



IJCAI 2024 workshop, August 3, 2024 - Jeju, South Korea



Multi-category Brain Tumor Segmentation via Multi-scale and Cross-category Relation Modeling

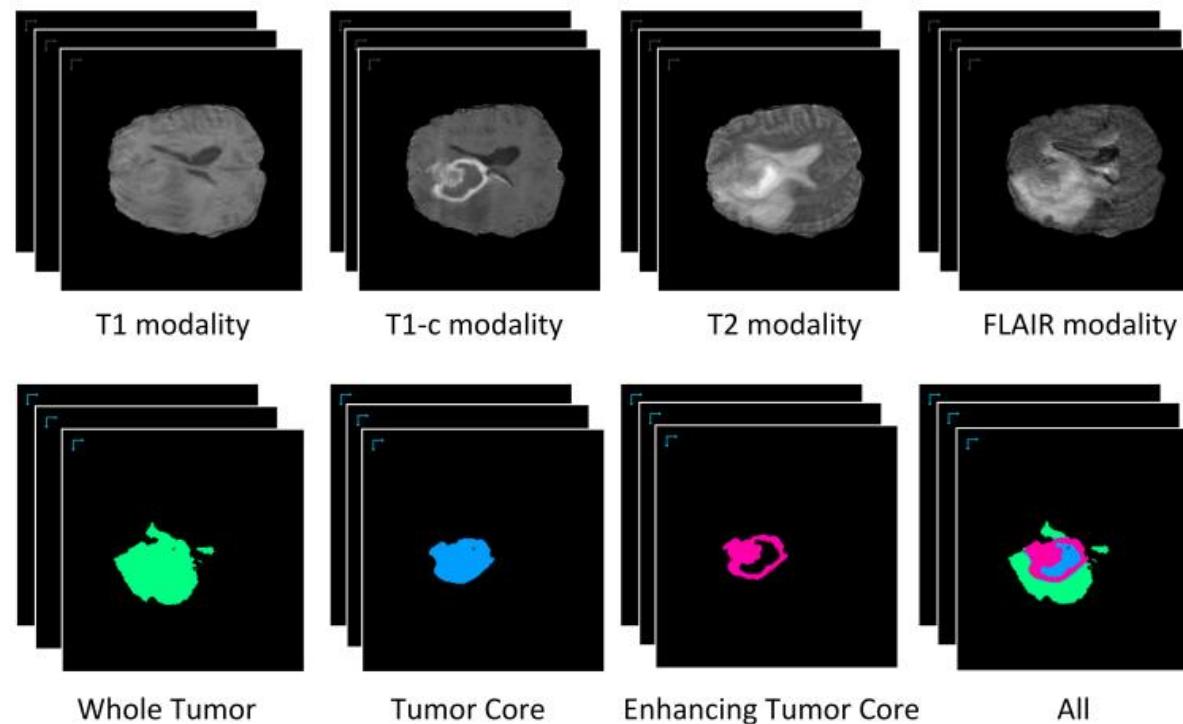
Dongzhe Li Baoyao Yang* Yuebin Xie Weide Zhan and Jingsong Lin

Guangdong University of Technology, Guangzhou

Background

■ Brain Tumor Segmentation Challenges

- Tumor location and boundary refinement
- Distinguish multiple categories of sub-regions



Related Work

◆ Traditional feature extraction methods in brain tumor segmentation

Element-wise addition or concatenation^{[1] [2] [3]}

Multi-scale feature aggregation^{[4] [5] [6]}

□ Limitations

- I. Fail to capture the complex spatial semantic relations in multi-category tumors
- II. Introduce redundant information

[1] O. Oktay et al. "Attention u-net: Learning where to look for the pancreas". in preprint arXiv:1804.03999. (2018)

[2] A. Myronenko et al.: "3d mri brain tumor segmentation using autoencoder regularization". in MICCAI Brainlesion Workshop. Springer, pp. 311–320. (2019)

[3] Z. Jiang, C. Ding, M. Liu, and D. Tao.: "Two-stage cascaded u-net: 1st place solution to brats challenge 2019 segmentation task". in MICCAI Brainlesion Workshop. Springer, pp. 231–241. (2020)

[4] Liang-Chieh Chen et al.: Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. In: Proceedings of the European Conference on Computer Vision (ECCV), (2018)

[5] M. Yang et al.: "Denseaspp for semantic segmentation in street scenes. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3684-3692 (2018)

[6] Shuanglang Feng et al.: "CPFNet: Context Pyramid Fusion Network for Medical Image Segmentation. In: IEEE Transactions on Medical Imaging, pp. 3008 - 3018 (2020)

Motivation

■ Different characteristics across different levels of the encoder

- Deep features carry richer semantic information → localization
- Shallow features contain more low-level visual information → capture subtle boundaries

■ Attention mechanisms in image segmentation

- Focus on the most relevant regions or features
- help the model better understand the correlation between different regions in the image

Idea

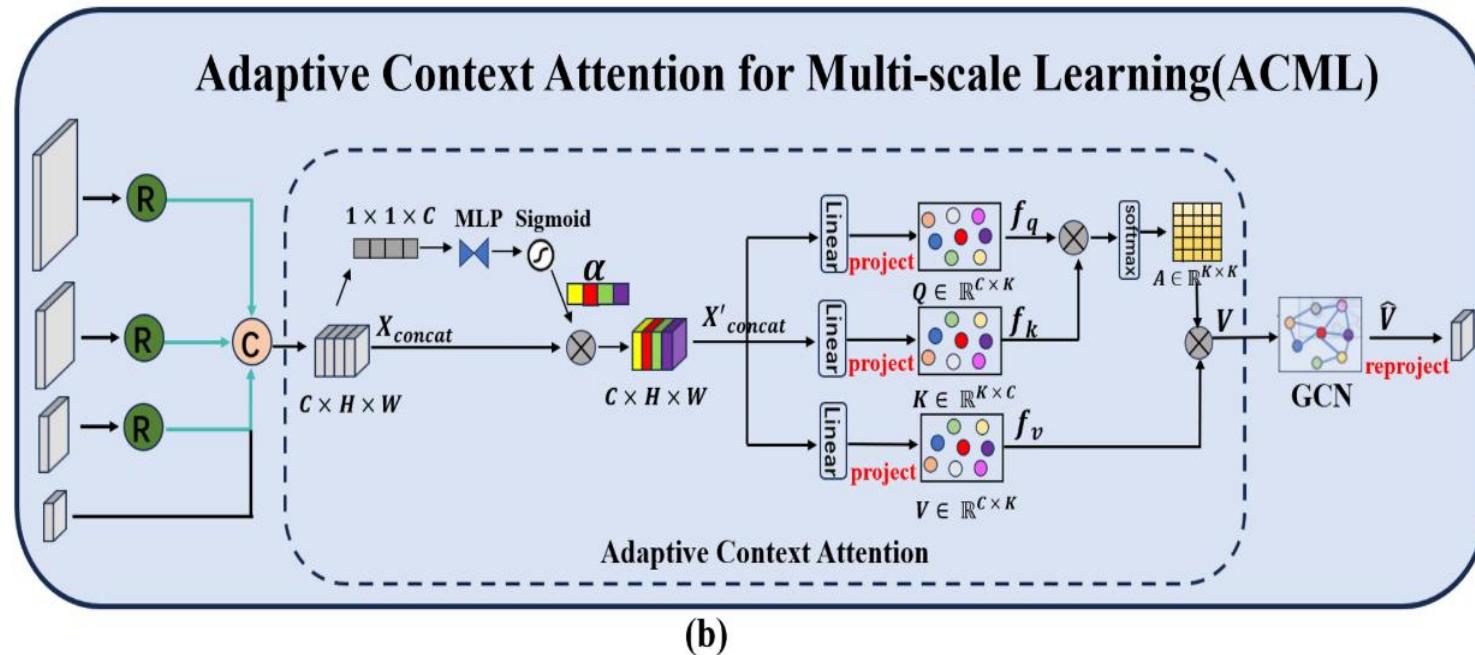
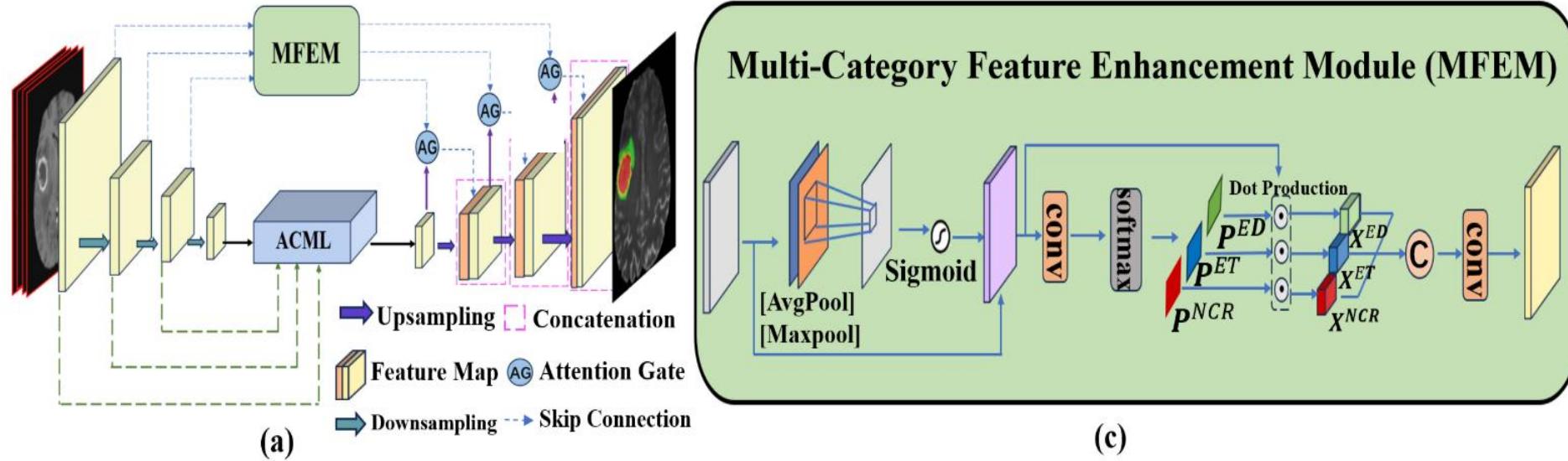
- *Aggregate multi-scale information and enhance features for multiple categories*

Multi-scale and Cross-category Relation Modeling

- *Dynamically combine local features with global dependencies*

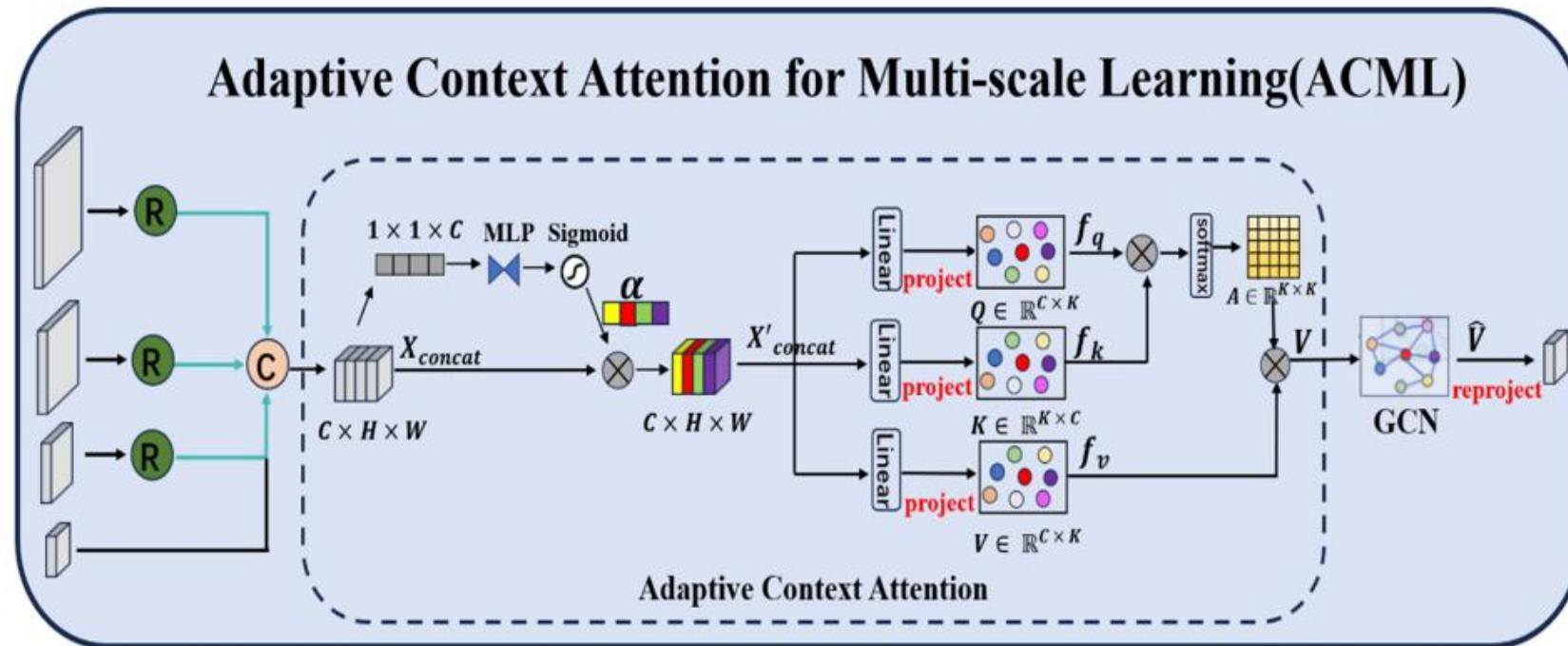
Capture positional information and category relations for localization and classification

Proposed Method



Proposed Method

➤ Adaptive Context Attention for Multi-scale Feature Learning (ACML):



- Scale attention α

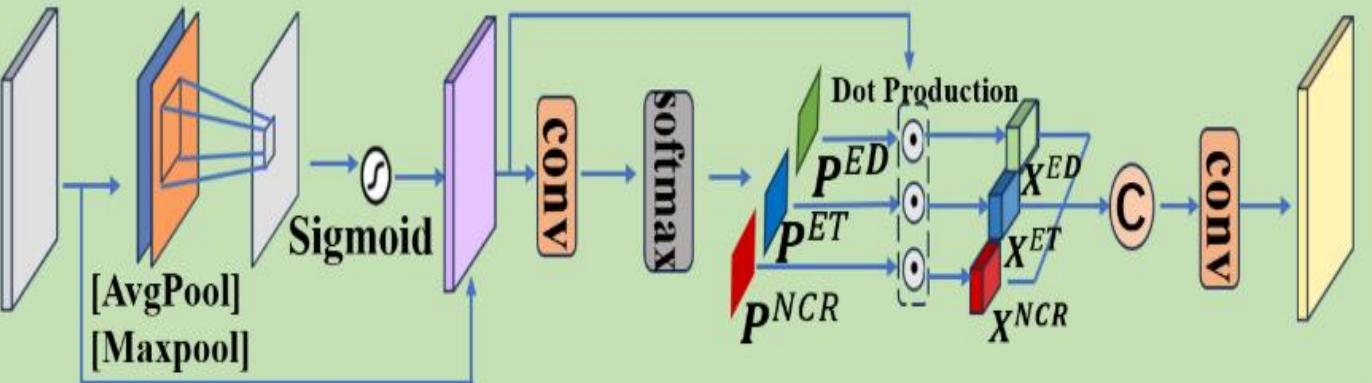
Capture the relations between different scales and gain global context

- Interaction coordinate space

Utilize Graph Convolutional Networks to model long-range dependence

Proposed Method

Multi-Category Feature Enhancement Module (MFEM)



➤ Multi-category Feature Enhancement Module (**MFEM**):

- Spatial attention

Perceive the spatial distribution between different regions

- Probability map:

$$P^{NCR, ED, ET} = \text{split}(\text{softmax}(\text{conv}(\tilde{X})))$$

Experiments

- Experimental Setting
 - **BraTS 2019 and BraTS 2020**
 - 335 and 369 annotated brain tumor samples
 - 4 modalities: T1, T1ce, T2, and FLAIR
 - **label 1:** Non-enhancing tumor core (NCR&NET), **label 2:** Peritumoral edema (ED), **label 4:** Enhancing tumor (ET), and **label 0:** Background
 - Each modality scan is sliced and cropped into dimensions of 160×160 . The datasets are split into 80% for training our model and 20% for testing.
 - Evaluation metrics:
 - **Dice score**
 - **95% Hausdorff distance (HD)**

Experiments

- Comparison Results with SOTA methods

Table 1. Result comparison of brain tumor segmentation methods on BraTS2019

Model	Dice(%) ↑				HD95(mm) ↓			
	ET	TC	WT	Ave	ET	TC	WT	Ave
TransBTS[10]	80.86	81.19	89.35	83.80	5.642	6.048	4.332	5.460
Nestedformer[9]	82.11	86.42	91.18	86.57	5.534	5.906	5.317	5.585
SF-Net[11]	80.08	82.33	88.61	83.67	4.787	7.440	7.288	6.505
ACM-Net[12]	80.63	87.15	88.08	85.28	4.564	7.774	3.862	5.400
Eoformer[13]	82.94	86.83	90.39	86.72	4.053	5.843	5.822	5.239
Ours	82.74	87.67	90.59	87.00	5.296	5.694	4.472	5.154

- Highest Dice scores for the Tumor Core region, indicating its efficacy in capturing the interrelations among brain tumor sub-region

Experiments

- Comparison Results with SOTA methods

Table 2. Result comparison of brain tumor segmentation methods on BraTS2020

Model	Dice(%) ↑				HD95(mm) ↓			
	ET	TC	WT	Ave	ET	TC	WT	Ave
TransBTS[10]	80.89	83.25	90.10	84.08	5.873	6.875	4.876	5.824
Nestedformer[9]	82.85	86.48	91.20	86.84	5.721	6.115	4.598	5.528
SF-Net[11]	81.10	83.84	89.01	84.65	4.305	7.661	7.720	6.562
ACM-Net[12]	82.42	87.75	90.08	86.75	4.492	7.624	3.956	5.375
Eoformer[13]	83.54	87.12	90.87	87.17	5.911	6.041	3.852	5.268
Ours	83.20	87.89	90.93	87.34	5.134	5.863	4.529	5.175

Experiments

■ Comparison Results Visualization

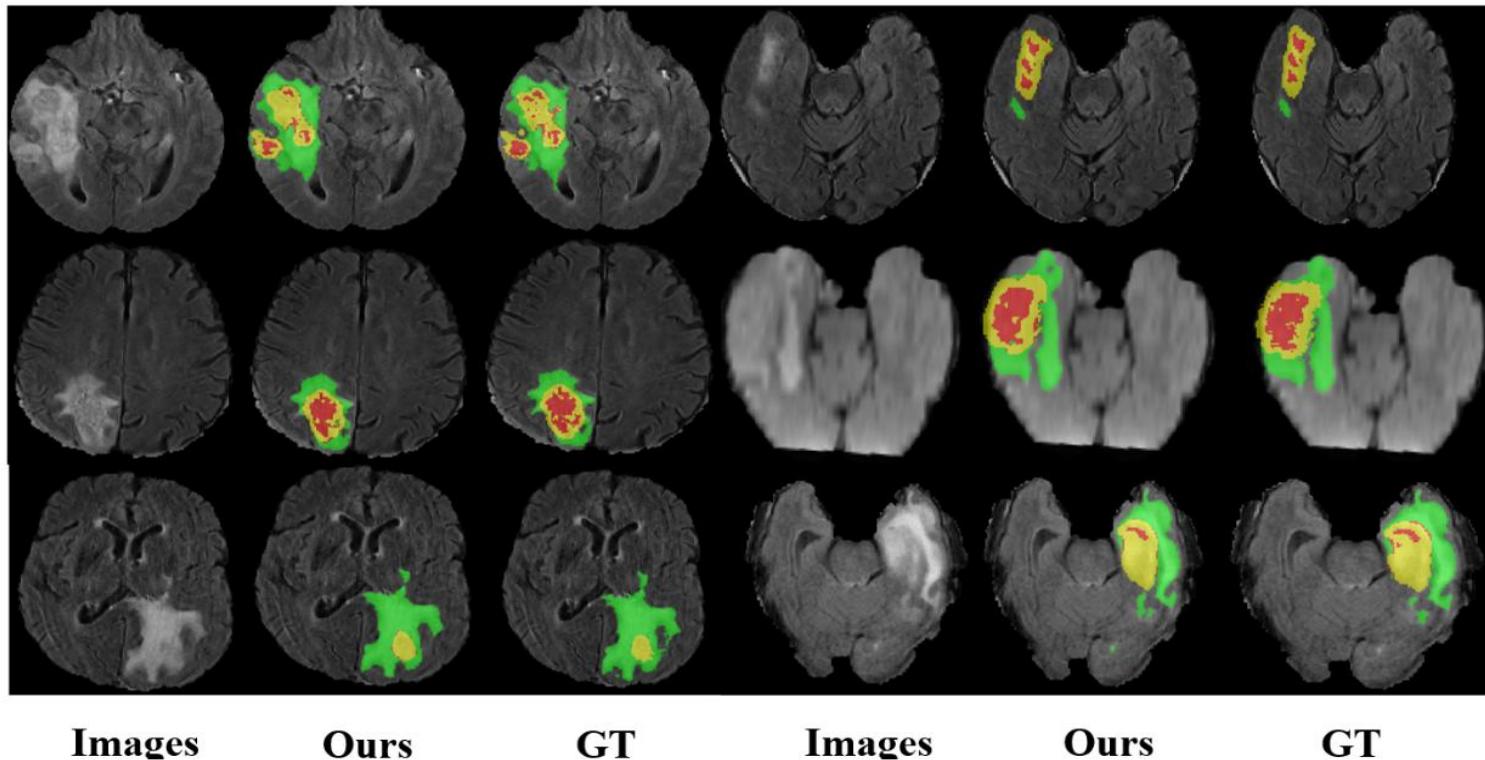


Fig. 2. The visual results on BraTS2019. Green, yellow and red colors indicate regions of ED, ET and NCR, respectively.

- Accurately segments the shape of brain tumors as well as their edges
- Even small necrotic(NCR) areas and enhanced tumor regions(ET) overlapping are correctly distinguished

Experiments

■ Ablation Study

Table 3. Ablation study of critical components on BraTS2019

Model	Dice(%)			
	ET	TC	WT	Ave
(1) Unet	79.10	81.00	87.93	82.67
(2) Unet+MFEM	80.80	85.34	90.17	85.43
(3) Unet+ACML	82.55	86.28	90.06	86.69
(4) Unet+ACML+MFEM	82.74	87.67	90.59	87.00

- Unet+ACML: Significant improvement in the segmentation of all categories, confirming that the global contextual information is effective
- Compared to Unet+MFEM , ET has shown a greater improvement in Unet+ACML over Unet, indicates that ACML can simultaneously focus on local features

Conclusion

- To capture enough global context and handle morphological heterogeneity relations across tumor categories, this paper presents a novel approach for adaptively capturing multi-scale and multi-category semantic relations.
- **Multi-category Brain Tumor Segmentation via Multi-scale and Cross-category Relation Modeling Framework is proposed:**
 - Adaptive Context Attention for Multi-scale Feature Learning
 - Multi-category Feature Enhancement Module
- Experiments among public datasets BraTS 2019 and BraTS 2020 validate the effectiveness of the proposed method.