Cross-site Validation of Deep-Learning Method for Segmenting Breast Cancers on MRI

Yu (Andy) Huang, PhD[1,2], Lukas Hirsch, MS[1], Roberto Lo Gullo, MD[2], Mary Hughes, MD[2], Jeffrey Reiner, MD[2], Nicole Saphier, MD[2], Kelly Myers, MD[3], Babita Panigrahi, MD[3], Emily Ambinder, MD[3], Philip Di Carlo, MD[3], Lars Grimm, MD[4], Dorothy Lowell, MD[4], Sora Yoon, MD[4], Sujata Ghate, MD[4], Danny F. Martinez, MS[2], Lucas C Parra, PhD[1], Elizabeth J. Sutton, MD[2]
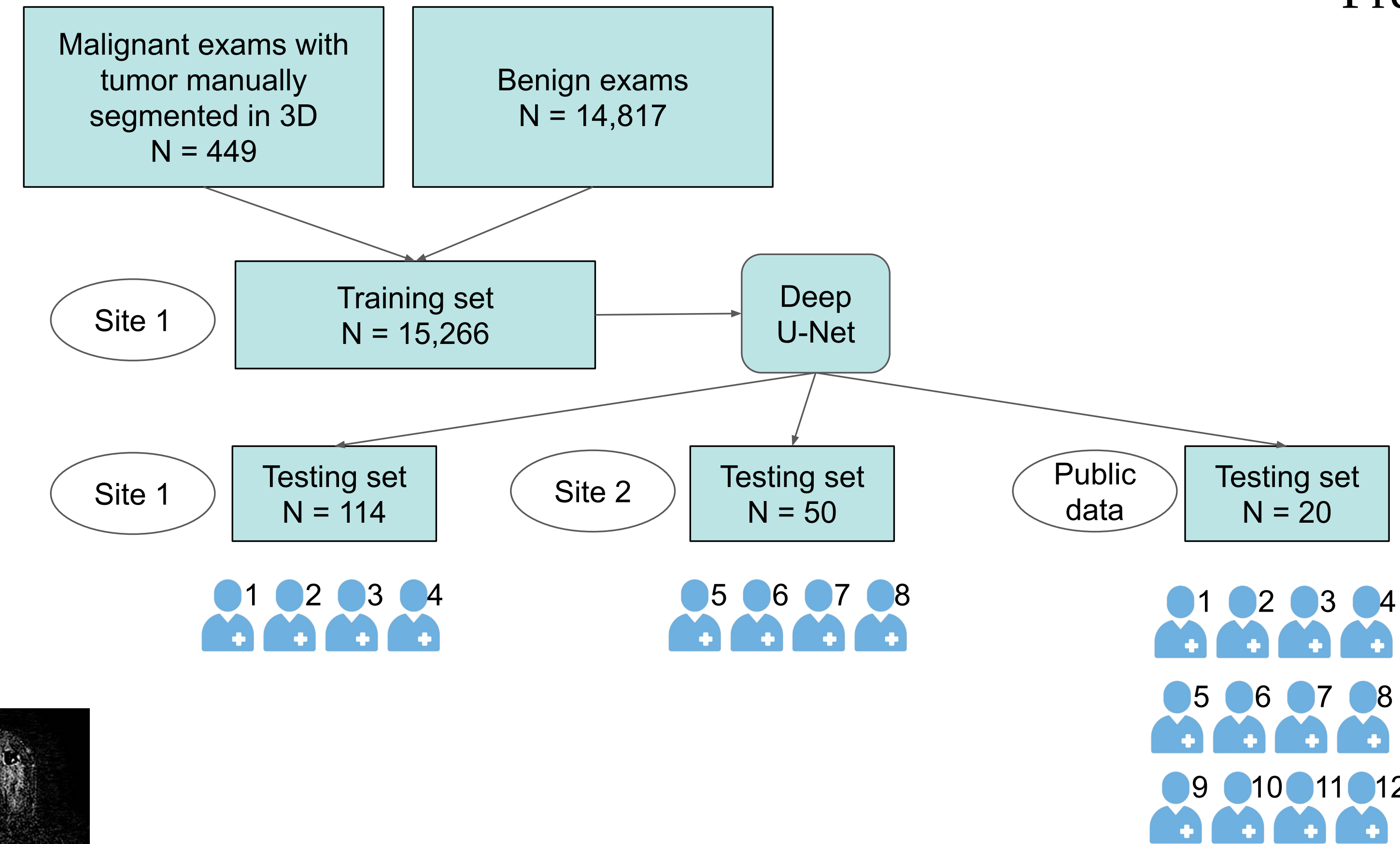
[1]Department of Biomedical Engineering, the City College of the City University of New York, New York, NY 10031; [2]Department of Radiology, Memorial Sloan Kettering Cancer Center, New York, NY 10065; [3]Department of Radiology and Radiological Science, Johns Hopkins Medicine, Baltimore, MD 21224; [4]Department of Radiology, Duke University School of Medicine, Durham, NC 27710

Abstract

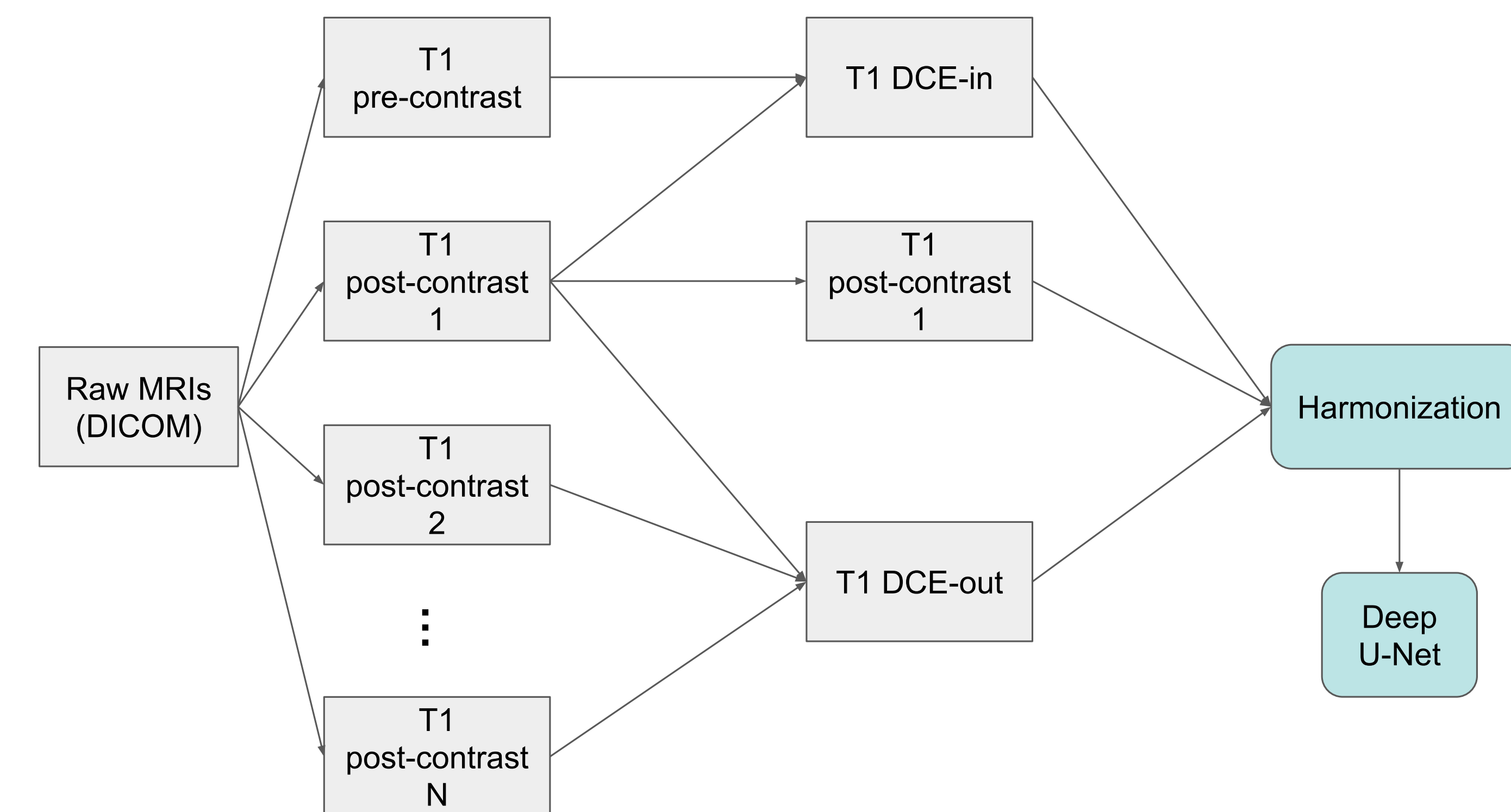The purpose of this work was to evaluate our previously developed deep neural network[1] for segmenting breast cancers on MRIs using data from two clinical sites. 449 contrast-enhanced breast MRI exams with tumors from Site 1 (MSKCC) were segmented in 3D. These exams were used together with another 14,817 benign exams from the same site to train a 3D U-Net. The trained network was tested on a different set of 114 exams from Site 1, and on a set of 50 exams from Site 2 (Duke University). At each site, we recruited four breast radiologists and asked each of them to manually segment one slice of breast tumors for all the MRIs in the testing set. Additionally, the network was tested on 20 exams from a public dataset from Site 2 which were also manually segmented by 12 radiologists. Results show that the network achieved radiologist-level performance in segmenting the breast tumors on the data from both sites, as well as on the public dataset. We hope that this work will provide some guidance on future efforts on cross-site applications of deep-learning methods for medical image data.
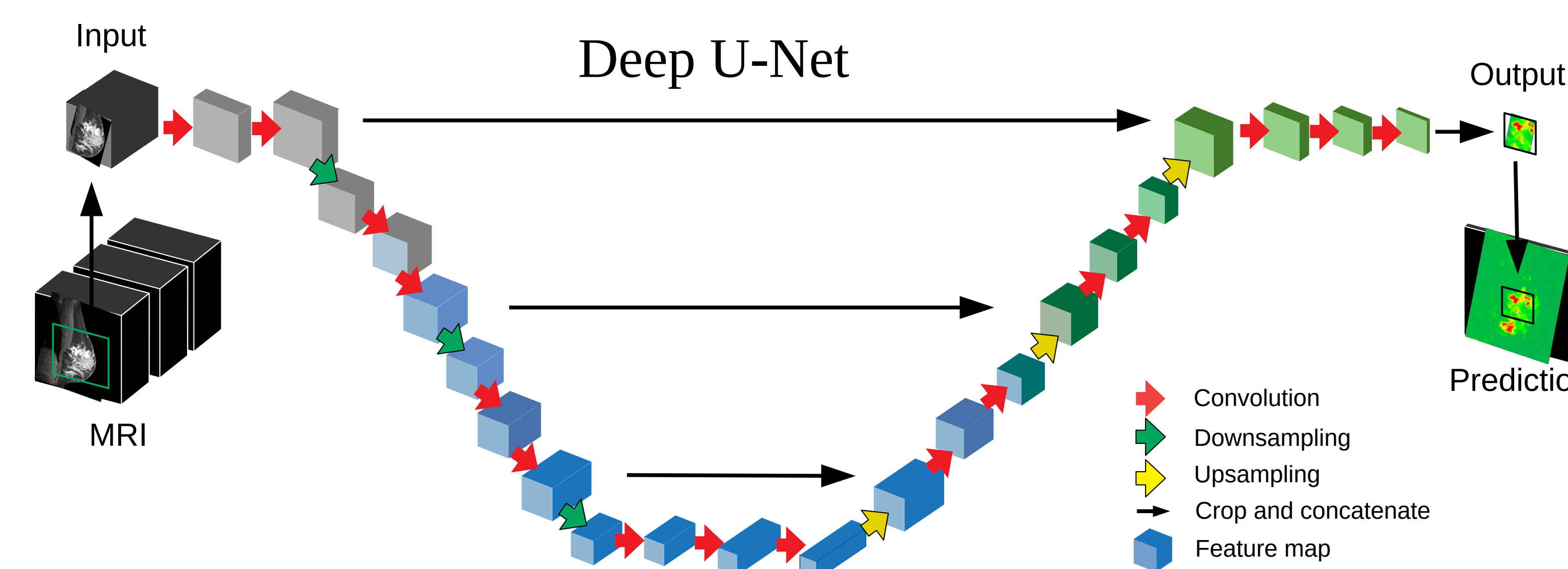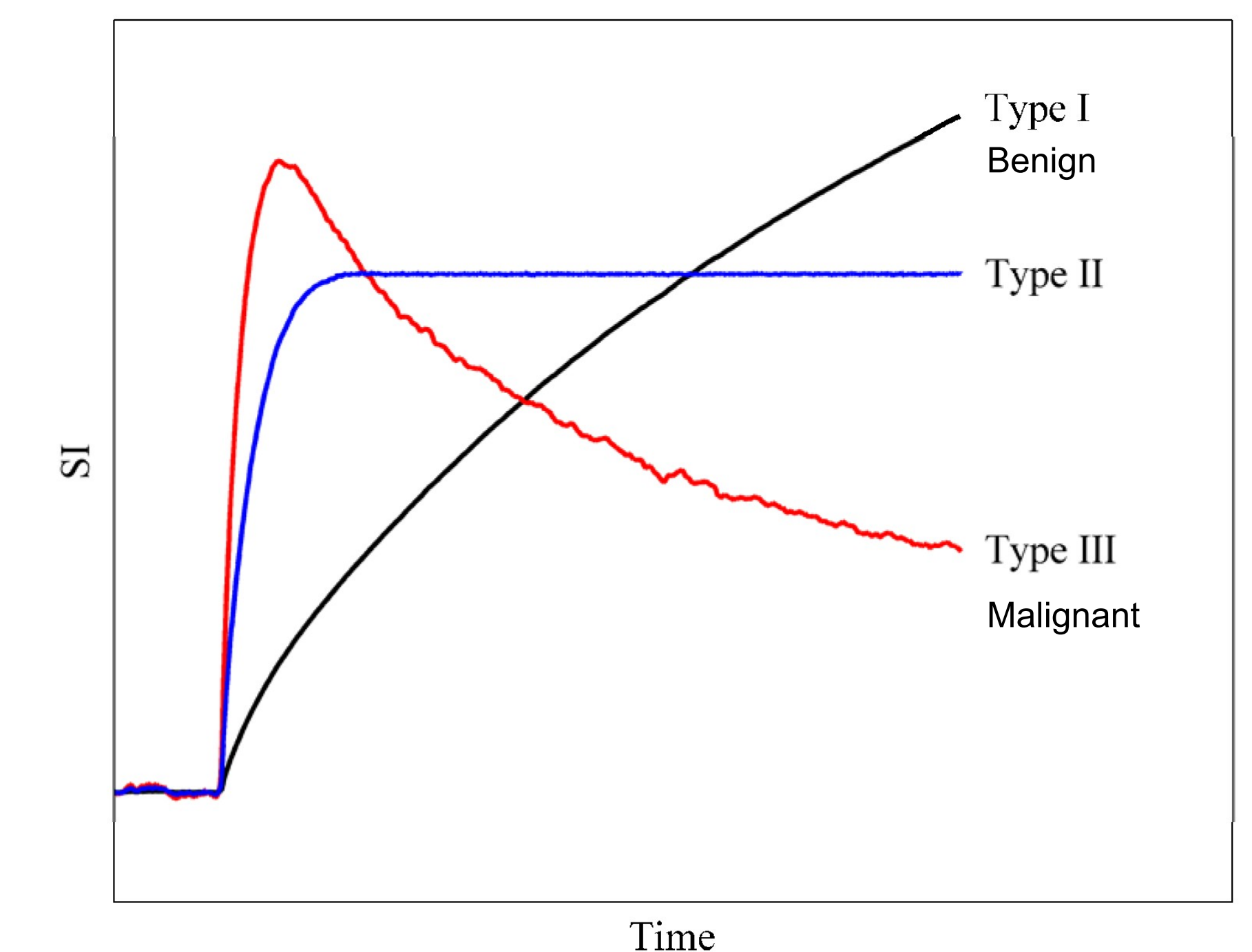
Method

Study Design

Pre-processing

Deep U-Net

Convolution
Downsampling
Upsampling
Crop and concatenate
Feature map

Dynamic Contrast Enhancement (DCE)

Results

One example from Site 2

T1 post | DCE-in | DCE-out
Machine | Ref1
| | Ref2 | Ref3 | Ref4

114 cases at Site 1

50 cases at Site 2

20 public cases

Discussions

- Deep U-Net trained with a large dataset shows feasibility of generalizing to unseen data at a different clinical site;
- Ongoing work shows generalization may also fail on testing data from a third clinical site;
- Fine-tuning may still be needed before cross-site deployment, which is cheap to perform on the new clinical site.

Reference: [1]. Hirsch, L., Huang, Y., et al., Radiologist-Level Performance by Using Deep Learning for Segmentation of Breast Cancers on MRI Scans. *Radiology: Artificial Intelligence*, 4, e200231 (2022)

Memorial Sloan Kettering Cancer Center

The City College of New York