Radiomics

Dawei Chang (張大衛)
MeDA Lab

Data Science Degree Program
National Taiwan University & Academia Sinica

Outline

- Introduction to Radiomics
- Introduction to Radiomic Feature Extraction
- Introduction to Radiomics Features
  - Filtering approach
  - Types of features
    - Intensity based
    - Texture based
    - Shape based
- Example Application

Introduction to Radiomics

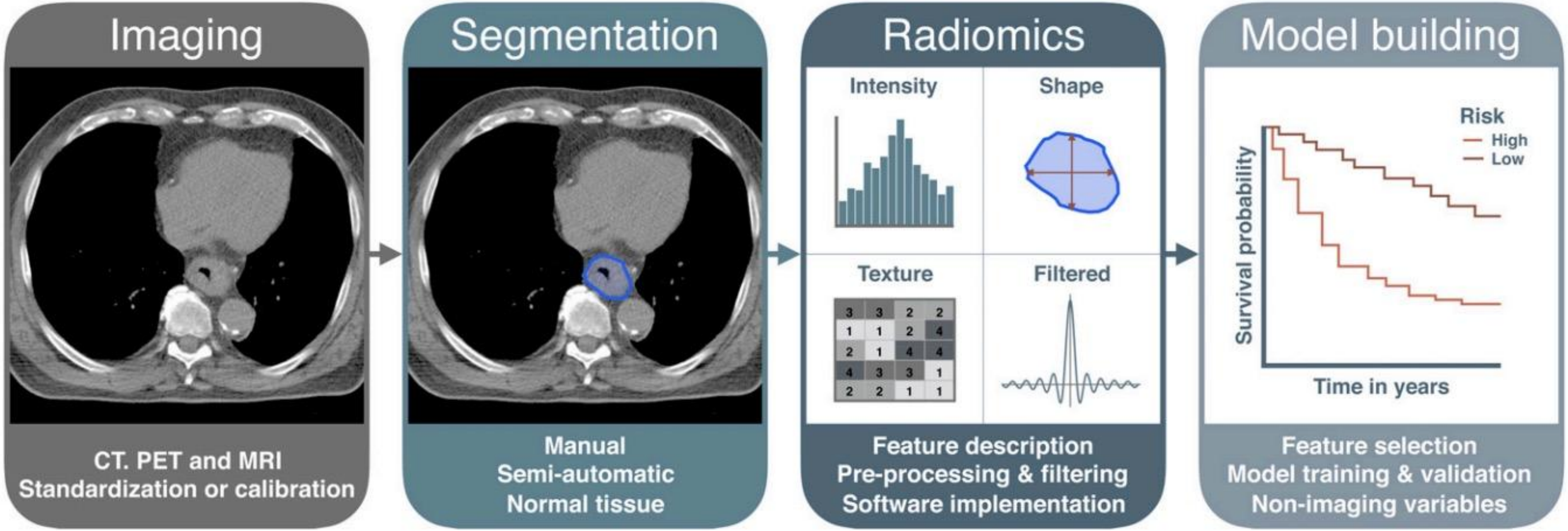
Motivation

- As long as extracting features of each sample data, we can build up a machine learning model accordingly for our task
- Feature extraction for medical image is not straightforward compared to non-image numerical data
    - E.g. Feature vectors from flattened images may have different size and lose the neighboring information

- Solutions:
    - Extract pre-defined features from medical images
        - Pre-defined radiomic features, radiomics
    - Extract features defined by model itself
        - Neuron network with convolutional layer

Radiomics

- Radiomics is a method to extract quantitative features from an interested region on a medical images for further analysis



Larue, R. T. H. M., Defraene, G., De Ruysscher, D., Lambin, P., & van Elmpt, W. (2017). Quantitative radiomics studies for tissue characterization: a review of technology and methodological procedures. *The British Journal of Radiology*, *90*(1070), 20160665.

Example Radiomics Feature

- Shape-based features
  - E.g. mesh volume, voxel volume, surface area, sphericity, elongation, ...
- Intensity-based features (first-order statistical features)
  - E.g. first-order: mean, first-order: median, first-order: energy, first-order: entropy, …
- Texture-based features (higher-order statistical features)
  - Gray level co-occurrence matrix (GLCM) derived features
  - Gray level run length matrix (GLRLM) derived features
  - Gray level size zone matrix (GLSZM) derived features
  - Neighbouring gray tone difference matrix (NGTDM) derived features
  - (Neighboring) gray level dependence matrix (NGLDM, GLDM) derived features
  - E.g. GLCM: joint entropy, GLRLM: gray-level non uniformity, NGTDM: busyness, ...

Potential Usage

- Prediction of treatment response and outcomes
  - E.g. For head-and-neck cancer, predict whether tumours will respond poorly to chemoradiotherapy by using MR
- Tumour staging
  - E.g. Predict lung tumor stage by using CT images
- Tissue identification
  - E.g. Classify pulmonary nodule as benign or malignant

Yip, S. S. F., & Aerts, H. J. W. L. (2016). Applications and limitations of radiomics. *Physics in Medicine and Biology*, *61*(13), R150–R166.

IBSI

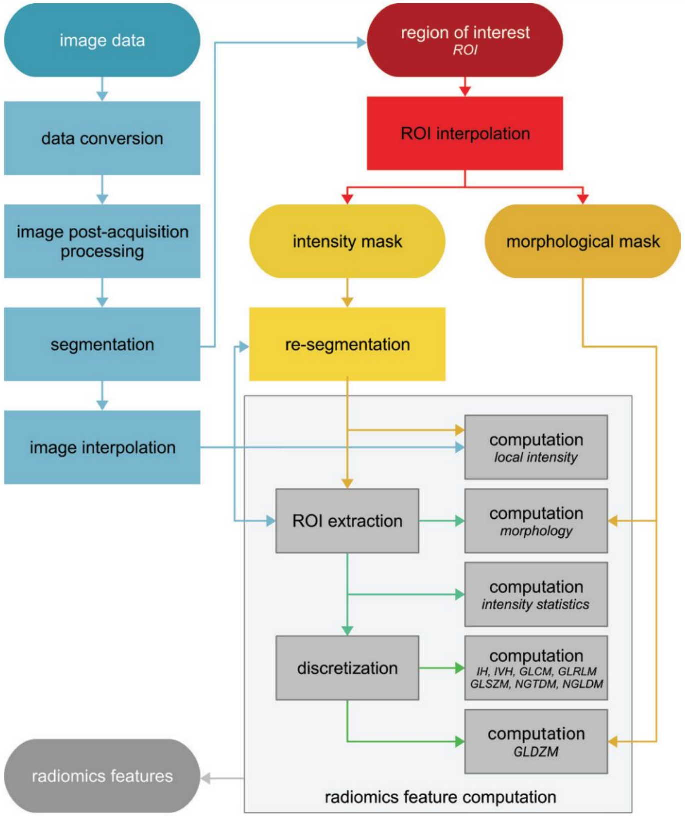
- IBSI stands for Image Biomarker Standardisation Initiative
- One of the main issues of radiomics is the reproducibility of features, due to many parameters from data acquisition, image processing, and feature extraction
- IBSI dedicates to establishing standardisation for extracting radiomic features, and hence increasing the reproducibility of radiomics-based research
- As a software developer, IBSI (IBSI Chapter 1) have established an standardisation of computation of common radiomic features, with benchmark data and reference values
- As a user, IBSI provides the reporting guideline to ensure the reproducibility
    - In short, you should report all the parameters that may cause the difference in features

https://theibsi.github.io/

Introduction to Radiomic Feature Extraction

Radiomic Feature Extraction

- Radiomic features are extracted from **an image** data with segmentation information (region of interest, **ROI**)
- Parameters and setting for each step in the workflow will affect the results of the output feature set; therefore, all the details should be reported while the radiomics analysis is presented
- For Python, one can extract the features easily by the package [PyRadiomics]



Zwanenburg, A., Vallières, M., Abdalah, M. A., Aerts, H. J. W. L., Andrearczyk, V., Apte, A., Ashrafinia, S., Bakas, S., Beukinga, R. J., Boellaard, R., Bogowicz, M., Boldrini, L., Buvat, I., Cook, G. J. R., Davatzikos, C., Depeursinge, A., Desseroit, M.-C., Dinapoli, N., Dinh, C. V., … Löck, S. (2020). The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*, *295*(2), 328–338.

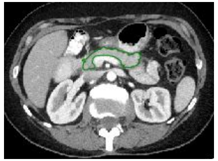
Extraction Workflow Breakdown

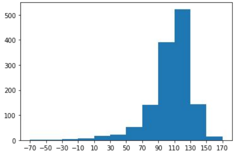
Image Data
- MR, PET, CT
- Stage, Tracer, Contrast Phase

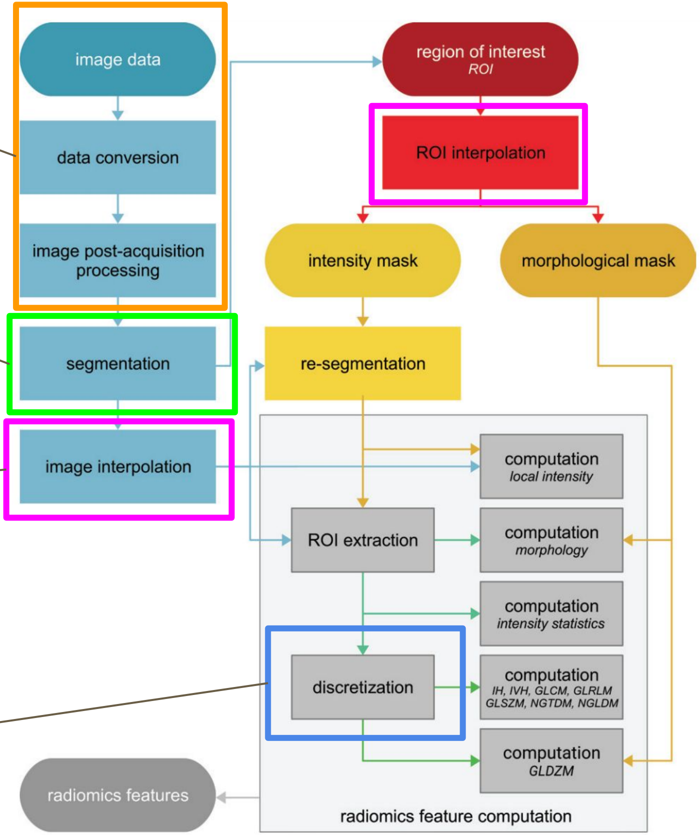
Segmentation
- Manual/Automatic Segmentation
- Radiologists, Segmentation Models

Interpolation (Resampling)
- Nearest Neighbor Interpolation
- Bilinear/Trilinear Interpolation

Discretization
- Bin Width, Bin Number





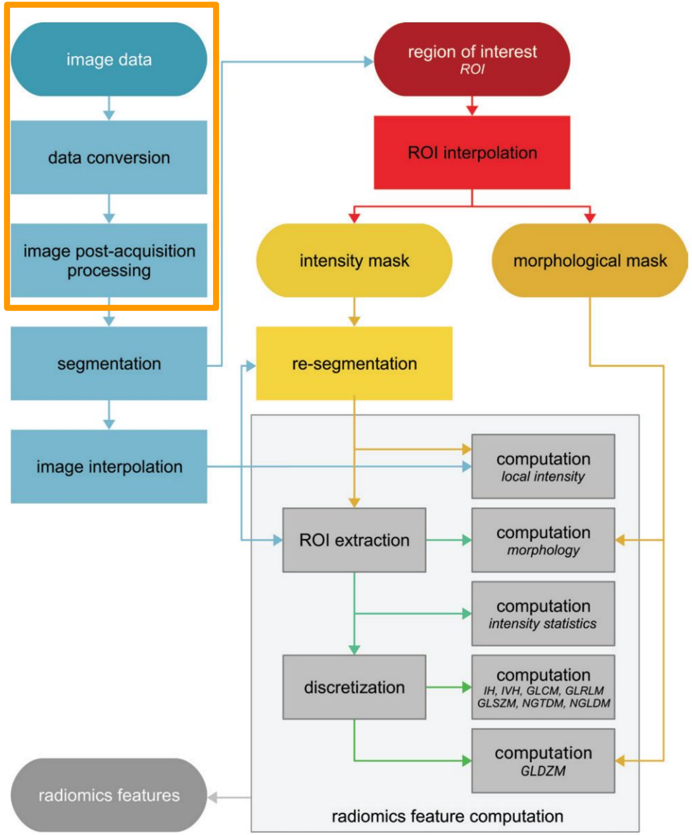


Example data and segmentation from PANCREAS-CT dataset in TCIA. https://wiki.cancerimagingarchive.net/display/Public/Pancreas-CT

Image Data

- Images is the most essential part of radiomics analysis
- Appropriate image modality and variation can cause difference of features and affect the difficulty for establishing model

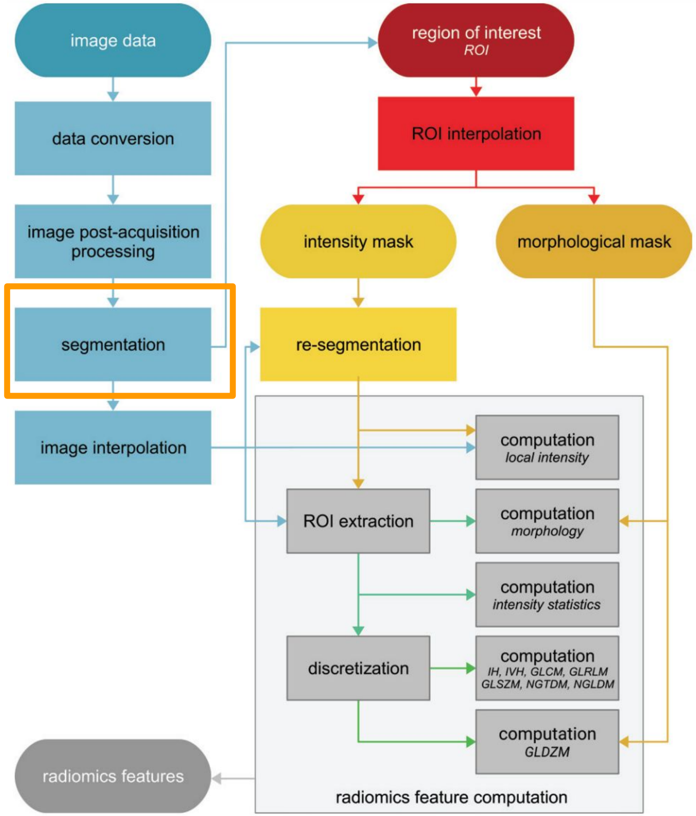
- Possible image modality
  - MRI
    - With stages as variation
  - PET
    - With tracer as variation
  - CT
    - With contrast phase as variation



Segmentation

- Segmentation is another crucial part of radiomics analysis
- ROI should be chosen as the part of images that is crucial to the study (e.g. tumor, a part of tissue)
- Quality and methodology of segmentation can cause the difference of extracted feature values

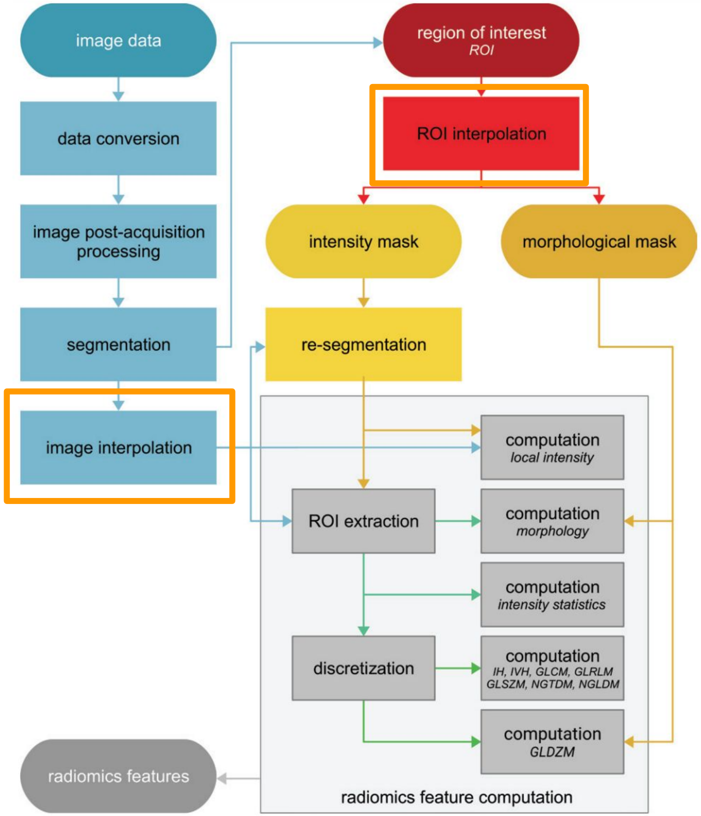
- Possible source of segmentation
  - Manually segmentation
  - Automatic segmentation by some rules or models



Interpolation (Resampling)

- Using different spacing of the images can produce different readiomic feature values
- To eliminate the variation caused by the spacing, the image and mask should be resampled into same spacing
- The interpolation method can also affect the resultant feature values

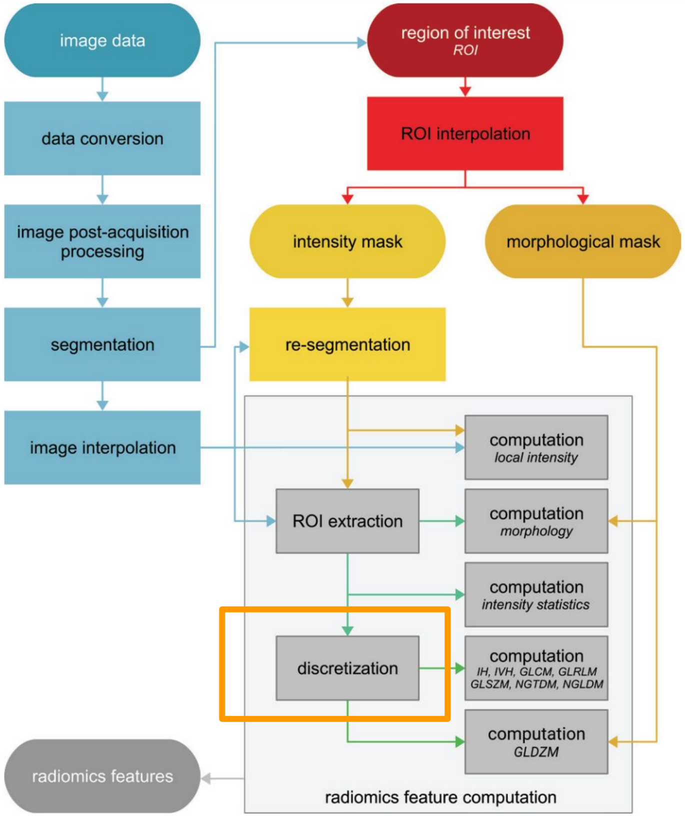
- Possible interpolation method
  - Nearest neighbor interpolation (typically for ROI)
  - Bilinear/Trilinear interpolation (typically for image)
  - Higher order polynomial interpolation



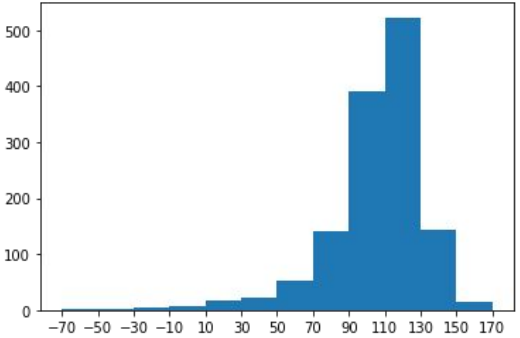
Discretization

- Some features (almost all of texture-based features) are computed by the gray level of each pixels, which means the intensities are divided into several bins of an histogram, and the intensities inside each bin have same gray level
- Smaller bin width (larger bin number) produces more sensitive features, which may have better ability to distinguish different texture but less stability

- Parameter for discretization
    - Bin number
    - Bin width

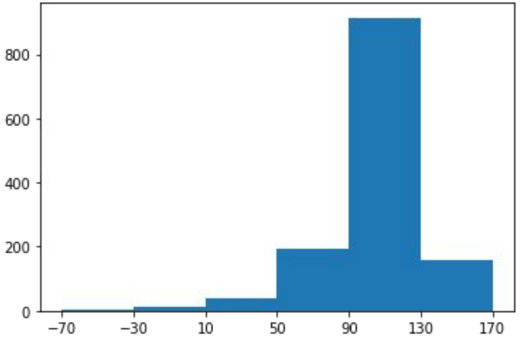


Discretization and Gray Level



Bin Width = 20

HU	Gray Level
-70 to -51	1
-50 to -30	2
...	...
150 to 170	12



Bin Width = 40

HU	Gray Level
-70 to -30	1
-30 to -10	2
...	...
130 to 170	6

Example data and segmentation from PANCREAS-CT dataset in TCIA. https://wiki.cancerimagingarchive.net/display/Public/Pancreas-CT

Introduction to Radiomic Features

Feature Class

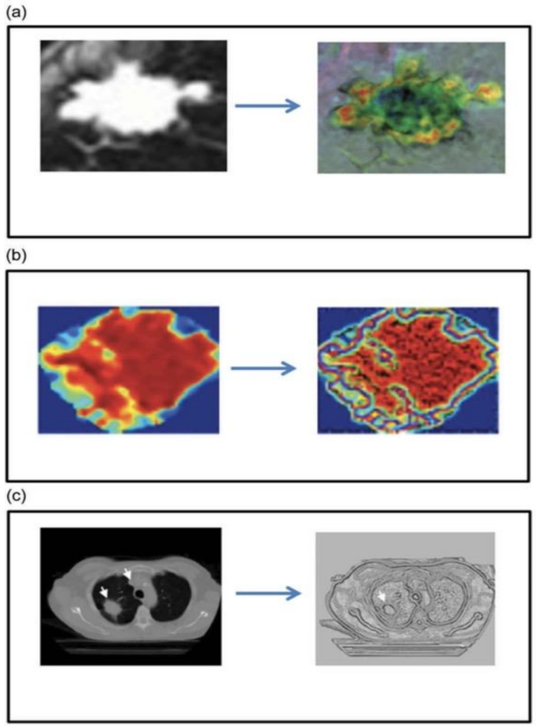
- Shape-based features
- Intensity-based features (first-order statistical features)
  - Features directly derived from intensity
  - Features derived from intensity histogram
- Texture-based features (higher-order statistical features)
  - Gray level co-occurrence matrix (GLCM) derived features
  - Gray level run length matrix (GLRLM) derived features
  - Gray level size zone matrix (GLSZM) derived features
  - Neighbouring gray tone difference matrix (NGTDM) derived features
  - (Neighboring) gray level dependence matrix (NGLDM, GLDM) derived features

Filtering Approach

Statistical Kernel
(e.g. median filter)

Edge Kernel
(e.g Laplacian of Gaussian filter)

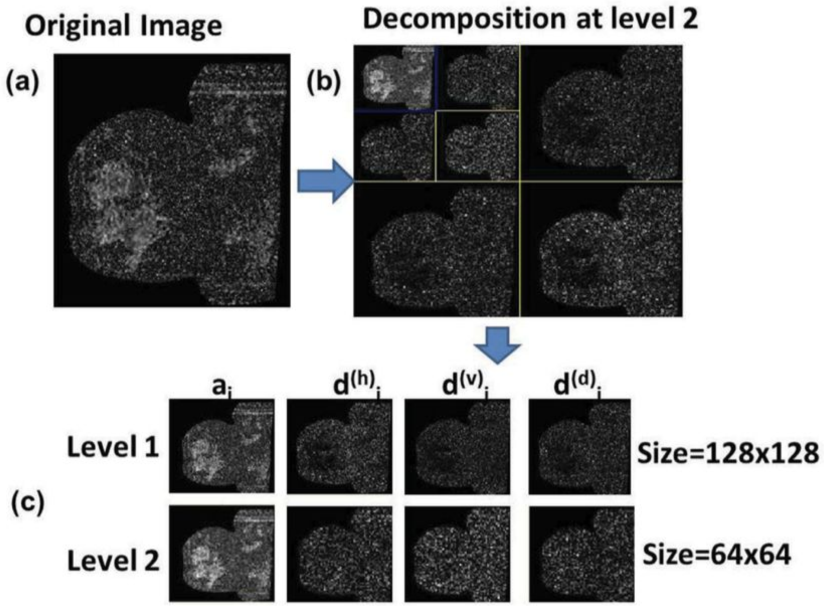
Special Kernel
(e.g. Fractal dimension filter)



Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

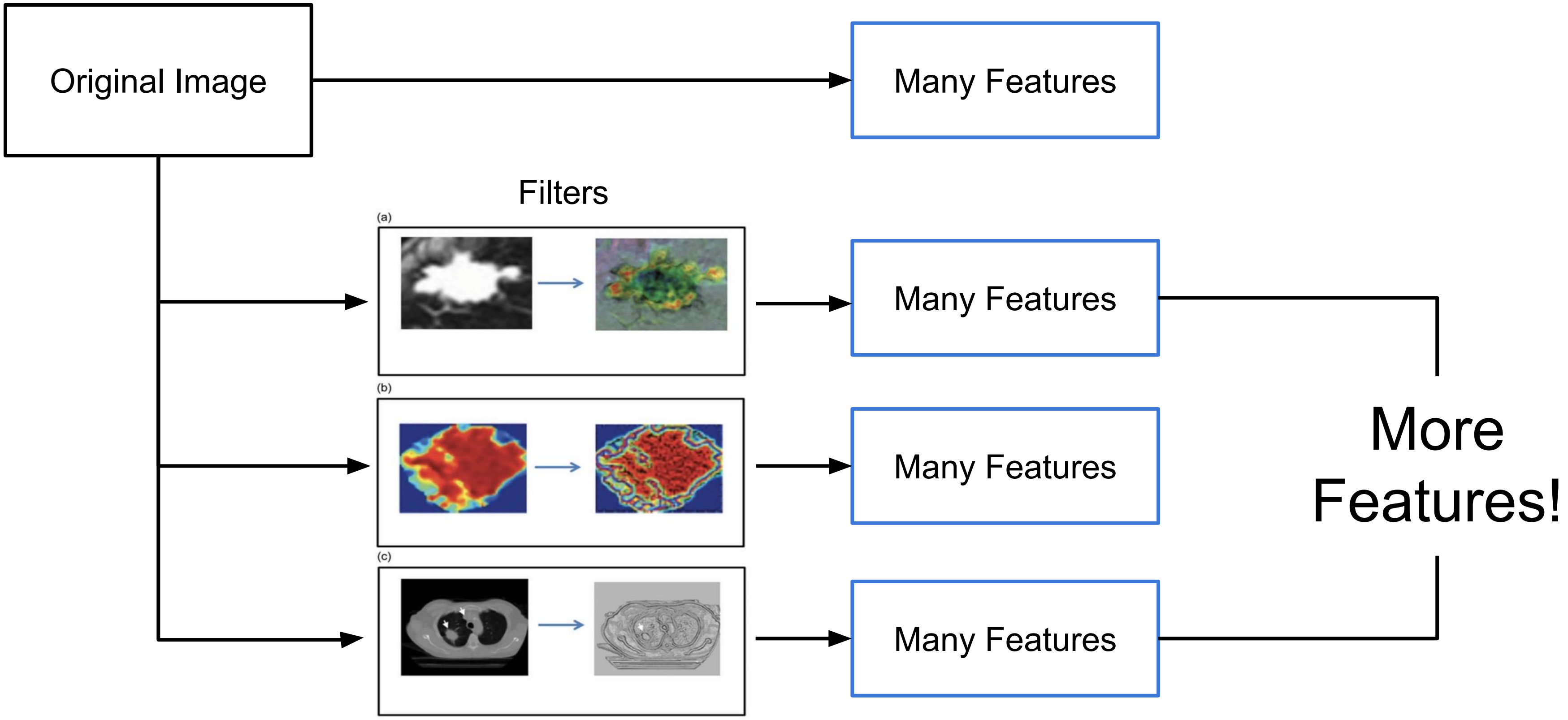
Multiresolution Image Scaling

- The picture below illustrates multiresolution scaling on an example breast MRI diffusion weighted image by wavelet



Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Feature × Filter



Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Shape-Based Features

Shape-Based Features

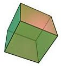
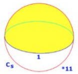
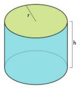
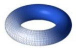
- Features related to the shape of ROI

- Related to measurement: mesh volume, voxel volume, surface area, …
- Related to appearance: elongation, flatness, distortion, ...

Example Shape-Based Feature

- Sphericity: measure how close a surface to the sphere

$$sphericity = \frac{\sqrt[3]{36\pi V^2}}{A}$$

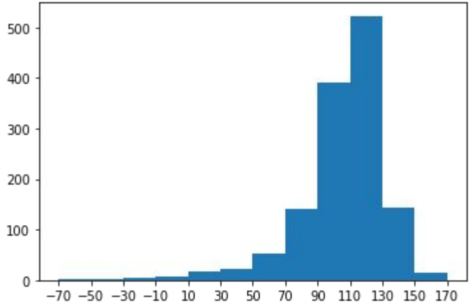
Name	Picture	Volume	Surface Area	Sphericity
Platonic Solids				
tetrahedron		$\frac{\sqrt{2}}{12}\,s^3$	$\sqrt{3}\,s^2$	$\left(\frac{\pi}{6\sqrt{3}}\right)^{\frac{1}{3}} \approx 0.671$
cube (hexahedron)		s^3	$6\,s^2$	$\left(\frac{\pi}{6}\right)^{\frac{1}{3}} \approx 0.806$
octahedron		$\frac{1}{3}\sqrt{2}\,s^3$	$2\sqrt{3}\,s^2$	$\left(\frac{\pi}{3\sqrt{3}}\right)^{\frac{1}{3}} \approx 0.846$
dodecahedron		$\frac{1}{4}\left(15+7\sqrt{5}\right)\,s^3$	$3\sqrt{25+10\sqrt{5}}\,s^2$	$\left(\frac{\left(15+7\sqrt{5}\right)^2\pi}{12\left(25+10\sqrt{5}\right)^{\frac{3}{2}}}\right)^{\frac{1}{3}} \approx 0.910$
icosahedron		$\frac{5}{12}\left(3+\sqrt{5}\right)\,s^3$	$5\sqrt{3}\,s^2$	$\left(\frac{\left(3+\sqrt{5}\right)^2\pi}{60\sqrt{3}}\right)^{\frac{1}{3}} \approx 0.939$
Round Shapes				
ideal cone $(h = 2\sqrt{2}r)$		$\frac{1}{3}\pi\,r^2 h = \frac{2\sqrt{2}}{3}\pi\,r^3$	$\pi\,r(r+\sqrt{r^2+h^2}) = 4\pi\,r^2$	$\left(\frac{1}{2}\right)^{\frac{1}{3}} \approx 0.794$
hemisphere (half sphere)		$\frac{2}{3}\pi\,r^3$	$3\pi\,r^2$	$\left(\frac{16}{27}\right)^{\frac{1}{3}} \approx 0.840$
ideal cylinder $(h = 2\,r)$		$\pi r^2 h = 2\pi\,r^3$	$2\pi r(r+h) = 6\pi\,r^2$	$\left(\frac{2}{3}\right)^{\frac{1}{3}} \approx 0.874$
ideal torus $(R = r)$		$2\pi^2 Rr^2 = 2\pi^2\,r^3$	$4\pi^2 Rr = 4\pi^2\,r^2$	$\left(\frac{9}{4\pi}\right)^{\frac{1}{3}} \approx 0.894$
sphere		$\frac{4}{3}\pi r^3$	$4\pi\,r^2$	1

https://en.wikipedia.org/wiki/Sphericity

Intensity-Based Features (First-Order Features)

First-Order Features (Intensity-Based Features)

- **Intensity-based features** sometimes are also called **first-order features**
- First-order features can be computed directly by the intensity, or by the intensity histogram (IH) in the ROI
- Note
  - In IBSI, “intensity-based features” only refer to the features directly computed by the intensity, and the features computed by IH are called “intensity histogram features”
  - To eliminate the ambiguity, we will used first-order features from this point



Example data and segmentation from PANCREAS-CT dataset in TCIA. https://wiki.cancerimagingarchive.net/display/Public/Pancreas-CT

Example First-Order Features (by Intensity)

- Energy: measure the magnitude

$$energy = \sum_{i=1}^{N_p} (I(i))^2$$

- Other Statistics:
  - related to ranking: max, min, medium, range, percentile, …
  - related to moment: mean, variance, skewness, kurtosis, …
  - related to deviation: standard deviation, root mean square, ...

Example First-Order Features (by IH)

- Entropy: measure the uncertainty or randomness

$$entropy = -\sum_{i=1}^{N_h} H(i)\log_2(H(i))$$

- Uniformity: measure the homogeneity

$$uniformity = \sum_{i=1}^{N_h} (H(i))^2$$

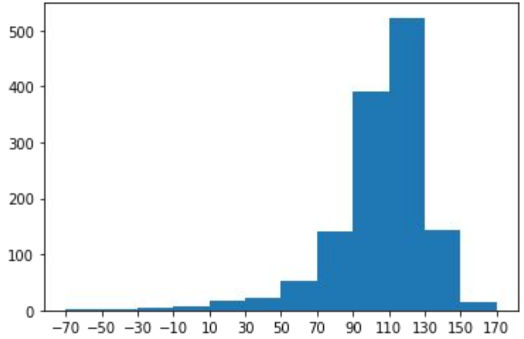
$$H(i) = \frac{\text{number of pixels(voxels) value in } i\text{-th bin}}{\text{total number of image pixels(voxels)}}$$

Texture-Based Features (Higher-Order Features)

Texture-Based Features (Higher-Order Features)

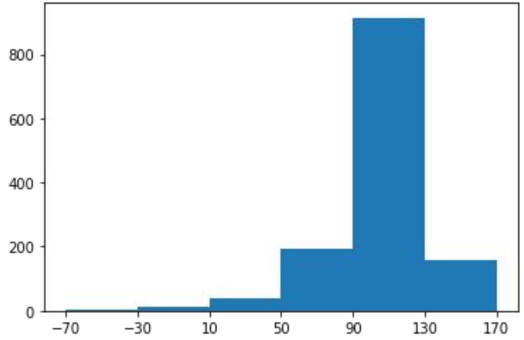
- **Texture-based features** sometimes are also called **higher-order features**
- Texture-based features are computed from the relationship between pixels, which can be seen as a kind of joint distribution
- The different kinds of relationship is usually summarized as different matrices
- In general, the image will be discretized first
  - I.e. each pixel will be assigned into a gray-level, which is equal to the numbering of bin that this pixel belongs to, for the computation of these matrices

Discretization and Gray Level



Bin Width = 20

HU	Gray Level
-70 to -51	1
-50 to -30	2
...	...
150 to 170	12



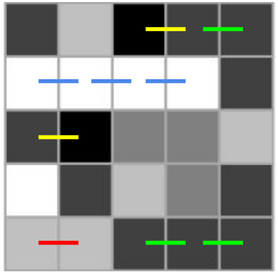
Bin Width = 40

HU	Gray Level
-70 to -30	1
-30 to 10	2
...	...
130 to 170	6

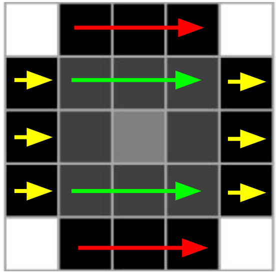
Example data and segmentation from PANCREAS-CT dataset in TCIA. https://wiki.cancerimagingarchive.net/display/Public/Pancreas-CT

Example Texture-Based Feature Classes

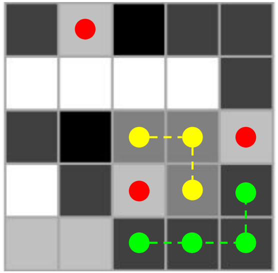
GLCM
Gray Level Co-occurrence Matrix



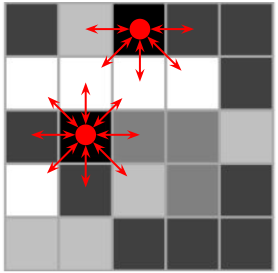
GLRLM
Gray Level Run Length Matrix



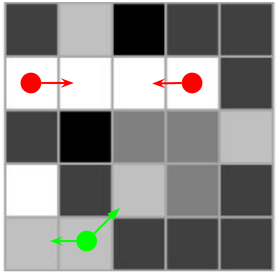
GLSZM
Gray Level Size Zone Matrix



NGTDM
Neighbouring Gray Tone Difference Matrix



NGLDM, GLDM
(Neighboring) Gray Level Dependence Matrix

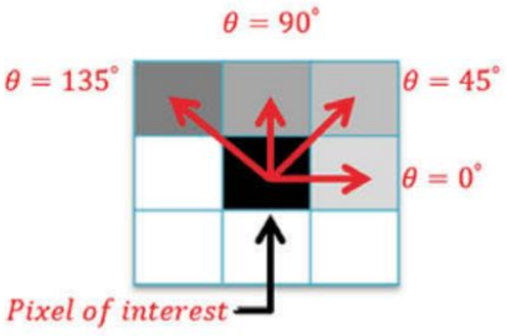


Texture-Based Features
Gray Level Co-Occurrence Matrix

Gray Level Co-Occurrence Matrix Features

- Features derived from (symmetrical) gray level co-occurrence matrix (GLCM)

- GLCM summarized the number of occurrence of the pairs that have specific pattern
- Given distance *d* and angle *θ*, each voxel derives a pair with two voxels
- The (*i*, *j*)-th element describe the number of pairs with two voxels values are *i* and *j*
- In PyRadiomics, each GLCM feature is computed as the average through every possible angles with given distance *d* (default = 1)



Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Example of Symmetrical GLCM Construction

$d = 1,\ \theta = 0°$ $d = 1,\ \theta = 45°$ $d = 2,\ \theta = 135°$

Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Example of Symmetrical GLCM Construction

$N_h = 5$, $d = 1$, $\theta = 0°$

2	1	5	3	2
2	2	4	4	1
4	3	4	2	2
3	2	4	4	1
1	4	4	2	3

I	1	2	3	4	5
1	0	1	0	3	1
2	1	4	3	4	0
3	0	3	0	2	1
4	3	4	2	6	0
5	1	0	1	0	0

Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Example of Symmetrical GLCM Construction

$N_h = 5$, $d = 1$, $\theta = 0°$

2	1	$\boxed{5}$	$\boxed{3}$	2
2	2	4	4	1
4	3	4	2	2
3	2	4	4	1
1	4	4	2	3

I	1	2	3	4	5
1	0	1	0	3	1
2	1	4	3	4	0
3	0	3	0	2	1
4	3	4	2	6	0
5	1	0	1	0	0

Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Example of Symmetrical GLCM Construction

$N_h = 5,\ d = 1,\ \theta = 0°$

2	1	5	3	2
2	2	4	4	1
4	3	4	2	2
3	2	4	4	1
1	4	4	2	3

x2

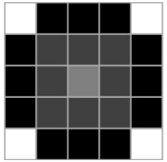
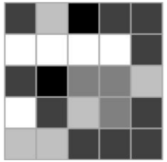
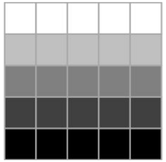
I	1	2	3	4	5
1	0	1	0	3	1
2	1	4	3	4	0
3	0	3	0	2	1
4	3	4	2	6	0
5	1	0	1	0	0

Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Example of Features Derived From GLCM

- Joint Entropy: measure the uncertainty or randomness in neighborhood intensity values

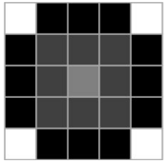
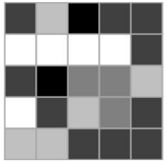
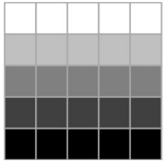
$$joint\ entropy = -\sum_{i=1}^{N_h}\sum_{j=1}^{N_h} \mathrm{GLCM}_{\mathrm{normed}}(i,j)\log_2\left(\mathrm{GLCM}_{\mathrm{normed}}(i,j)\right)$$

			
Joint Entropy $_{\mathrm{GLCM}}$ (d = 1, θ = 0°)	2.846	3.909	2.322
Joint Entropy $_{\mathrm{GLCM}}$ (d = 1, θ = 45°)	2.750	3.858	3.000
Joint Entropy $_{\mathrm{GLCM}}$ (d = 1, θ = 90°)	2.846	3.972	3.000
Joint Entropy $_{\mathrm{GLCM}}$ (d = 1, θ = 135°)	2.750	3.781	3.000

Example of Features Derived From GLCM

- Contrast: measure the gap or disparity among neighboring voxels

$$contrast = \sum_{i=1}^{N_h} \sum_{j=1}^{N_h} |i-j|^2 \mathrm{GLCM}_{\mathrm{normed}}(i,j)$$

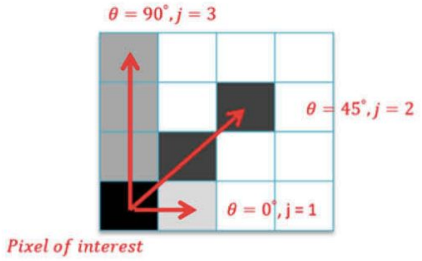
			
Contrast $_{\mathrm{GLCM}}$ (d = 1, θ = 0°)	3.600	2.400	0.000
Contrast $_{\mathrm{GLCM}}$ (d = 1, θ = 45°)	1.750	5.563	1.000
Contrast $_{\mathrm{GLCM}}$ (d = 1, θ = 90°)	3.600	4.850	1.000
Contrast $_{\mathrm{GLCM}}$ (d = 1, θ = 135°)	1.750	4.188	1.000

Texture-Based Features
Gray Level Run Length Matrix

Gray Level Run Length Matrix Features

- Features derived from gray level run length matrix (GLRLM)

- GLRLM describe the information of consecutive lined voxel bin (run length)
- The (i, j)-th element describe the number of voxels in bin i with run length j
- Given angle θ. If for a voxel in intensity bin i, it has longest consecutive same intensity bin for j voxels, then we say this voxel has run length j
- In pyradiomics, each feature is computed as the average through every possible angles



Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Example of Symmetrical GLRLM Construction

$N_h = 5, \theta = 0°$

2	1	5	3	2
2	2	4	4	1
4	3	4	2	2
3	2	4	4	1
1	4	4	2	3

I\j	j=1	j=2	j=3	j=4	j=5
I=1	4	0	0	0	0
I=2	4	2	0	0	0
I=3	4	0	0	0	0
I=4	2	3	0	0	0
I=5	1	0	0	0	0

Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Example of Features Derived From GLRLM

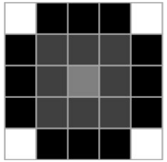
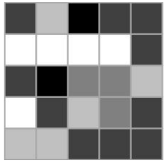
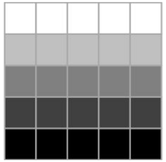
- Short Run Emphasis (SRE): great value indicates more shorter run length and more fine texture

$$SRE = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_r} \frac{\mathrm{GLRLM}(i,j)}{j^2}}{\sum_{i=1}^{N_h} \sum_{j=1}^{N_r} \mathrm{GLRLM}(i,j)}$$

- Long Run Emphasis (LRE): great value indicates more longer run length and more coarse structural texture

$$LRE = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_r} \mathrm{GLRLM}(i,j) j^2}{\sum_{i=1}^{N_h} \sum_{j=1}^{N_r} \mathrm{GLRLM}(i,j)}$$

Example of Features Derived From GLRLM

			
SRE_{GLRLM} (d = 1, θ = 0°)	0.791	0.760	0.040
SRE_{GLRLM} (d = 1, θ = 45°)	0.857	0.935	1.000
SRE_{GLRLM} (d = 1, θ = 90°)	0.791	0.898	1.000
SRE_{GLRLM} (d = 1, θ = 135°)	0.857	0.886	1.000
LRE_{GLRLM} (d = 1, θ = 0°)	2.882	2.882	25.000
LRE_{GLRLM} (d = 1, θ = 45°)	1.571	1.261	1.000
LRE_{GLRLM} (d = 1, θ = 90°)	2.882	1.409	1.000
LRE_{GLRLM} (d = 1, θ = 135°)	1.571	1.667	1.000

Texture-Based Features
Gray Level Size Zone Matrix

Gray Level Size Zone Matrix Features

- Features derived from gray level size zone matrix (GLSZM)

- GLSZM describe the information of consecutive area in gray level
- The (i, j)-th element describe the number of areas with gray level i and area j
- Different from GLCM and GLRLM, GLSZM is independent to rotation

Example of GLSZM Construction

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(5, 1)

https://pyradiomics.readthedocs.io/en/latest/features.html

Example of GLSZM Construction

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(3, 5)

https://pyradiomics.readthedocs.io/en/latest/features.html

Example of Features Derived From GLSZM

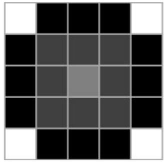
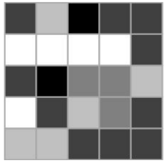
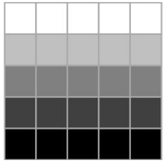
- Small Area Emphasis (SAE): great value indicates more smaller area region and more fine texture

$$SAE = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_r} \frac{\mathrm{GLSZM}(i,j)}{j^2}}{\sum_{i=1}^{N_h} \sum_{j=1}^{N_r} \mathrm{GLSZM}(i,j)}$$

- Large Area Emphasis (LAE): great value indicates more area region and more coarse texture

$$LAE = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_r} \mathrm{GLSZM}(i,j) j^2}{\sum_{i=1}^{N_h} \sum_{j=1}^{N_r} \mathrm{GLSZM}(i,j)}$$

Example of Features Derived From GLSZM

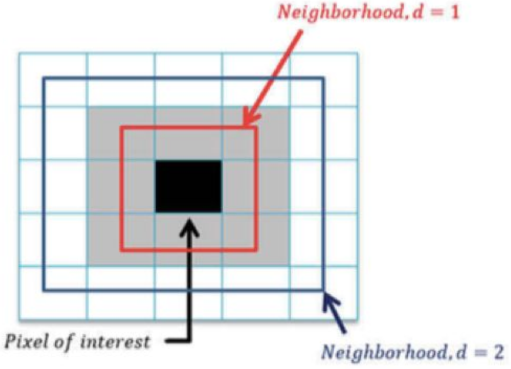
			
$\text{SAE}_{\text{GLSZM}}$	0.718	0.584	0.040
$\text{LAE}_{\text{GLSZM}}$	30.429	7.727	25.000

Texture-Based Features
Neighboring Gray Tone Difference Matrix

Neighboring Gray Tone Difference Matrix Features

- Features derived from neighbouring gray tone difference matrix (NGTDM)

- NGTDM is a table containing:
  - Number and proportion of voxels with certain gray level
  - Quantification of the difference between certain gray level voxels and their neighbors
- The *i*-th element of NGTDM describe the summary of difference between an *i*-th gray level voxels and average gray level of its neighborhood



Parekh, V., & Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, *1*(2), 207–226.

Example of NGTDM Construction

d = 1

$$\mathbf{I} = \begin{bmatrix} 1 & 2 & 5 & 2 \\ 3 & 5 & 1 & 3 \\ 1 & 3 & 5 & 5 \\ 3 & 1 & 1 & 1 \end{bmatrix}$$

i	n_i	p_i (n_i/n_{total})	s_i
1	6	0.375	13.35
2	2	0.125	2.00
3	4	0.25	3.03
4	0	0.00	0.00
5	4	0.25	10.075

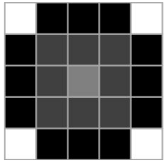
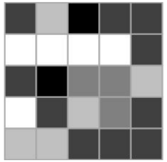
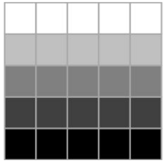
$$s_3 = \boxed{|3 - 12/5|} + \boxed{|3 - 18/5|} + \boxed{|3 - 20/8|} + \boxed{|3 - 5/3|} = 3.03$$

https://pyradiomics.readthedocs.io/en/latest/features.html

Example of Features Derived From NGTDM

- Complexity: high value when image has many primitive components

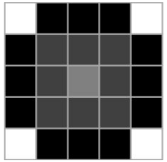
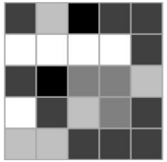
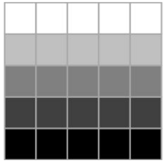
$$Complexity = \frac{1}{N_{v,p}} \sum_{i=1}^{N_h} \sum_{j=1}^{N_h} |i-j| \frac{p_i s_i + p_j s_j}{p_i + p_j}, \text{ where } p_i \neq 0, p_j \neq 0$$

			
Complexity $_{\text{NGTDM}}$	30.527	30.747	8.952

Example of Features Derived From NGTDM

- Strength: high value when primitives are easily defined and visible

$$Strength = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_h} (p_i + p_j)(i - j)^2}{\sum_{i=1}^{N_g} s_i}, \text{ where } p_i \neq 0, p_j \neq 0$$

			
Strength $_{\text{NGTDM}}$	1.298	1.112	6.383

Texture-Based Features
Gray Level Dependence Matrix

Gray Level Dependence Matrix Features

- Features derived from gray level dependence matrix (GLDM)

- GLDM describes the dependency between gray level.
- Given a threshold α. If a region is given, a voxel with gray level t is said to be dependent to center voxel with level u if $|t - u| < \alpha$
- The (i, j) element of GLDM represent the number of cases that j voxels are dependent to the center voxel with i gray level.

- Note
  - In IBSI, this matrix is called neighboring gray level dependence matrix (NGLDM)

Example of GLDM Construction

$\alpha = 0, d = 1$

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix} \qquad \mathbf{P} = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 2 & 3 & 0 \\ 1 & 4 & 4 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}$$

https://pyradiomics.readthedocs.io/en/latest/features.html

Example of GLDM Construction

$\alpha = 0$, $d = 1$

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix}$$

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 2 & 3 & 0 \\ 1 & 4 & 4 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}$$

https://pyradiomics.readthedocs.io/en/latest/features.html

Example of GLDM Construction

$\alpha = 0,\ d = 1$

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix} \qquad \mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix}$$

$$\mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix} \qquad \mathbf{I} = \begin{bmatrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{bmatrix}$$

$$\mathbf{P} = \begin{bmatrix} 0 & 1 & 2 & 1 \\ 1 & 2 & 3 & 0 \\ 1 & 4 & 4 & 0 \\ 1 & 2 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{bmatrix}$$

https://pyradiomics.readthedocs.io/en/latest/features.html

Example of Features Derived From GLDM

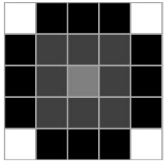
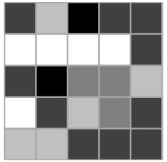
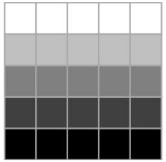
- Small Dependence Emphasis (SDE): great value indicates smaller dependence and less homogeneous texture

$$SDE = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_d} \frac{\text{GLDM}(i,j)}{j^2}}{\sum_{i=1}^{N_h} \sum_{j=1}^{N_d} \text{GLDM}(i,j)}$$

- Large Dependence Emphasis (LDE): great value indicates larger dependence and more homogeneous texture

$$LDE = \frac{\sum_{i=1}^{N_h} \sum_{j=1}^{N_d} \text{GLDM}(i,j) j^2}{\sum_{i=1}^{N_h} \sum_{j=1}^{N_d} \text{GLDM}(i,j)}$$

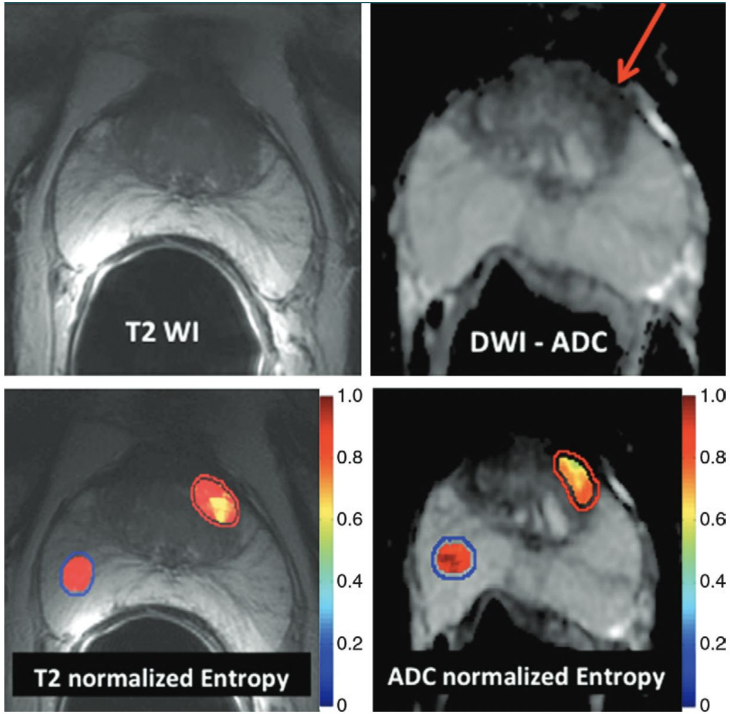
Example of Features Derived From GLDM

			
SDE_{GLDM}	0.278	0.350	0.167
LDE_{GLDM}	9.960	6.360	7.000

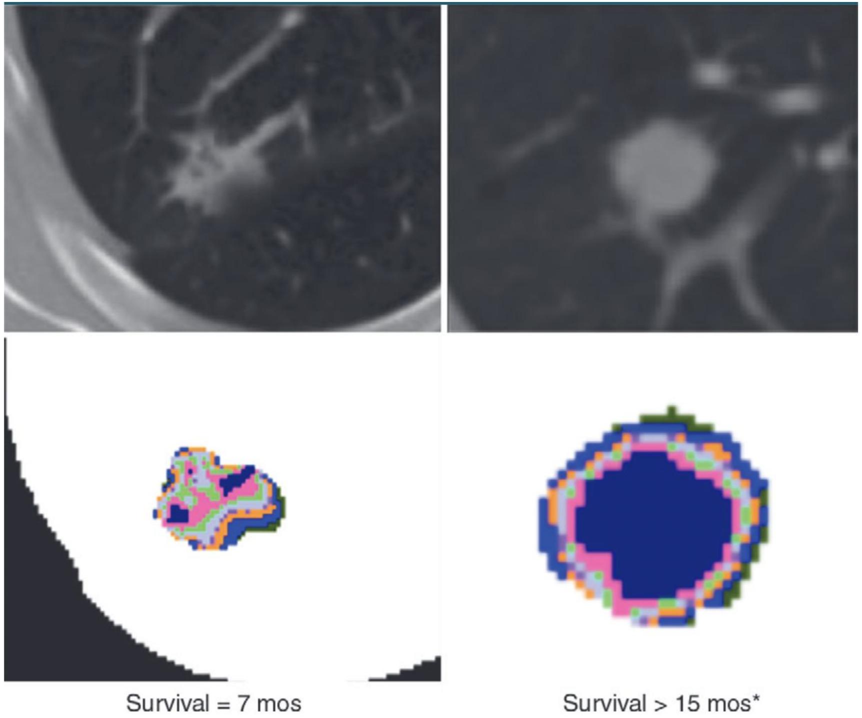
Example Application

Volume VOI and Voxel-wise VOI

Voxel-wise analysis (5 by 5 by 1 window)
Find prostate cancer on T2-weighted MRI



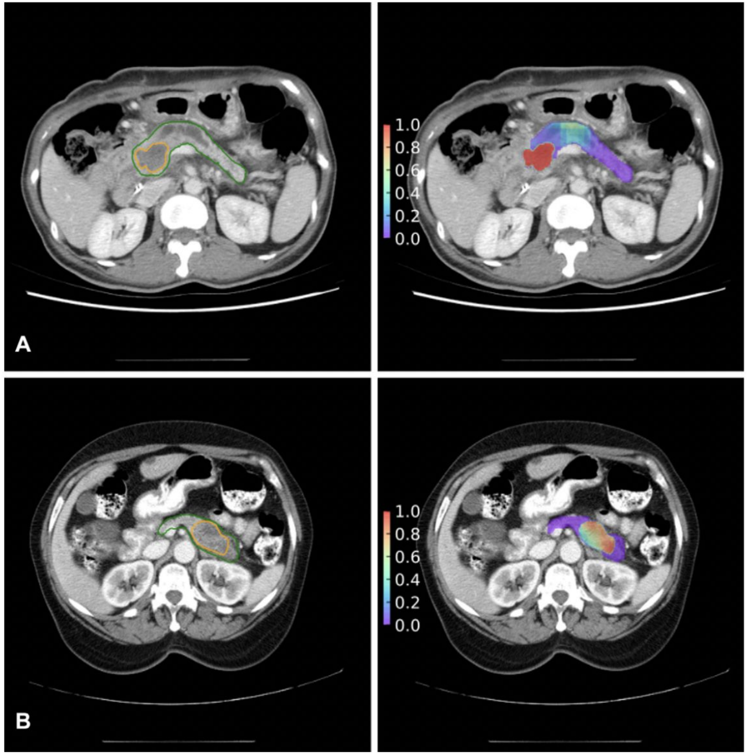
Volume analysis
Prognosis on lung CT



Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images Are More than Pictures, They Are Data. *Radiology*, *278*(2), 563–577.

Distinguish Tissue of Pancreatic Cancer on CT

- Given region of interest (ROI)
- Two-stage analysis
  - Patch-based analysis: crop ROI patches, extract radiomic features, and then classify patches into cancerous or non-cancerous by XGBoost
  - Patient-based analysis: use rule-based method to aggregate the results from patch-based analysis into patient-based analysis
- Results:
  - Patch-based: 94.3% sensitivity, 87.4% specificity
  - Patient-based: 94.7% sensitivity, 95.4 specificity



Chen, P.-T., Chang, D., Yen, H., Liu, K.-L., Huang, S.-Y., Roth, H., Wu, M.-S., Liao, W.-C., & Wang, W. (2021). Radiomic Features at CT Can Distinguish Pancreatic Cancer from Noncancerous Pancreas. *Radiology. Imaging Cancer*, *3*(4), e210010.

Explainability

Local model	
Feature	**Gain value**
First order: median	670.4
NGTDM: busyness	654.5
GLCM: cluster shade	425.5
First order: 90 percentile	405.2
First order: skewness	169.8
GLDM: gray level non-uniformity	161.1
GLDM: large dependence low gray level emphasis	123.2
First order: interquartile range	88.5
GLDM: dependence non-uniformity	80.0
GLCM: correlation	74.9
GLRLM: run length non-uniformity normalized	69.7
First order: mean	65.3
GLCM: sum entropy	62.8
GLRLM: run entropy	59.4

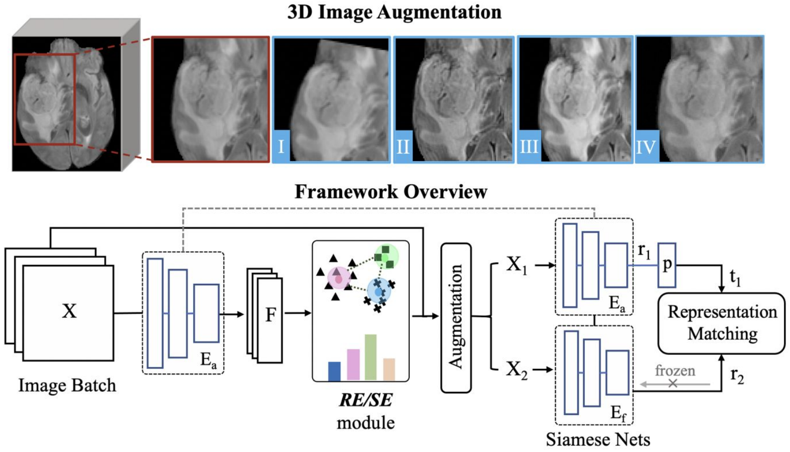
Features related to intensity
(cancerous < non-cancerous)

Features related to heterogeneity
(cancerous > non-cancerous)

Chen, P.-T., Chang, D., Yen, H., Liu, K.-L., Huang, S.-Y., Roth, H., Wu, M.-S., Liao, W.-C., & Wang, W. (2021). Radiomic Features at CT Can Distinguish Pancreatic Cancer from Noncancerous Pancreas. *Radiology. Imaging Cancer*, *3*(4), e210010.

Handicraft Radiomic Features & Deep Learning Features

- Use self-supervised learning to extract "deep-learning-based" features
- Mix traditional radiomic features and deep-learning-based features can get better performance



Methods	*BraTS*		*Lung cancer staging*	
	Sensitivity/Specificity		Overall/Minor-class accuracy	
	Full labels	*50% labels*	*Full labels*	*50% labels*
Trad. radiomics	0.888/0.697	**0.848**/0.697	0.490/0.375	0.481/0.325
Rubik's cube [25]	0.744/0.526	0.680/0.486	0.459/0.325	0.433/0.275
3DSiam	0.844/0.407	0.808/0.526	0.459/0.300	0.445/0.300
3DSiam+SE	0.848/0.513	0.824/0.566	0.471/0.350	0.443/0.325
3DSiam+RE	0.868/0.486	0.828/0.605	0.459/0.375	0.445/0.325
Trad.+3DSiam	0.904/0.645	0.804/0.566	0.495/0.350	0.486/0.350
Trad.+3DSiam+SE	0.916/**0.711**	**0.848/0.763**	**0.538**/0.375	**0.519**/0.350
Trad.+3DSiam+RE	**0.920/0.711**	0.804/0.763	0.524/**0.425**	0.502/**0.40**
Supervised learning	0.888/0.711	0.804/0.566	0.526/0.375	0.467/0.325

Li, H., Xue, F.-F., Chaitanya, K., Luo, S., Ezhov, I., Wiestler, B., Zhang, J., & Menze, B. (2021). Imbalance-Aware Self-supervised Learning for 3D Radiomic Representations. *Medical Image Computing and Computer Assisted Intervention – MICCAI 2021*, 36–46.

Other Resource

Resource

- Related python package
  - pyradiomics
- Notebook demo
  - Extraction of radiomic features by pyradiomics:
    https://colab.research.google.com/drive/1J7TQf5_ROSxofVH9XIHu1OjRCzVyxJva?usp=sharing
  - Classification by radiomic features:
    https://colab.research.google.com/drive/18vaGYTHbJfqBH1OHkL8oA0Fe0zxHaNss?usp=sharing

Videos

- MeDA Lab
  - 影像也可以變數值資料？ 影像組學簡介，特徵抽取超詳細介紹！
    - 上集 : https://youtu.be/jE_uHDsTcVE
    - 下集 : https://youtu.be/HwF0tpzMrNs

Thank You!