MedLSAM: Localize and Segment Anything Model for 3D Medical Images

Updated on 2023.06.28

Key Features

This repository provides the official implementation of MedLSAM: Localize and Segment Anything Model for 3D Medical Images

- **First Fully-Automatic Medical Adaptation of SAM**: MedLSAM is the first complete medical adaptation of the Segment Anything Model (SAM). The primary goal of this work is to significantly reduce the annotation workload in medical image segmentation.
- **Foundation Model for 3D Medical Image Localization**: MedLAM: MedLSAM introduces MedLAM as a foundational model for the localization of 3D medical images.
- **Segment Any Anatomy Target Without Additional Annotation**: MedLSAM is designed to segment any anatomical target in 3D medical images without the need for further annotations, contributing to the automation and efficiency of the segmentation process.

Details

> The Segment Anything Model (SAM) has recently emerged as a groundbreaking model in the field of image segmentation. Nevertheless, both the original SAM and its medical adaptations necessitate slice-by-slice annotations, which directly increase the annotation workload with the size of the dataset. We propose MedLSAM to address this issue, ensuring a constant annotation workload irrespective of dataset size and thereby simplifying the annotation process. Our model introduces a few-shot localization framework capable of localizing any target anatomical part within the body. To achieve this, we develop a Localize Anything Model for 3D Medical Images (MedLAM), utilizing two self-supervision tasks: relative distance regression (RDR) and multi-scale similarity (MSS) across a comprehensive dataset of 14,012 CT scans. We then establish a methodology for accurate segmentation by integrating MedLAM with SAM. By annotating only six extreme points across three directions on a few templates, our model can autonomously identify the target anatomical region on all data scheduled for annotation. This allows our framework to generate a 2D bounding box for every slice of the image, which are then leveraged by SAM to carry out segmentations. We conducted experiments on two 3D datasets covering 38 organs and found that MedLSAM matches the performance of SAM and its medical adaptations while requiring only minimal extreme point annotations for the entire dataset. Furthermore, MedLAM has the potential to be seamlessly integrated with future 3D SAM models, paving the way for enhanced performance.

Extreme points templates

MedLAM

SAM

Query datasets

Projected 3D bboxes

3D Segmentation results
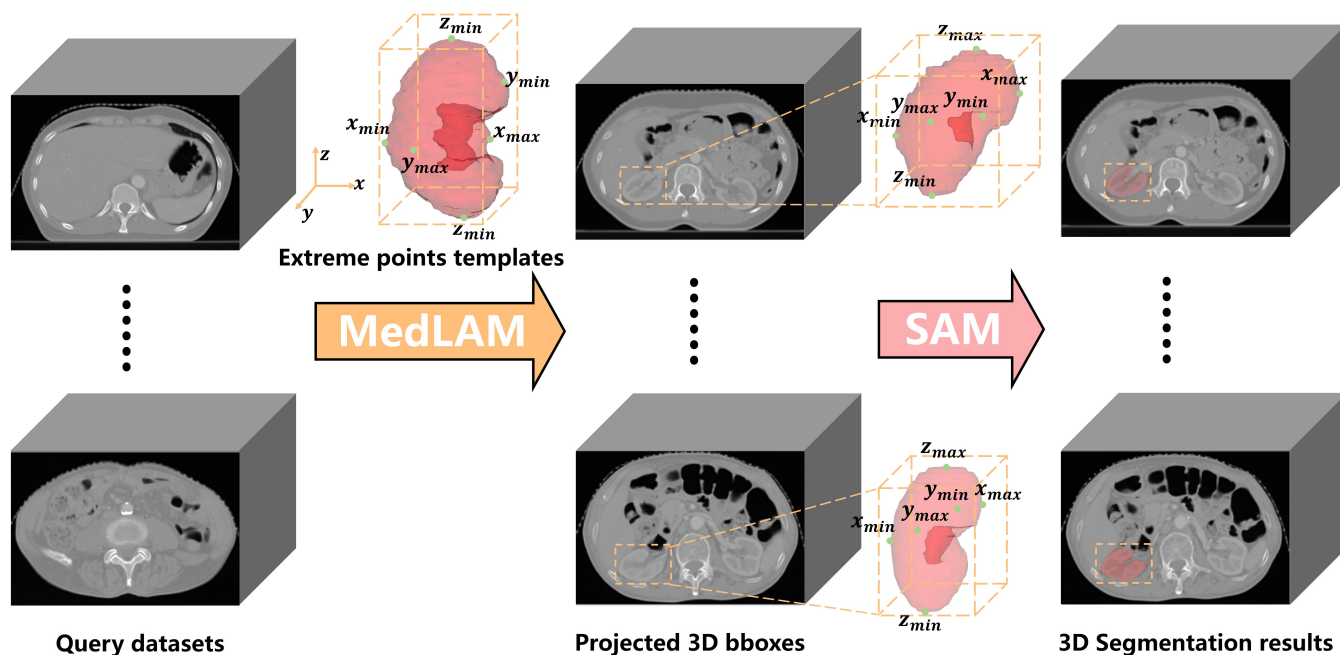
Get Started

Main Requirements

```
torch>=1.11.0
tqdm
nibabel
scipy
SimpleITK
monai
```

Installation

1. Create a virtual environment `conda create -n medlsam python=3.10 -y` and activate it `conda activate medlsam`
2. Install Pytorch
3. `git clone https://github.com/openmedlab/MedLSAM`
4. Enter the MedSAM folder `cd MedLSAM` and run `pip install -e .`

Download Model

Download MedLAM checkpoint, SAM checkpoint, MedSAM checkpoint and place them at `checkpoint/medlam.pth`, `checkpoint/sam_vit_b_01ec64.pth` and `checkpoint/medsam_20230423_vit_b_0.0.1.pth`

Inference

GPU requirement

We recommend using a GPU with 8GB or more memory for inference.

Data preparation

- [StructSeg Task1 HaN OAR](#)
- [WORD](#) (Request for access is required to download this dataset.)

Note: You can also download other CT datasets and place them any place you want. MedLSAM will **automaticly** apply the preprocessing procedure during the inference time, so please do **not** normalize the original CT images!!!

After downloading the datasets, you should sort the data into "support" and "query" groups. This does not require moving the actual image files. Rather, you need to create separate lists of file paths for each group.

For each group ("support" and "query"), create a .txt file listing the paths to the image files, and another .txt file listing the paths to the corresponding label files. Ensure that the ordering of images and labels aligns in both lists. These lists will be used to direct MedLSAM to the appropriate files during the inference process.

Config preparation

MedLAM and **MedLSAM** load their configurations from a .txt file. The structure of the file is as follows:

```
[data]
support_image_ls      =  config/data/StructSeg_HaN/support_image.txt
support_label_ls      =  config/data/StructSeg_HaN/support_label.txt
query_image_ls        =  config/data/StructSeg_HaN/query_image.txt
query_label_ls        =  config/data/StructSeg_HaN/query_label.txt
gt_slice_threshold    = 10
fg_class              =
[1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22]
seg_save_path         = result/npz/StructSeg
seg_png_save_path     = result/png/StructSeg

[vit]
net_type              = vit_b

[weight]
medlam_load_path   = checkpoint/medlam.pth
vit_load_path  = checkpoint/medsam_20230423_vit_b_0.0.1.pth
```

Each of the parameters is explained as follows:

- `support_image_ls`: The path to the list of support image files.
- `support_label_ls`: The path to the list of support label files.
- `query_image_ls`: The path to the list of query image files.
- `query_label_ls`: The path to the list of query label files.
- `gt_slice_threshold`: The threshold value for ground truth slice selection.
- `fg_class`: The list of foreground class to be used for localization and segmentation. This could be a list of integers indicating the class labels. You can only select a part of them as target classes.
- `seg_save_path`: The path to save the segmentation results in .npz format, **only required for MedLSAM**.

- `seg_png_save_path`: The path to save the segmentation results in .png format, **only required for MedLSAM**.
- `net_type`: The type of vision transformer model to be used, **only required for MedLSAM**. By default, this is set to vit_b.
- `medlam_load_path`: The path to load the pretrained MedLAM model weights.
- `vit_load_path`: The path to load the pretrained vision transformer model weights, **only required for MedLSAM**.

Inference

- MedLAM (Localize any anatomy target)

```
CUDA_VISIBLE_DEVICES=0 python MedLAM_Inference.py --config_file
path/to/your/test_medlam_config.txt
```

Example:

```
CUDA_VISIBLE_DEVICES=0 python MedLAM_Inference.py --config_file
config/test_config/test_structseg_medlam.txt
```

- MedLSAM (Localize and segment any anatomy target)

```
CUDA_VISIBLE_DEVICES=0 python MedLSAM_Inference.py --config_file
path/to/your/test_medlsam_config.txt
```

Example:

```
CUDA_VISIBLE_DEVICES=0 python MedLSAM_Inference.py --config_file
config/test_config/test_structseg_medlam_medsam.txt
```

Results

- MedLAM (Localize any anatomy target): MedLAM automatically calculates and saves the mean Intersection over Union (IoU) along with the standard deviation for each category in a .txt file. These files are stored under the result/iou directory.
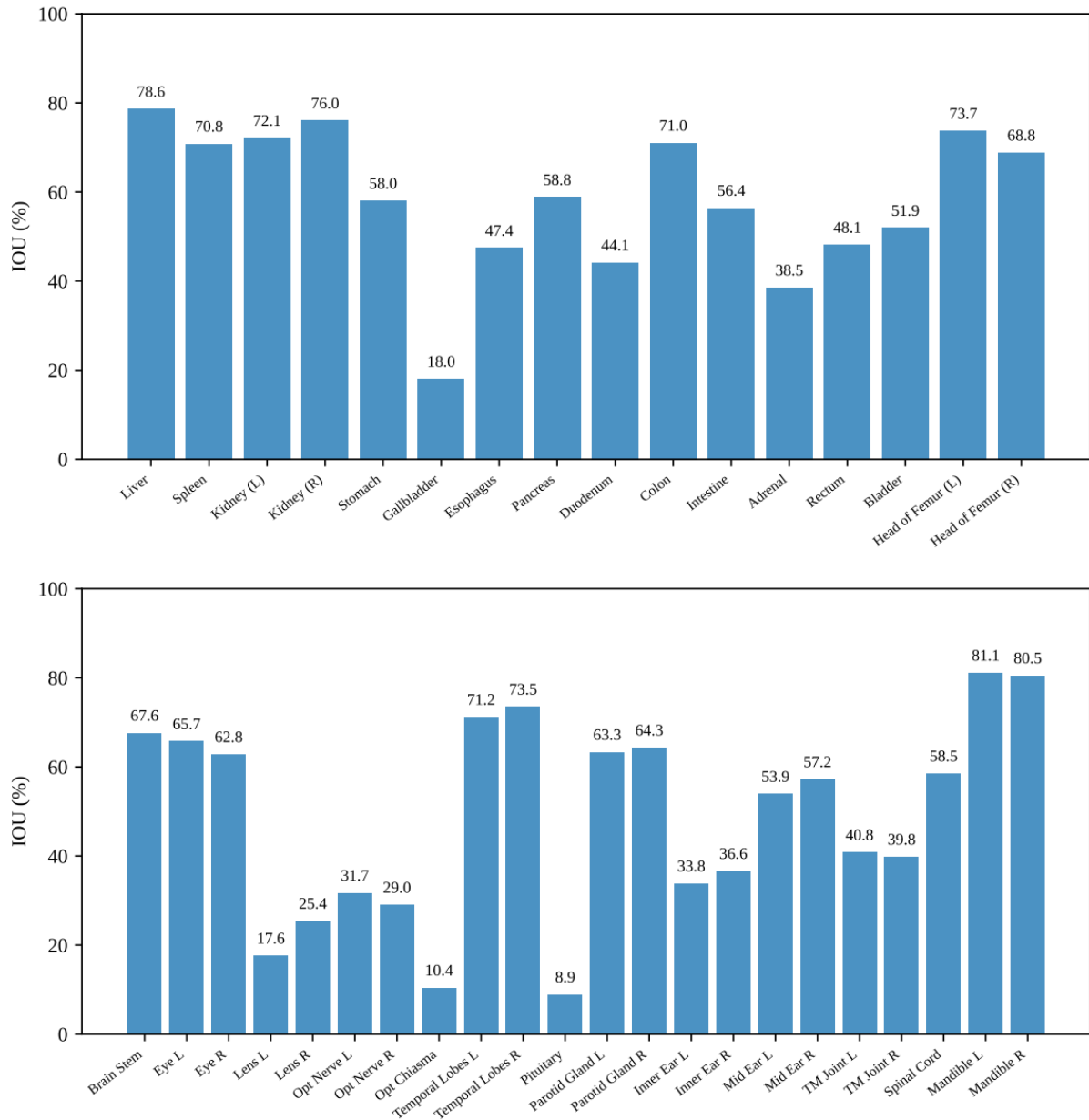
Fig. 5: The mean IOU score of each organ in the WORD (top) and StructSeg (bottom) dataset.

- MedLSAM (Localize and segment any anatomy target): MedLSAM automatically calculates and saves the mean Dice Similarity Coefficient (DSC) along with the standard deviation for each category in a .txt file. These files are stored under the result/dsc directory.

Table 2: DSC (mean±std %) evaluation of 3D head-and-neck organs segmentation in the StructSeg Task1 dataset. The table compares the performance of SAM and MedSAM as a segmentation basis within the MedLsAM framework, along with results from manually assisted localizations.

Localization	MedLAM		Manual	
Organs	SAM	MedSAM	SAM	MedSAM
Brain Stem	53.5 ± 5.5	$\mathbf{64.7 \pm 6.3}$	65.2 ± 3.7	$\mathbf{72.8 \pm 3.3}$
Eye L	$\mathbf{63.9 \pm 6.1}$	61.1 ± 6.1	$\mathbf{67.6 \pm 5.0}$	66.8 ± 5.6
Eye R	$\mathbf{66.3 \pm 5.3}$	63.4 ± 5.2	$\mathbf{69.5 \pm 4.6}$	67.6 ± 4.9
Lens L	22.2 ± 7.5	$\mathbf{16.5 \pm 3.1}$	$\mathbf{21.4 \pm 9.5}$	15.9 ± 2.8
Lens R	20.6 ± 6.9	$\mathbf{13.6 \pm 2.8}$	$\mathbf{20.7 \pm 10.8}$	13.8 ± 3.5
Opt Nerve L	$\mathbf{31.4 \pm 9.5}$	29.7 ± 13.2	$\mathbf{32.4 \pm 12.9}$	22.2 ± 17.9
Opt Nerve R	$\mathbf{34.6 \pm 8.9}$	32.1 ± 12.2	32.2 ± 15.9	$\mathbf{36.4 \pm 13.1}$
Opt Chiasma	$\mathbf{29.0 \pm 10.0}$	28.9 ± 16.0	$\mathbf{37.9 \pm 14.9}$	25.3 ± 14.7
Temporal Lobes L	25.4 ± 16.3	$\mathbf{72.3 \pm 4.8}$	37.7 ± 20.2	$\mathbf{78.2 \pm 6.4}$
Temporal Lobes R	19.9 ± 20.2	$\mathbf{67.7 \pm 8.6}$	34.4 ± 21.2	$\mathbf{76.5 \pm 7.6}$
Pituitary	$\mathbf{36.2 \pm 21.1}$	28.5 ± 16.1	$\mathbf{36.6 \pm 17.0}$	29.1 ± 15.4
Parotid Gland L	7.1 ± 6.5	$\mathbf{44.2 \pm 10.3}$	27.8 ± 10.0	$\mathbf{50.7 \pm 11.0}$
Parotid Gland R	8.1 ± 8.6	$\mathbf{44.7 \pm 8.5}$	30.2 ± 9.9	$\mathbf{49.4 \pm 10.5}$
Inner Ear L	$\mathbf{51.7 \pm 16.5}$	48.1 ± 13.7	$\mathbf{56.4 \pm 16.1}$	54.7 ± 12.4
Inner Ear R	$\mathbf{63.8 \pm 10.3}$	43.5 ± 18.8	$\mathbf{60.5 \pm 15.8}$	44.5 ± 16.9
Mid Ear L	$\mathbf{64.1 \pm 11.8}$	27.7 ± 14.1	$\mathbf{74.4 \pm 7.7}$	39.5 ± 14.4
Mid Ear R	$\mathbf{64.7 \pm 10.6}$	33.2 ± 14.2	$\mathbf{74.4 \pm 7.7}$	45.0 ± 9.6
TM Joint L	$\mathbf{54.0 \pm 8.3}$	34.9 ± 12.4	$\mathbf{62.8 \pm 11.5}$	37.8 ± 16.1
TM Joint R	$\mathbf{58.5 \pm 7.9}$	43.6 ± 11.1	$\mathbf{64.5 \pm 14.8}$	46.6 ± 11.9
Spinal Cord	$\mathbf{9.5 \pm 3.8}$	9.4 ± 3.5	$\mathbf{40.4 \pm 6.3}$	32.7 ± 6.4
Mandible L	$\mathbf{48.3 \pm 5.1}$	9.0 ± 3.8	$\mathbf{85.3 \pm 2.4}$	11.1 ± 4.6
Mandible R	$\mathbf{43.5 \pm 5.3}$	2.6 ± 2.9	$\mathbf{80.4 \pm 2.5}$	12.9 ± 7.8
Average	$\mathbf{39.6 \pm 7.6}$	37.5 ± 7.1	$\mathbf{50.6 \pm 6.3}$	42.3 ± 7.8

Table 3: DSC (mean±std %) evaluation of 3D head-and-neck organs segmentation in the WORD dataset. The table compares the performance of SAM and MedSAM as segmentation basis within the MedLSAM framework, along with results from manually assisted localizations.

Localization	MedLAM		Manual	
Organs	SAM	MedSAM	SAM	MedSAM
Liver	55.5 ± 9.5	$\mathbf{67.1 \pm 8.7}$	$\mathbf{84.5 \pm 6.3}$	76.3 ± 6.7
Spleen	$\mathbf{60.9 \pm 12.4}$	40.3 ± 20.8	$\mathbf{87.2 \pm 7.3}$	60.2 ± 19.4
Kidney (L)	$\mathbf{78.5 \pm 12.1}$	66.2 ± 13.1	$\mathbf{92.0 \pm 4.4}$	72.3 ± 7.3
Kidney (R)	$\mathbf{83.0 \pm 11.0}$	62.3 ± 9.7	$\mathbf{92.9 \pm 2.3}$	67.1 ± 6.7
Stomach	$\mathbf{43.0 \pm 13.1}$	35.2 ± 15.9	$\mathbf{79.4 \pm 8.6}$	62.2 ± 14.6
Gallbladder	$\mathbf{33.2 \pm 25.0}$	27.8 ± 22.7	$\mathbf{72.9 \pm 9.7}$	67.3 ± 10.5
Esophagus	$\mathbf{24.6 \pm 11.2}$	22.1 ± 13.8	$\mathbf{68.2 \pm 6.8}$	48.2 ± 13.5
Pancreas	$\mathbf{34.1 \pm 12.1}$	28.4 ± 11.0	$\mathbf{64.0 \pm 11.8}$	56.7 ± 9.2
Duodenum	$\mathbf{22.9 \pm 13.2}$	18.5 ± 9.8	$\mathbf{59.7 \pm 11.9}$	42.3 ± 11.3
Colon	$\mathbf{18.5 \pm 6.3}$	12.8 ± 8.3	$\mathbf{42.9 \pm 9.0}$	22.0 ± 9.9
Intestine	$\mathbf{35.5 \pm 9.3}$	20.2 ± 8.8	$\mathbf{60.2 \pm 7.2}$	31.3 ± 9.1
Adrenal	$\mathbf{3.8 \pm 5.2}$	3.4 ± 3.1	$\mathbf{18.7 \pm 12.1}$	17.3 ± 10.0
Rectum	$\mathbf{37.6 \pm 11.1}$	29.9 ± 12.9	$\mathbf{74.4 \pm 5.5}$	53.7 ± 12.9
Bladder	$\mathbf{68.4 \pm 21.3}$	62.4 ± 17.8	$\mathbf{85.8 \pm 11.6}$	73.7 ± 10.5
Head of Femur (L)	$\mathbf{74.5 \pm 7.7}$	48.2 ± 15.6	$\mathbf{89.3 \pm 4.7}$	62.1 ± 12.8
Head of Femur (R)	$\mathbf{71.5 \pm 5.3}$	46.6 ± 10.1	$\mathbf{87.9 \pm 3.6}$	65.5 ± 7.0
Average	$\mathbf{45.9 \pm 11.2}$	37.7 ± 12.4	$\mathbf{72.5 \pm 3.1}$	54.9 ± 6.7

To do list

- ☐ Support scribble prompts

🙋‍♀️ Feedback and Contact

- lyc745307452@gmail.com

🛡️ License

This project is under the CC-BY-NC 4.0 license. See LICENSE for details.

🙏 Acknowledgement

- A lot of code is modified from MedSAM.
- We highly appreciate all the challenge organizers and dataset owners for providing the public dataset to the community.
- We thank Meta AI for making the source code of segment anything publicly available.

📝 Citation

If you find this repository useful, please consider citing this paper:

```
@article{Lei2023medlam,
  title={MedLSAM: Localize and Segment Anything Model for 3D Medical
```

```
  Images},
    author={Wenhui Lei, Xu Wei, Xiaofan Zhang, Kang Li, Shaoting Zhang},
    journal={arXiv preprint arXiv:},
    year={2023}
  }
```