types of ablation studies: (1) removing CDO (i.e. No CDO) and (2) removing EIL and CDO (i.e. No EIL&CDO). Table 2 shows the quantitative results of ablation studies. Experimental results proved both EIL and CDO contribute to tumor segmentation on BraTS 2021 dataset and liver MRI dataset.

Discussion of Learnt Reliability Coefficients η . Table 3 and Table 4 show the learned reliability coefficients η on BraTS 2021 dataset and liver MRI dataset. The higher η value, the higher reliability of the corresponding region segmentation. As shown in Table 3, the Flair modality provides the highest reliability for ED segmentation. And both the T2 modality and T1ce modality provide relatively high reliability for ET and NCR segmentation. As shown in Table 4, the T1ce modality provides the highest reliability for hemangioma and HCC segmentation. These reliability values are the same as clinical experience [1, 3].

4 Conclusion

In this paper, we proposed a novel DDPM-based framework for tumor segmentation under the condition of multi-modality medical images. The EIL and CDO

enable our EI-DDPM to capture the reliability of different modality medical images with respect to different tumor regions. It provides a way of deploying contextual discounted DST to parse the reliability of multi-modality medical images. Extensive experiments prove the superiority of EI-DDPM for tumor segmentation on multi-modality medical images, which has great potential to aid in clinical diagnosis. The weakness of EI-DDPM is that it takes around 13 s to predict one segmentation image. In future work, we will focus on improving sampling steps in parallel DDPM paths to speed up EI-DDPM.

Acknowledgements. This work is partly supported by the China Scholarship Council (No. 202008370191)

References

1. Baid, U., et al.: The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification. arXiv preprint [arXiv:2107.02314](https://arxiv.org/abs/2107.02314) (2021)
2. Bakas, S., et al.: Advancing the cancer genome atlas glioma MRI collections with expert segmentation labels and radiomic features. *Sci. Data* **4**(1), 1–13 (2017)
3. Balogh, J., et al.: Hepatocellular carcinoma: a review. *Journal of hepatocellular carcinoma*, pp. 41–53 (2016)
4. Baltrušaitis, T., Ahuja, C., Morency, L.P.: Multimodal machine learning: a survey and taxonomy. *IEEE Trans. Pattern Anal. Mach. Intell.* **41**(2), 423–443 (2018)
5. Chen, J., et al.: Transunet: transformers make strong encoders for medical image segmentation. arXiv preprint [arXiv:2102.04306](https://arxiv.org/abs/2102.04306) (2021)
6. Denoëux, T., Kanjanatarakul, O., Sriboonchitta, S.: A new evidential k-nearest neighbor rule based on contextual discounting with partially supervised learning. *Int. J. Approximate Reasoning* **113**, 287–302 (2019)
7. Ho, J., Jain, A., Abbeel, P.: Denoising diffusion probabilistic models. *Adv. Neural. Inf. Process. Syst.* **33**, 6840–6851 (2020)
8. Huang, L., Ruan, S., Decazes, P., Denoëux, T.: Lymphoma segmentation from 3d PET-CT images using a deep evidential network. *Int. J. Approximate Reasoning* **149**, 39–60 (2022)
9. Kim, B., Oh, Y., Ye, J.C.: Diffusion adversarial representation learning for self-supervised vessel segmentation. arXiv preprint [arXiv:2209.14566](https://arxiv.org/abs/2209.14566) (2022)
10. Lian, C., Ruan, S., Denoëux, T., Li, H., Vera, P.: Joint tumor segmentation in PET-CT images using co-clustering and fusion based on belief functions. *IEEE Trans. Image Process.* **28**(2), 755–766 (2018)
11. Lim, K.Y., Mandava, R.: A multi-phase semi-automatic approach for multisequence brain tumor image segmentation. *Expert Syst. Appl.* **112**, 288–300 (2018)
12. Menze, B.H., et al.: The multimodal brain tumor image segmentation benchmark (brats). *IEEE Trans. Med. Imaging* **34**(10), 1993–2024 (2014)
13. Mercier, D., Quost, B., Denoëux, T.: Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Inf. Fusion* **9**(2), 246–258 (2008)
14. Nichol, A.Q., Dhariwal, P.: Improved denoising diffusion probabilistic models. In: *International Conference on Machine Learning*, pp. 8162–8171. PMLR (2021)
15. Qu, T., et al.: M3net: a multi-scale multi-view framework for multi-phase pancreas segmentation based on cross-phase non-local attention. *Med. Image Anal.* **75**, 102232 (2022)

16. Raju, A., et al.: Co-heterogeneous and adaptive segmentation from multi-source and multi-phase ct imaging data: a study on pathological liver and lesion segmentation. In: European Conference on Computer Vision. pp. 448–465. Springer (2020)
17. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
18. Shafer, G.: A mathematical theory of evidence, vol. 42. Princeton University Press (1976)
19. Wang, J., et al.: Tensor-based sparse representations of multi-phase medical images for classification of focal liver lesions. *Pattern Recogn. Lett.* **130**, 207–215 (2020)
20. Wolleb, J., Sandkühler, R., Bieder, F., Valmaggia, P., Cattin, P.C.: Diffusion models for implicit image segmentation ensembles. In: International Conference on Medical Imaging with Deep Learning, pp. 1336–1348. PMLR (2022)
21. Zhang, L., et al.: Robust pancreatic ductal adenocarcinoma segmentation with multi-institutional multi-phase partially-annotated CT scans. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12264, pp. 491–500. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59719-1_48
22. Zhao, J., et al.: United adversarial learning for liver tumor segmentation and detection of multi-modality non-contrast MRI. *Med. Image Anal.* **73**, 102154 (2021)
23. Zhou, Y., et al.: Hyper-pairing network for multi-phase pancreatic ductal adenocarcinoma segmentation. In: Shen, D., et al. (eds.) MICCAI 2019. LNCS, vol. 11765, pp. 155–163. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-32245-8_18