Automatic Segmentation with Radiomic Analysis of Sarcoma in Mice

M. D. Holbrook[1], S. J. Blocker[1], Y. M. Mowery[2], A. Badea[1], Y. Qi[1], E. S. Xu[2], D.G. Kirsch[2], G.A. Johnson[1], C. T. Badea[1]

[1] Center for In Vivo Microscopy, Department of Radiology, Duke University Medical Center, Durham, NC 27710

[2] Department of Radiation Oncology, Duke University Medical Center, Durham, Durham, NC 27710

Duke

Introduction

- We are involved in a co-clinical trial studying synergy between immune checkpoint inhibitors and radiotherapy.
- Small animal imaging enhances the simulation of clinical practice.

Mouse model:
- $p53^{fl/fl}$ mouse model [1]
- Hind limb sarcoma generated by delivery of Adeno-Cre followed by carcinogen 3-methylcholanthrene

Imaging Protocol:
- T2-weighted and two T1-weighted (before and after contrast injection)
- Radiation therapy (RT, 20 Gy) on a small animal irradiator
- One week later, the mice were re-imaged
- Tumor was surgically removed by amputating the tumor-bearing hind limb
- Mice were followed for up to 6 months

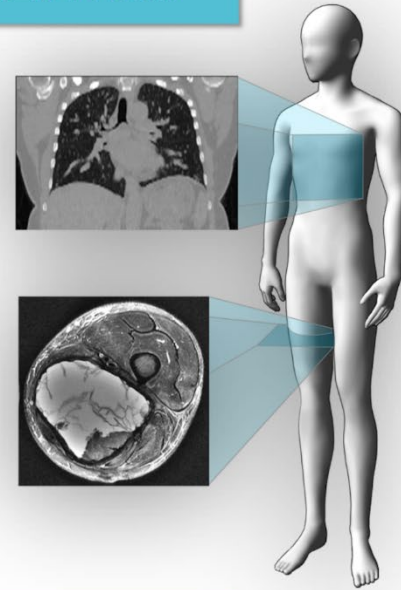More information on this work has been recently published [2].

Goal:
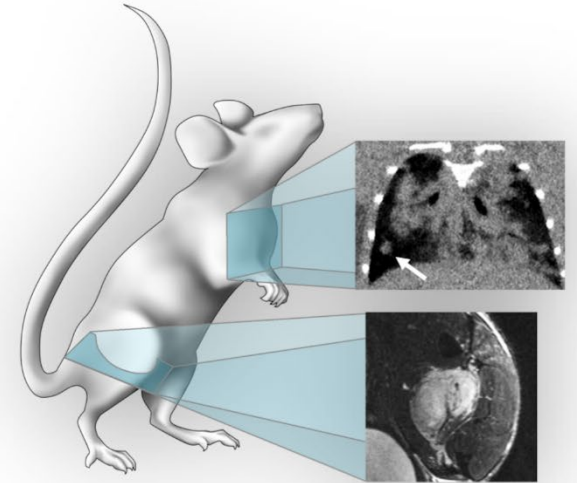- **Build and evaluate a pipeline for segmentation and analysis of preclinical sarcoma MRI images.**

[1] C. L. Lee *et al.*, JCI Insight, 2019.
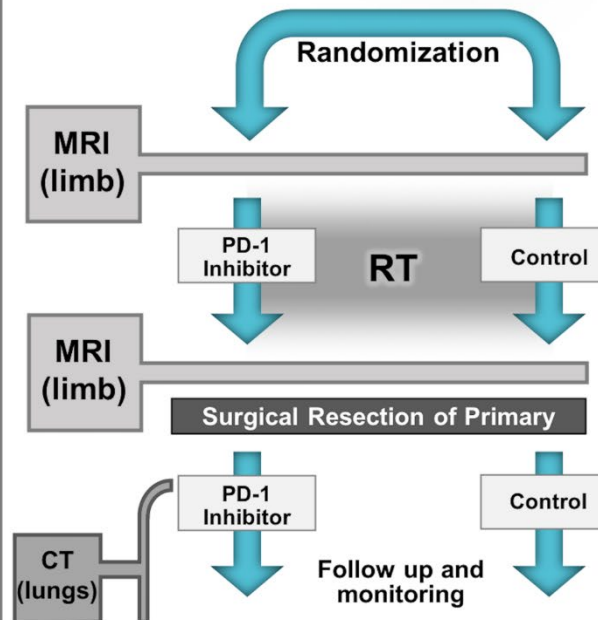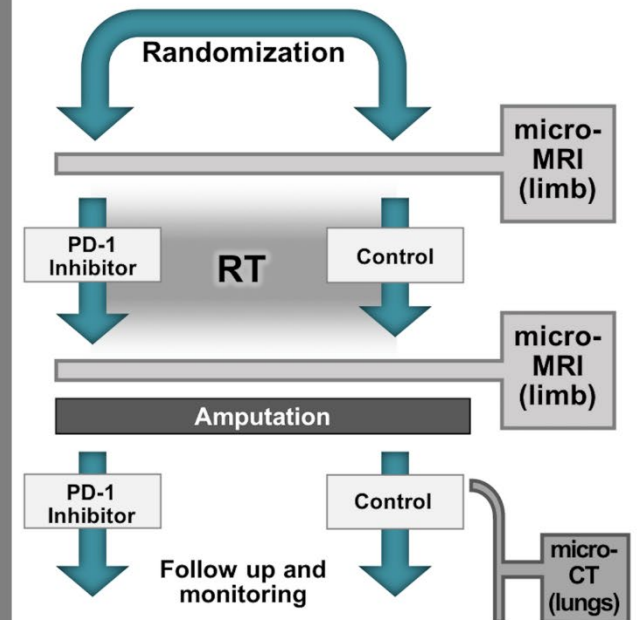[2] M. Holbrook *et al.*, Tomography, 2020.
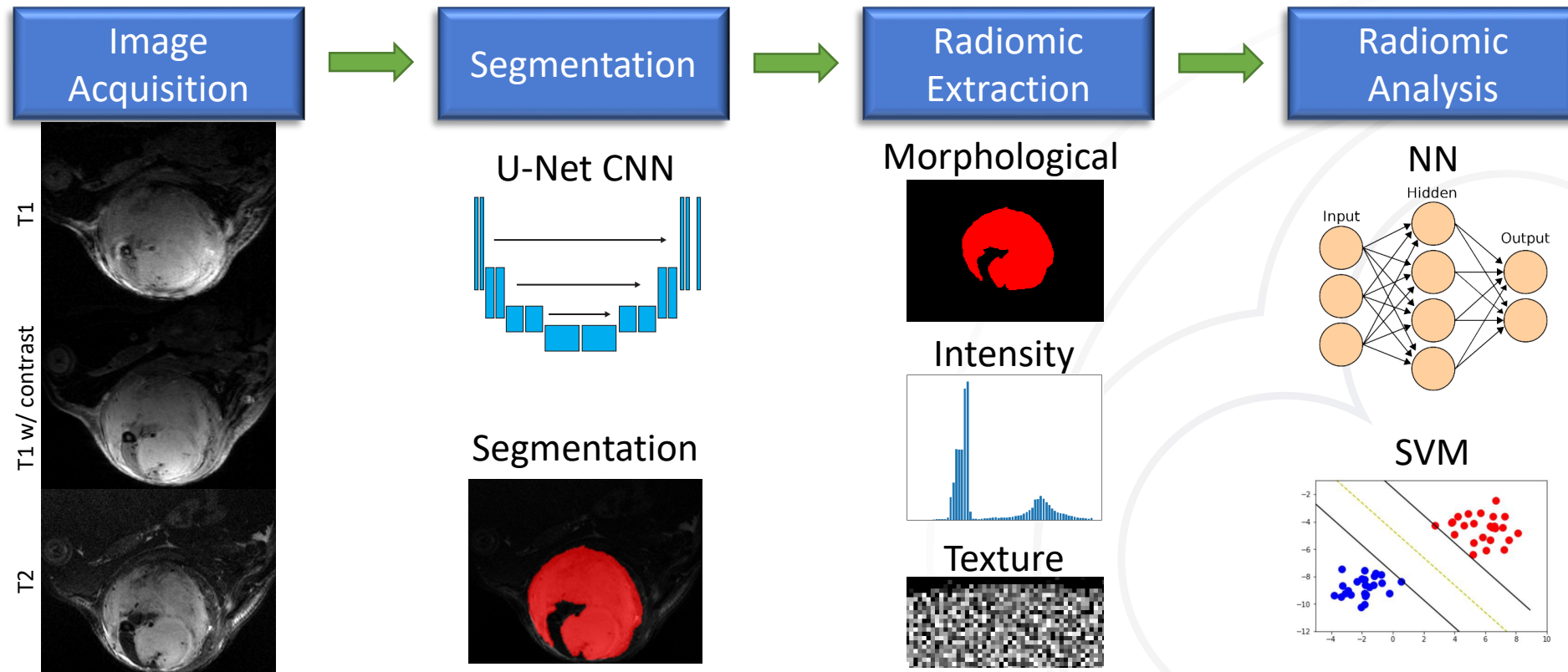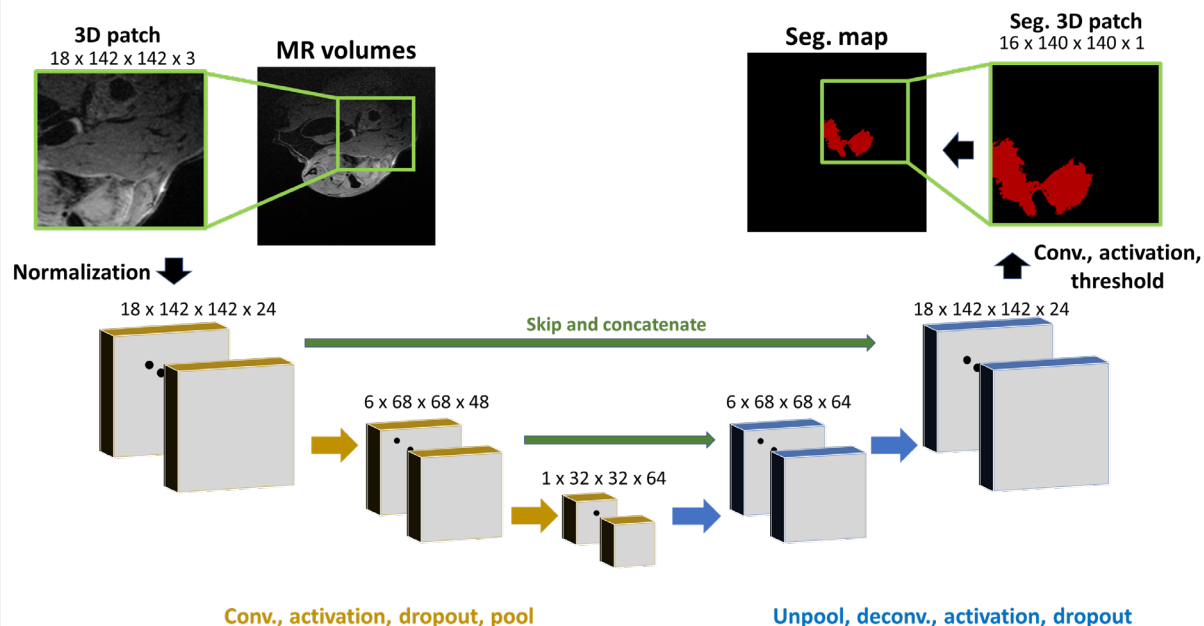
Image processing pipeline

Outline of the image processing pipeline created in this work. Three MR image contrasts are acquired and passed through a convolutional neural network (CNN) to segment the tumor. Segmentation maps are used to calculate a battery of radiomic features, including those based on morphology, gray-level intensities, and texture. These features are analyzed to in terms of changes with radiation therapy (RT), and machine learning tools are employed to predict primary tumor recurrence within the study population.

U-net Structure and training

CNN architecture used for tumor segmentation. Volumetric patches are extracted from each MR image and reassembled after processing. The network consists of an encoder (gold) and decoder (blue) with skip connections between them. All kernels are 3x3x3 and use a stride of 1x1x1.

Segmentation results

Best performing network:
- Based on multi-contrast images
- Contained skip connections
- Cross entropy cost function

Metrics after K-folds:
- Dice: 0.8422 ± 0.0187
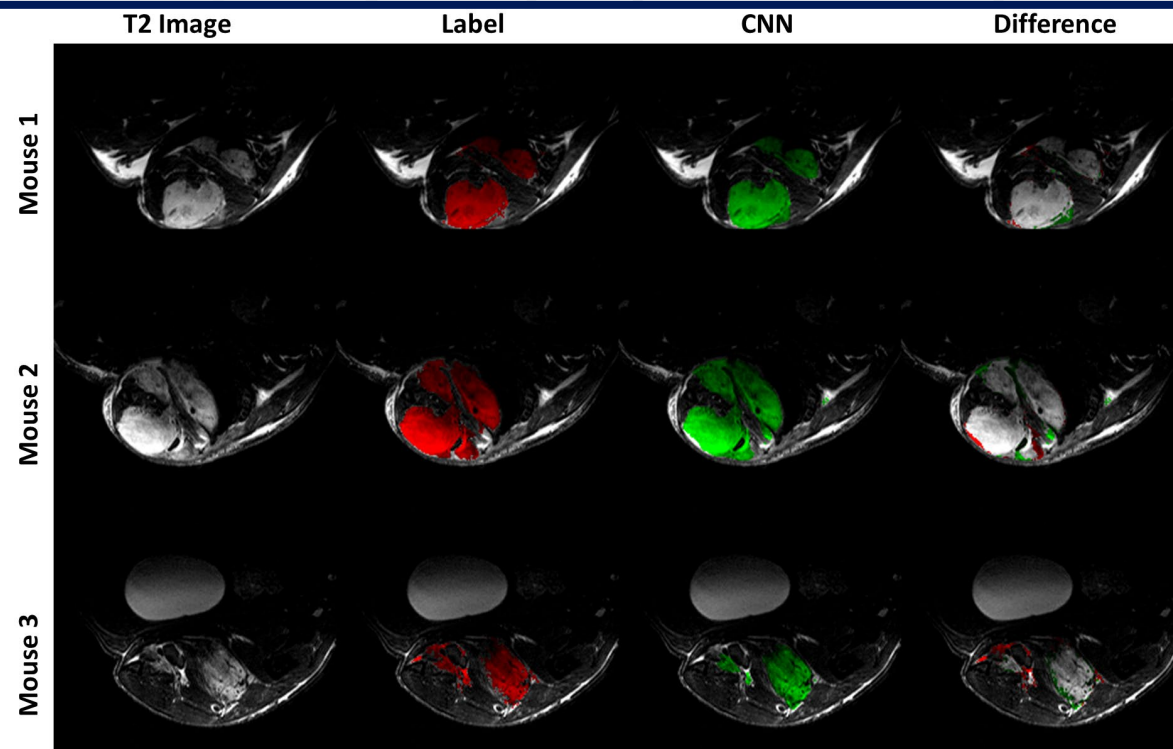- VOE: 0.9933 ± 0.0009

[1] Kikinis R *et al.*, 2014.

Segmentation results are shown comparing the label (red) and predictions (green) for three different tumors. This network was train using k-folds validation and run as an ensemble. Errors are shown in the difference column: red for false negatives and green for false positives. The time required to segment a volume is 0.56 seconds

8 network configurations tested:
- Using T2 vs multi-contrast images
- With and without skip connections
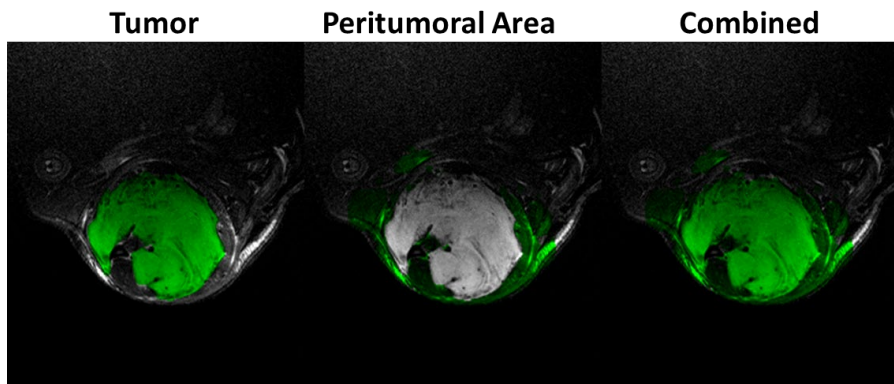- Dice vs cross entropy cost functions

Training data:
- All image volumes are 280 x 280 x 60, with 100-µm in-plane resolution and 300-µm slices.
- 70 tumor segmentations were manually created in 3D Slicer [1].
- Data was split into training (70%), validation (20%), and test (10%) sets.
- The best performing network was re-trained with 5-fold cross validation and employed in ensemble to process all volumes.
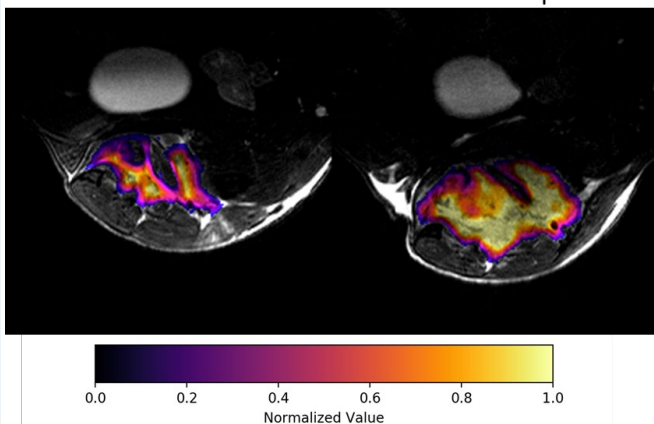
Radiomic feature extraction

- Features were extracted from 3 regions using PyRadiomics [1]:

Tumor **Peritumoral Area** **Combined**

- Each region yields 107 features (321 for 3 MR contrasts).
- Feature selection was performed via mRMR [1].
- **42** mice with pre- and post-RT scans were analyzed.
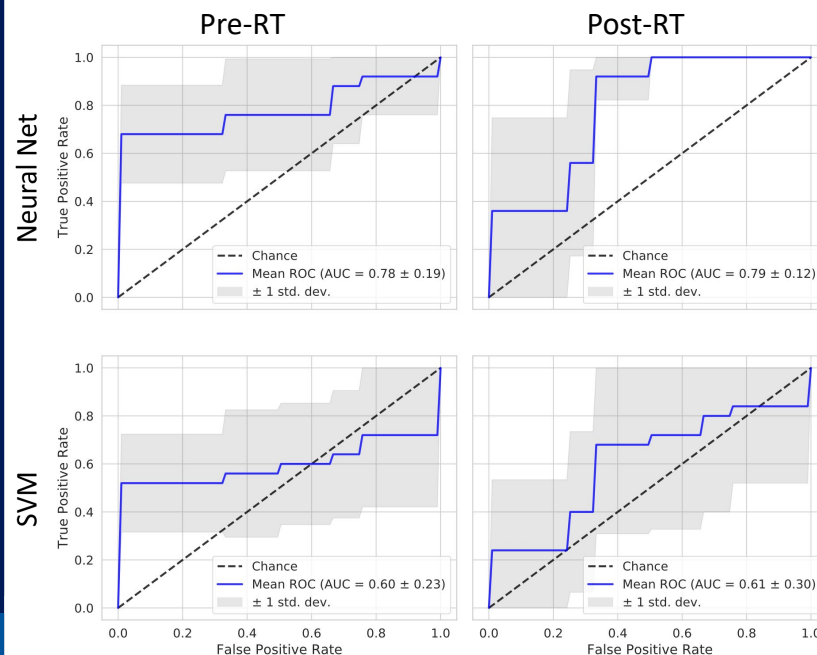
Change in features with RT

GLDM: Dependence Non-Uniformity
Pre RT   One-week post RT

Normalized Value

Multiple t-tests were used to find significantly different radiomic features from images captured before and after RT. 76 significant features were found, well divided across shape and contrast images (shape: 11, T1: 23, T1 with contrast: 19, T2: 23). An example, dependence non-uniformity from a T2 image, is shown here. Note differences in tumor shape and feature intensity.

[1] van Griethuysen JJM *et al.*, Cancer Research. 2017.

Prediction of primary tumor recurrence

Pre-RT   Post-RT

Receiver operating curves showing the accuracy of a NN and an SVM for predicting primary tumor recurrence in this population are shown. These curves pertain to data calculated in the peritumoral area, and the 10 most significant radiomic features were used. The greatest predictive power was found after RT (AUC: 0.79).

Conclusions

- We have developed and demonstrated the performance of a new automatic segmentation and radiomic analysis pipeline built on machine learning models which greatly increase the processing consistency and speed.
- The best performing segmentation network (Dice: 0.8422 ± 0.0187) has a U-net structure with skip connections and was trained with multi-modal MRI.
- Prediction of primary tumor recurrence in the study population points to radiomic features of interest which will be examined in future work.
- The methods developed here will form a vital role in ongoing pre-clinical trials.