VoxTell: Free-Text Promptable Universal 3D Medical Image Segmentation

Supplementary Material

Contents

A. Implementation Details

A.1. Training Configuration

All experiments were trained under a consistent configuration to ensure reproducibility and fair comparisons across ablations and baselines. For ablation studies, a single GPU with a batch size of 2 was used, while the final VoxTell model was trained on 64 NVIDIA A100 GPUs (40GB each) with a batch size of 128, achieving 64× higher throughput. We leverage the nnU-Net [77] framework, which provides a robust 3D segmentation training pipeline. Each model was trained for 2,000 epochs, following the convention of 250 iterations per epoch, using stochastic gradient descent (SGD) with an initial learning rate of 1×10^{-4}, decayed according to a polynomial schedule. Input volumes were processed as 192^3 voxel patches. Standard nnU-Net data augmentation strategies were applied, with the exception of mirroring along the left–right axis to avoid ambiguities in laterality-sensitive anatomical structures.

A.2. Model Architecture

The vision backbone is a ResEncL encoder [78] comprising six layers, producing hierarchical feature maps with channel dimensions `[32, 64, 128, 256, 320, 320]`. The prompt decoder is a standard transformer using cross attention with six layers, eight heads, and a query dimension of 2048. Cross-modal integration is performed via a two-layer MLP vision–text adapter with hidden dimension 2048, which aligns the dimensions of textual features produced by the prompt decoder with visual features. The text encoder is a frozen Qwen3-Embedding-4B [231] model, converting natural language prompts into fixed embeddings. During training, all text embeddings are precomputed for efficiency. We use the following instruction prompt:

"Instruct: Given an anatomical term query, retrieve the precise anatomical entity and location it represents. Query: [text input]."

Initial experimentation revealed that this prompt ensures consistent mapping of anatomical and pathological terms to similar embeddings.

A.3. Training Strategy

Each training image is queried with three textual prompts as this is the maximum that fits on an NVIDIA A100 GPU (40GB) during training with batch size 2 per GPU. We use two positive prompts corresponding to structures present in the volume, and one negative prompt corresponding to an absent structure. The negative prompt is critical for teaching the model to output empty masks when the target is not present. Foreground structures are oversampled at an 85% probability to increase the frequency of patches containing segmentation targets. For positive text prompts, we randomly sample synonyms and rephrasings for each concept as prompts, while emphasizing the main term (*e.g.* "liver" over "hepatic organ"), by selecting the default name with 25% probability and a rephrased variant with 75%. The segmentation objective combines Dice loss with binary cross-entropy (BCE) loss to optimize both volumetric and pixel-wise accuracy. Deep supervision is applied at all five decoder scales, using nnU-Net's default weights $\lambda_s = [1, 1/2, 1/4, 1/8, 1/16]$. Vision–language feature fusion is applied at all scales, ensuring repeated cross-modal interactions throughout the decoder and enhancing alignment between textual prompts and visual features.

B. Dataset Details

B.1. Large-Scale Multi-Modality Dataset

To train VoxTell, we assembled a comprehensive multi-modality 3D medical imaging corpus by aggregating 158 publicly available datasets, encompassing over 62,000 volumetric scans (≈4 TB) and 1,087 distinct anatomical and pathological concepts. The collection spans Computed Tomography (CT), multi-sequence Magnetic Resonance Imaging (MRI), and Positron Emission Tomography (PET), covering both healthy and pathological anatomies across the brain, thorax, abdomen, pelvis, musculoskeletal system, and vasculature.

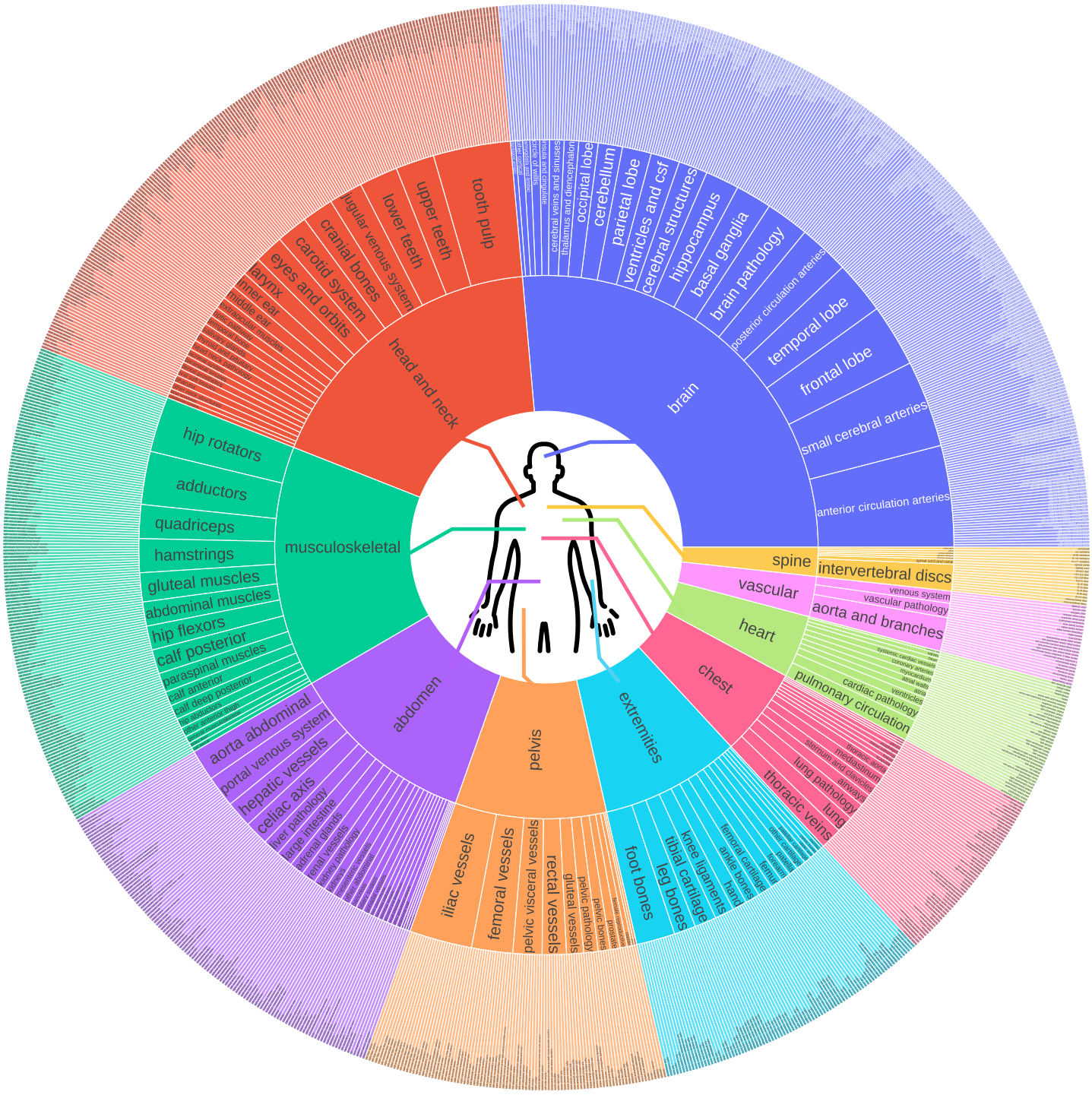

Figure 6. **Comprehensive overview of the anatomical and pathological concepts** included in our multi-modality 3D medical imaging dataset. The dataset aggregates 158 public sources, covering CT, MRI, and PET scans of over 62,000 volumetric images across brain, head and neck, thorax, abdomen, pelvis, musculoskeletal, extremities and vascular systems. The visualization highlights both coarse and fine-grained structures, ranging from major organs to substructures and lesions, demonstrating the semantic diversity leveraged for language-conditioned 3D segmentation in VoxTell.

Name	Images	Categories	Modality	Target	Link
Decatlon Task 2 [181]	20	1	MRI	Left Atrium	http://medicaldecathlon.com
Decatlon Task 3 [181]	131	2	CT	Liver, L. Tumor	http://medicaldecathlon.com
Decatlon Task 4 [181]	260	2	MRI	Hippocampus	http://medicaldecathlon.com
Decatlon Task 5 [181]	32	3	MRI	Prostate	http://medicaldecathlon.com
Decatlon Task 6 [181]	63	1	CT	Lung Cancer	http://medicaldecathlon.com
Decatlon Task 7 [181]	281	2	CT	Pancreas, P. Tumor	http://medicaldecathlon.com
Decatlon Task 8 [181]	303	2	CT	Hepatic Vessel, H. Tumor	http://medicaldecathlon.com
Decatlon Task 9 [181]	41	1	CT	Spleen	http://medicaldecathlon.com
Decatlon Task 10 [181]	126	1	CT	Colon Cancer	http://medicaldecathlon.com
ISLES2015 [142]	28	1	MRI	Stroke Lesions	http://www.isles-challenge.org/ISLES2015
BTCV [105]	30	13	CT	13 Abdominal Organs	https://www.synapse.org/Synapse:syn3193805/wiki/217789
RibSeg [83, 221]	654	25	CT	Ribs	https://ribfrac.grand-challenge.org
AortaSeg24 [75]	50	23	CT	Aorta	https://aortaseg24.grand-challenge.org
LIDC [17]	1010	1	CT	Lung Lesion	https://www.cancerimagingarchive.net/collection/lidc-idri
CTPelvic1k [127]	1106	3	CT	Pelic Bones	https://github.com/MIRACLE-Center/CTPelvic1K
Promise [124]	50	1	MRI	Prostate	https://zenodo.org/records/8026660
Duke Liver [136]	310	1	MRI	Liver	https://zenodo.org/records/7774566
ACDC [26]	200	3	MRI	Cardiac Structures	https://www.creatis.insa-lyon.fr/Challenge/acdc
AbdOrgSegm [45, 62, 62, 170, 172, 219]	63	8	CT	Abdominal Organs	https://zenodo.org/records/1169361
CHAOS [89, 90]	60	4	MR	liver, right kidney, left kidney, spleen	https://chaos.grand-challenge.org/Data
OpenMind Tissue [28, 197]	3644	32	MRI, CT	Brain Structures	https://github.com/BBillot/SynthSeg
StructSeg Task 1 [109]	50	22	CT	23 Head and Neck Structures	https://structseg2019.grand-challenge.org
StructSeg Task 2 [109]	50	1	CT	Nasopharynx Cancer	https://structseg2019.grand-challenge.org
StructSeg Task 3 [109]	50	6	CT	6 Thoracic Structures	https://structseg2019.grand-challenge.org
StructSeg Task 4 [109]	50	1	CT	Lung Tumor	https://structseg2019.grand-challenge.org
COVID-19-20 [10, 173]	199	1	CT	Covid	https://covid-segmentation.grand-challenge.org/COVID-19-20
SegTHOR [103]	40	4	CT	Heart, Aorta, Trachea, Esophagus	https://competitions.codalab.org/competitions/21145
FETA [155]	40	7	MRI	Brain Regions	https://fetachallenge.github.io/pages/Data_description
ISLES2022 [72]	250	1	MRI	Stroke Lesion	https://zenodo.org/records/7153326
LGGMRISeg [34]	110	1	MRI	Brain Tumor	https://www.kaggle.com/datasets/mateuszbuda/lgg-mri-segmentation/data
NIH-Pan [172]	82	1	CT	Pancreas	https://wiki.cancerimagingarchive.net/display/Public/Pancreas-CT
M-CRIB [8]	10	100	MRI	Neonatal Brain Atlas	https://osf.io/4vthr
CAP [86]	1637	1	MRI	Left Ventricle	https://www.cardiacatlas.org/lv-segmentation-challenge
AtriaSeg2018 [217]	154	2	MRI	Left Atrium	https://www.cardiacatlas.org/atriaseg2018-challenge/atria-seg-data
Spine-Mets [157]	55	17	CT	Vertebrae	https://www.cancerimagingarchive.net/collection/spine-mets-ct-seg
VerSe2019 [130, 177, 178]	80	26	CT	Vertebrae	https://osf.io/jtfa5
VerSe2020 [130, 177, 178]	61	25	CT	Vertebrae	https://verse2020.grand-challenge.org
WMHSegChallenge [98]	60	1	MRI	White Matter Hyperintensities	https://dataverse.nl/dataset.xhtml?persistentId=doi:10.34894/AECRSD
NCI-ISBI-Prostate [29]	60	3	MRI	Prostate Zones	http://doi.org/10.7937/K9/TCIA.2015.zF0vlOPv
OASIS [143]	436	3	MRI	Brain Regions	https://sites.wustl.edu/oasisbrains/home/oasis-1
BraTS24 Task 1 Glioma [47]	1350	4	MRI	Glioblastoma	https://www.synapse.org/Synapse:syn53708249/wiki/627500
BraTS2024 Task 2 Africa [3]	60	3	MRI	Brain Tumor	https://www.synapse.org/Synapse:syn53708249/wiki/627501
BraTS24 Task 3 Meningioma [100]	500	1	MRI	Brain Tumor	https://www.synapse.org/Synapse:syn53708249/wiki/627503
BraTS24 Task 4 Brain Metastases [149]	652	3	MRI	Brain Metastases	https://www.synapse.org/Synapse:syn53708249/wiki/627504
BraTS24 Task 5 Pediatric [91]	261	4	MRI	Brain Tumor	https://www.synapse.org/Synapse:syn53708249/wiki/627505
LNQ2023 [52]	393	1	CT	Mediastinal Lymph Nodes	https://lnq2023.grand-challenge.org/data
MediaLymph [33]	15	1	CT	Mediastinal Lymph Nodes	https://github.com/dbouget/ct_mediastinal_structures_segmentation
MediaStruct [32]	15	12	CT	Mediastinal Structures	https://github.com/dbouget/ct_mediastinal_structures_segmentation
CT Lymph Nodes [171]	176	1	CT	Mediastinal Lymph Nodes	https://www.cancerimagingarchive.net/collection/ct-lymph-nodes
MAMA MIA [59]	1506	1	MRI	Breast Tumor	https://www.synapse.org/Synapse:syn60868042/wiki/628716
ATM2022 [228]	0	1	CT	Airway Tree	https://atm22.grand-challenge.org
Atlas Bourgogne [161]	60	2	MRI	Liver, Tumor	https://atlas-challenge.u-bourgogne.fr
Couinaud [229]	193	8	CT	Liver Segments	https://www.kaggle.com/datasets/louisgv/couinaud-liver-segmentation
CURVAS [165]	60	3	CT	Pancreas, Kidney, Liver	https://curvas.grand-challenge.org/curvas-dataset
MMs [35]	300	3	MRI	Cardiac Structures	https://www.ub.edu/mnms
Emidec [102]	100	4	MRI	Cardiac Structures	https://emidec.com
Kipa22 [69]	70	4	CT	Kidney, Vessel, Tumor	https://kipa22.grand-challenge.org
MrBrains18 [97]	30	9	MRI	Brain Structures	https://mrbrains18.isi.uu.nl/index.html
OrCaScore [211]	32	3	CT	Heart Calcifications	https://orcascore.grand-challenge.org
TriALS Task 2 [54]	240	2	CT	Liver, L. Tumor	https://www.synapse.org/Synapse:syn53285416
Parse22 [133]	100	1	CT	Pulmonary Artery	https://parse2022.grand-challenge.org/Parse2022
PDDCA [164]	47	9	CT	Head and Neck Structures	https://www.imagenglab.com/newsite/pddca
Aortic Dissection [145]	39	2	CT	True and False Aortic Lumen	https://figshare.com/ndownloader/articles/22269091/versions/1
ProstateEdgeCases [87]	131	5	CT	Bladder, Prostate, Rectum	https://doi.org/10.7937/QSTF-ST65
SKI10 [194]	100	4	MRI	Cartilage, Bone	https://ski10.grand-challenge.org
Spider [193]	447	19	MRI	Lumbar Spine	https://zenodo.org/records/10159290
VALDO Task 2 [187]	72	1	MRI	Cerebral Microbleed	https://valdo.grand-challenge.org/Task2
BONBID-HIE [22]	85	1	MRI	Brain Lesion	https://zenodo.org/records/1060276
UPENN-GBM [20]	147	3	MRI	Brain Edema and Tumor	https://www.cancerimagingarchive.net/collection/upenn-gbm
LNDb [156]	236	1	CT	Lung Nodule	https://lndb.grand-challenge.org
ReMIND [85]	213	1	MRI	Brain Resection	https://www.cancerimagingarchive.net/collection/remind
ProstateX [125]	140	2	MRI	Prostate Tumor	https://www.cancerimagingarchive.net/collection/prostatex
Prostate158 [2]	188	1	MRI	Prostate, P. Tumor	https://zenodo.org/records/6481141
TotalSegmentatorMRI [206]	616	50	MRI	Whole Body Structures	https://zenodo.org/records/14710732
RiderLungCT [234]	65	1	CT	Lung Cancer	https://www.cancerimagingarchive.net/collection/rider-lung-ct
LiverMets [182]	171	4	CT	Liver Metastases	https://doi.org/10.7937/QXK2-QG03
HippocampusSubfield [117]	20	14	MRI	Hippocampus Structures	https://plus.figshare.com/ndownloader/articles/26075713/versions/1
PANORAMA [9]	482	6	CT	Pancreas Structures	https://panorama.grand-challenge.org
TOM500 [185]	400	9	MRI	Eye Structures	https://springernature.figshare.com/ndownloader/files/49499655

Table 4. **Overview of training datasets part 1.** Overview of the 158 datasets used for model training (Part 1/2), covering names, image counts, number of categories, modalities, targets, and access links.

Name	Images	Categories	Modality	Target	Link
WAW-TACE [23]	114	1	CT	Hepatocellular Carcinoma	https://zenodo.org/records/12741586
LiverHccSeg [64]	34	2	MRI	Hepatocellular carcinoma	https://zenodo.org/records/8179129
Enhance-PET [56, 180]	1597	132	CT, PET	Whole Body Organs	https://github.com/ENHANCE-PET/MOOSE
LLD-MMRI [131, 141]	3984	1	MRI	Liver Lesion	https://huggingface.co/datasets/wanglab/LLD-MMRI-MedSAM2
AIIB23 [111, 153]	120	1	CT	Airway	https://codalab.lisn.upsaclay.fr/competitions/13238
BTCV-Cervix [104]	30	4	CT	Abdominal Organs	https://www.synapse.org/Synapse:syn3193805/wiki/217752
CARE25-MyoPS [51, 160, 241]	449	5	MRI	Cardiac Structures, C. Edema and Scars	https://zmic.org.cn/care_2025/track2
CARE25-WHS [58, 240, 242]	86	7	CT, MRI	Cardiac Structures	https://zmic.org.cn/care_2025/track3
cSeg2022 [188]	13	3	MRI	Brain Regions	https://tarheels.live/cseg2022
ImageCAS [226]	1000	1	CT	Coronary Artery	https://www.kaggle.com/datasets/xiaoweixumedicalai/imagecas/data
ToothFairy 3 [30, 31, 132]	532	77	CT	Dental Structures	https://ditto.ing.unimore.it/toothfairy3
FUMPE [144]	35	1	CT	Pulmonary Embolism	https://www.kaggle.com/datasets/andrewmvd/pulmonary-embolism-in-ct-images
SLAWT [88]	10	1	CT	Left Atrium, Atrium Wall	https://www.doc.ic.ac.uk/~rkarim/la_lv_framework/wall/index.html
LiverCirrMRI [81]	676	1	MRI	Liver	https://osf.io/cuk24
3D-IRCADb-01 [183]	20	31	CT	Abdominal Organs, Liver Tumor	https://www.ircad.fr/research/data-sets/liver-segmentation-3d-ircadb-01
ImageTBAD [223]	100	3	CTA	Aorta True Lumen, False Lumen, Thrombus	https://www.kaggle.com/datasets/xiaoweixumedicalai/imagetbad
ImageCHD [218]	110	7	CTA	Heart, Vascular Structures	https://www.kaggle.com/datasets/xiaoweixumedicalai/imagechd
MS-Brain-MRI-Lesion [152]	60	1	MRI	Multiple Sclerosis Lesion	https://data.mendeley.com/datasets/8bctsm8jz7/1
ImageALCAPA [225]	30	7	CT	Cardiac Structures	https://www.kaggle.com/datasets/xiaoweixumedicalai/imagealcapa
ImageTAAD [184]	120	35	CT	Abdominal Structures	https://www.kaggle.com/datasets/xiaoweixumedicalai/imagetaad
Crossmoda2022 [179]	210	2	MRI	vestibular Schwannoma, Cochlea	https://zenodo.org/records/6504722
Atlas22 [123]	655	1	MRI	lesion	https://bids.neuroimaging.io
PulmonaryVessels [43]	106	2	CT	Pulmonary vessels	https://www.kaggle.com/datasets/xiaoweixumedicalai/mytest
HOCMvalvesSeg [239]	27	7	CT	Cardiac structures	https://www.kaggle.com/datasets/xiaoweixumedicalai/hocmvalvesseg
CTSpine1K [49]	1005	25	CT	Vertebrae	https://huggingface.co/datasets/alexanderdann/CTSpine1K
Han-Seg23-MR [159]	42	30	MRI	Head Neck Structures	https://zenodo.org/records/7442914#.ZBwp7XbMJaR
HNTSMRG24 [196]	150	2	MRI	Head Neck Lesions	https://zenodo.org/records/11199559
IVDM3Seg [39]	16	1	MRI	Intervertebral Discs	https://ivdm3seg.weebly.com
KiTS23 [71]	489	4	CT	Kidney, K. Tumor, K. Cyst	https://github.com/neheller/kits23
AutoPet2 [61]	1014	1	CT, PET	Lesions	https://autopet-ii.grand-challenge.org
AutoPet3 [60, 76]	597	1	CT, PET	Lesions	https://autopet-iii.grand-challenge.org
AMOS [82]	360	16	CT	Abdominal Organs	https://zenodo.org/record/7155725#.Y0OOCOxBztM
RHUH-GBM [38]	120	3	MRI	Brain Tumor	https://www.cancerimagingarchive.net/collection/rhuh-gbm
AutoPet4 [99]	670	1	CT	Lesions	https://fdat.uni-tuebingen.de/records/qwsry-7t837
CPTAC-HNSCC [176]	245	5	CT	Head Neck Lesions	https://doi.org/10.7937/PFEC-T641
BrainTRGammaKnife [203, 205]	76	1	MRI	Brain Lesions	https://www.cancerimagingarchive.net/collection/brain-tr-gammaknife
LAScarQS24 Task 1 [113–116]	60	2	MRI	Left Atrium, Atrial Scars	https://zmic.org.cn/care_2024/track2
LAScarQS24 Task 2 [113–116]	130	1	MRI	Left Atrium	https://zmic.org.cn/care_2024/track2
QUADRA-HC [67]	96	58	CT	Whole Body Organs	https://zenodo.org/records/16686025
PancreasSegMRI T1 [232]	385	1	MRI	Pancreas	https://osf.io/kysnj
TopCoW24 [222]	250	13	CT	Vessel Components of CoW	https://topcow24.grand-challenge.org
MRISegmentator	540	45	MRI	Whole Body Organs	https://github.com/rsummers11/MRISegmenter
TopBrain-CT [222]	25	40	CT	Brain Structures	https://topbrain2025.grand-challenge.org
TopBrain-MR [222]	25	42	MRI	Brain Structures	https://topbrain2025.grand-challenge.org
LISA25 [106]	79	8	MRI	Brain Structures	https://www.synapse.org/Synapse:syn65670170/wiki/631438
Lower Extremity Muscles [13]	78	126	MRI	Musculoskeletal Structures	https://digitalcommons.du.edu/visiblehuman
CMRxMotion [202]	138	3	MRI	Cardiac Structures	https://cmr.miccai.cloud
BreastDivider [168]	200	2	MRI	Left Breast, Right Breast	https://huggingface.co/datasets/Bubenpo/BreastDividerDataset
TotalSegmentatorV2 [207]	1228	120	CT	Whole Body Structures	https://zenodo.org/records/10047292
Hecktor2022 [12]	524	2	CT, PET	Head Neck Tumor	https://hecktor.grand-challenge.org
Instance2022 [119]	100	1	CT	Intracranial Hemorrhage	https://instance.grand-challenge.org
MS Ljubljana [57, 107]	264	1	MRI	Muliple Sclerosis Lesion	https://lit.fe.uni-lj.si/en/research/resources/3D-MR-MS/
FLARE2022 [140]	50	13	CT	Abdominal Organs	https://flare22.grand-challenge.org
SegRap23 Task 1 [135]	120	45	CT	Head Neck Structures	https://drive.google.com/drive/folders/115mzmN1ZRIewnSR2QFDwW_-RkNM0LC9D
SegA [162]	56	1	CT	Aortic Vessel Tree	https://figshare.com/ndownloader/articles/14806362/versions/1
WORD [122, 134]	120	17	CT	Abdominal Organs	https://github.com/HiLab-git/WORD
AbdomenCT1K [138]	996	4	CT	Liver, Kidney, Spleen, Pancreas	https://zenodo.org/records/7860267
DAP-ATLAS [80]	533	143	CT	Abdominal Organs	https://www.synapse.org/#!Synapse:syn52287632.1/datasets
TORG [166]	140	4	CT	Liver, Bladder, Lungs, Kidneys, Bone, Brain	https://www.cancerimagingarchive.net/collection/ct-org
HanSeg-CT [158, 159]	42	30	CT	Head Neck Structures	https://zenodo.org/records/7442914
MU-Glioma-Post [224]	593	4	MRI	Brain Tumor	https://www.cancerimagingarchive.net/collection/mu-glioma-post
ACRIN-HN [93]	67	1	CT, PET	Head Neck Lesion	https://www.cancerimagingarchive.net/collection/acrin-hnscc-fdg-pet-ct
Head-and-Neck-PET-CT [192]	34	1	CT, PET	Head Neck Lesion	https://www.cancerimagingarchive.net/collection/head-neck-pet-ct
NSCLC-Radiogenomics [21]	69	1	CT, PET	Lung Lesion	https://www.cancerimagingarchive.net/collection/nsclc-radiogenomics
Soft-Tissue-Sarcoma [191]	42	1	CT, PET	Soft Tissue Sarcoma	https://www.cancerimagingarchive.net/collection/soft-tissue-sarcoma
TCGA-Lung [7]	5	1	CT, PET	Lung Tumor	https://www.cancerimagingarchive.net/collection/tcga-luad
MSCMRSeg [243]	25	3	MRI	Cardiac Structures	https://zmiclab.github.io/zxh/0/mscmrseg19/
DeepLesion [220]	1093	1	CT	Lesion	https://nihcc.app.box.com/v/DeepLesion/
Covid19CTLung [137]	10	1	CT	Covid	https://zenodo.org/records/3757476
PANTHER Task 2 [27]	50	2	MRI	Pancreas, P. Tumor	https://zenodo.org/records/15192302
NSCLC-PleuralEffusion [6, 96]	78	1	CT	Pleural Effusion	https://www.cancerimagingarchive.net/analysis-result/plethora
NSCLC-Radiomics [4]	78	1	CT	Lung Tumor	https://www.cancerimagingarchive.net/collection/nsclc-radiomics
MedSeg Liver Segments [1]	50	9	CT	Liver Segments	https://www.medseg.ai/database/liver-segments-50-cases
MedSeg Vasculature Brain [1]	1	73	MRI	Brain Vessels	https://www.medseg.ai/database/brain-vasculature
MedSeg Vasculature Abdomen [1]	1	43	CT	Abdominal Vessels	https://www.medseg.ai/database/vasculature-of-the-abdomen
MedSeg Vasculature Neck [1]	1	19	CT	Abdominal Vasculature	https://www.medseg.ai/database/vasculature-of-the-neck
MedSeg Vasculature Pelvis [1]	1	74	CT	Pelvic Vasculature	https://www.medseg.ai/database/vasculature-of-the-pelvis
MedSeg Musulature Pelvis [1]	1	67	CT	Pelvic Musculature	https://www.medseg.ai/database/musculature-of-the-pelvis
MedSeg Brain Ventricle [1]	10	1	MRI	Brain Ventricles	https://www.medseg.ai/database/lateral-ventricles-50-mri-cases

Table 5. **Overview of training datasets part 2.** Overview of the 158 datasets used for model training (Part 2/2), covering names, image counts, number of categories, modalities, targets, and access links.

Dataset targets range from large organs (*e.g.* liver, lungs, heart, brain) to fine-grained substructures (*e.g.* hippocampal subfields, aortic segments, vertebral bodies) and pathological lesions (*e.g.* tumors, vascular anomalies, white matter hyperintensities). This breadth yields rich semantic diversity suitable for learning language-conditioned representations.

To give a sense of anatomical coverage, the dataset includes comprehensive labeling of brain structures (*e.g.* cortex, white matter, basal ganglia, ventricles), cardiac structures (atria, ventricles, valves), thoracic, abdominal and pelvic organs (lung, liver, pancreas, kidneys, bladder), musculoskeletal structures (muscles, ligaments, bones), and vascular networks (aorta, venous system, pulmonary vessels), along with diverse pathologies including tumors, lesions, and thrombi. An overview of the structures from coarse to fine is shown in Fig. 6, whereas all full list of all concepts is given in Tab. 10).

While we cannot publicly share the full dataset due to licensing constraints, comprehensive details for all 158 datasets, including the number of images, categories, imaging modalities, target structures, and access links, are provided in Appendix Tables 4–5. For model development and ablation studies, we reserve 5% of each dataset as a held-out validation split to ensure consistent and unbiased evaluation. All ablation experiments are performed exclusively on this subset, preserving test integrity. To enable a fair comparison with the current state-of-the-art model SAT, we also benchmark a variant of SAT on this same training and validation split, providing it with the same text encoder we use (Qwen3-Embedding-4B), which outperforms SAT's original encoder, shown in Tab. 2. This ensures that improvements are attributable to the model rather than differences in text embeddings or training dataset.

B.2. In-House Radiotherapy Dataset

To assess the models capability of interpreting real-world clinical language, we curated an in-house cohort of 203 patients who underwent stereotactic body radiotherapy (SBRT) for either primary or secondary lung tumors at Heidelberg University Hospital. For each patient, planning CT scans were available, acquired either with or without contrast agent. All scans contained clinically approved, expert-annotated gross tumor volume (GTV) contours, which served as the reference segmentations.

In addition to the imaging data, textual descriptions of the target structures were extracted from the corresponding radiology reports, ensuring a one-to-one correspondence between each annotated lesion and its clinical description. These paired image–text samples constitute a multimodal dataset representative of real-world radiotherapy planning workflows. Example CT image slices and corresponding textual descriptions are shown in Fig. 1d. The textual findings were translated into English using a large language model (gpt-oss-120b), ensuring the translations preserve the semantic precision of the original reports. Representative examples include:

- "Suspicious for cavitary bronchogenic carcinoma in the right apical segment of the upper lobe."
- "Histologically confirmed adenocarcinoma NSCLC in the right upper lobe with broad-based pleural contact."
- "Peripheral bronchogenic carcinoma in the right lower lobe with pleural contact."
- "Bronchogenic carcinoma in the left upper lobe with fibrotic streaky consolidations."
- "Suspicious for bronchogenic carcinoma in the left apical upper lobe."
- "Round pulmonary nodule in the left upper lobe associated with known squamous cell carcinoma."

This dataset was held out entirely during training and used exclusively for independent evaluation.

B.3. Instance-Focused Findings Dataset

While the large-scale semantic dataset enables comprehensive anatomical understanding, current models often fail on *localized, instance-specific* or *context-dependent* queries, an issue also noted in prior work, benchmarking models on the ReXGroundingCT dataset [19]. They demonstrated that no current model can reliably handle truly instance-specific prompts, and even after fine-tuning on the ReXGroundingCT training split, performance remained suboptimal. To address this gap, we curate an *instance-focused dataset* specifically designed to support reasoning over fine-grained, spatially grounded prompts such as "spiculated tumor in the left lower lobe" or "cluster of HCC lesions in Couinaud segment 5". For fair comparison, we fine-tune both the current state-of-the-art ReXGroundingCT baseline, SAT [238], and our proposed model on this extended dataset.
Our instance dataset is constructed through three complementary pathways:

1. **Conversion of semantic lesion datasets to instance-level form:** We reformulate existing semantic lesion segmentation datasets by converting them into *instance annotations* conditioned on anatomical location. Specifically, we employ TotalSegmentator [207] to extract lung lobes and liver sub-segments (Couinaud segments), as well as left and right kidney masks, which then serve as contextual anchors for generating localized textual prompts for lung, liver and kidney lesions.

2. **Integration of location-rich public datasets:**

Dataset	Images	Modality	Target	Link
Decathlon Task 3 [16, 181]	131	CT	Liver Tumor	http://medicaldecathlon.com
Decathlon Task 6 [16, 181]	63	CT	Lung Lesion	http://medicaldecathlon.com
Decathlon Task 8 [16, 181]	303	CT	Hepatic Tumor	http://medicaldecathlon.com
LIDC [17]	1010	CT	Lung Lesion	https://www.cancerimagingarchive.net/collection/lidc-idri
StructSeg Task4 [110]	50	CT	Lung Cancer	https://structseg2019.grand-challenge.org
COVID-19-20 [173]	199	CT	COVID-19	https://covid-segmentation.grand-challenge.org/COVID-19-20
Atlas Bourgogne [161]	60	MRI	Liver Tumor	https://atlas-challenge.u-bourgogne.fr
TriALS [55]	240	CT	Liver Lesion	https://www.synapse.org/Synapse:syn53285416/wiki/625814
HCC Tace [147]	65	CT	Hepatocellular Carcinoma	https://www.cancerimagingarchive.net/collection/hcc-tace-seg/
RiderLung [235]	58	CT	Non-Small Cell Lung Carcinoma	https://www.cancerimagingarchive.net/collection/rider-lung-ct/
Colorectal Liver Mets [182]	171	CT	Colorectal Liver Metastases	https://www.cancerimagingarchive.net/collection/colorectal-liver-metastases
HCC Tace MRI [65]	34	MRI	Hepatocellular Carcinoma	https://zenodo.org/records/8179129
BrainGammaKnife [204]	76	MRI	Brain Tumor	https://www.cancerimagingarchive.net/collection/brain-tr-gammaknife
RexGroundingCT Train [18]	2,992	CT	Chest Findings	https://huggingface.co/datasets/rajpurkarlab/ReXGroundingCT
RADCURE-Tumor [210]	3,199	CT	Head Neck Tumor	https://www.cancerimagingarchive.net/collection/radcure
MSWAL [214]	484	CT	Abdominal Lesion	https://huggingface.co/datasets/zhaodongwu/MSWAL/tree/main
NSCLC Pleural Effusion [96]	78	CT	Pleural Effusion	https://www.cancerimagingarchive.net/analysis-result/plethora
NSCLC Radiomics [5]	503	CT	Lung Lesions	https://www.cancerimagingarchive.net/collection/nsclc-radiomics
RexGroundingCT Test [18]	50	CT	Chest Findings	https://huggingface.co/datasets/rajpurkarlab/ReXGroundingCT

Table 6. **Datasets used for instance-specific training and benchmarking on ReXGroundingCT.** The collection spans multiple organs and modalities (CT, MRI), integrating semantic lesion datasets reformulated into instance-level form, location-rich public datasets with spatial metadata (*e.g.* RADCURE, BrainGammaKnife), and the manually annotated ReXGroundingCT benchmark. Together, these sources enable fine-grained, text-conditioned localization of clinically described findings within complex anatomical contexts.

We incorporate publicly available TCIA[1] datasets across brain [204] and head–neck [210] domains that include original DICOM metadata. These metadata elements (*e.g.* study description, series body part, and slice positioning) are reformulated into *structured location prompts* per segmented object.

3. **Linking free-text findings to spatial annotations:** Complementing the above, we leverage the ReXGroundingCT [18] dataset, a large-scale, manually annotated benchmark providing *pixel-level 3D segmentations* aligned with corresponding *radiology report findings* from the CT-RATE [68] corpus. Unlike the above (semi-)synthetic datasets, this dataset enables explicit grounding of free-text clinical descriptions to precise 3D regions. We use the official training and validation splits respectively.

The full list of datasets used is given in Tab. 6. In contrast to standard semantic segmentation datasets, which assign voxels to pre-defined classes, our instance-focused dataset is designed to facilitate query-driven localization allowing models to identify specific, user-defined findings within complex anatomical contexts. Within the *VoxTell* architecture, this capability is reinforced through *iterative image–text fusion* during training, promoting the emergence of spatially grounded, instance-aware representations. By moving beyond class-level semantics, this complementary dataset supports clinically meaningful spatial reasoning, directly linking descriptive radiology language to voxel-level anatomical understanding.

[1] https://www.cancerimagingarchive.net/

C. Vocabulary Construction Details

We construct a unified vocabulary across 158 heterogeneous 3D medical segmentation datasets using an iterative pipeline comprising three components: (1) semantic label expansion, (2) cross-dataset harmonization, and (3) human expert validation. The final vocabulary contains 1,087 unified concepts and 9,682 rewritten labels. Figure 7 illustrates the pipeline. For the expansion and harmonization stages, we use Anthropic Claude Sonnet 4 [15] with extended thinking enabled.

1. **Semantic Label Expansion:**
For each dataset, we first consolidate instance-level annotations into semantic base labels (*e.g.* merging individual tumor instances into a single tumor concept). These base labels are then processed together with publicly available metadata, including published papers of the dataset, challenge documentation, dataset websites, and public ontologies when available. The instruction set in Fig. 9 generates:

- single-label alternatives for each semantic base label,
- combined-label alternatives for clinically meaningful groupings (*e.g.* combining individual rib labels into "rib cage"),
- a minimum of five validated variants per concept.

For example, a base label "kidney" is expanded to include variants such as "renal parenchyma", "renal organ", and "kidney tissue". Positional terms (*e.g.* "left", "anterior") must always include the anatomical structure name to ensure self-contained descriptions. Combined labels are generated only for established clinical groupings, such as merging left and right kidney labels into "kidneys" or "bilateral kidneys". Organ labels by default refer to the complete structure including any pathologies unless

Complexity	Dataset	Images	Modality	Prompt	TotalSegmentator	BiomeParse	Text3DSam	SegVol	BiomedParseV2	SAT	VoxTell
Healthy (Known Concept)	Pediatric-CT-SEG [84]	21	CT	gallbladder	78.91	8.74	7.25	71.51	84.63	80.25	85.28
				duodenum	53.16	2.68	6.47	51.56	59.05	62.17	66.97
				right adrenal gland	41.27	0.38	0.01	37.99	40.52	46.84	48.52
				left adrenal gland	42.53	1.17	2.01	45.27	44.32	52.71	52.57
				urinary bladder	83.45	38.17	35.63	81.41	81.47	82.59	79.96
				bones	90.21	0.74	2.35	83.8	0.01	8.64	93.62
				left breast	23.41	0	0	0	0	39.86	41.52
				right breast	24.01	0	0	0	0	39.91	45.07
				esophagus	61.67	8.17	25.46	59.11	64.16	66.48	65.96
				heart	81.99	2.9	68.02	4.49	85.69	90.29	48.95
				left kidney	94.17	22.86	59.28	93.05	96.22	96.12	95.31
				right kidney	94.45	18.77	28.35	92.17	94.36	94.26	94.7
				large intestine	69.54	3.25	7.29	6.52	5.69	11.98	63.88
				liver	95.61	35.62	76.64	95.76	96.28	95.65	95.56
				left lung	86.38	0	77.16	87.23	80.49	87.73	87.65
				right lung	86.13	0	79.83	89.77	84.57	88.22	87.65
				pancreas	78.1	11.66	14.3	72.62	78.68	78.81	77.6
				rectum	0	0.69	0	25.99	0	68.76	48.24
				skin	41.44	0.01	0	5.45	0	69.11	63.16
				small intestine	72.49	3.15	8.29	28.60	23.46	71.13	70.00
				spinal canal	73.77	0.04	0	4.02	0.25	79.71	87.3
				spleen	92.1	36.2	64.27	88.32	92.99	93.1	92.85
				stomach	90.82	14.47	48.37	82.79	78.18	77.87	85.23
	AeroPath [186]	27	CT	lung	96.44	0	68.14	97.28	54.27	96.93	96.99
				trachea	74.02	0.01	62.58	80.06	70.92	79.02	82.31
	VEELA2025 [189]	20	CT	liver	95.19	17.01	77.88	94.25	94.67	94.6	94.47
				hepatic vessels	37.77	2.59	5.81	29.76	6.76	9.37	41.17
				portal vessels	15.49	0	3.05	0	3.75	8.99	46.04
	RADCURE [210]	22	CT	brainstem	-	3.22	2.12	69.61	65.12	82.57	82.27
				left cochlea	-	0	0	0	0	18.15	19.58
				right cochlea	-	0	0	0	0	21.49	23.31
				esophagus	-	25.98	12.51	40.42	55.36	63.71	61.63
				left eye	-	1.73	0	40.64	11.54	87.47	85.2
				right eye	-	1	0	28.31	3.72	86.14	83.12
				left lacrimal gland	-	0	0	0	9.99	28.07	0
				right lacrimal gland	-	0.01	0	0.06	18.36	32.07	0
				left submandibular gland	-	0.71	0.71	27.91	59.66	76.65	77.14
				right submandibular gland	-	0.88	0.43	31.89	59.02	71.91	71.47
				larynx	-	0.72	0	2.79	17.15	47.14	39.55
				left lens	-	0.07	0	1.19	0	62.59	74.32
				right lens	-	0.03	0	0.88	0	63.66	74.74
				lips	-	0	0.23	20.8	13.05	26.62	25.39
				mandible	-	0.07	0.03	82.83	0	86.87	91.55
				superior constrictor muscle	-	0.79	0	0	0.16	32.95	3.37
				middle constrictor muscle	-	0.94	0	0	0	18.63	1.73
				inferior constrictor muscle	-	2.8	0.01	0	0.04	1.56	5.11
				left optic nerve	-	0.39	0.58	5.49	1.03	63.41	66.58
				right optic nerve	-	0.09	0.13	14.3	4.97	58.08	67.32
				chiasma opticum	-	0	0	0	0	39.49	40.14
				left parotid gland	-	0.12	1.72	1.14	59.85	82.84	83.51
				right parotid gland	-	0.15	0.04	5.03	54.75	82.9	83.71
				spinal cord	-	0.11	13.74	56.98	61.9	68.91	69.87
	HVSMR-2.0 [154]	60	CMR	left ventricle	-	26.63	34.9	0.05	17.8	53.54	46.91
				right ventricle	-	16.21	18.96	5.01	10.42	43.39	55.12
				left atrium	-	12.12	26.32	0	20.49	47.31	68.23
				right atrium	-	6.04	21.19	0.02	10.74	47.85	50.39
				aorta	-	1.31	5.9	5.76	44.78	53.71	70.68
				pulmonary artery	-	1.5	4.51	0	12.73	39.73	48.43
				superior vena cava	-	0.2	4.23	0.32	4.88	36.19	47.63
				inferior vena cava	-	0.93	9.18	3.63	23.03	25.3	34.4
Pathological (Known Concept)	ARCIN NSCLC [94]	44	PET	fdg-avid tumor	-	2.73	12.21	0	0.47	77.13	83.24
	ISBI MS [37]	42	MR (FLAIR)	MS lesion	-	9.57	0.21	0	2.03	13.68	72.71
	Adrenal-ACC-Ki67 [148]	52	CT	tumor	-	52.91	0	51.08	55.86	0.09	77.23
	HCC-TACE-SEG [147]	65	CT	hepatocellular carcinoma	-	41.25	0.52	58.35	70.37	62.28	73.24
	Pengwin [129]	100	CT	sacrum	-	1.05	15.24	64.55	3.21	95.33	96.48
				left hipbone	-	0.14	33.93	0.02	0.69	96.45	98.24
				right hipbone	-	0.21	25.09	1.75	0.15	96.37	98.05
	BrainMetShare [66]	105	MR (T1)	lesion	-	9.61	0.1	0	18.66	22.16	48.19
Cross-Modality	QIN Breast [118]	36	PET	fdg-avid breast cancer	-	0.75	0	0.05	0	58.26	72.27
	PANTHER Task 1 [27]	92	MR (T1)	pancreas tumor	-	1.39	4.52	14.64	18.24	19.25	35.66
	Soft-Tissue-Sarcoma [191]	102	MR (T1/T2)	soft tissue lesion	-	36.20	12.83	0.83	7.68	10.64	40.34
Unknown Concept	FedBCa (Center2) [36]	47	MR (T2)	carcinoma	-	10.37	2.04	0.08	2.69	9.56	25.76
	MedSeg Esophageal [1]	1	CT	esophageal cancer	-	17.86	7.28	32.24	16.56	0	69.07

Table 7. **Overview of held-out test datasets.** This table presents the full set of evaluation results across all datasets, including healthy and pathological known concepts, as well as generalization to unseen modalities and classes. Dataset metadata is also included. For partially annotated datasets such as Pediatric-CT-SEG and RADCURE-Structures, only the subset containing all annotated classes were used.

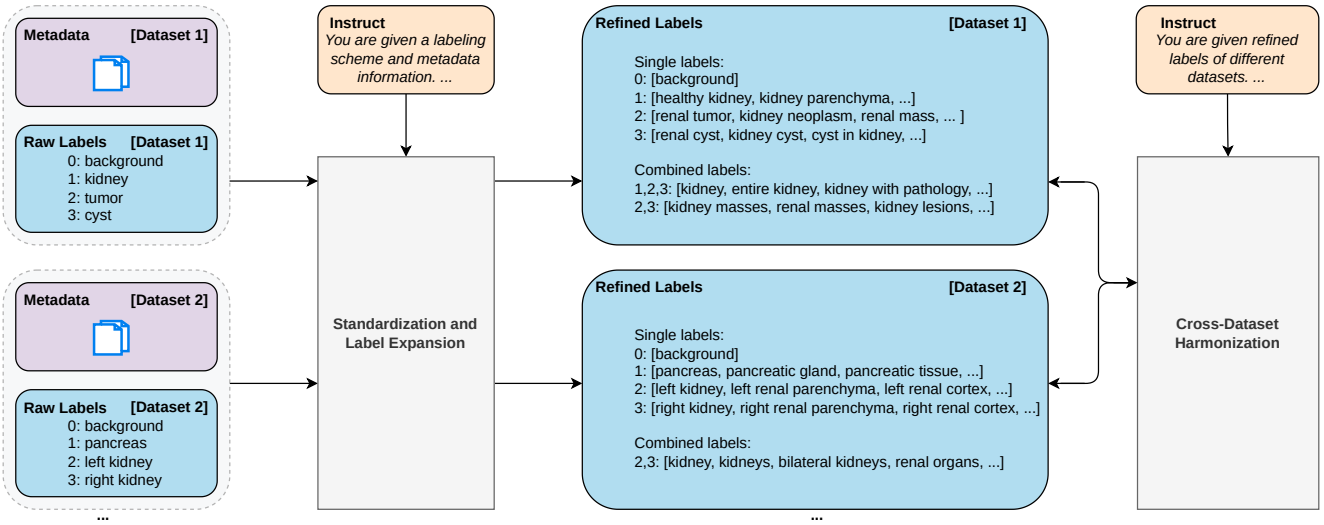

Figure 7. **Iterative pipeline for large-scale label standardization.** Each dataset provides raw labels and metadata, which are processed by a *Semantic Label Expansion* module to generate structured and semantically consistent label variants. A *Cross-Dataset Harmonization* module analyzes these expanded label sets to identify inconsistent or overlapping label definitions and proposes standardized mappings across datasets. Experts manually review proposed mappings, refine the expanded labels when necessary, and ensure that final label definitions remain faithful to the original dataset semantics.

explicitly specified otherwise (*e.g.* "liver parenchyma only").

2. **Cross-Dataset Harmonization:**
Identical label strings may denote different anatomical targets or coverage definitions across datasets. To resolve such inconsistencies, an LLM-based harmonization stage processes expanded label sets together with metadata and ontology-derived descriptors. The conflict-detection prompt in Fig. 8 outputs structured JSON specifying:

- conflict presence and severity (none / minor / major),
- anatomical discrepancy description,
- recommended harmonization action,
- affected datasets.

Major conflicts include differences in organ identity (*e.g.* "left ventricle" referring to myocardial wall versus blood-filled chamber), inclusion of pathologies (*e.g.* "liver" with or without tumors), or whole-organ versus subregion definitions. Minor conflicts reflect boundary variations or substructure inclusion differences. For instance, if "kidney" refers to healthy parenchyma in one dataset but includes tumors and cysts in another, these are maintained as distinct concepts with explicit descriptions. Ontology terms support this process by providing standardized anatomical references.

3. **Human Expert Review:**
All major-conflict cases and all final label variants are reviewed by a human expert, which verifies anatomical correctness, alignment with dataset semantics as described in published papers, and consistency within hierarchical groupings (*e.g.* ensuring "thoracic spine" variants align with individual vertebra labels). The reviewer can accept, modify, or reject any LLM-generated suggestion. All modifications are documented to ensure consistency across the final vocabulary. This validation step ensures that the vocabulary remains faithful to established medical terminology and prevents the introduction of spurious or anatomically inaccurate terms.

D. Text Embedding Model Selection

To identify an effective text encoder for our framework, we evaluate several state-of-the-art text embedding models from the MTEB[2] benchmark under consistent training and validation conditions. Given the large number of parameters in most text encoders and to preserve their pretrained knowledge, we freeze all encoder weights during training to efficiently apply it in downstream tasks. Candidate models are grouped by scale: small (300M–600M), medium (4B), and large (7B–8B). All encoders remain frozen during training, and instruction-tuned variants use the same instruct stated above.

As shown in Table 8, segmentation performance generally improves with model capacity. Small-scale models such as EmbeddingGemma and SAT Text Encoder exhibit

[2] http://mteb-leaderboard.hf.space/?benchmark_name=MTEB%28Medical%2C+v1%29

Text Embedding Model	Model Size	Dice
EmbeddingGemma [195]	300M	56.77
SAT Text Encoder [238]	450M	60.43
Qwen3-Embedding-0.6B [231]	600M	**62.24**
Qwen3-Embedding-4B [231]	4B	**62.55**
Jasper En Vision-Language v1 [227]	7B	58.42
SFR-Embedding-2 [146]	7B	62.08
GTE-Qwen2-7B-instruct [120]	7B	62.44
Linq-Embed-Mistral [92]	7B	62.78
E5-Mistral-7b-instruct [200, 201]	7B	**62.88**
Qwen3-Embedding-8B [231]	8B	62.45

Table 8. **Comparison of text embedding models.** We evaluate the top open-source text encoders from the MTEB benchmark [151] alongside the **SAT Text Encoder** [238], using identical training and validation data. Models are grouped by size (small, medium, large). Across all groups, **Qwen3-Embedding** shows consistently strong performance, achieving the highest overall Dice. Performance saturates below 63%, so we adopt the medium-scale variant to balance accuracy and computational efficiency.

noticeably lower Dice scores, while large-scale encoders (7B–8B) yield strong results but demand more than 24 GB of GPU memory, often far exceeding the capabilities of standard hospital hardware. Among medium-scale candidates, **Qwen3-Embedding-4B** attains performance comparable to the best large models, offering an optimal balance between accuracy and computational feasibility. Consequently, we adopt Qwen3-Embedding-4B as the default text encoder in VoxTell.

E. Extended Results

Results per prompt. Table 7 presents a comprehensive overview of segmentation performance across all held-out test datasets, extending the results reported in the main paper (Tab. 1). For healthy anatomical structures with known concepts, such as those in Pediatric-CT-SEG, AeroPath, VEELA2025, RADCURE, and HVSMR-2.0, VoxTell consistently achieves the highest Dice scores across the majority of organs and structures, often substantially outperforming prior methods such as TotalSegmentator, BiomeParse, Text3DSam, SegVol, BiomedParseV2, and SAT. Notably, VoxTell maintains strong performance on challenging, small, or laterality-sensitive structures such as adrenal glands, breast tissue, and the spinal canal, where competing methods frequently fail or produce near-zero predictions.

On pathological datasets with known concepts, including ARCIN NSCLC, ISBI MS, Adrenal-ACC-Ki67, HCC-TACE-SEG, Pengwin, and BrainMetShare, VoxTell demonstrates robust segmentation of tumors and lesions across CT, PET and MR modalities, achieving substantial improvements over previous approaches that often struggle with heterogeneous appearances or low contrast. In cross-modality scenarios, such as QIN Breast, PANTHER Task 1, and Soft-Tissue-Sarcoma, VoxTell exhibits superior generalization, successfully segmenting structures unseen during training despite modality shifts. Finally, for unknown concepts, including FedBCa and MedSeg Esophageal, VoxTell significantly outperforms all baseline methods, demonstrating its capacity for zero-shot anatomical understanding and reliable localization of rare or unseen structures. Overall, these extended results confirm that VoxTell achieves state-of-the-art performance across a diverse range of anatomical, pathological, and cross-modal segmentation tasks, highlighting its robustness, generalization, and practical applicability in clinical imaging.

Zero-shot segmentation on rare body regions. To probe the limits of open-vocabulary generalization, we evaluated VoxTell and prior text-promptable segmentation models on the Stanford Knee MRI dataset [50]. The knee region is sparsely represented in the training set and includes cartilage and meniscus structures absent from training data. As Table 9 shows, all models struggle on these highly out-of-distribution (OOD) structures, which differ both spatially and visually from familiar anatomy, though VoxTell is able to segment a few structures with lower performance.

Despite low absolute Dice scores, VoxTell partially recognizes key anatomical components such as femoral and tibial cartilage. This demonstrates the frontier of zero-shot medical segmentation: while VoxTell generalizes across modalities and related anatomical concepts (Tab. 3), completely unseen regions remain challenging. These findings suggest a promising path for combining vision–language pretraining with lightweight few-shot or report-level fine-tuning to close remaining gaps.

Structure	Patellar Cartilage	Femoral Cartilage	Lateral Tibial Cartilage	Medial Meniscus	Lateral Meniscus	Medial Tibial Cartilage	Mean Dice
BioMedParse	0.44	0.54	0.22	0.47	0.72	0.29	0.45
Text3DSam	0.0	0.08	0.09	0.97	0.1	0.31	0.26
SegVol	0.0	0.0	0.0	0.0	0.0	0.0	0.0
BioMedParseV2	0.0	0.15	0.0	0.0	0.0	0.0	0.02
SAT	0.01	1.34	0.02	0.0	0.0	0.0	0.23
VoxTell	0.07	46.45	2.97	0.04	0.0	14.78	10.72

Table 9. **Limitations of zero-shot segmentation on rare body regions.** This evaluation highlights VoxTell and prior text-promptable models on knee MRI (SKM-TEA dataset [50]), a modality and body region very rarely seen during training. Most prompts correspond to fully out-of-distribution (OOD) anatomical structures, making it extremely challenging for the model to interpolate. As shown, VoxTell achieves partial recognition of some structures but, like other methods, struggles to generalize to entirely novel spatial and visual patterns, illustrating the limitations discussed in the main paper.

```
Analyze if the label '{label_name}' has conflicting definitions across datasets - meaning it refers to different
    anatomical structures or body parts.
{dataset_definitions}
Focus: Does this label text refer to different parts of the body or include/exclude different anatomical components
    ?
Conflicts to detect:
1. Different anatomical structures entirely (e.g. 'mass' = lung mass in one dataset, liver mass in another)
2. Different inclusion of sub-structures (e.g. 'brain' with CSF vs without CSF)
3. Whole organ vs part of organ (e.g. 'liver' = whole organ vs 'liver' = 'parenchyma only')
4. Different pathology inclusions (e.g. 'kidney' with tumors vs without)
Not conflicts:
- Terminology variations for same structure ('hepatic' vs 'liver')
- Different alternative phrasings in the alternatives list
- Minor wording differences
Output (Json only):
{{
    'has_conflict': true/false,
    'conflict_severity': 'none'/'minor'/'major',
    'conflict_description': 'What anatomical difference exists, or 'No conflicts'',
    'recommendations': ['action 1', 'action 2'],
    'affected_datasets': ['Dataset001', ...]
}}
Severity:
- 'none': Same anatomical structure across all datasets
- 'minor': Small inclusion differences (e.g. organ edge definitions)
- 'major': Different anatomical structures or major component differences
JSON only, no other text.
```

Figure 8. **Instructions for cross-dataset conflict detection.** Label definitions are analyzed using these instructions to identify cases where identical text refers to different anatomical structures across datasets. The instructions distinguish genuine semantic conflicts (*e.g.* different organs, inclusion/exclusion of pathologies) from benign terminology variations. The structured output provides conflict severity assessment and resolution recommendations for systematic harmonization before multi-dataset training.

```
You are given a labeling scheme and metadata information.
The labeling has a format like the following example:

'labels': {
        'background': 0,
        'liver tumor': 1,
        'liver': [1, 2],
        ...
    },
...

where a text label is mapped to one or to a list of label ids.
Given the segmentation dataset metadata and label definitions, generate comprehensive and diverse alternative
     labels for every segmentation label in the dataset.json in a SINGLE, complete JSON response.
IMPORTANT: The output JSON must use numerical label IDs as keys, not the original label names. Map each ID to a
     list of alternative text labels.
The output format must be a single, complete JSON object:
{'single_labels': {
    '0': ['background'],
    '1': ['alternative_name_1', 'alternative_name_2', ...],
    '2': ['alternative_name_1', 'alternative_name_2', ...],
    ...
},
'combined_labels': {
    '1,2': ['combined_name_1', 'combined_name_2', ...],
    // Include meaningful combinations using numeric IDs
}}

Use the given information in the metadata to diversity the labels in a useful manner.
Critical for multi-dataset training:
1. Every alternative label must be fully self-contained and unambiguous across datasets
2. Positional terms (e.g. 'anterior', 'left') must never be used in isolation-always pair them with the anatomical
     structure (e.g. 'anterior hippocampus')
3. Always include the anatomical/structural name in every alternative (e.g. 'anterior hippocampus' not just '
     anterior')
4. For subregions or subdivisions, always maintain the parent structure in the label (e.g. 'hippocampal head' NOT
     just 'head')
5. Any abbreviation must be domain-specific and widely recognized (e.g. 'L. hippocampus' for left hippocampus is
     acceptable)
6. Labels must be uniquely identifiable even when used across multiple different dataset contexts
7. Don't assign duplicate labels to different structures or combinations (e.g. assigning 'liver' to both '2' and
     '1,2' is not allowed)
8. Organ labels always mean the whole organ including its components (e.g. lesions, tumors, vessels)
9. Critical: For grouped instances (e.g. 'mediastinal lymph nodes': [1, 2, 3, ...]):
   - Do not create individual entries for these IDs in 'single_labels'
   - Instead, add them directly to 'combined_labels' with the comma-separated IDs as the key
   - Example: For {'mediastinal lymph nodes': [1, 2, 3, ..., 56]}, add '1,2,3,...,56': ['mediastinal lymph nodes',
     ...] to 'combined_labels'
   - These IDs should not appear individually in 'single_labels'
Diversity requirements:
1. For each label, provide at least 5 diverse yet clinically accurate alternatives
2. Include variations in terminology across different medical contexts:
   - Formal anatomical terminology (e.g. 'hepatic tissue')
   - Clinical shorthand used in practice (e.g. 'HCC' for 'hepatocellular carcinoma')
   - Descriptive terms used in radiology reports
   - Common terms used when discussing with patients
3. Consider different ways to describe the same structure:
   - Positional descriptions with structure names (e.g. 'superior cerebellar peduncle', not just 'superior')
   - Functional descriptions where relevant, always with the structure name
   - Size/shape-based descriptions where appropriate, always with the structure name
Quality control:
1. Every alternative must be clinically accurate and actually used in medical contexts
2. Avoid artificially creating diversity by using uncommon or imprecise terminology
3. Only include terms that medical professionals would recognize and use
4. Maintain anatomical precision - alternatives must refer to the exact same structure
5. For combined labels, only create meaningful clinical groupings that would be referred to together. Do not
     combine labels with the background label.
6. Do not create alternatives for the background label.
Consistency guidelines:
1. Maintain parallel structure across related anatomical terms
2. Use consistent terminology for laterality (left/right) and positioning
3. When creating alternatives for a series of related structures, ensure naming patterns are consistent
4. Unqualified organ names (e.g. 'liver') naturally refer to the complete organ including any pathologies, unless
     explicitly specified otherwise

Verification:
Before returning your response:
- Ensure every label ID from dataset.json is represented in single_labels
- Verify that combined_labels ONLY reference label IDs that actually exist in single_labels
- Check that no non-existent label IDs are included in any combined label keys
- Confirm that each label has at least 5 diverse alternatives
- Verify that all alternatives accurately represent the original label
Only return the json, nothing else.
```

Figure 9. **Instructions for label standardization and expansion.** Each dataset's raw label definitions and metadata are processed using these instructions to generate a standardized JSON output containing: (1) diverse alternative labels for each label ID, ensuring anatomical precision and self-contained descriptions (*e.g.* "anterior hippocampus" not "anterior"), and (2) meaningful combined labels representing valid anatomical groupings. The instructions enforce at least 5 clinically accurate alternatives per label and prevent cross-dataset ambiguity to support robust text-conditioned segmentation.

Region	Subregion	Structures (Count)
Brain	Cerebral structures	brainstem (7360), brain (4330), left cerebral white matter (3654), right cerebral white matter (3654), left cerebral cortex (3644), right cerebral cortex (3644), white matter (519), gray matter (489), white matter hyperintensities (60), deep gray matter (40), cortical gray matter (30), white matter lesions (30)
	Ventricles and csf	cerebrospinal fluid (4133), left lateral ventricle (3654), right lateral ventricle (3654), left inferior lateral ventricle (3644), right inferior lateral ventricle (3644), 3rd ventricle (3644), 4th ventricle (3644), left brain lateral ventricle (158), right brain lateral ventricle (158), lateral ventricles (80), third ventricle (10), fourth ventricle (10)
	Basal ganglia	left caudate (3654), right caudate (3654), left putamen (3654), right putamen (3654), left pallidum (3654), right pallidum (3654), left accumbens area (3644), right accumbens area (3644), left caudate nucleus (158), right caudate nucleus (158), left lentiform nucleus (158), right lentiform nucleus (158), basal ganglia (30), left nucleus accumbens (10), right nucleus accumbens (10)
	Thalamus and diencephalon	left thalamus (3654), right thalamus (3654), left ventral dc (3644), right ventral dc (3644), left ventral diencephalon (10), right ventral diencephalon (10)
	Hippocampus	left hippocampus (3932), right hippocampus (3932), anterior hippocampus (260), posterior hippocampus (260), left hippocampus subiculum (20), left hippocampus cornu ammonis 2 (20), left hippocampus cornu ammonis 1 (20), left hippocampus cornu ammonis 4 and gyrus dentatus (20), left hippocampus entorhinal cortex (20), left hippocampus cornu ammonis 3 (20), left hippocampus tail (20), right hippocampus subiculum (20), right hippocampus cornu ammonis 2 (20), right hippocampus cornu ammonis 1 (20), right hippocampus cornu ammonis 4 and gyrus dentatus (20), right hippocampus entorhinal cortex (20), right hippocampus cornu ammonis 3 (20), right hippocampus tail (20)
	Amygdala and limbic	left amygdala (3654), right amygdala (3654), corpus callosum (10)
	Cerebellum	left cerebellum white matter (3644), right cerebellum white matter (3644), left cerebellum cortex (3644), right cerebellum cortex (3644), cerebellum (70), cerebellar vermis superior posterior (10), cerebellar vermis anterior (10), cerebellar vermis inferior posterior (10), left cerebellar hemisphere (10), right cerebellar hemisphere (10)
	Frontal lobe	left caudal anterior cingulate (10), left caudal middle frontal (10), left lateral orbitofrontal (10), left medial orbitofrontal (10), left paracentral lobule (10), left pars opercularis (10), left pars orbitalis (10), left pars triangularis (10), left precentral gyrus (10), left rostral anterior cingulate (10), left rostral middle frontal (10), left superior frontal (10), left frontal pole (10), right caudal anterior cingulate (10), right caudal middle frontal (10), right lateral orbitofrontal (10), right medial orbitofrontal (10), right paracentral lobule (10), right pars opercularis (10), right pars orbitalis (10), right pars triangularis (10), right precentral gyrus (10), right rostral anterior cingulate (10), right rostral middle frontal (10), right superior frontal (10), right frontal pole (10)
	Temporal lobe	left temporal lobe (120), right temporal lobe (120), left temporal lobes (50), right temporal lobes (50), left banks superior temporal sulcus (10), left entorhinal (10), left fusiform gyrus (10), left inferior temporal (10), left middle temporal (10), left parahippocampal gyrus (10), left superior temporal (10), left temporal pole (10), left transverse temporal (10), right banks superior temporal sulcus (10), right entorhinal (10), right fusiform gyrus (10), right inferior temporal (10), right middle temporal (10), right parahippocampal gyrus (10), right superior temporal (10), right temporal pole (10), right transverse temporal (10)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
	Parietal lobe	left inferior parietal (10), left postcentral gyrus (10), left posterior cingulate (10), left precuneus (10), left superior parietal (10), left supramarginal gyrus (10), right inferior parietal (10), right postcentral gyrus (10), right posterior cingulate (10), right precuneus (10), right superior parietal (10), right supramarginal gyrus (10)
	Occipital lobe	left cuneus (10), left lateral occipital (10), left lingual gyrus (10), left pericalcarine (10), right cuneus (10), right lateral occipital (10), right lingual gyrus (10), right pericalcarine (10)
	Insula and cingulate	left isthmus cingulate (10), left insula (10), right isthmus cingulate (10), right insula (10)
	Other cortical	extra-axial space (40), left unknown cortex (10), right unknown cortex (10)
	Anterior circulation arteries	right internal carotid artery (833), left internal carotid artery (833), anterior communicating artery (300), right anterior cerebral artery (250), left anterior cerebral artery (250), right middle cerebral artery (250), left middle cerebral artery (250), third segment of anterior cerebral artery (a2) (250), right a1 and a2 of anterior cerebral artery (50), left a1 and a2 of anterior cerebral artery (50), right a3 of anterior cerebral artery (50), left a3 of anterior cerebral artery (50), third a2 of anterior cerebral artery (50), third a3 of anterior cerebral artery (50), right m1 of middle cerebral artery (50), left m1 of middle cerebral artery (50), right m2 of middle cerebral artery (50), right m3 of middle cerebral artery (50), left m2 of middle cerebral artery (50), left m3 of middle cerebral artery (50), paraopthalmic part of right internal carotid artery (1), posterior communicating part of right internal carotid artery (1), anterior choroidal part of the right internal carotid artery (1), right carotid terminus (1), right m1 (1), right a1 (1), anterior communicating artery partially fenestrated (1), right a2 (1), left paraopthalmic part of the internal carotid (1), left posterior communicating part of the internal carotid (1), left anterior choroidal artery part of the internal carotid (1), left carotid terminus (1), left a1 (1), left a2 (1), left m1 (1), insular segment of the right middle cerebral artery (1), insular branches of the left middle cerebral artery (m3/m4) (1), right anterior insular part of middle cerebral artery (m3/m4) (1), right posterior insular part of the middle cerebral artery (1), right superior terminal branch of the middle cerebral artery (1), right parietal cortical branches of the insular part of the middle cerebral artery (1), left m2-branches (1), left m3/m4 (1)
	Posterior circulation arteries	basilar artery (302), right posterior cerebral artery (250), left posterior cerebral artery (250), right superior cerebellar artery (51), left superior cerebellar artery (51), right vertebral artery (50), left vertebral artery (50), right p1 and p2 of posterior cerebral artery (50), left p1 and p2 of posterior cerebral artery (50), right p3 and p4 of posterior cerebral artery (50), left p3 and p4 of posterior cerebral artery (50), right anterior inferior cerebellar artery (50), left anterior inferior cerebellar artery (50), right posterior inferior cerebellar artery (50), left posterior inferior cerebellar artery (50), right p1 (1), right p2 (1), left p1 (1), left p2 (1), left p3 (1), left p4/p5 (1)
	Circle of willis	right posterior communicating artery (301), left posterior communicating artery (301), right anterior choroidal artery (51), left anterior choroidal artery (51)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
	Small cerebral arteries	right ophthalmic artery (50), left ophthalmic artery (50), left opthalmic artery (1), right opthalmic artery (1), right lenticolstriate arteries (1), left lenticulostriate arteries (1), right posteromedial central arteries (1), left posteromedial central arteries (1), right paramedian artery branch (1), right medial frontobasal artery (1), right lateral frontobasal artery (1), left medial orbitofrontal artery (1), left polar frontal artery (1), right polar frontal artery (1), left pericallosal artery (1), left callosomarginal artery (a3/4) (1), right callosomarginal artery (a3/4) (1), right anterior internal frontal artery (1), left anterior internal frontal artery (1), right intermediate internal frontal artery (1), left intermediate internal frontal artery (1), right prefrontal artery (1), right artery of precentral gyrus (1), right artery of central sulcus (1), right polar temporal artery (1), right temporal branch (1), right anterior temporal artery branch (1), right middle temporal artery branch (1), right posterior parietal artery (1), right posterior temporal artery branch of the mca (1), right parietal branches of the middle cerebral artery (1), right posterior temporal branches of the pca (1), right medial occipital artery (1), right precuneal branches (1), right calcarine branches (1), right parietooccipital branches (1)
	Cerebral veins and sinuses	superior sagittal sinus (26), vein of galen (25), straight sinus (25), internal cerebral veins (25), right basal vein of rosenthal (25), left basal vein of rosenthal (25)
	Meningeal vessels	right middle meningeal artery (26), left middle meningeal artery (26)
	Brain pathology	enhancing tumor (4681), edema (2840), necrosis/non-enhancing tumor core (2183), non-enhancing tumor core (1943), surrounding non-enhancing hyperintensity (1943), resection cavity (1943), brain lesion (655), gross tumor volume (500), stroke (453), multiple sclerosis lesion (366), necrotic tumor (294), non-enhancing tumor (261), cystic component (261), brain tumor (186), hie lesion (85), cerebral microbleed (72), infarction (30)
Head and neck	Eyes and orbits	left eye (3505), right eye (3505), left lens (3505), right lens (3505), left eyeball (533), right eyeball (533), eyeball globe (400), orbital fat (400), left anterior eyeball (84), right anterior eyeball (84), left posterior eyeball (84), right posterior eyeball (84)
	Extraocular muscles	eye superior rectus muscle (400), eye lateral rectus muscle (400), eye medial rectus muscle (400), eye inferior rectus muscle (400), eye superior oblique muscle (400)
	Optic pathway	left optic nerve (3636), right optic nerve (3636), chiasma opticum (3636), optic nerve (400)
	Lacrimal system	left lacrimal gland (3419), right lacrimal gland (3419), lacrimal gland (400)
	Facial structures	left cheek (533), right cheek (533)
	Inner ear	left cochlea (3539), right cochlea (3539), cochlea (210), left vestibular semicircular canal (120), right vestibular semicircular canal (120), left inner ear (50), right inner ear (50)
	Middle ear	left middle ear (170), right middle ear (170), left tympanic cavity (120), right tympanic cavity (120), left eustachian tube bone (120), right eustachian tube bone (120)
	Temporal bone	left internal auditory canal (120), right internal auditory canal (120), left mastoid process (120), right mastoid process (120)
	Cranial bones	mandible (3466), skull (3358), lower jawbone (532), upper jawbone (532), left inferior alveolar canal (532), right inferior alveolar canal (532), left mandibular incisive canal (532), right mandibular incisive canal (532), lingual canal (532), left mandible (170), right mandible (170), left temporomandibular joint (170), right temporomandibular joint (170)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
	Paranasal sinuses	nasal cavity (533), left maxillary sinus (532), right maxillary sinus (532)
	Upper teeth	upper right central incisor (532), upper right lateral incisor (532), upper right canine (532), upper right first premolar (532), upper right second premolar (532), upper right first molar (532), upper right second molar (532), upper right third molar (wisdom tooth) (532), upper left central incisor (532), upper left lateral incisor (532), upper left canine (532), upper left first premolar (532), upper left second premolar (532), upper left first molar (532), upper left second molar (532), upper left third molar (wisdom tooth) (532)
	Lower teeth	lower left central incisor (532), lower left lateral incisor (532), lower left canine (532), lower left first premolar (532), lower left second premolar (532), lower left first molar (532), lower left second molar (532), lower left third molar (wisdom tooth) (532), lower right central incisor (532), lower right lateral incisor (532), lower right canine (532), lower right first premolar (532), lower right second premolar (532), lower right first molar (532), lower right second molar (532), lower right third molar (wisdom tooth) (532)
	Tooth pulp	upper right central incisor pulp (532), upper right lateral incisor pulp (532), upper right canine pulp (532), upper right first premolar pulp (532), upper right second premolar pulp (532), upper right first molar pulp (532), upper right second molar pulp (532), upper right third molar (wisdom tooth) pulp (532), upper left central incisor pulp (532), upper left lateral incisor pulp (532), upper left canine pulp (532), upper left first premolar pulp (532), upper left second premolar pulp (532), upper left first molar pulp (532), upper left second molar pulp (532), upper left third molar (wisdom tooth) pulp (532), lower left central incisor pulp (532), lower left lateral incisor pulp (532), lower left canine pulp (532), lower left first premolar pulp (532), lower left second premolar pulp (532), lower left first molar pulp (532), lower left second molar pulp (532), lower left third molar (wisdom tooth) pulp (532), lower right central incisor pulp (532), lower right lateral incisor pulp (532), lower right canine pulp (532), lower right first premolar pulp (532), lower right second premolar pulp (532), lower right first molar pulp (532), lower right second molar pulp (532), lower right third molar (wisdom tooth) pulp (532)
	Dental prosthetics	bridge (532), crown (532), implant (532)
	Oral soft tissues	lips (3419), oral cavity (204), buccal mucosa (84)
	Salivary glands	left parotid gland (3636), right parotid gland (3636), left submandibular gland (3586), right submandibular gland (3586)
	Thyroid and pituitary	left thyroid (2226), right thyroid (2226), thyroid gland (1446), pituitary gland (254)
	Pharynx	pharynx (532), pharyngeal constrictor muscle (120), nasopharynx cancer (50)
	Larynx	larynx (3455), larynx glottis (204), larynx supraglottic (204), arytenoid cartilage (84), cricopharyngeal inlet (84)
	Carotid system	left common carotid artery (1931), right common carotid artery (1761), left carotid artery (98), right carotid artery (98), right external carotid artery (25), left external carotid artery (25), right superficial temporal artery (25), left superficial temporal artery (25), right maxillary artery (25), left maxillary artery (25), common carotid artery (1), internal carotid artery (1), external carotid artery (1)
	Jugular venous system	right internal jugular vein (533), left internal jugular vein (533), internal jugular vein (1), external jugular vein (1), posterior auricular vein (1), retromandibular vein (posterior branch) (1), facial vein (1), transverse cervical veins (1), superficial temporal veins (1), middle temporal vein (1), deep superficial veins (1), retromandibular vein (anterior branch) (1)
	Subclavian system	brachiocephalic vein (1), subclavian vein (1)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
	Other neck structures	head neck lymph node (245)
	Head neck pathology	head neck primary tumor (674), head neck metastatic lymph nodes (674), head neck lesion (245), vestibular schwannoma (210)
Chest	Lung	left lung (4404), right lung (4404), lung (3427)
	Lung lobes	left lung lower lobe (3454), right lung lower lobe (3454), right lung middle lobe (3454), left lung upper lobe (3454), right lung upper lobe (3454)
	Airways	trachea (2018), bronchi (583), airways (162)
	Ribs left	1st left rib (4108), 2nd left rib (4108), 3rd left rib (4108), 4th left rib (4108), 5th left rib (4108), 6th left rib (4108), 7th left rib (4108), 8th left rib (4108), 9th left rib (4108), 10th left rib (4108), 11th left rib (4108), 12th left rib (4108), 13th left rib (1597)
	Ribs right	1st right rib (4108), 2nd right rib (4108), 3rd right rib (4108), 4th right rib (4108), 5th right rib (4108), 6th right rib (4108), 7th right rib (4108), 8th right rib (4108), 9th right rib (4108), 10th right rib (4108), 11th right rib (4108), 12th right rib (4108), 13th right rib (1597), ribs (654)
	Sternum and clavicles	left clavicula (3974), right clavicula (3974), sternum (3454), rib cartilage (1761), sternum manubrium (533)
	Scapulae	left scapula (3974), right scapula (3974)
	Breast	left breast (1086), right breast (1086)
	Breast pathology	breast lesion (1506)
	Mediastinum	thymus (886), mediastinal tissue (533), mediastinal lymph node (393), mediastinal lymph nodes (192)
	Esophagus	esophagus (7549), cervical esophagus (84)
	Thoracic aorta	left subclavian artery (1931), brachiocephalic artery (1812), right subclavian artery (1761), subclavian artery (1)
	Thoracic veins	left brachiocephalic vein (1761), right brachiocephalic vein (1761), brachiocephalic veins (15), subclavian and carotid arteries (15), azygos (15), azygos vein (2), vertebral artery (1)
	Pleura	pleura lesion (245), pleural effusion (78)
	Lung pathology	lung lesion (1075), lung nodule (481), covid (209), lung tumor (113), non-small cell lung cancer (78), fdg-avid non-small cell lung cancer (69), fdg-avid lung adenocarcinoma (5)
Heart	Heart	heart (2855), heart tissue (533)
	Atria	left atrium (2516), right atrium (2386), left atrial appendage (1228)
	Ventricles	right ventricle (2800), left ventricular cavity (2503), left ventricle (703), right ventricular cavity (279)
	Atrial walls	left atrium wall (154), left atrium endocardium (154), atrium wall (10)
	Ventricular walls	left ventricle wall (3234)
	Myocardium	left ventricular myocardium (849), myocardium (700), excised myocardium (27)
	Valves	aortic valve (27), mitral valve (27)
	Coronary arteries	coronary artery (1000), left coronary artery (30), right coronary artery (30)
	Pulmonary circulation	pulmonary artery trunk (1713), pulmonary vein (1349), pulmonary artery (593), pulmonary artery tree (316), left atrium and pulmonary veins (137), pulmonary arterial tree (15)
	Systemic cardiac vessels	superior vena cava (1289), aorta with arch branches (110), aorta with branches and coronary arteries (27)
	Cardiac pathology	myocardial infarction (100), no reflow (100), left atrial scars (60), calcification in left anterior descending (32), calcification in left circumflex (32), calcification in right coronary artery (32)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
Abdomen	Liver parenchyma	liver (8669), liver without tumor or vessels (20)
	Liver segments	liver segment 1 (243), liver segment 2 (243), liver segment 3 (243), liver segment 5 (243), liver segment 6 (243), liver segment 7 (243), liver segment 8 (243), liver segment 4 (193), liver segment 4a (50), liver segment 4b (50)
	Liver pathology	liver lesion (4469), hepatic tumor (303), hepatocellular carcinoma (179), colorectal liver metastases (171), liver cancer (131), liver tumor (114), liver cyst (20)
	Pancreas	pancreas (7934), pancreas parenchyma (482), pancreatic duct (482)
	Pancreas pathology	pancreatic ductal adenocarcinoma lesion (482), pancreas tumor (423)
	Pancreas vessels	pancreatic veins (482), pancreatic arteries (482)
	Spleen	spleen (6823)
	Biliary system	gallbladder (5606), common bile duct (482), biliary system (20)
	Stomach and duodenum	stomach (5840), duodenum (4073)
	Small intestine	jejunum and ileum (3600), small bowel (50)
	Large intestine	colon (3506)
	Colon pathology	colon cancer primaries (126)
	Kidneys	left kidney (5666), right kidney (5603), kidney (3972)
	Kidney pathology	left kidney cyst (1228), right kidney cyst (1228), kidney tumor (559), kidney cyst (489), kidney masses (489)
	Adrenal glands	left adrenal gland (5423), right adrenal gland (5423), adrenal glands (140), left adrenal tumor (20), right adrenal tumor (20)
	Aorta abdominal	aorta (5057), descending aorta (135), aortic arch (135), abdominal aorta (120), abdominal aorta segment 1 (120), ascending aorta (segment) (120), aortic arch false lumen (120), abdominal aorta false lumen (120), abdominal aorta segment 1 false lumen (120), ascending aorta false lumen (120), aortic intimal tear (120), aorta vessel tree (56), thoracic aorta (30)
	Celiac axis	celiac trunk (533), splenic artery (121), celiac trunk / artery of the gut (120), celiac artery (50), coeliac trunk (1), splenic vein (1), left gastric artery (1), right gastric artery (1), short gastric veins (1), right gastroepiploic vein (1)
	Hepatic vessels	hepatic vessels (494), hepatic artery (120), common hepatic artery (1), gastroduodenal artery (1), hepatic artery proper (1), right hepatic artery (1), left hepatic artery (1), separate segment 5/6 vein (variant) (1), right hepatic vein (segment 7 only here, variant) (1), middle hepatic vein (1), left hepatic vein (1)
	Portal venous system	portal vein and splenic vein (2947), portal splenic vein (1597), portal veins (191), portal vein (21), venous system (20), right portal vein (1), left portal vein (1), jejunal branches of the superior mesenteric vein (1), ileal branches of the superior mesenteric vein (1), superior posterior pancreaticoduodenal vein (1), right colic vein (1), middle colic vein (1), pancreaticoduodenal veins (plexus) (1)
	Mesenteric vessels	superior mesenteric artery (171), inferior mesenteric artery (2), inferior mesenteric vein (2), superior mesenteric vein (1)
	Renal vessels	left renal artery (171), right renal artery (171), renal artery (70), renal vein (70), left renal vein (1), right renal vein (1)
	Inferior vena cava	inferior vena cava (5016)
	Gonadal vessels	left gonadal vein (2), right gonadal vein (1), a double/fenestrated right gonadal (1)
	Other abdominal vessels	left ascending lumbar vein (1), hemiazygos vein (1)
	Abdominal wall	abdominal tissue (533), fat (533)
	Other abdominal	stones (20), artery (20), metal (20), metastasectomy site (20)
Pelvis	Bladder	urinary bladder (5098), bladder (126)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
	Prostate	prostate (3003), prostate peripheral zone (269), prostate central gland (237), prostate transitional zone (32)
	Female reproductive	cervix (949), uterus (583), prostate and uterus (360)
	Gonads	gonads (886)
	Rectum	rectum (1281)
	Pelvic bones	sacrum (5720), left hip bone (5720), right hip bone (5720), left femoral head (604), right femoral head (604)
	Iliac vessels	left iliac artery (4514), right iliac artery (4514), left iliac vena (4514), right iliac vena (4514), right common iliac artery (172), left common iliac artery (172), right internal iliac artery (51), left internal iliac artery (51), right external iliac artery (51), left external iliac artery (51), right common iliac vein (2), left common iliac vein (1), right external iliac vein (1), left external iliac vein (1), right internal iliac vein (1), left internal iliac vein (1), right inferior epigastric artery (1), left inferior epigastric artery (1), right epigastric vein (1), left epigastric vein (1), right superficial epigastric vein(s) (1), left superficial epigastric vein(s) (1), middle sacral artery (1), middle sacral vein (1), right deep circumflex iliac artery (1), left deep circumflex iliac artery (1)
	Gluteal vessels	right superior gluteal vein (1), left superior gluteal vein (1), right inferior gluteal vein (1), left inferior gluteal vein (1), right inferior gluteal artery (1), left inferior gluteal artery (1), left superior gluteal artery (1)
	Femoral vessels	right femoral artery (121), left femoral artery (121), right common femoral artery (1), left common femoral artery (1), right common femoral vein (1), left common femoral vein (1), right femoral vein (1), left femoral vein (1), right deep femoral vein (1), right deep femoral artery (1), left deep femoral artery (1), right great saphenous vein (1), left great saphenous vein (1), left lateral circumflex vein (1), left lateral circumflex artery (1), right lateral circumflex artery (1), descending branch of the right lateral circumflex artery (1)
	Pelvic visceral vessels	right inferior vesical veins (1), left inferior vesical veins (1), right obturator veins (1), left obturator veins (1), right internal pudendal artery and vein (1), left internal pudendal artery and vein (1), right obturator artery (1), left obturator artery (1), right pubic veins (1), left ureter (1), right ureter (1), vesical plexus vein branch (only left side shown) (1)
	Rectal vessels	superior rectal artery (1), left colic artery (1), superior sigmoid artery (part of the sigmoid branches) (1), superior rectal vein (with branches) (1), sigmoid veins (1), left colic vein (1), left sacral veins (1), inferior rectal veins (1), right inferior rectal veins (1), left inferior rectal vein (1)
	Pelvic pathology	fdg-avid tumor (1159), psma-avid tumor (597), prostate tumor (188), clinically insignificant prostate lesion (140), clinically significant prostate lesion (140)
Spine	Cervical vertebrae	c1 vertebra (4504), c2 vertebra (4504), c3 vertebra (4504), c4 vertebra (4504), c5 vertebra (4504), c6 vertebra (4504), c7 vertebra (4504)
	Thoracic vertebrae	t9 vertebra (5546), t10 vertebra (5546), t11 vertebra (5546), t12 vertebra (5546), t7 vertebra (5099), t8 vertebra (5099), t1 vertebra (4559), t2 vertebra (4559), t3 vertebra (4559), t4 vertebra (4559), t5 vertebra (4559), t6 vertebra (4559), t13 vertebra (80)
	Lumbar vertebrae	l1 vertebra (5546), l2 vertebra (5546), l3 vertebra (5546), l4 vertebra (5546), l5 vertebra (5546), l6 vertebra (2743)
	Sacral	s1 vertebra (1228)
	Spinal cord and canal	spinal cord (5433), spinal canal (1383)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
	Intervertebral discs	intervertebral discs (616), l5 s1 disc (447), l4 l5 disc (447), l3 l4 disc (447), l2 l3 disc (447), l1 l2 disc (447), t12 l1 disc (447), t11 t12 disc (447), t10 t11 disc (447), t9 t10 disc (447), intervertebra disc (16)
	Other spine	vertebrae (616)
Extremities	Upper arm	left humerus (3974), right humerus (3974)
	Forearm	left radius (1597), right radius (1597), left ulna (1597), right ulna (1597)
	Hand	carpal left (1597), carpal right (1597), left metacarpal (1597), right metacarpal (1597), left fingers (1597), right fingers (1597)
	Femur	left femur (3974), right femur (3974), femur (100), left femur bone (78), right femur bone (78)
	Patella	left patella (1597), right patella (1597), left patella bone (78), right patella bone (78)
	Leg bones	left tibia (1597), right tibia (1597), left fibula (1597), right fibula (1597), tibia (100), left tibia bone (78), right tibia bone (78), left fibula bone (78), right fibula bone (78)
	Ankle bones	left tarsal (1597), right tarsal (1597), left calcaneous bone (78), right calcaneous bone (78), left talus bone (78), right talus bone (78)
	Foot bones	left metatarsal (1597), right metatarsal (1597), left toes (1597), right toes (1597), left cuboid bone (78), right cuboid bone (78), left intermediate cuneiform bone (78), right intermediate cuneiform bone (78), left lateral cuneiform bone (78), right lateral cuneiform bone (78), left medial cuneiform bone (78), right medial cuneiform bone (78), left navicular bone (78), right navicular bone (78), left phalanges bone (78), right phalanges bone (78)
	Pelvic bones	left pelvis bone (78), right pelvis bone (78)
	Femoral cartilage	femoral articular cartilage (100), left femur distal cartilage (78), right femur distal cartilage (78), left femur head cartilage (78), right femur head cartilage (78)
	Patellar cartilage	left patella cartilage (78), right patella cartilage (78)
	Tibial cartilage	tibial articular cartilage (100), left tibia distal cartilage (78), right tibia distal cartilage (78), left tibia lateral cartilage (78), right tibia lateral cartilage (78), left tibia medial cartilage (78), right tibia medial cartilage (78)
	Other cartilage	left pelvis acetabulum cartilage (78), right pelvis acetabulum cartilage (78), left talus cartilage (78), right talus cartilage (78)
	Knee ligaments	left acl ligament (78), right acl ligament (78), left lcl ligament (78), right lcl ligament (78), left mcl ligament (78), right mcl ligament (78), left pcl ligament (78), right pcl ligament (78)
Musculo-skeletal	Paraspinal muscles	left autochthon (4514), right autochthon (4514), left longissimus thoracis (97), right longissimus thoracis (97), right spinalis thoracis (1), left spinalis thoracis (1), right ilicostalis lumborum (1), left iliocostalis lumborum (1)
	Abdominal muscles	right quadratus lumborum (1), left quadratus lumborum (1), right rectus abdominis (1), left rectus abdominis (1), right transversus abdominis (1), left transversus abdominis (1), right external oblique (1), left external oblique (1), right internal oblique (1), left internal oblique (1)
	Gluteal muscles	left gluteus maximus (4611), right gluteus maximus (4611), left gluteus medius (4611), right gluteus medius (4611), left gluteus minimus (4071), right gluteus minimus (4071), left gluteus maximus muscle (78), right gluteus maximus muscle (78), left gluteus medius muscle (78), right gluteus medius muscle (78), left gluteus minimus muscle (78), right gluteus minimus muscle (78)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
	Hip flexors	left iliopsoas (4610), right iliopsoas (4610), right psoas major muscle (78), left psoas major muscle (78), right illiacus muscle (78), left illiacus muscle (78), right psoas (1), left psoas (1), right iliacus (1), left iliacus (1)
	Hip rotators	right piriformis muscle (78), left piriformis muscle (78), right superior gemellus muscle (78), left superior gemellus muscle (78), right inferior gemellus muscle (78), left inferior gemellus muscle (78), right obturator internus muscle (78), left obturator internus muscle (78), right obturator externus muscle (78), left obturator externus muscle (78), left quadratis femoris muscle (78), right quadratus femoris muscle (78), right piriformis (1), left piriformis (1), right superior gemellus (1), left gemellus superior (1), right inferior gemellus (1), left gemellus inferior (1), right obturator internus (1), left obturator internus (1), right obturator externus (1), left obturator externus (1), right quadratus femoris (1)
	Hip abductors	right tensor fasciae latae muscle (78), left tensor fasciae latae muscle (78), right tensor fascia lata (1), left tensor fascia lata (1)
	Pelvic floor	right ischiocavernosus (1), left ischiocavernosus (1)
	Quadriceps	right rectus femoris muscle (78), left rectus femoris muscle (78), right vastus intermedius muscle (78), left vastus intermedius muscle (78), right vastus lateralis muscle (78), left vastus lateralis muscle (78), left vastus medialis muscle (78), right vastus medialis muscle (78), right rectus femoris (1), left rectus femoris (1), right vastus intermedius (1), left vastus intermedius (1), right vastus lateralis (1), left vastus lateralis (1)
	Other anterior thigh	right sartorius muscle (78), left sartorius muscle (78), right sartorius (1), left sartorius (1)
	Adductors	right pectineus muscle (78), left pectineus muscle (78), right gracilis muscle (78), left gracilis muscle (78), right adductor longus muscle (78), left adductor longus muscle (78), right adductor brevis muscle (78), left adductor brevis muscle (78), right adductor magnus muscle (78), left adductor magnus muscle (78), right pectineus (1), left pectineus (1), right gracilis (1), left gracilis (1), right adductor longus (1), left adductor longus (1), right adductor brevis (1), left adductor brevis (1), right adductor minimus (1), left adductor minimus (1), right adductor magnus (1), left adductor magnus (1)
	Hamstrings	right biceps femoris long head muscle (78), left biceps femoris long head muscle (78), right biceps femoris short head muscle (78), left biceps femoris short head muscle (78), right semitendinosus muscle (78), left semitendinosus muscle (78), right semimembranosus muscle (78), left semimembranosus muscle (78), right biceps femoris (long head) (1), left biceps femoris (long head) (1), right semitendinosus (1), left semitendinosus (1), right semimembranosus (1), left semimembranosus (1)
	Calf posterior	left gastrocnemius lateral muscle (78), right gastrocnemius lateral muscle (78), left gastrocnemius medial muscle (78), right gastrocnemius medial muscle (78), left soleus muscle (78), right soleus muscle (78), left plantaris muscle (78), right plantaris muscle (78), left popliteus muscle (78), right popliteus muscle (78)
	Calf anterior	left tibialis anterior muscle (78), right tibialis anterior muscle (78), left extensor digitorum longus muscle (78), right extensor digitorum longus muscle (78), left extensor hallucis longus muscle (78), right extensor hallucis longus muscle (78)
	Calf deep posterior	left tibialis posterior muscle (78), right tibialis posterior muscle (78), left flexor digitorum longus muscle (78), right flexor digitorum longus muscle (78), left flexor hallucis longus muscle (78), right flexor hallucis longus muscle (78)
	Calf lateral	left peroneus longus muscle (78), right peroneus longus muscle (78)

(continued on next page)

(continued from previous page)

Region	Subregion	Structures (Count)
	General musculoskeletal	bones (2070), skin (886), muscles (533)
Vascular	Aorta and branches	aortic arch (135), descending aorta (135), abdominal aorta (120), innominate artery (120), innominate artery segment 1 (120), innominate artery segment 1 false lumen (120), right common iliac artery false lumen (120), left common iliac artery false lumen (120), left common carotid artery false lumen (120), left subclavian artery false lumen (120), descending aorta false lumen (120), ascending aorta (101), aorta zone 0 (50), aorta zone 1 (50), aorta zone 2 (50), aorta zone 3 (50), aorta zone 4 (50), aorta zone 5 (50), aorta zone 6 (50), aorta zone 7 (50), aorta zone 8 (50), aorta zone 9 (50), thoracic aorta (30)
	Venous system	inferior vena cava (5016), superior vena cava (1289), portal vein (21), azygos vein (2), splenic vein (1)
	Vascular pathology	true lumen (100), false lumen (100), false lumen thrombus (100), false aortic lumen (39), true aortic lumen (39), pulmonary embolism (35)
General pathology	Tumors and lesions unspecified	lesion (1093), malignant melanoma tumor (670), fdg-avid soft tissue sarcoma (42), tumor (unspecified) (20)

Table 10. **Summary of anatomical structures.** Overview of all 1078 distinct anatomical labels from public training datasets. Each entry lists the structure name and its training case frequency (in parentheses), organized by anatomical region and subregion.

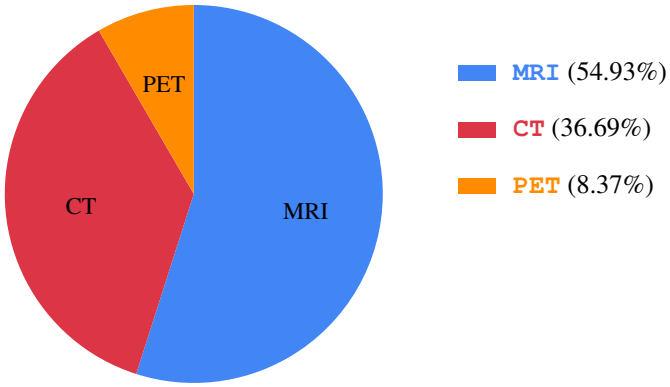

Figure 10. Distribution of medical imaging modalities in the semantic train datasets.

Region	Masks	Labels
Chest	181,095	74
Brain	155,141	280
Head and Neck	131,434	177
Spine	135,132	41
Abdomen	106,661	118
Extremities	55,330	80
Musculoskeletal	54,504	150
Pelvis	52,323	93
Heart	27,950	36
Vascular	8,723	34
General Pathology	1,825	4
Total	**910,118**	**1,087**

Table 11. **Summary of unique label masks by anatomical region.** Aggregate counts of label masks and corresponding labels across all 158 semantic datasets used to train VoxTell. Each dataset defines its own label dictionary; the number of label masks is computed as the number of labels multiplied by the number of scans, regardless of whether a structure is present in every scan.