Supplementary materials of “Hi-gMISnet: Generalized medical image segmentation using DWT based multilayer fusion and dual mode attention into high resolution pGAN”

This supplementary material contains two sections. The first section provides a detailed description of the datasets used in our paper. And second section presents detailed comparative results on each dataset used in the paper.

1. Datasets

In our work, we employed nine distinct datasets encompassing diverse modalities, including CT scan, MRI, optical microscopy, colonoscopy image, ultrasound, X-ray, dermatological photograph, and retinal fundus image. Subsequent sections provide comprehensive information regarding these datasets.

1.1. PhysioNet ICH dataset for stroke lesion segmentation

This dataset [1] contains 82 CT scans, including 36 scans for patients diagnosed with intracranial hemorrhage with the following types: Intraventricular, Intraparenchymal, Subarachnoid, Epidural, and Subdural. Each CT scan for every patient comprises approximately 30 slices, each with a slice thickness of 5 mm. Radiologists have meticulously annotated the regions of intracranial hemorrhage in every slice. As our goal is focused on segmenting the stroke lesions within the CT scan images, we exclusively considered slices that contained lesions. We performed a random split of the dataset into training, validation, and test sets, adhering to an 8:1:1 ratio. The size of all images was fixed to 256×256 pixels in our experiments.

1.2. BUSI dataset for breast lesion segmentation

The Breast Ultrasound Images (BUSI) [2] dataset comprises a rich collection of 780 ultrasound images acquired from 600 female patients. This dataset has 437 images representing benign lesions, 210 images showcasing malignant lesions, and 133 images depicting normal cases. In this work, we used 647 images containing lesions, excluding the 133 normal images for the segmentation problem. The ground truth masks were created by expert radiologists. These remaining 647 images were divided in an 8:1:1 ratio, respectively, for the training, validation, and test sets. The resolution of all images was set to 256×256 pixels in our experiments.

1.3. CVC-ClinicDB dataset for polyp segmentation

The CVC-ClinicDB dataset [3] comprises video frames extracted from colonoscopies, specifically showcasing a diverse range of polyp formations. Expert radiologists have thoughtfully provided ground truth annotations, consisting of masks that delineate the area occupied by each polyp within the images. This comprehensive dataset comprises 612 images and we split these images into training, validation, and test sets in an 8:1:1 ratio. All the images were also resized to a consistent resolution of 288×384 pixels in our experiments.

1.4. MoNuSeg dataset for nuclei segmentation

The MoNuSeg dataset [4] features H&E stained images captured at a $40\times$ magnification, that are obtained from various organs and patients. The training and test splitting for this dataset is provided by the dataset organizers. The training set, consisting of 30 images from seven distinct organs, comprises approximately 21,623 annotations marking nuclear boundaries. The test set, similarly encompassing 14 images across seven organs, contains approximately 7,223 annotations outlining nuclear boundaries. The training set is randomly distributed into training and validation, with 23 images for training and 7 images for validation. The size of all images was fixed to 512×512 pixels in our experiments.

1.5. GLAS dataset for gland segmentation

The Gland Segmentation (GLAS) dataset [5] comprises H&E stained microscopy images of T3 or T42 colorectal adenocarcinoma with ground truth annotations from pathologists. This dataset encompasses a total of 165 images thoughtfully divided by the dataset organizers into 85 for training and 80 for testing. Within the 85 training images, it is further subdivided into two distinct sets: a training set comprising 72 images and a validation set comprising 13 images. Given the diverse sizes of the images within the dataset, in our experiment, we standardized all the images to 512×512 pixels.

1.6. ISIC-2018 dataset for skin lesion segmentation

The ISIC-2018 challenge dataset [6] offers a substantial collection of dermoscopy images aimed at advancing the field of automated diagnosis of melanoma, a type of skin cancer. The Task-1 training dataset of the ISIC-2018 challenge is used in this work. This training dataset is comprised of 2594 images, each accompanied by the corresponding ground truth annotation. These images were randomly distributed into training, validation, and testing sets. The dataset was randomly partitioned into three subsets for training, validation, and testing. The training set consisted of 2074 images, while the validation and testing sets each contained 260 images. All the images were standardized to 128×128 pixels.

1.7. DRIVE dataset for retinal vessel segmentation

The Digital Retinal Images for Vessel Extraction (DRIVE) dataset [7] serves as a valuable resource for retinal vessel segmentation. Comprising a total of 40 color fundus images, this dataset includes 7 cases with abnormal pathology. The images were acquired through a diabetic retinopathy screening program in the Netherlands using a Canon CR5 non-mydriatic 3CCD camera with a 45 degree field of view (FOV). Each image was captured using 8 bits per color plane at 768 by 584 pixels. Originally, the dataset was evenly split into train and test sets (20 images each) under the guidance of its organizers. Both the training and test sets are provided with manually segmented ground truth annotations, diligently created by an ophthalmological expert. In this work, we randomly selected 16 images from the training set for model training, and 4 images for validation purposes. The dimension of all images was set to 512×512 pixels.

1.8. Montgomery dataset for lung segmentation

The Montgomery County dataset [8] comprises a collection of 138 frontal chest X-rays originating from Montgomery County's Tuberculosis screening program. Among these, 80 cases are normal chest X-rays, while 58 cases exhibit indications of TB. These X-ray images were captured utilizing a stationary Eureka X-ray machine (CR) and are provided in a 12-bit gray-level format. The images come in two distinct sizes, with dimensions of either 4020×4892 or 4892×4020 pixels. To ensure consistency, we standardized the dimensions of all images to 256×256 pixels. In our experiment, we divided the entire dataset into subsets for training, validation, and testing, each containing 110, 14, and 14 images, respectively.

1.9. PROMISE12 dataset for prostate segmentation

This dataset was first published in a MICCAI Grand Challenge [9] and it contains a 'training set', a 'live challenge test set', and a 'test set'. The 'training set' comprises 50 cases, while the 'live challenge test set' contains 20 cases, and the 'test set' encompasses 30 cases. These cases are transversal T2-weighted MR images of different prostate sizes and appearances. We extracted all slices from each case and retained only those images containing the prostate. Subsequently, we employed the 'training set' for model training, the 'live challenge test set' for validation, and the 'test set' for model testing. To ensure uniformity, we standardized the dimensions of all images to 256×256 pixels.

2. Comparative Results

Detailed comparative results containing all the performance metrics (F1, IoU, Precision, Recall, HD, and ASSD) with the existing methods on each dataset (refer to section 1) are shown in tables 1 - 9.

Table 1: Quantitative performance comparison among different methods on the Physionet PhysioNet ICH dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	64.85±1.99	54.68±1.38	75.25±5.56	62.83±1.87	29.36±2.31	4.37±0.25
UNet++ [11]	70.23±0.43	59.73±0.70	76.30 ±1.88	68.40±0.92	31.66±3.96	3.96±1.04
ResUNet++ [12]	56.91±2.69	46.26±2.87	71.39±1.09	52.43±3.43	52.55±1.55	11.72±2.83
MultiResUNet [13]	72.64±2.28	61.35±2.04	71.57±1.62	76.28±2.49	28.73±3.44	3.81±0.27
CE-Net [14]	70.69±1.36	58.94±1.20	78.48 ±1.19	66.39±2.76	16.66±0.44	2.25±0.13
PraNet [15]	70.38±2.10	57.78±1.75	65.25±2.75	80.00±1.03	22.56±1.03	5.49±1.69
CPFNet [16]	70.33±1.80	59.25±1.40	75.81±3.57	68.21±0.79	25.89±3.68	5.06±3.02
EANet [17]	76.93±0.40	65.49±0.17	78.57±0.47	76.60±1.38	24.14±5.13	3.11±0.41
COMA-Net [18]	74.15±0.66	62.03±0.62	73.44±1.89	77.85±3.72	25.29±4.49	3.35±0.79
HTC-Net [19]	70.85±0.69	58.86±1.37	70.38±3.20	77.75±2.97	41.06±13.54	4.62±0.83
EGAN [20]	78.52±1.43	66.45±1.57	82.70±1.33	76.88±0.36	22.39±4.66	2.57±0.64
nnU-Net [21]	73.10±0.67	62.48±0.58	80.31±0.93	70.53±1.80	18.26±4.99	2.54±0.83
Hi-gMISnet	79.53±0.41	68.11±0.15	82.09±2.74	79.12±1.37	24.46±7.45	1.85±0.41

Table 2: Quantitative performance comparison among different methods on the BUSI dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	77.79±1.68	67.99±1.68	85.98±2.42	76.27±3.44	35.40±0.85	3.66±0.36
UNet++ [11]	80.61±0.72	71.20±0.79	86.00 ±0.59	80.87±1.26	30.84±3.93	3.19±1.15
ResUNet++ [12]	78.89±1.05	68.36±1.82	83.52±2.77	79.94±0.43	31.86±4.88	3.18±0.70
MultiResUNet [13]	81.03±0.79	71.77±0.74	80.93±1.75	83.57±2.11	37.92±9.06	3.53±0.07
CE-Net [14]	85.95±0.30	77.52±0.47	87.78 ±0.12	87.36±0.51	24.29±0.31	1.58±0.02
PraNet [15]	86.26±0.68	78.02±0.79	88.63±0.53	87.65±0.49	22.11±0.82	1.73±0.08
CPFNet [16]	83.81±1.78	75.15±2.08	88.28±1.92	84.11±1.89	22.74±2.14	1.63±0.12
EANet [17]	84.63±0.47	75.95±0.43	86.84±0.26	86.92±0.51	25.18±1.76	1.89±0.27
COMA-Net [18]	87.28±0.06	79.29±0.03	91.31±0.74	86.64±0.71	20.12±1.07	1.19±0.07
HTC-Net [19]	83.76±0.63	74.64±0.58	84.14±1.46	87.43±1.67	29.24±3.42	2.72±0.55
EGAN [20]	82.99±0.31	74.11±0.58	86.62±1.06	85.10±1.35	25.42±1.03	2.19±0.15
nnU-Net [21]	85.36±0.35	77.15±0.49	88.21±0.22	87.35±0.30	20.68±1.05	1.57±0.09
Hi-gMISnet	88.68±0.15	81.17±0.15	92.91±0.11	87.20±0.27	17.64±1.86	0.95±0.03

Table 3: Quantitative performance comparison among different methods on the CVC-ClinicDB dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	86.41±1.24	79.72±1.08	91.64±0.94	84.22±2.09	33.75±1.95	2.30±1.21
UNet++ [11]	88.67±0.31	82.16±0.34	91.72±1.33	88.77±1.07	32.60±5.21	1.95±0.39
ResUNet++ [12]	87.81±0.56	82.27±0.90	89.76±0.38	87.48±0.64	26.79±3.77	2.62±0.82
MultiResUNet [13]	87.62±0.81	81.48±0.98	87.12±0.76	89.26±1.24	31.02±6.19	0.72±0.03
CE-Net [14]	91.33±0.95	86.5±0.76	93.52±0.59	90.63±1.40	16.62±1.06	1.06±0.09
PraNet [15]	93.64±0.55	89.46±0.58	92.99±1.08	95.39±1.22	16.64±1.14	1.33±0.39
CPFNet [16]	93.15±0.84	88.43±1.09	95.58±0.34	92.14±1.32	13.38±1.55	0.71±0.19
EANet [17]	90.93±0.53	85.80±0.56	92.88±0.48	89.93±0.60	23.59±4.64	1.54±0.07
COMA-Net [18]	94.38±0.32	90.16±0.49	95.04±0.97	94.38±0.23	12.72±1.67	0.63±0.10
HTC-Net [19]	89.27±0.76	82.60±1.33	90.36±0.45	89.92±1.75	31.32±1.52	2.24±0.13
EGAN [20]	90.93±0.71	85.39±0.71	92.92±0.34	90.36±0.37	22.91±1.61	1.83±1.17
nnU-Net [21]	73.10±0.67	62.48±0.58	73.44±1.89	77.85±3.72	25.29±4.49	3.35±0.79
Hi-gMISnet	94.19±0.04	90.68±0.12	94.58±0.25	94.25±0.23	13.66±0.88	1.44±0.15

Table 4: Quantitative performance comparison among different methods on the MoNuSeg dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	78.74±0.07	65.03±0.11	74.61±1.36	83.82±1.59	36.04±0.48	0.72±0.03
UNet++ [11]	77.58±0.25	64.50±1.32	74.83±1.48	81.18±1.72	35.34±0.46	0.69±0.02
ResUNet++ [12]	78.24±0.36	64.36±0.48	75.19±2.28	82.36±2.05	34.96±0.29	0.72±0.01
MultiResUNet [13]	78.85±0.74	65.18±0.99	74.50±2.37	84.42±1.71	35.45±0.54	0.71±0.04
CE-Net [14]	79.45±0.11	66.01±0.15	74.26±1.92	83.49±1.87	35.71±1.23	0.69±0.02
PraNet [15]	73.33±0.31	57.98±0.39	66.20±0.89	82.57±0.59	34.45±0.52	0.79±0.03
CPFNet [16]	79.59±0.24	66.19±0.32	74.46±1.31	86.12±1.54	35.01±0.67	0.69±0.005
EANet [17]	78.98±0.24	65.36±0.33	75.76±1.410	83.68±1.88	38.44±0.85	0.76±0.02
COMA-Net [18]	81.53±0.20	68.90±0.29	78.14±0.79	85.63±0.55	35.48±0.32	0.62±0.01
HTC-Net [19]	75.93±0.95	61.68±1.15	69.22±2.41	86.23±1.64	36.85±1.08	0.97±0.07
EGAN [20]	67.15±2.93	50.89±3.31	60.47±2.79	80.64±2.01	35.88±1.15	1.17±0.15
nnU-Net [21]	81.22±0.14	68.85±0.19	80.92±0.58	82.94±0.41	38.54±1.08	0.72±0.01
Hi-gMISnet	82.50±0.08	70.31±0.11	80.11±1.22	85.54±1.25	36.35±0.77	0.66±0.05

Table 5: Quantitative performance comparison among different methods on the GLAS dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	91.15±0.16	84.33±0.20	91.13±0.79	91.94±1.07	82.52±0.58	1.69±0.07
UNet++ [11]	91.73±0.19	85.30±0.33	92.32±0.84	92.83±0.85	81.75±0.98	1.78±0.13
ResUNet++ [12]	86.37±0.21	77.12±0.32	86.39±0.34	88.05±0.39	91.16±3.03	3.51±0.10
MultiResUNet [13]	91.45±0.14	84.93±0.21	91.68±0.57	92.06±0.77	90.45±8.76	1.76±0.05
CE-Net [14]	92.26±0.06	86.23±0.08	92.11±0.29	93.05±0.29	79.11±1.49	1.84±0.17
PraNet [15]	92.97±0.34	87.39±0.56	92.74±0.19	93.76±0.77	70.01±3.57	1.46±0.09
CPFNet [16]	92.09±0.16	85.96±0.28	92.08±0.35	92.86±0.33	76.32±1.25	1.77±0.07
EANet [17]	90.90±0.47	84.03±0.74	90.45±0.56	92.37±0.96	83.12±3.11	2.09±0.10
COMA-Net [18]	93.13±0.10	87.65±0.12	92.91±0.44	93.87±0.34	71.83±3.58	1.41±0.16
HTC-Net [19]	90.52±0.24	83.59±0.36	90.54±0.33	91.89±0.21	75.57±1.70	1.81±0.10
EGAN [20]	86.71±0.20	77.47±0.32	86.44±0.39	88.61±0.34	90.52±1.12	2.89±0.05
nnU-Net [21]	90.62±0.37	83.79±0.54	88.65±0.36	93.42±0.41	80.79±3.12	2.33±0.14
Hi-gMISnet	93.25±0.12	87.83±0.16	94.01±0.74	93.06±0.85	71.04±3.54	1.27±0.05

Table 6: Quantitative performance comparison among different methods on the ISIC-2018 dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	88.85±0.11	81.34±0.14	90.29±0.32	90.33±0.27	9.70±0.26	0.47±0.01
UNet++ [11]	89.12±0.09	81.71±0.14	92.46±0.74	88.51±0.64	9.33±0.05	0.45±0.01
ResUNet++ [12]	88.95±0.07	81.49±0.05	91.52±0.32	89.09±0.31	9.38±0.01	0.45±0.01
MultiResUNet [13]	87.74±0.23	79.76±0.14	89.52±0.51	89.26±0.74	10.34±0.19	0.53±0.02
CE-Net [14]	89.29±0.04*	81.98±0.04*	91.51±0.14*	89.70±0.19*	18.46±0.27*	0.79±0.04*
PraNet [15]	89.46±0.11	82.05±0.13	90.96±0.34	90.23±0.27	8.87±0.18	0.38±0.02
CPFNet [16]	89.44±0.06	82.11±0.07	91.86±0.22	89.49±0.21	9.00±0.13	0.41±0.01
EANet [17]	88.37±0.08	80.77±0.12	91.05±1.06	88.76±0.93	10.21±0.28	0.51±0.03
COMA-Net [18]	89.63±0.08	82.40±0.09	91.04±0.36	90.59±0.28	8.72±0.06	0.39±0.01
HTC-Net [19]	88.23±0.57	80.27±0.83	87.63±1.08	91.65±0.28	12.03±3.71	0.47±0.06
EGAN [20]	88.75±0.07	81.81±0.08	91.15±0.23	89.98±0.34	9.37±0.31	0.58±0.02
nnU-Net [21]	87.93±0.18	80.49±0.21	94.06±0.42	85.87±0.21	10.26±0.16	0.63±0.02
Hi-gMISnet	90.40±0.02	84.23±0.06	91.79±0.61	92.10±0.55	8.44±0.29	0.33±0.01

$*$ Due to CE-Net's architectural limitations, 256 × 256 image size is used for evaluation

Table 7: Quantitative performance comparison among different methods on the DRIVE dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	77.95±1.44	63.92±1.91	78.59±2.48	77.89±0.59	40.00±1.46	0.75±0.11
UNet++ [11]	80.71±0.14	67.69±0.20	83.11 ±0.17	78.93±0.33	32.36±1.43	0.48±0.02
ResUNet++ [12]	76.53±0.79	62.05±1.00	82.98±1.01	71.49±1.99	40.83±0.70	0.93±0.05
MultiResUNet [13]	76.16±0.26	61.54±0.32	76.48±0.32	76.51±0.68	45.04±1.53	1.02±0.04
CE-Net [14]	75.07±0.56	60.12±0.56	72.64±0.59	78.17±1.02	48.15±0.38	0.97±0.05
PraNet [15]	61.60±0.12	44.57±0.13	56.64±0.24	67.87±0.07	68.93±1.61	2.31±0.04
CPFNet [16]	74.97±0.62	60.00±0.79	70.50±1.65	80.55±0.82	37.09±1.37	0.75±0.03
EANet [17]	79.02±0.05	65.34±0.08	78.26±0.86	80.31±0.78	33.80±0.86	0.58±0.02
FR-UNet[22]	79.47±0.38	65.97±0.53	81.84±0.89	77.77±0.24	36.01±1.26	0.62±0.02
COMA-Net [18]	75.59±1.05	60.80±1.35	69.81±3.86	83.21±2.89	31.88±0.73	0.58±0.01
HTC-Net [19]	77.79±0.83	63.69±1.11	82.01±1.42	74.56±2.19	40.37±3.79	0.82±0.13
EGAN [20]	72.33±0.59	56.72±0.72	75.21±0.56	70.21±0.97	51.35±1.28	1.30±0.06
nnU-Net [21]	79.59±0.02	66.94±0.04	81.19±0.27	79.11±0.28	29.63±1.26	0.54±0.01
Hi-gMISnet	81.65±0.09	69.01±0.12	79.23±0.70	84.59±0.67	31.18±0.69	0.43±0.01

Table 8: Quantitative performance comparison among different methods on the Montgomery dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	97.28±0.12	94.74±0.23	97.11±0.28	97.49±0.04	11.18±1.84	0.07±0.01
UNet++ [11]	97.98±0.07	96.06±0.13	98.00±0.03	97.99±0.12	6.68±0.34	0.03±0.01
ResUNet++ [12]	97.59±0.11	95.31±0.21	97.60±0.05	97.60±0.24	9.79±1.98	0.04±0.01
MultiResUNet [13]	97.51±0.16	95.14±0.29	96.49±0.38	98.56±0.13	16.64±6.39	0.05±0.01
PraNet [15]	97.96±0.06	96.01±0.11	97.91±0.05	98.02±0.07	7.01±0.23	0.03±0.01
CE-Net [14]	98.06±0.03	96.20±0.06	98.41±0.17	97.73±0.13	6.79±0.62	0.03±0.01
CPFNet [16]	97.91±0.04	95.92±0.09	98.09±0.12	97.75±0.07	6.75±0.37	0.03±0.01
FR-UNet[22]	98.07±0.21	96.23±0.38	81.84±0.89	77.77±0.24	36.01±1.26	0.62±0.02
EANet [17]	97.98±0.13	96.05±0.25	98.40±0.25	97.59±0.39	7.79±3.16	0.03±0.01
COMA-Net [18]	98.14±0.02	96.35±0.04	97.88±0.18	98.41±0.14	5.69±0.15	0.02±0.01
HTC-Net [19]	97.85±0.27	95.84±0.46	97.39±0.39	98.37±0.36	8.49±0.86	0.07±0.04
EGAN [20]	97.72±0.09	95.56±0.17	97.95±0.17	97.52±0.35	7.64±0.84	0.04±0.01
nnU-Net [21]	97.83±0.02	96.21±0.05	98.19±0.07	97.78±0.06	4.89±0.15	0.03±0.01
Hi-gMISnet	98.48±0.01	97.02±0.01	98.62±0.15	98.35±0.16	8.68±2.56	0.02±0.01

Table 9: Quantitative performance comparison among different methods on the PROMISE12 dataset

Method	F1-score (%)	IoU-score (%)	Precision (%)	Recall (%)	HD	ASSD
UNet [10]	88.97±0.26	81.68±0.47	88.47±1.27	91.62±0.99	7.30±0.35	0.46±0.08
UNet++ [11]	80.61±0.72	71.20±0.79	86.00 ±0.59	80.87±1.26	30.84±3.93	3.19±1.15
ResUNet++ [12]	85.64±0.35	76.97±0.43	84.89±0.58	89.58±0.53	9.73±0.89	0.59±0.03
MultiResUNet [13]	87.83±0.34	79.92±0.61	84.58±0.99	93.56±0.65	8.76±1.89	0.44±0.02
CE-Net [14]	89.63±0.18	82.48±0.23	88.63 ±0.30	92.42±0.07	6.32±0.10	0.31±0.01
PraNet [15]	90.12±0.04	83.17±0.11	88.75±0.48	93.11±0.57	5.66±0.05	0.28±0.01
CPFNet [16]	89.34±0.11	81.96±0.22	89.51±0.89	90.97±1.06	6.14±0.26	0.30±0.02
FR-UNet[22]	86.92±0.94	78.92±1.24	81.84±0.89	77.77±0.24	36.01±1.26	0.62±0.02
EANet [17]	89.88±0.09	82.83±0.14	89.41±0.48	92.10±0.63	6.55±0.20	0.32±0.01
COMA-Net [18]	90.22±0.09	83.29±0.16	89.56±0.15	92.53±0.19	5.80±0.17	0.27±0.01
HTC-Net [19]	89.18±0.31	82.01±0.37	86.90±0.44	93.61±0.76	7.05±0.07	0.41±0.02
EGAN [20]	85.47±0.25	76.16±0.45	85.23±0.46	88.94±0.81	8.81±0.04	0.52±0.02
nnU-Net [21]	89.51±0.12	82.45±0.16	87.78±0.24	93.24±0.06	6.11±0.05	0.34±0.01
Hi-gMISnet	90.79±0.08	84.09±0.10	91.16±0.67	91.89±0.58	5.62±0.13	0.24±0.01

References

[1] M. Hssayeni, M. Croock, A. Salman, H. Al-khafaji, Z. Yahya, and B. Ghoraani. Computed tomography images for intracranial hemorrhage detection and segmentation. *Intracranial Hemorrhage Segmentation Using A Deep Convolutional Model. Data*, 5(1):14, 2020.

[2] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy. Dataset of breast ultrasound images. *Data brief*, 28:104863, 2020.

[3] J. Bernal, F. J. Sánchez, G. Fernández-Esparrach, D. Gil, C. Rodríguez, and F. Vilariño. Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. *Comput. Med. Imaging Graph.*, 43:99–111, 2015.

[4] N. Kumar, R. Verma, S. Sharma, S. Bhargava, A. Vahadane, and A. Sethi. A dataset and a technique for generalized nuclear segmentation for computational pathology. *IEEE Trans. Med. Imaging*, 36(7):1550–1560, 2017.

[5] K. Sirinukunwattana, J. P. Pluim, H. Chen, X. Qi, P.-A. Heng, Y. B. Guo, L. Y. Wang, B. J. Matuszewski, E. Bruni, U. Sanchez, et al. Gland segmentation in colon histology images: The glas challenge contest. *Med. Image Anal.*, 35:489–502, 2017.

[6] N. Codella, V. Rotemberg, P. Tschandl, M. E. Celebi, S. Dusza, D. Gutman, B. Helba, A. Kalloo, K. Liopyris, M. Marchetti, et al. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1902.03368*, 2019.

[7] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken.

Ridge-based vessel segmentation in color images of the retina. *IEEE Trans. on Med. Imaging*, 23(4):501–509, 2004.

[8] S. Jaeger, S. Candemir, S. Antani, Y.-X. J. Wáng, P.-X. Lu, and G. Thoma. Two public chest x-ray datasets for computer-aided screening of pulmonary diseases. *Quant. Imaging Med. Surg.*, 4(6):475, 2014.

[9] G. Litjens, R. Toth, W. Van De Ven, C. Hoeks, S. Kerkstra, B. Van Ginneken, G. Vincent, G. Guillard, N. Birbeck, J. Zhang, et al. Evaluation of prostate segmentation algorithms for mri: the promise12 challenge. *Med. image analysis*, 18(2):359–373, 2014.

[10] O. Ronneberger, P. Fischer, and T. Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 234–241. Springer, 2015.

[11] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang. Unet++: A nested u-net architecture for medical image segmentation. In *Deep learning in medical image analysis and multimodal learning for clinical decision support*, pages 3–11. Springer, 2018.

[12] D. Jha, P. H. Smedsrud, M. A. Riegler, D. Johansen, T. De Lange, P. Halvorsen, and H. D. Johansen. Resunet++: An advanced architecture for medical image segmentation. In *2019 IEEE International Symposium on Multimedia (ISM)*, pages 225–2255. IEEE, 2019.

[13] N. Ibtehaz and M. S. Rahman. Multiresunet: Rethinking the u-net architecture for multimodal biomedical image segmentation. *Neural networks*, 121:74–87, 2020.

[14] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, and J. Liu. Ce-net: Context encoder network for 2d medical image segmentation. *IEEE Trans. Med. Imaging*, 38(10):2281–2292, 2019.

[15] D.-P. Fan, G.-P. Ji, T. Zhou, G. Chen, H. Fu, J. Shen, and L. Shao. Pranet: Parallel reverse attention network for polyp segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 263–273. Springer, 2020.

[16] S. Feng, H. Zhao, F. Shi, X. Cheng, M. Wang, Y. Ma, D. Xiang, W. Zhu, and X. Chen. Cpfnet: Context pyramid fusion network for medical image segmentation. *IEEE Trans. Med. Imaging*, 39(10):3008–3018, 2020.

[17] K. Wang, X. Zhang, X. Zhang, Y. Lu, S. Huang, and D. Yang. Eanet: Iterative edge attention network for medical image segmentation. *Pattern Recognit.*, 127:108636, 2022.

[18] S. Ahmed and M. K. Hasan. Coma-net: Towards generalized medical image segmentation using complementary attention guided bipolar refinement modules. *Biomed. Signal Process. Control*, 86:105198, 2023.

[19] H. Tang, Y. Chen, T. Wang, Y. Zhou, L. Zhao, Q. Gao, M. Du, T. Tan, X. Zhang,

and T. Tong. Htc-net: A hybrid cnn-transformer framework for medical image segmentation. *Biomed. Signal Process. Control*, 88:105605, 2024.

[20] S. Innani, P. Dutande, U. Baid, V. Pokuri, S. Bakas, S. Talbar, B. Baheti, and S. C. Guntuku. Generative adversarial networks based skin lesion segmentation. *Sci. Reports*, 13(1):13467, 2023.

[21] F. Isensee, P. F. Jaeger, S. A. Kohl, J. Petersen, and K. H. Maier-Hein. nnu-net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. methods*, 18(2):203–211, 2021.

[22] W. Liu, H. Yang, T. Tian, Z. Cao, X. Pan, W. Xu, Y. Jin, and F. Gao. Full-resolution network and dual-threshold iteration for retinal vessel and coronary angiograph segmentation. *IEEE J. Biomed. Health Informatics*, 26(9):4623–4634, 2022.