MRI Breast Segmentation Dataset Comparison

Bikem Çamli

May 29, 2023

Duke-Breast-Cancer-MRI

Dataset Overview

- **Collection:** Single-institutional, retrospective collection of 922 biopsy-confirmed breast cancer patients, for a decade
- **Quality:** Annotations provided by 8 radiologists, further reviewed and modified for accuracy.
- **Number of Images:** 773,888 DICOM images.

Dataset Components

- Demographic, clinical, pathology, treatment, outcomes, and genomic data.
- Pre-operative dynamic contrast enhanced (DCE)-MRI: T1-weighted and post-contrast sequences.
- Locations of lesions in DCE-MRI: Annotations by radiologists.
- Imaging features from DCE-MRI: 529 computer-extracted features: size, shape, texture, and enhancement of both the tumor and the surrounding tissue.

Duke-Breast-Cancer-MRI

Image Annotations

- Drawn by 8 radiologists using a GUI developed in MATLAB.
- 3D boxes around areas of mass and non-mass enhancement.
- The MRI sequences that were involved in annotation were: (a) pre-contrast, (b) first post-contrast, and (c) subtracted

Two annotation phases:

- For 271 patients, 6 radiologists annotated
  - Up to five lesions were annotated per patient.
  - Biopsied tumor selection involved reviewing radiology and pathology reports.
  - Largest tumor selected in multiple biopsies
- For 651 patients 4 radiologists annotated
  - A modified annotation procedure was implemented
  - Radiologists were given the locations of the biopsies, instructed to annotate the largest biopsied lesion.

ACRIN 6698/I-SPY2

Dataset Overview

- **Collection:** From 10 instution collection of a 406 women with invasive breast cancer, for 3 years
- **Quality:** Image quality control system consists of three sequential but independent assessment stages: protocol compliance, image quality and usability, and ROI confidence
- **Number of Images:** 2,911,334 images.

Dataset Components

- Patient demographic, clinical and outcome data files for limited set of patients
- T2-weighted imaging, diffusion-weighted imaging (DWI), and dynamic contrast-enhanced (DCE)
- Manual DWI Whole-Tumor Segmentation
- Test-Retest Data to allow evaluation of repeatability and reproducibility of new DWI metrics and analysis techniques

ACRIN 6698/I-SPY2

Image Annotations

- Manually delimited tumor segmentations from the primary study analysis (for all studies rated as analyzable in the QC evaluation)
- Region definition was done at the UCSF processing lab using in-house software tools
- ROIs defined for published primary analysis
- Tumor was identified on post-contrast DCE subtraction images and then localized on the ADC map
- Multi-slice, whole-tumor regions of interest (ROIs) were manually defined by selecting regions with low ADC and hyperintensity on a high b-value DWI
- The segmentations are provided both as DICOM SEG objects and as DICOM MRI objects on TCIA

I-SPY 2 Breast Cancer Trial

Dataset Overview

- **Collection:** From 22 health center collection of a 719 patients with invasive breast cancer, along 2010-2016
- **Quality:** Over 95 % of the DCE imaging data met acceptance criteria for analysis of functional tumor volume (FTV)
- **Number of Images:** 5,586,493 images.

Dataset Components

- T2-weighted imaging, diffusion-weighted imaging (DWI), and dynamic contrast-enhanced (DCE)
- Early-treatment (T1, optional test/retest visit), Mid-treatment (T2), Post-treatment (T3)
- Derived objects from the DCE acquisitions, enhancement maps and functional tumor volume (FTV) analysis mask

I-SPY 2 Breast Cancer Trial

Image Annotations

- FTV Analysis Masks:
  - Bit-encoded segmentations
  - Encode masking steps in primary FTV analysis
  - DICOM SEG objects in a separate series
- Segmentation Details:
  - Pre-contrast background thresholding
  - Minimum percent enhancement thresholding
  - Manually defined rectangular volume of interest (VOI) for enhancing tumor analysis
  - Manually defined "OMIT" regions to exclude non-tumor enhancing regions

Dataset Comparison

Dataset	Duke	ACRIN 6698	ISPY2
Modalities	MR, SEG	MR, SEG	MR, SEG
# of Participants	922	385	719
# of Studies	922	1,123	2,688
# of Series	5,161	18,747	32,411
Images	773,888	2,911,334	5,586,493
Image Size	368.4 GB	842 GB	1.6 TB