Interactive Medical Image Segmentation: A Benchmark Dataset and Baseline

Junlong Cheng[1,2] **Bin Fu**[2,4] **Jin Ye**[2] **Guoan Wang**[2] **Tianbin Li**[2] **Haoyu Wang**[2]
Ruoyu Li[1] He Yao[1] Junren Cheng[1] Jinwen Li[3] Yanzhou Su[2] Min Zhu[1*] Junjun He[2†]
[1]College of Computer Science, Sichuan University
[2]Shanghai AI Laboratory, China
[3]School of Computer Science and Technology, Xinjiang University
[4] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences
{chengjunlong, fubin, yejin, wangguoan, litianbin, wanghaoyu}@pjlab.org.cn,
{ruoyuli, 2024323040035, chengjunren}@stu.scu.edu.cn, 107556522204@stu.xju.edu.cn,
yanzhou.su@outlook.com, hejunjun.su@pjlab.org.cn, zhumin@scu.edu.cn

Appendix A: Demo and Code

Our code, model weights, and dataset are available at: https://github.com/uni-medical/IMIS-Bench.

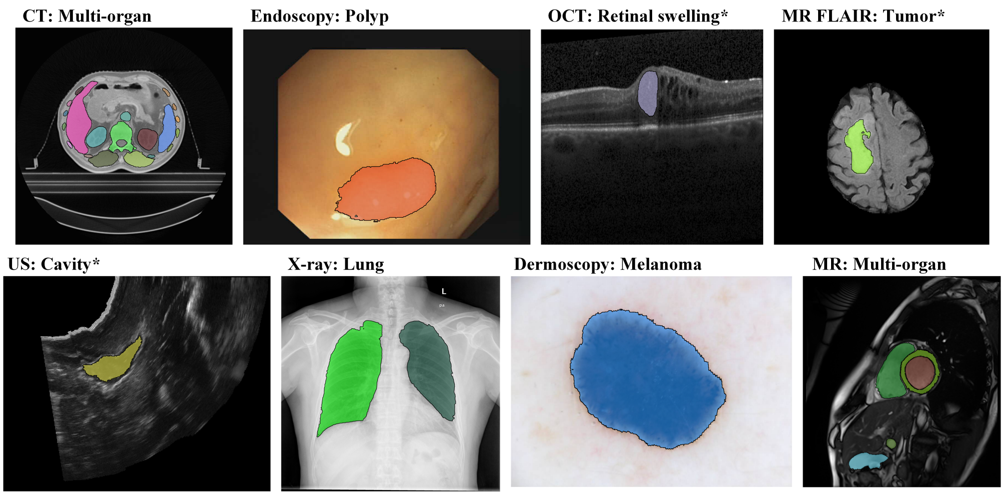

Figure 1: **Example predictions of IMIS-Net across different modalities and segmentation tasks**. "*" Indicates that the corresponding image modality or segmentation task was not included in our training plan. Our model demonstrates its versatility by effectively handling multiple medical image modalities and performing various segmentation tasks, even on those that it has not previously encountered.

Appendix B: IMed-361M Information and Availability

We have compiled 110 medical image segmentation datasets into a comprehensive, large-scale, multimodal, high-quality dataset named IMed-361M, making it openly accessible for interactive

*Corresponding author.
†Project lead.

medical image segmentation. This dataset includes over 6.4 million images, 87.6 million ground truth annotations, and 273.4 million interactive masks (IMask), encompassing 14 image modalities and 204 segmentation targets. Fig.2 presents representative samples, while detailed category information is provided in Tab.1, which forms the basis for IMIS-Net's text prompts. The modality and category details of these open-source and private datasets are displayed in Tab.2.

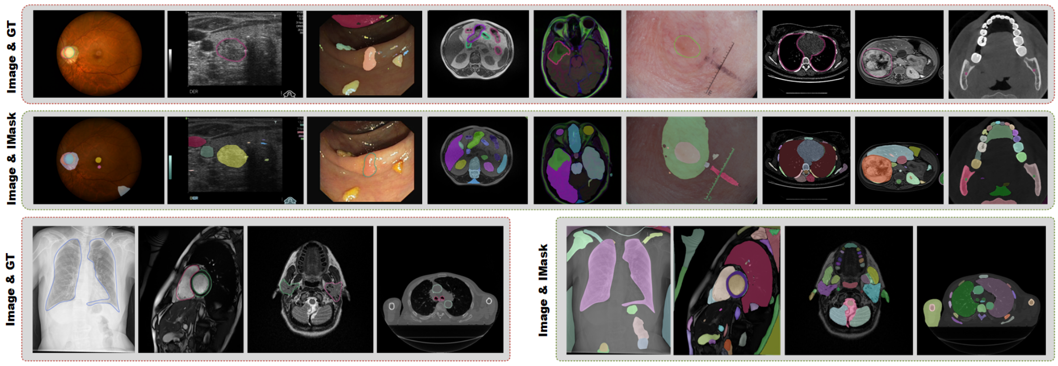

Figure 2: **IMed-361M:** A comprehensive dataset of multimodal medical images encompassing nearly all human organs and lesions, with interactive masks offering detailed, dense annotations.

Table 1: IMed-361M dataset contains six anatomical categories: A (Abdomen), S (Skeleton), H (Head & Neck), T (Thorax), P (Pelvis), and L (Lesions). The symbol $*$ indicates that the target has left and right parts.

A01: Adrenal gland $*$	H07: Optic chiasm	S07: Scapula $*$	L01: Lung infections
A02: Aorta	H08: Pituitary gland	S08: Cervical spine (C1-C7)	L02: Liver tumor
A03: Autochthonous muscles$*$	H09: Brain stem	S09: Lumbar spine (L1-L6)	L03: Kidney tumor
A04: Colon	H10: Temporal lobe $*$	S10: Thoracic spine (T1-T13)	L04: Kidney cyst
A05: Duodenum	H11: Parotid gland $*$	T01: Esophagus	L05: Pleural effusion
A06: Gallbladder	H12: Ear ($*$, Inner, Middle)	T02: Atrium $*$	L06: Myocardial edema
A07: Iliac artery $*$	H13: Temporomandibular $*$	T03: Myocardium $*$	L07: Myocardial scars
A08: Iliac vein $*$	H14: Mandible $*$	T04: Ventricle $*$	L08: Necrosis
A09: Iliopsoas $*$	H15: Thyroid gland	T05: Lower lobe $*$	L09: Edema
A10: Inferior vena cava	H16: Submandibular gland $*$	T06: Middle lobe $*$	L10: Non enhancing tumor
A11: Kidney $*$	H17: Oral cavity	T07: Upper lobe $*$	L11: Enhancing tumor
A12: Liver	H18: Eustachian tube $*$	T08: Pulmonary artery	L12: Necrotic tumor core
A13: Pancreas	H19: Hippocampus $*$	T09: Trachea	L13: Peritumoral edema
A14: Portal and splenic veins	H20: Mastoid $*$	T10: Lung $*$	L14: Myocardial infarction
A15: Small intestine	H21: Tympanic cavity $*$	T11: Heart	L15: No reflow
A16: Spleen	H22: Semicircular canal $*$	T12: Bronchus $*$	L16: Brain aneurysm
A17: Stomach	H23: Optic cup	T13: Breast $*$	L17: Neuroblastoma
A18: Spinal cord	H24: Optic disc	T14: Ascending aorta	L18: Prostate AFMS
A19: Rectum	H25: Larynx glottis	P01: Gluteus maximus $*$	L19: Hypoxic-ischemic
A20: Portal veins	H26: Larynx	P02: Gluteus medius $*$	L20: Breast tumor
A21: Large bowel	H27: Pharyngeal constrictor	P03: Gluteus minimus $*$	L21: Glioma
H01: Brain	S01: Clavicle $*$	P04: Bladder	L22: Thyroid nodule
H02: Face	S02: Femur $*$	P05: Prostate and uterus	L23: Skin lesion
H03: Airway	S03: Hip $*$	P06: Prostate	L24: Polyp
H04: Eye $*$	S04: Humerus $*$	P07: Testicle	P10: Prostatic urethra
H05: Crystalline lens $*$	S05: Rib (L&R, 1-12)	P08: Prostate peripheral zone	
H06: Optic nerve $*$	S06: Sacrum	P09: Prostate transition zone	

Division of Datasets. Tab.2 summarizes the datasets used for training and testing our model. Each dataset was divided into 90% for training and 10% for testing, with 3D datasets split along the volume dimension. To ensure evaluation reliability, we limited the test set of each dataset to a maximum of 3,000 images for final model assessment. Additionally, we evaluated the model's zero-shot capability using three external datasets: SegThor [1], TotalSegmentatorMRI [2], and ISLES [3]. Therefore, our training and test sets share the same data distribution, including modality and category.

Table 2: **Training and Test Datasets.** The following datasets were collected for training and validating IMIS-Net. Processed non-private datasets will be made publicly available for research purposes.

Dataset	Segmentation target	Modality	Category
SegRap2023 [3]	A18; H01,04-20,25-26; T01,09	CT	45
AbdomenAtlasMini1.0 [4]	A02,06,10-13,16-17	CT	9
AbdomenCT1K [5]	A11-13,16	CT	4
AMOS2022 [6]	A01-02,05-06,10-13,16-17; P04-05; T01	CT	15
BTCV [7]	A01-02,06,10-14,16-17; T01	CT	13
Colorectal_Liver_Metastases [8]	A12; L02	CT	2
Continuous_Registration_task1 [9]	T10	CT	1
COVID-19 CT scans [10, 11]	T10; L01	CT	1
CTSpine1K_Full [12]	S-08-10	CT	25
Finding-lungs-in-cTdata_3d [13]	T10	CT	1
FLARE21 [14]	A11-13,16	CT	4
FLARE22 [15]	A01-02,05-06,10-13,16-17; T01	CT	13
HCC-TACE-Seg [4]	A02,12,20; L02	CT	4
KiTS [16]	A11; L03	CT	2
KiTS2021 [17]	A11; L03-04	CT	3
KiTS2023 [18]	A11; L03-04	CT	3
Learn2Reg2022_AbdomenCTCT	A01-02,06,10-14,16-17; T01	CT	13
Learn2Reg2022_AbdomenMRCT	A11-12,16	CT	4
LITS [19]	A12; L02	CT	2
LUNA16 [20]	T10	CT	1
MMWHS [21, 22, 23]	T02-04,08,14	CT	7
MSD_Liver [24]	A12; L02	CT	2
MSD_Spleen [24]	A16	CT	1
PleThora	T10; L05	CT	2
Prostate-AnatomicaLEdge-Cases [24]	A19; S-02; P04,06	CT	5
SLIVER07 [25]	A12	CT	1
STACOM_SLAWT [26]	T02	CT	1
Totalsegmentator [2]	A01-17; H01-02; S-01-10; P01-04; T01-09	CT	104
VESSEL2012 [5]	T10	CT	1
Sz_cxr [27]	T10	X-Ray	1
WORD [28]	A01,04-06,11-13,15-17,19; S-02; P04; T01	CT	16
SegRap2023 [3]	A18; H01,04-20,25-26; T01,09	CTA	45
CMRxMotions	T03-04	MR-CMR	3
Myops2020 [29, 30]	T03-04; L06-07	MR	5
BraTS2013 [31, 32]	L08-11	MR-FLAIR	4
BraTS2015 [31, 32]	L08-11	MR-FLAIR	4
BraTS2018 [31, 33, 34]	L11-13	MR-FLAIR	3
BraTS2019 [31, 33, 34]	L11-13	MR-FLAIR	3
BraTS2020 [31, 33, 34]	L11-13	MR-FLAIR	3
BraTS2021 [33, 34, 35]	L11-13	MR-FLAIR	3
BraTS2023_GLI [6]	L08-09,11	MR-FLAIR	3
BraTS2023_MEN	L09-11	MR-FLAIR	3
BraTS2023_MET	L09-11	MR-FLAIR	3
BraTS2023_PED	L09-11	MR-FLAIR	3
BraTS2023_SSA	L09-11	MR-FLAIR	3
BraTS-TCGA-GBM [7]	L11-13	MR-FLAIR	3
BraTS-TCGA-LGG	L11-13	MR-FLAIR	3
SPPIN2023 [8]	L17	MR-TIGD	1
ATLAS2023	A12; L02	MR-T1W	2
CHAOS_Task_4 [36, 37, 38]	A11-12,16	MR-T1W	4
Learn2Reg2022_AbdomenMRCT	A11-12,16	MR-T1W	4
MSD_Prostate [24]	P08-09	MR-T2W	2
Myops2020 [29, 30]	T03-04,06-07	MR-T2W	5
CHAOS_Task_4 [36, 37, 38]	A11-12,16	MR-T2W	4
ISBI-MR-Prostate-2013 [9]	P06,08	MR-T2W	2
Prostate_MRI [24]	P06	MR-T2W	1

Continued on next page

[3] https://www.isles-challenge.org/
[4] https://www.cancerimagingarchive.net/collection/hcc-tace-seg/
[5] https://zenodo.org/records/8055066
[6] https://www.synapse.org/Synapse:syn51156910/wiki/621282
[7] https://www.cancerimagingarchive.net/analysis-result/brats-tcga-gbm/
[8] https://github.com/myrthebuser/SPPIN2023
[9] https://www.cancerimagingarchive.net/analysis-result/isbi-mr-prostate-2013/

Dataset	Segmentation target	Modality	Category
Table 2 : Continued from previous page			
PROSTATEx-Seg-HiRes [10]	P06	MR-T2W	1
PROSTATEx-Seg-Zones	P08-10; L18	MR-T2W	4
u-RegPro [11]	P06	MR-T2W	1
ADAM2020	L16	MR-TOF	2
ACDC [39]	T03-04	MR	3
AMOS2022 [6]	A01-02,05-06,10-13,16-17; P04-05; T01	MR	15
EMIDEC [40]	T03-04; L14-15	MR	4
Heart_Seg_MRI [41]	T02	MR	1
MMWHS [21, 22, 23]	T02-04,08,14	MR	7
Mnms2	T03-04	MR	3
MSD_Heart [42]	T02	MR	1
PROMISE12 [43]	P06	MR	1
CETUS2014 [12]	T04	US	1
CuRIOUS2022_tumor	L19	US	1
TDSC-ABUS2023 [13]	L20	US	1
u-RegPro	P06	US	1
BraimMRI [44]	L21	MR	1
Brain-MRI	L21	MR	1
UW-Madison	A15,17,21	MR	3
DDTI	L22	US	1
Drishti-GS_Cup [45, 46]	H23	Fundus	1
Drishti-GS_Od [45, 46]	H24	Fundus	1
Gamma [47, 48, 49]	H23-24	Fundus	2
ADAM []	H24	Fundus	1
PAPILA [50]	H23-24	Fundus	2
Refuge2 [51, 52]	H23-24	Fundus	2
Rim-one-DL [53]	H23-24	Fundus	2
Finding-lungs-in-cTdata_2d [13]	T10	CT	1
Isic2016_task1 [14]	L23	Dermoscopy	1
Isic2017_task1	L23	Dermoscopy	1
Isic2018_task1	L23	Dermoscopy	1
Ph2 [54]	L23	Dermoscopy	1
Cvc_clinicdb [55]	L24	Endoscopy	1
Endovis15 [56]	L24	Endoscopy	1
Hyper-kvasir-segmented-images	L24	Endoscopy	1
Kvasir_seg [57]	L24	Endoscopy	1
Kvasir_seg_aliyun	L24	Endoscopy	1
Kvasircapsule_seg [58]	L24	Endoscopy	1
Sun_seg	L24	Endoscopy	1
Private1	A02; T02,04	CTA	3
Private2	A19; S-02; P04	CT	4
Private3	A01-17; H01-02; S-01-10; P01-04; T01-09	CT	104
Private4	A18; H09-10	CT	4
Private5	A01-17; H01; S-01-10; P01-04; T01-09	CT	104
Private6	H01	CT	1
Private7	H11, 13-14	CT	10
Private8	H15-17,25,27	CT	6
Private9	T01,09-11,18	CT	6
Private10	A11-12,16-17	CT	5
Private11	A01-17; H01-02; S-01-10; P01-04; T01-09	CT	104
Private12	A19; S-02; P04,06,07	CT	6
Private13	T13	CT	2
Private14	T02,04,09,12	CT	7
Private15	A01-19; H01-14; S-01-10; P01-04; T01-09	CT	130

Table 3: Access links to the raw data in the IMed-361M dataset, where "✔" indicates that we are able to distribute the processed data to researchers for further analysis.

Dataset Name	Official Link	Release
SegRap2023 [3]	https://segrap2023.grand-challenge.org/	✔
AbdomenAtlasMini1.0 [4]	https://huggingface.co/datasets/AbdomenAtlas/AbdomenAtlas1.0MiniBeta	—
AbdomenCT1K [5]	https://abdomenct-1k-fully-supervised-learning.grand-challenge.org/	✔
AMOS2022 [6]	https://amos22.grand-challenge.org/	✔
BTCV [7]	https://www.synapse.org/#!Synapse:syn3193805/wiki/217752	✔
Colorectal_Liver_Metastases [8]	https://www.cancerimagingarchive.net/collection/colorectal-liver-metastases/	✔
Continuous_Registration_task1 [9]	https://continuousregistration.grand-challenge.org/rules/	✔
COVID-19 CT scans [10, 11]	https://tianchi.aliyun.com/dataset/dataDetail?dataId=90014	✔

[10]https://www.cancerimagingarchive.net/analysis-result/prostatex-seg-hires/

[11]https://muregpro.github.io/

[12]https://www.creatis.insa-lyon.fr/Challenge/CETUS/

[13]https://tdsc-abus2023.grand-challenge.org/

[14]https://challenge.isic-archive.com/data/

Dataset Name	Official Link	Release
CTSpine1K_Full [12]	https://github.com/ICT-MIRACLE-lab/CTSpine1K	✔
Finding-lungs-in-cTdata_3d [13]	https://www.kaggle.com/datasets/kmader/finding-lungs-in-ct-data	✔
FLARE21 [14]	https://flare.grand-challenge.org/FLARE21/	✔
FLARE22 [15]	https://flare22.grand-challenge.org/	✔
HCC-TACE-Seg	https://www.cancerimagingarchive.net/collection/hcc-tace-seg/	✔
KiTS [16]	https://kits19.grand-challenge.org/home/	✔
KiTS2021 [17]	https://kits21.kits-challenge.org/	✔
KiTS2023 [18]	https://kits-challenge.org/kits23/	✔
Learn2Reg2022_AbdomenCTCT	https://paperswithcode.com/dataset/learn2reg	✔
Learn2Reg2022_AbdomenMRCT	https://paperswithcode.com/dataset/learn2reg	✔
LITS [19]	https://www.kaggle.com/datasets/andrewmvd/liver-tumor-segmentation	—
LUNA16 [20]	https://luna16.grand-challenge.org/	✔
MMWHS [21, 22, 23]	https://zmiclab.github.io/zxh/0/mmwhs/	✔
MSD_Liver [24]	http://medicaldecathlon.com/	✔
MSD_Spleen [24]	http://medicaldecathlon.com/	✔
PleThora	https://www.cancerimagingarchive.net/analysis-result/plethora/	✔
Prostate-AnatomicaLEdge-Cases [24]	https://www.cancerimagingarchive.net/collection/prostate-anatomical-edge-cases/	✔
SLIVER07 [25]	https://sliver07.grand-challenge.org/	—
STACOM_SLAWT [26]	https://www.doc.ic.ac.uk/~rkarim/la_lv_framework/wall/datasets.html	—
Totalsegmentator [2]	https://totalsegmentator.com/	✔
VESSEL2012 [15]	https://vessel12.grand-challenge.org/	✔
Sz_cxr [27]	https://www.kaggle.com/datasets/yoctoman/shcxr-lung-mask	✔
WORD [28]	https://github.com/hilab-git/word	✔
CMRxMotions	http://cmr.miccai.cloud/	✔
Myops2020 [29, 30]	https://zmiclab.github.io/zxh/0/myops20/	✔
BraTS2013 [31, 32]	https://www.smir.ch/BRATS/Start2013	✔
BraTS2015 [31, 32]	https://www.smir.ch/BRATS/Start2015	✔
BraTS2018 [31, 33, 34]	https://www.med.upenn.edu/sbia/brats2018/data.html	✔
BraTS2019 [31, 33, 34]	https://www.med.upenn.edu/cbica/brats-2019/	✔
BraTS2020 [31, 33, 34]	https://www.med.upenn.edu/cbica/brats2020/data.html	✔
BraTS2021 [33, 34, 35]	https://www.med.upenn.edu/cbica/brats2021/	✔
BraTS2023_GLI	https://www.synapse.org/Synapse:syn51156910/wiki/621282	✔
BraTS2023_MEN	https://www.synapse.org/Synapse:syn51514106	✔
BraTS2023_MET	https://www.synapse.org/Synapse:syn51514107	✔
BraTS2023_PED	https://www.synapse.org/Synapse:syn51514108	✔
BraTS2023_SSA	https://www.synapse.org/Synapse:syn51514109	✔
BraTS-TCGA-GBM	https://www.cancerimagingarchive.net/analysis-result/brats-tcga-gbm/	✔
BraTS-TCGA-LGG	https://www.cancerimagingarchive.net/analysis-result/brats-tcga-lgg/	✔
SPPIN2023	https://github.com/myrthebuser/SPPIN2023	✔
ATLAS2023	https://atlas-challenge.u-bourgogne.fr	✔
CHAOS_Task_4 [36, 37, 38]	https://chaos.grand-challenge.org/Publications/	✔
Learn2Reg2022_AbdomenMRCT	https://paperswithcode.com/datasets	✔
MSD_Prostate [24]	http://medicaldecathlon.com/	✔
CHAOS_Task_4 [36, 37, 38]	https://chaos.grand-challenge.org/Publications/	✔
ISBI-MR-Prostate-2013	https://www.cancerimagingarchive.net/analysis-result/isbi-mr-prostate-2013/	✔
Prostate_MRI [24]	https://www.cancerimagingarchive.net/collection/prostate-mri/	✔
PROSTATEx-Seg-HiRes	https://www.cancerimagingarchive.net/analysis-result/prostatex-seg-hires/	✔
PROSTATEx-Seg-Zones	https://www.cancerimagingarchive.net/collection/prostatex/	✔
u-RegPro	https://muregpro.github.io/	✔
ADAM2020	https://amd.grand-challenge.org/	✔
ACDC [39]	https://www.creatis.insa-lyon.fr/Challenge/acdc/	✔
AMOS2022 [6]	https://amos22.grand-challenge.org/	✔
EMIDEC [40]	http://emidec.com/	✔
Heart_Seg_MRI [41]	https://tianchi.aliyun.com/dataset/dataDetail?dataId=90148	✔
MMWHS [21, 22, 23]	https://zmiclab.github.io/zxh/0/mmwhs/	✔
Mnms2	https://www.ub.edu/mnms-2/	✔
MSD_Heart [42]	http://medicaldecathlon.com/	✔
PROMISE12 [43]	https://promise12.grand-challenge.org/	✔
CETUS2014	https://www.creatis.insa-lyon.fr/Challenge/CETUS/	✔
CuRIOUS2022_tumor	https://curious2022.grand-challenge.org/curious2022/	—
TDSC-ABUS2023	https://tdsc-abus2023.grand-challenge.org/	✔
BraimMRI [44]	https://tianchi.aliyun.com/dataset/127459	✔
Brain-MRI	https://tianchi.aliyun.com/dataset/127583	✔
UW-Madison	https://www.kaggle.com/competitions/uw-madison-gi-tract-image-segmentation/data	—
DDTI	https://www.kaggle.com/datasets/dasmehdixtr/ddti-thyroid-ultrasound-images	—
Drishti-GS_Cup [45, 46]	https://www.kaggle.com/datasets/lokeshsaipureddi/drishtigs-retina-dataset-for-onh-segmentation	✔
Drishti-GS_Od [45, 46]	https://www.kaggle.com/datasets/lokeshsaipureddi/drishtigs-retina-dataset-for-onh-segmentation	✔
Gamma [47, 48, 49]	https://gamma.grand-challenge.org/	✔
ADAM []	https://amd.grand-challenge.org/	✔
PAPILA [50]	https://figshare.com/articles/dataset/PAPILA/14798004/1?file=28454352	✔
Refuge2 [51, 52]	https://refuge.grand-challenge.org/	—
Rim-one-DL [53]	https://github.com/miag-ull/rim-one-dl	✔
Finding-lungs-in-cTdata_2d [13]	https://www.kaggle.com/datasets/kmader/finding-lungs-in-ct-data	✔
Isic2016_task1	https://challenge.isic-archive.com/landing/2016/	✔
Isic2017_task1	https://challenge.isic-archive.com/landing/2017/	✔
Isic2018_task1	https://workshop2018.isic-archive.com/	✔
Ph2 [54]	https://www.fc.up.pt/addi/ph2%20database.html	—
Cvc_clinicdb [55]	https://tianchi.aliyun.com/dataset/dataDetail?dataId=93690	✔
Endovis15 [56]	https://polyp.grand-challenge.org/	✔
Hyper-kvasir-segmented-images [16]	https://datasets.simula.no/hyper-kvasir/	✔
Kvasir_seg [57]	https://opendatalab.com/Kvasir-Sessile_dataset	✔
Kvasir_seg_aliyun	https://tianchi.aliyun.com/dataset/84385/notebook	✔
Kvasircapsule_seg [58]	https://www.kaggle.com/datasets/debeshjha1/kvasircapsuleseg	✔
Sun_seg	https://github.com/GewelsJI/VPS/blob/main/docs/DATA_PREPARATION.md	✔

[15] https://zenodo.org/records/8055066

[16] https://www.nature.com/articles/s41597-020-00622-y

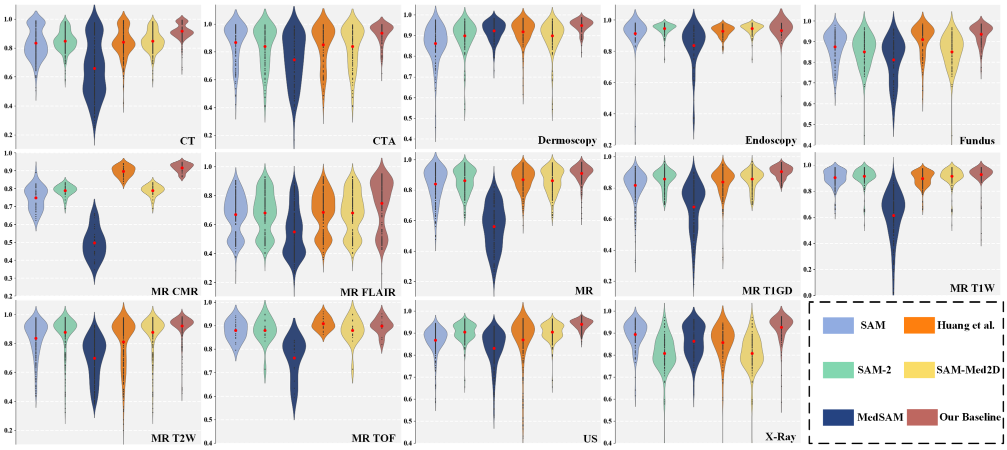

Figure 3: Comparison of segmentation performance of different methods in 14 medical image modalities, where the red points represent the means.

Appendix C: Additional Experimental Analysis

Comparative experiments with other models on different modalities. Fig.3 shows the segmentation performance across 14 medical image modalities, with the red dots representing the average values. The methods proposed by MedSAM and Huang et al. utilize bounding box prompts, while the other models are evaluated based on better performance achieved from either bounding box prompts or three-click inputs. It can be observed that for medical modalities similar to natural images (e.g., dermoscopy and endoscopy), the performance of SAM and SAM-2 is comparable to the fine-tuned models, which validates the effectiveness of large-scale pretraining data. Additionally, our baseline model performs excellently across 12 modalities, with Dice scores exceeding 90%, and shows significant stability in modalities such as CTA, dermoscopy, ultrasound, and X-ray. These results suggest that directly using SAM or SAM-2 as a solution for medical image modalities with similar characteristics to natural images is feasible. However, for modalities that significantly differ from natural images, fine-tuning the base model can significantly improve segmentation performance.

Visualization of interactive segmentation. We assess interactive segmentation performance using Dice scores. The minimum enclosing bounding box of the ground truth serves as the model's prompt input. Fig.4 presents the prediction results generated by different models based on a single bounding box prompt. Due to the detailed spatial information provided by the bounding box, the Dice scores of various models are mostly above 0.8. In practical applications, this singular interactive approach may not directly meet user needs; hence, the IMIS method supports correcting predictions by providing additional click interactions. Our IMIS-Net still achieves high Dice segmentation. As shown in Fig.5 and Fig.6, we visualize the results of 5 simulated interactive experiments for SAM, SAM-2, and IMIS-Net.

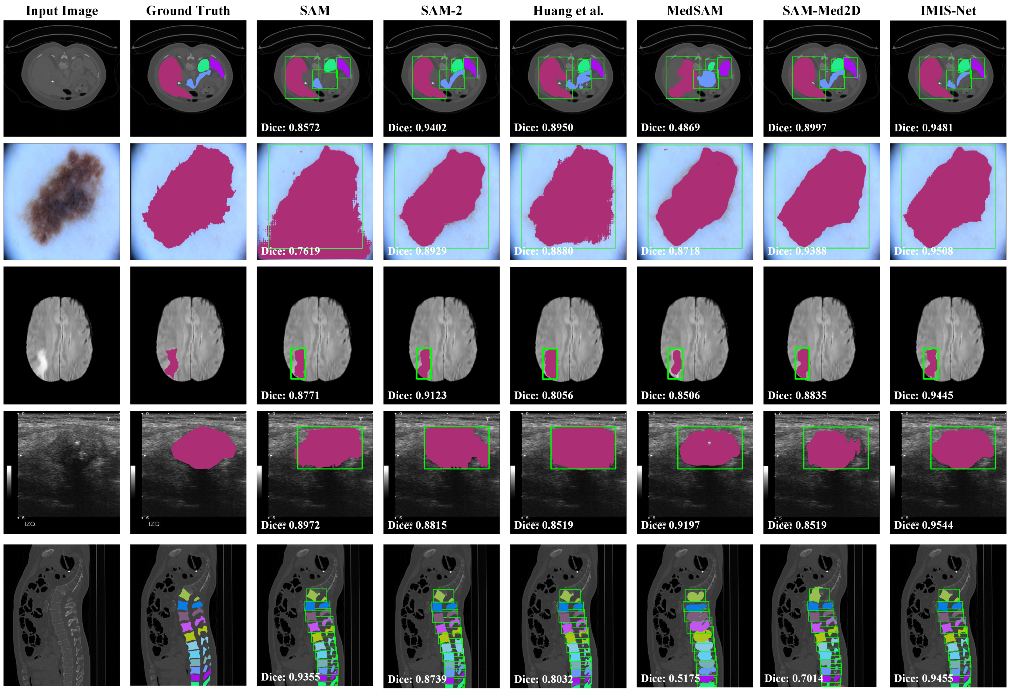

Figure 4: Simulate interactive segmentation results with identical bounding box coordinates for all model inputs.

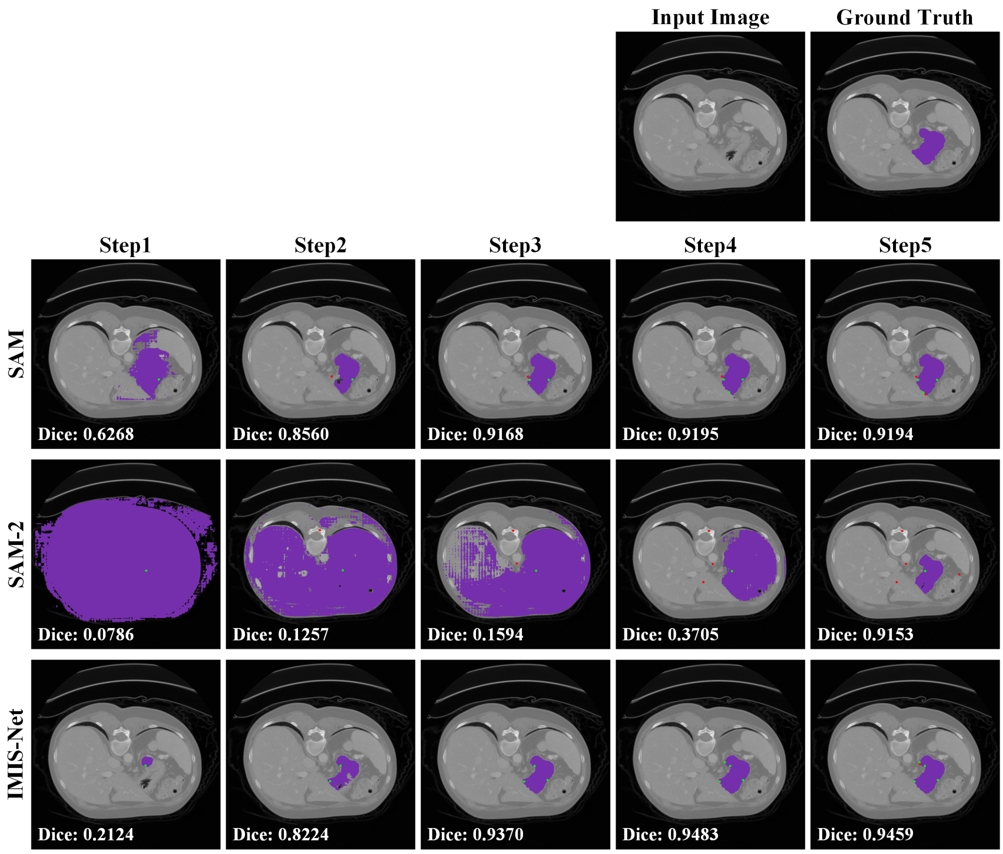

Figure 5: An interactive segmentation example of the stomach in CT images. SAM and SAM-2 typically require more prompts to achieve better results, while IMIS-Net achieves comparable performance with fewer interactions.

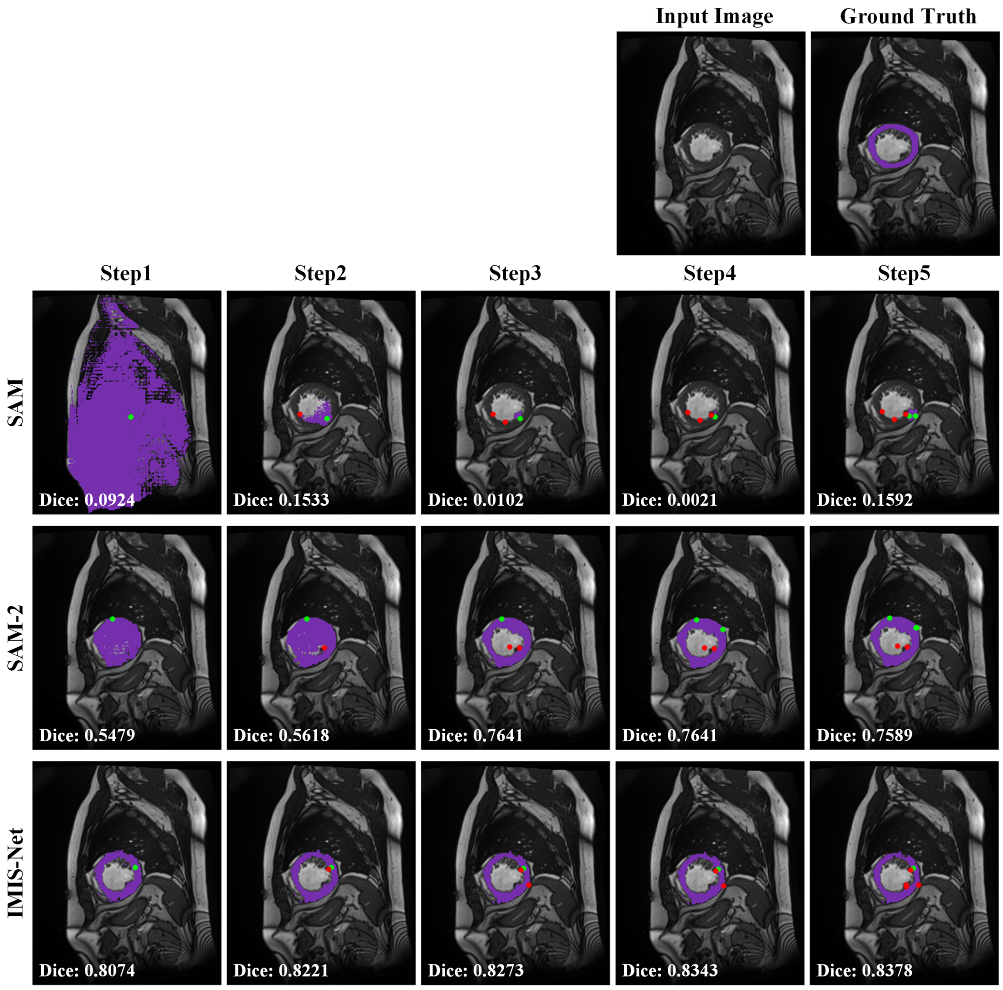

Figure 6: An interactive segmentation example of the cardiac myocardium in MR images. SAM performs poorly when dealing with annular myocardium, while SAM-2 and IMIS-Net are able to obtain predictions of the target area through multiple interactions. Our network consistently outperforms other methods.

References

[1] Zoé Lambert, Caroline Petitjean, Bernard Dubray, and Su Kuan. Segthor: Segmentation of thoracic organs at risk in ct images. In 2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA), pages 1–6. IEEE, 2020.

[2] Jakob Wasserthal, Hanns-Christian Breit, Manfred T. Meyer, Maurice Pradella, Daniel Hinck, Alexander W. Sauter, Tobias Heye, Daniel T. Boll, Joshy Cyriac, Shan Yang, Michael Bach, and Martin Segeroth. TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images. Radiology: Artificial Intelligence, 5(5), sep 2023.

[3] Xiangde Luo, Jia Fu, Yunxin Zhong, Shuolin Liu, Bing Han, Mehdi Astaraki, Simone Bendazzoli, Iuliana Toma-Dasu, Yiwen Ye, Ziyang Chen, et al. Segrap2023: A benchmark of organs-at-risk and gross tumor volume segmentation for radiotherapy planning of nasopharyngeal carcinoma. arXiv preprint arXiv:2312.09576, 2023.

[4] Chongyu Qu, Tiezheng Zhang, Hualin Qiao, Yucheng Tang, Alan L Yuille, Zongwei Zhou, et al. Abdomenatlas-8k: Annotating 8,000 ct volumes for multi-organ segmentation in three weeks. Advances in Neural Information Processing Systems, 36, 2024.

[5] Jun Ma, Yao Zhang, Song Gu, Cheng Zhu, Cheng Ge, Yichi Zhang, Xingle An, Congcong Wang, Qiyuan Wang, Xin Liu, et al. Abdomenct-1k: Is abdominal organ segmentation a solved problem? IEEE Transactions on Pattern Analysis and Machine Intelligence, 44(10):6695–6714, 2021.

[6] Yuanfeng Ji, Haotian Bai, Chongjian Ge, Jie Yang, Ye Zhu, Ruimao Zhang, Zhen Li, Lingyan Zhanng, Wanling Ma, Xiang Wan, et al. Amos: A large-scale abdominal multi-organ benchmark for versatile medical image segmentation. Advances in Neural Information Processing Systems, 35:36722–36732, 2022.

[7] Bennett Landman, Zhoubing Xu, J Igelsias, Martin Styner, T Langerak, and Arno Klein. Miccai multi-atlas labeling beyond the cranial vault–workshop and challenge. In Proc. MICCAI Multi-Atlas Labeling Beyond Cranial Vault—Workshop Challenge, volume 5, page 12, 2015.

[8] Jack Martin, Angelica Petrillo, Elizabeth C Smyth, Nadeem Shaida, Samir Khwaja, HK Cheow, Adam Duckworth, Paula Heister, Raaj Praseedom, Asif Jah, et al. Colorectal liver metastases: Current management and future perspectives. World Journal of Clinical Oncology, 11(10):761, 2020.

[9] Kasper Marstal, F Berendsen, Niels Dekker, Marius Staring, and Stefan Klein. The continuous registration challenge: evaluation-as-a-service for medical image registration algorithms. In 2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019), pages 1399–1402. IEEE, 2019.

[10] O Paiva. Helping radiologists to help people in more than 100 countries! coronavirus cases. CORONACASES. ORG, Ed., ed, 2020.

[11] Ma Jun, Ge Cheng, Wang Yixin, An Xingle, Gao Jiantao, Yu Ziqi, Zhang Minqing, Liu Xin, Deng Xueyuan, Cao Shucheng, et al. Covid-19 ct lung and infection segmentation dataset. 2020.

[12] Yang Deng, Ce Wang, Yuan Hui, Qian Li, Jun Li, Shiwei Luo, Mengke Sun, Quan Quan, Shuxin Yang, You Hao, et al. Ctspine1k: A large-scale dataset for spinal vertebrae segmentation in computed tomography. arXiv preprint arXiv:2105.14711, 2021.

[13] Hugo J Kuijf, J Matthijs Biesbroek, Jeroen De Bresser, Rutger Heinen, Simon Andermatt, Mariana Bento, Matt Berseth, Mikhail Belyaev, M Jorge Cardoso, Adria Casamitjana, et al. Standardized assessment of automatic segmentation of white matter hyperintensities and results of the wmh segmentation challenge. IEEE transactions on medical imaging, 38(11):2556–2568, 2019.

[14] Jun Ma, Yao Zhang, Song Gu, Xingle An, Zhihe Wang, Cheng Ge, Congcong Wang, Fan Zhang, and Yu Wang. Fast and low-gpu-memory abdomen ct organ segmentation: The flare challenge. Medical Image Analysis, 82:102616, 2022.

[15] Jun Ma, Yao Zhang, Song Gu, Cheng Ge, Shihao Ma, Adamo Young, Cheng Zhu, Kangkang Meng, Xin Yang, Ziyan Huang, Fan Zhang, Wentao Liu, and YuanKe Pan. Unleashing the strengths of unlabeled data in pan-cancer abdominal organ quantification: the flare22 challenge. arXiv preprint arXiv:2308.05862, 2023.

[16] Nicholas Heller, Niranjan Sathianathen, Arveen Kalapara, Edward Walczak, Keenan Moore, Heather Kaluzniak, Joel Rosenberg, Paul Blake, Zachary Rengel, Makinna Oestreich, et al. The kits19 challenge data: 300 kidney tumor cases with clinical context, ct semantic segmentations, and surgical outcomes. arXiv preprint arXiv:1904.00445, 2019.

[17] Zhongchen Zhao, Huai Chen, and Lisheng Wang. A coarse-to-fine framework for the 2021 kidney and kidney tumor segmentation challenge. In International Challenge on Kidney and Kidney Tumor Segmentation, pages 53–58. Springer, 2021.

[18] Andriy Myronenko, Dong Yang, Yufan He, and Daguang Xu. Automated 3d segmentation of kidneys and tumors in miccai kits 2023 challenge. In International Challenge on Kidney and Kidney Tumor Segmentation, pages 1–7. Springer, 2023.

[19] Patrick Bilic, Patrick Ferdinand Christ, Eugene Vorontsov, Grzegorz Chlebus, Hao Chen, Qi Dou, Chi-Wing Fu, Xiao Han, and Pheng-Ann Heng. The liver tumor segmentation benchmark (lits), 2019.

[20] Arnaud Arindra Adiyoso Setio, Alberto Traverso, Thomas De Bel, Moira SN Berens, Cas Van Den Bogaard, Piergiorgio Cerello, Hao Chen, Qi Dou, Maria Evelina Fantacci, Bram Geurts, et al. Validation, comparison, and combination of algorithms for automatic detection of pulmonary nodules in computed tomography images: the luna16 challenge. Medical image analysis, 42:1–13, 2017.

[21] Xiahai Zhuang. Multivariate mixture model for myocardial segmentation combining multi-source images. IEEE transactions on pattern analysis and machine intelligence, 41(12):2933–2946, 2018.

[22] Xiahai Zhuang and Juan Shen. Multi-scale patch and multi-modality atlases for whole heart segmentation of mri. Medical image analysis, 31:77–87, 2016.

[23] Xinzhe Luo and Xiahai Zhuang. x-metric: An n-dimensional information-theoretic framework for groupwise registration and deep combined computing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022.

[24] Michela Antonelli, Annika Reinke, Spyridon Bakas, Keyvan Farahani, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, Olaf Ronneberger, Ronald M Summers, et al. The medical segmentation decathlon. Nature communications, 13(1):4128, 2022.

[25] Tobias Heimann, Bram Van Ginneken, Martin A Styner, Yulia Arzhaeva, Volker Aurich, Christian Bauer, Andreas Beck, Christoph Becker, Reinhard Beichel, György Bekes, et al. Comparison and evaluation of methods for liver segmentation from ct datasets. IEEE transactions on medical imaging, 28(8):1251–1265, 2009.

[26] Rashed Karim, Lauren-Emma Blake, Jiro Inoue, Qian Tao, Shuman Jia, R James Housden, Pranav Bhagirath, Jean-Luc Duval, Marta Varela, Jonathan M Behar, et al. Algorithms for left atrial wall segmentation and thickness–evaluation on an open-source ct and mri image database. Medical image analysis, 50:36–53, 2018.

[27] Sergii Stirenko, Yuriy Kochura, Oleg Alienin, Oleksandr Rokovyi, Yuri Gordienko, Peng Gang, and Wei Zeng. Chest x-ray analysis of tuberculosis by deep learning with segmentation and augmentation. In 2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO), pages 422–428, 2018.

[28] Xiangde Luo, Wenjun Liao, Jianghong Xiao, Jieneng Chen, Tao Song, Xiaofan Zhang, Kang Li, Dimitris N Metaxas, Guotai Wang, and Shaoting Zhang. Word: A large scale dataset, benchmark and clinical applicable study for abdominal organ segmentation from ct image. Medical Image Analysis, 82:102642, 2022.

[29] Xinzhe Luo and Xiahai Zhuang. X-metric: An n-dimensional information-theoretic framework for groupwise registration and deep combined computing. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022.

[30] Junyi Qiu, Lei Li, Sihan Wang, Ke Zhang, Yinyin Chen, Shan Yang, and Xiahai Zhuang. Myops-net: Myocardial pathology segmentation with flexible combination of multi-sequence cmr images. Medical Image Analysis, 84:102694, 2023.

[31] Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). IEEE transactions on medical imaging, 34(10):1993–2024, 2014.

[32] Michael Kistler, Serena Bonaretti, Marcel Pfahrer, Roman Niklaus, and Philippe Büchler. The virtual skeleton database: An open access repository for biomedical research and collaboration. J Med Internet Res, 15(11):e245, Nov 2013.

[33] Spyridon Bakas, Hamed Akbari, Aristeidis Sotiras, Michel Bilello, Martin Rozycki, Justin S Kirby, John B Freymann, Keyvan Farahani, and Christos Davatzikos. Advancing the cancer genome atlas glioma mri collections with expert segmentation labels and radiomic features. Scientific data, 4(1):1–13, 2017.

[34] Spyridon Bakas, Mauricio Reyes, Andras Jakab, Stefan Bauer, Markus Rempfler, Alessandro Crimi, Russell Takeshi Shinohara, and Christoph Berger. Identifying the best machine learning algorithms for brain tumor segmentation, progression assessment, and overall survival prediction in the brats challenge, 2019.

[35] Ujjwal Baid, Satyam Ghodasara, Suyash Mohan, Michel Bilello, Evan Calabrese, Errol Colak, Keyvan Farahani, Jayashree Kalpathy-Cramer, and Felipe C. Kitamura. The rsna-asnr-miccai brats 2021 benchmark on brain tumor segmentation and radiogenomic classification, 2021.

[36] A. Emre Kavur, N. Sinem Gezer, Mustafa Barış, Sinem Aslan, Pierre-Henri Conze, Vladimir Groza, Duc Duy Pham, Soumick Chatterjee, and Philipp Ernst. CHAOS Challenge - combined (CT-MR) healthy abdominal organ segmentation. Medical Image Analysis, 69:101950, April 2021.

[37] Ali Emre Kavur, M. Alper Selver, Oğuz Dicle, Mustafa Barış, and N. Sinem Gezer. Chaos - combined (ct-mr) healthy abdominal organ segmentation challenge data. April 2019.

[38] A. Emre Kavur, Naciye Sinem Gezer, Mustafa Barış, Yusuf Şahin, Savaş Özkan, Bora Baydar, Ulaş Yüksel, Çağlar Kılıkçıer, Şahin Olut, Gözde Bozdağı Akar, Gözde Ünal, Oğuz Dicle, and M. Alper Selver. Comparison of semi-automatic and deep learning based automatic methods for liver segmentation in living liver transplant donors. Diagnostic and Interventional Radiology, 26:11–21, January 2020.

[39] Olivier Bernard, Alain Lalande, Clement Zotti, Frederick Cervenansky, Xin Yang, Pheng-Ann Heng, Irem Cetin, Karim Lekadir, Oscar Camara, Miguel Angel Gonzalez Ballester, and Gerard Sanroma. Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: Is the problem solved? IEEE Transactions on Medical Imaging, 37(11):2514–2525, 2018.

[40] Alain Lalande, Zhihao Chen, Thibaut Pommier, Thomas Decourselle, Abdul Qayyum, Michel Salomon, Dominique Ginhac, Youssef Skandarani, Arnaud Boucher, Khawla Brahim, et al. Deep learning methods for automatic evaluation of delayed enhancement-mri. the results of the emidec challenge. Medical Image Analysis, 79:102428, 2022.

[41] Catalina Tobon-Gomez, Arjan J Geers, Jochen Peters, Jürgen Weese, Karen Pinto, Rashed Karim, Mohammed Ammar, Abdelaziz Daoudi, Jan Margeta, Zulma Sandoval, et al. Benchmark for algorithms segmenting the left atrium from 3d ct and mri datasets. IEEE transactions on medical imaging, 34(7):1460–1473, 2015.

[42] Amber L Simpson, Michela Antonelli, Spyridon Bakas, Michel Bilello, Keyvan Farahani, Bram Van Ginneken, Annette Kopp-Schneider, Bennett A Landman, Geert Litjens, Bjoern Menze, et al. A large annotated medical image dataset for the development and evaluation of segmentation algorithms. arXiv preprint arXiv:1902.09063, 2019.

[43] Geert Litjens, Robert Toth, Wendy Van De Ven, Caroline Hoeks, Sjoerd Kerkstra, Bram Van Ginneken, Graham Vincent, Gwenael Guillard, Neil Birbeck, Jindang Zhang, et al. Evaluation of prostate segmentation algorithms for mri: the promise12 challenge. Medical image analysis, 18(2):359–373, 2014.

[44] Aliyun Tianchi. braimmri dataset. https://tianchi.aliyun.com/dataset/127459, 2022.

[45] Jayanthi Sivaswamy, S Krishnadas, Arunava Chakravarty, G Joshi, A Syed Tabish, et al. A comprehensive retinal image dataset for the assessment of glaucoma from the optic nerve head analysis. JSM Biomedical Imaging Data Papers, 2(1):1004, 2015.

[46] Jayanthi Sivaswamy, SR Krishnadas, Gopal Datt Joshi, Madhulika Jain, and A Ujjwaft Syed Tabish. Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation. In 2014 IEEE 11th international symposium on biomedical imaging (ISBI), pages 53–56. IEEE, 2014.

[47] Huazhu Fu, Jun Cheng, Yanwu Xu, Damon Wing Kee Wong, Jiang Liu, and Xiaochun Cao. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. IEEE Transactions on Medical Imaging, 37(7):1597–1605, 2018.

[48] José Ignacio Orlando, Huazhu Fu, João Barbosa Breda, Karel Van Keer, Deepti R Bathula, Andrés Diaz-Pinto, Ruogu Fang, Pheng-Ann Heng, Jeyoung Kim, JoonHo Lee, et al. Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. Medical image analysis, 59:101570, 2020.

[49] Huazhu Fu, Fei Li, Xu Sun, Xingxing Cao, Jingan Liao, Jose Ignacio Orlando, Xing Tao, Yuexiang Li, Shihao Zhang, Mingkui Tan, et al. Age challenge: angle closure glaucoma evaluation in anterior segment optical coherence tomography. Medical Image Analysis, 66:101798, 2020.

[50] Oleksandr Kovalyk, Juan Morales-Sánchez, Rafael Verdú-Monedero, Inmaculada Sellés-Navarro, Ana Palazón-Cabanes, and José-Luis Sancho-Gómez. Papila: Dataset with fundus images and clinical data of both eyes of the same patient for glaucoma assessment. Scientific Data, 9(1):291, 2022.

[51] José Ignacio Orlando, Huazhu Fu, João Barbosa Breda, Karel van Keer, Deepti R. Bathula, Andrés Diaz-Pinto, Ruogu Fang, Pheng-Ann Heng, Jeyoung Kim, and JoonHo Lee. Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. Medical Image Analysis, 59:101570, 2020.

[52] Fei Li, Diping Song, Han Chen, Jian Xiong, Xingyi Li, Hua Zhong, Guangxian Tang, Sujie Fan, Dennis SC Lam, Weihua Pan, et al. Development and clinical deployment of a smartphone-based visual field deep learning system for glaucoma detection. NPJ digital medicine, 3(1):123, 2020.

[53] Francisco José Fumero Batista, Tinguaro Diaz-Aleman, Jose Sigut, Silvia Alayon, Rafael Arnay, and Denisse Angel-Pereira. Rim-one dl: A unified retinal image database for assessing glaucoma using deep learning. Image Analysis and Stereology, 39(3):161–167, 2020.

[54] Teresa Mendonça, Pedro M Ferreira, Jorge S Marques, André RS Marcal, and Jorge Rozeira. Ph 2-a dermoscopic image database for research and benchmarking. In 2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC), pages 5437–5440. IEEE, 2013.

[55] Jorge Bernal, F Javier Sánchez, Gloria Fernández-Esparrach, Debora Gil, Cristina Rodríguez, and Fernando Vilariño. Wm-dova maps for accurate polyp highlighting in colonoscopy: Validation vs. saliency maps from physicians. Computerized medical imaging and graphics, 43:99–111, 2015.

[56] Jorge Bernal, Nima Tajkbaksh, Francisco Javier Sanchez, Bogdan J Matuszewski, Hao Chen, Lequan Yu, Quentin Angermann, Olivier Romain, Bjørn Rustad, Ilangko Balasingham, et al. Comparative validation of polyp detection methods in video colonoscopy: results from the miccai 2015 endoscopic vision challenge. IEEE transactions on medical imaging, 36(6):1231–1249, 2017.

[57] Debesh Jha, Pia H Smedsrud, Michael A Riegler, Pål Halvorsen, Thomas de Lange, Dag Johansen, and Håvard D Johansen. Kvasir-seg: A segmented polyp dataset. In International Conference on Multimedia Modeling, pages 451–462. Springer, 2020.

[58] Debesh Jha, Nikhil Kumar Tomar, Sharib Ali, Michael A. Riegler, Håvard D. Johansen, Dag Johansen, Thomas de Lange, and Pål Halvorsen. Nanonet: Real-time polyp segmentation in video capsule endoscopy and colonoscopy, 2021.