MICCAI 2024 Marrakesh MOROCCO

ORAL

TinyU-Net: Lighter yet Better U-Net with Cascaded Multi-Receptive Fields

Junren Chen[1], Rui Chen[2], Wei Wang[3], Junlong Cheng[1], Lei Zhang[1*], Liangyin Chen[1*]

MICCAI

INTRODUCTION

Limited-resource Settings

Low Resource Settings
- Edge Devices
- Mobile Devices
- Embedded Devices
- Point-of-Care Devices

⋮

Health Inequity ☹

Well-resourced Settings

Giving hope to people who do not have access to medical imaging in low resource Settings.

Health Equity ☺

How to Chase High Performance While Being Lightweight?

Low Parameters and Computations

Weak Feature Representation

Suboptimal Performance

Lightweight AI Models

Integrate

Low Resource Settings

Heavyweight AI Models

Integrate

Optimal Performance

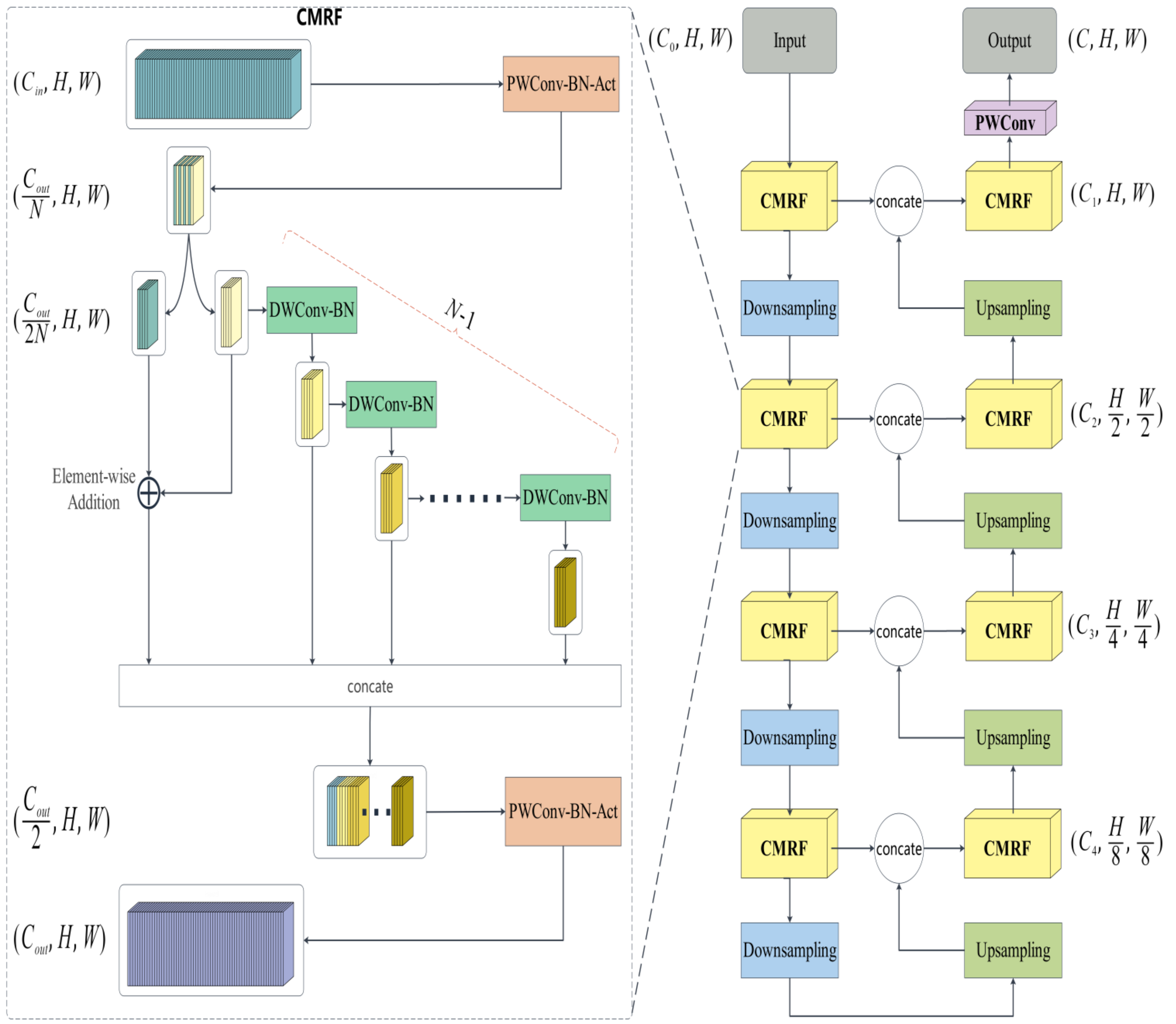
Strong Feature Representation

High Parameters and Computations

Contributions

- Proposing the CMRF module, which enhances feature representation by fusing multi-receptive fields using a cost-effective cascading strategy.
- Applying the CMRF module to other models consistently improves segmentation performance while reducing parameters and computational complexity.
- Proposing TinyU-Net, a CMRF-based model with only 0.48M parameters, achieving exceptional performance in medical image segmentation with low parameters and computational complexity.

METHODOLOGY



MOTIVATION

Standing on the Shoulders of Giants

Cheap Operation

Depthwise convolution (DWConv) and pointwise convolution (PWConv) have low parameter overhead.

Cost-effective Way

Perform convolution operations on part of the feature maps due to redundancies across different channels.

Receptive Field

As the network deepens, the receptive field expands.

Feature Augmentation

The fusion of feature via element-wise addition can improve feature representation.

RESULTS

Table 1. Comparative quantitative results on ISIC2018 dataset.

Model	Params (M)	FLOPs (G)	IoU (%) Mean	IoU (%) Lesion	IoU (%) BG	Dice (%) Mean	Dice (%) Lesion	Dice (%) BG
U-Net [3]	24.89	112.91	82.69	76.14	89.23	90.38	86.45	94.31
SegNet [26]	29.44	80.34	83.64	77.53	89.74	91.06	87.35	94.59
DeepLabV3+ [27]	5.81	13.22	83.43	77.35	89.50	91.02	87.23	94.46
TransUNet [4]	93.23	64.48	84.66	78.99	90.32	91.75	88.26	94.91
Swin-Unet [5]	27.15	15.46	83.39	77.28	89.49	90.98	87.19	94.45
UNeXt [7]	1.47	1.15	82.85	76.39	89.31	90.50	86.61	94.35
ConvUNeXt [10]	3.51	14.51	83.96	77.95	89.96	91.24	87.61	94.71
U-Lite [11]	0.88	1.52	84.13	78.12	90.13	91.29	87.72	94.81
CMUNeXt [12]	3.15	14.84	84.62	78.82	90.42	91.62	88.16	94.97
FBSNet [13]	0.60	5.76	84.03	78.12	89.93	91.35	87.72	94.70
TinyU-Net	0.48	3.33	85.49	80.04	90.94	92.18	88.91	95.25

Table 2. Comparative quantitative results on NCP dataset. GL, GGO, LF, and BG denote consolidation, ground-glass opacity, lung field, and background, respectively.

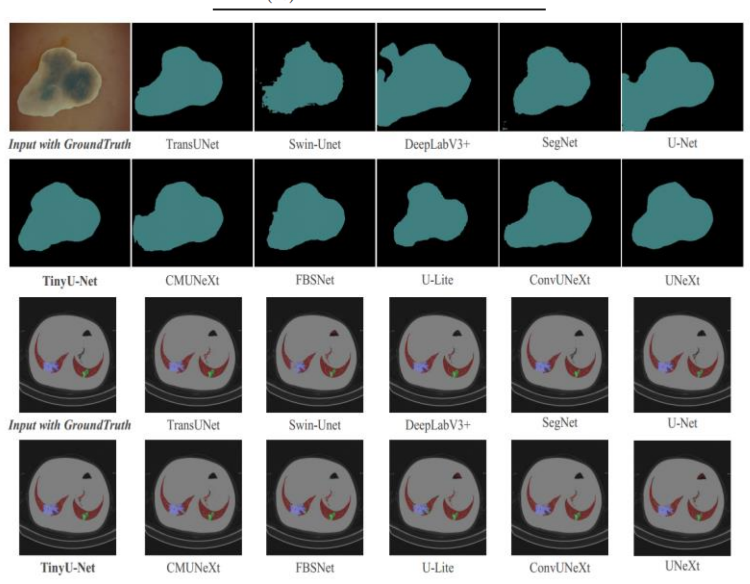
Model	Params (M)	FLOPs (G)	IoU (%) Mean	IoU (%) CL	IoU (%) GGO	IoU (%) LF	IoU (%) BG	Dice (%) Mean	Dice (%) CL	Dice (%) GGO	Dice (%) LF	Dice (%) BG
U-Net [3]	24.89	112.93	78.78	65.21	56.68	93.65	99.58	86.95	78.94	72.35	96.72	99.79
SegNet [26]	29.45	80.49	76.04	63.40	48.44	92.75	99.57	84.73	77.60	65.27	96.24	99.79
DeepLabV3+ [27]	5.81	13.22	77.28	64.73	52.41	92.52	99.46	85.80	78.59	68.78	96.11	99.73
TransUNet [4]	93.23	64.51	79.50	65.60	58.98	93.82	99.58	87.50	79.22	74.19	96.81	99.79
Swin-Unet [5]	27.15	15.46	72.98	54.74	46.33	91.48	99.35	82.32	70.75	63.32	95.55	99.67
UNeXt [7]	1.47	1.15	76.33	62.19	51.03	92.62	99.47	85.05	76.69	67.58	96.17	99.74
ConvUNeXt [10]	3.51	14.52	79.36	66.97	57.18	93.70	99.59	87.38	80.22	72.76	96.75	99.80
U-Lite [11]	0.88	1.52	75.66	60.00	50.52	92.61	99.51	84.51	75.00	67.13	96.16	99.75
CMUNeXt [12]	3.15	14.84	79.55	67.44	57.44	93.73	99.59	87.52	80.55	72.97	96.77	99.80
FBSNet [13]	0.61	5.80	77.92	64.76	53.83	93.48	99.59	86.26	78.61	69.99	96.63	99.79
TinyU-Net	0.48	3.34	80.27	68.75	58.80	93.93	99.59	88.05	81.48	74.05	96.87	99.80

Table 3. Ablation results (IoU (%)) for CMRF on ISIC2018 and NCP datasets. GL, GGO, and LF denote consolidation, ground-glass opacity, and lung field, respectively.

Model	Feature extraction block	Params (M)	FLOPs (G)	ISIC2018 Lesion	NCP CL	NCP GGO	NCP LF
SegNet [26]	Original	29.44	80.34	77.53	63.40	48.44	92.75
	CMRF	**0.64**	**7.02**	**78.42**	**63.66**	**50.60**	**92.80**
CMUNeXt [12]	Original	3.15	14.84	78.82	67.44	57.44	93.73
	CMRF	**1.97**	**12.23**	**79.16**	**67.48**	**57.58**	**93.74**

Table 4. Ablation results (mIoU (%)) for the number of DWConv-BN blocks on NCP datasets.

N	1	2	4	8	16
mIoU (%)	79.24	79.58	79.77	**80.27**	79.56



CONCLUSION

- Problems in existing lightweight models :
  - Weak feature representation due to low parameters and computational complexity.
- To solve the above problem, we propose the following methods:
  - **CMRF**: We introduce a novel CMRF module that fuses information from multi-receptive fields in a layer based on a cost-friendly cascading strategy to improve feature representation.
  - **TinyU-Net**: Building upon our CMRF, we further propose a lightweight TinyU-Net, a simple yet effective U-shaped network, for medical image segmentation.
- Our TinyU-Net achieves state-of-the-art segmentation performance with a small number of parameters and FLOPs.
- We believe that the proposed methods can adapt well to limited-resource settings, thereby fostering health equity.