This is lung data from TCIA NSCLC Radiogenomics

Refer to: <https://wiki.cancerimagingarchive.net/display/Public/NSCLC+Radiogenomics>

**radiogenomics\_PET.mat**: this file is extracted from PET image.

nameArray: patient number

imageSUV: ROI manually draw by Xiaxia and the pixel number in the ROI

xyzStep: the resolution for each pixel in xyz direction. (mm)

SUVmax: the maximum SUV value of the ROI(imageSUV)

SUVmean: the mean SUV of the ROI (imageSUV)

SUV95: 95% percentile of SUV in the ROI(imageSUV)

SUV95\_40mean: the mean SUV of the ROI (imageSUV95\_40)

imageSUV95\_40: ROI automatically drawn using SUV95\*40% as a threshold

Position: ROI coordinate position

**radiogenomics\_CT.mat**: this file is extracted from CT image.

nameArray: patient number

imageHU: ROI manually draw by Xiaxia and the pixel number in the ROI

xyzStep: the resolution for each pixel in xyz direction. (mm)

HUmax: the maximum HU value of the ROI(imageHU)

HUmean: the mean HU of the ROI (imageHU)

HUmean40: the mean HU of the ROI (imageHU40)

imageHU\_pos: each pixel in the ROI (imageHU) is minus the minimum value of the ROI (imageHU) to make it positive

imageHU40: ROI automatically drawn using HUmax\*40% as a threshold

**NSCLCR label detail.xlsx**: this file contains all the information for each patient. You can do classification according to it.

**heterogeneity.xlsx**: this file contains heterogeneity value calculated by Moran’s I from PET ROI(imageSUV95\_40);

**label for reference.xlsx:** this file is classification label for reference

**radiogenomicsFeature.xlsx**: radiomics features calculated from PET ROI(imageSUV95\_40) and CT ROI(imageHU);

changes:

10/9/2018 Some changes were made in recurrent label. The no recurrent patients whose follow up time less than 365 days were left out. Their follow up time is too short to draw conclusions.