Radiomics in Head and Neck Cancer

Severin Langberg

December 6, 2018

OUTLINE

1) **Personalized cancer medicine:**
   - Radiomics: Motivation and principles
2) **A head and neck cancer study:**
   - Predicting patient treatment outcome

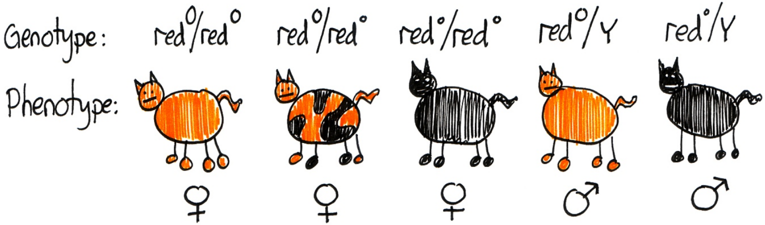
Personalized medicine:

- Predict patient response → tailor therapies

Impersonalized medicine:

Personalized medicine:

- Phenotypic differences in medical images



1

[1] http://cubocube.com/dashboard.php?a=1181b=1260c=103

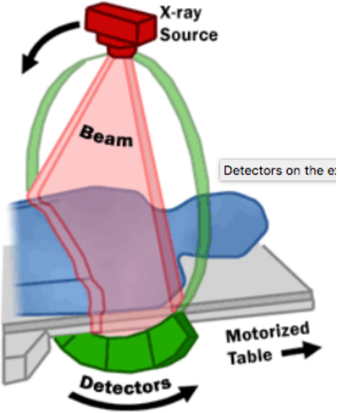
Imaging for precision medicine:
Characterization and quantification of disease stage

- Non-invasive
- Different modalities in clinical practice

- Heterogenous aquisition protocols
- Qualitative not quantitative

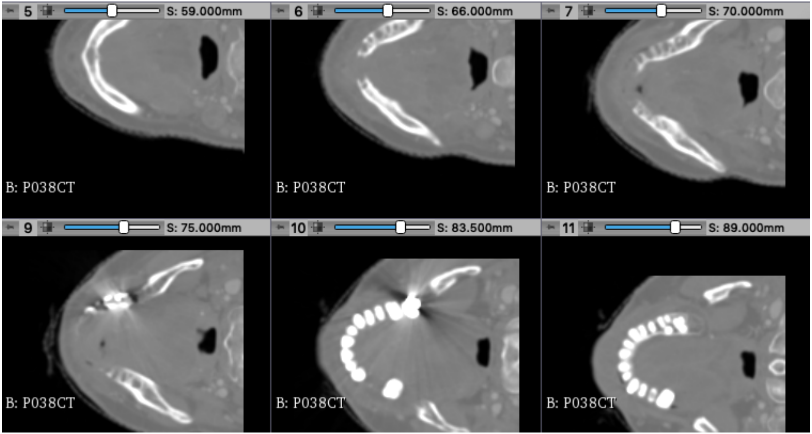
Computed Tomography (CT)

[2]

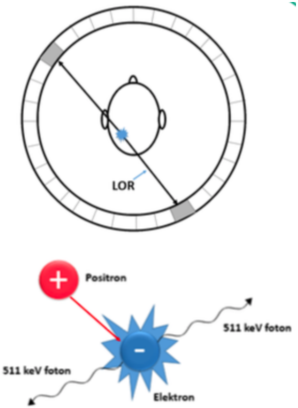


- Displays bodily structures
- Measures tissue X-ray absorption

[2] U.S. Food and Drug Administration. www.fda.gov. 11/27/2018

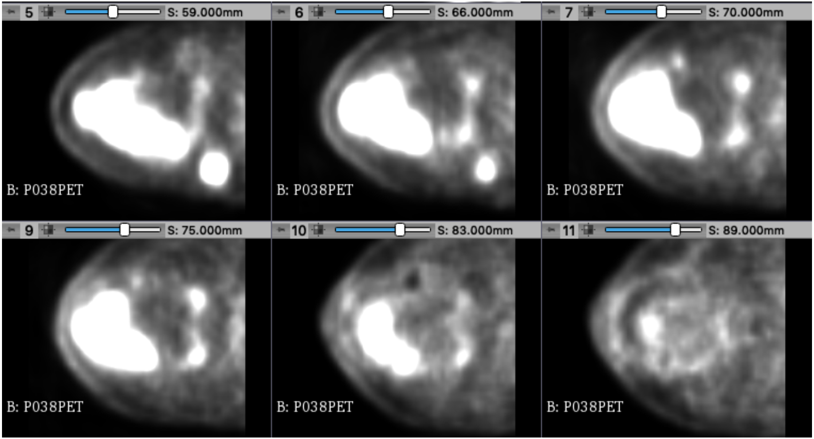

5
S: 59.000mm
B: P038CT
6
S: 66.000mm
B: P038CT
7
S: 70.000mm
B: P038CT
9
S: 75.000mm
B: P038CT
10
S: 83.500mm
B: P038CT
11
S: 89.000mm
B: P038CT

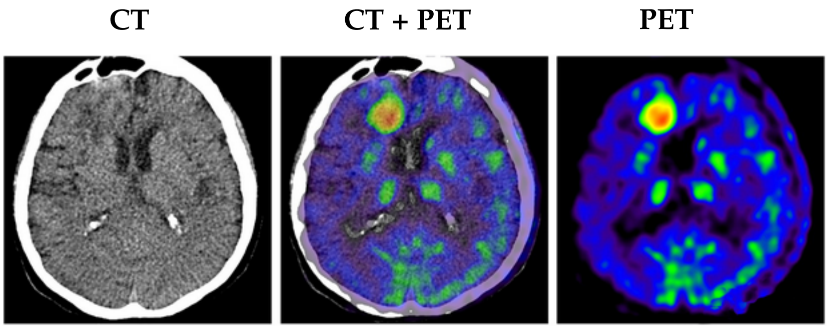
Positron Emission Tomography (PET)

[3] 

- Captures metabolic activity
- Inject radioactive tracers $\rightarrow$ record photon emission

[3] U.S. Food and Drug Administration. www.fda.gov. 11/27/2018

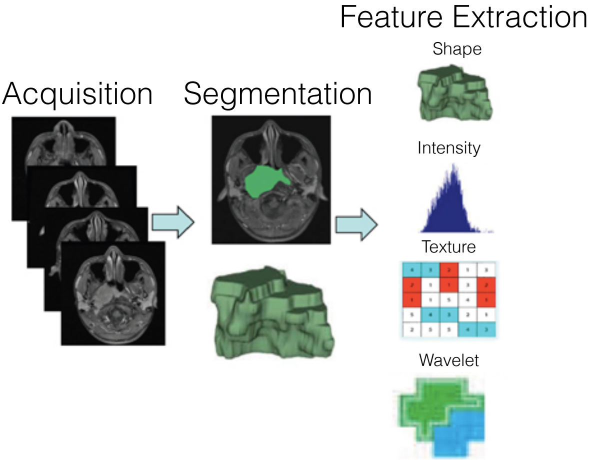

5 S: 59.000mm
B: P038PET
6 S: 66.000mm
B: P038PET
7 S: 70.000mm
B: P038PET
9 S: 75.000mm
B: P038PET
10 S: 83.000mm
B: P038PET
11 S: 89.000mm
B: P038PET



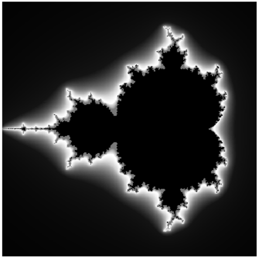
[4]

[4] Rutten et al. 2007. PET/CT of skull base meningiomas using 2-18F-fluoro-L-tyrosine: initial report.

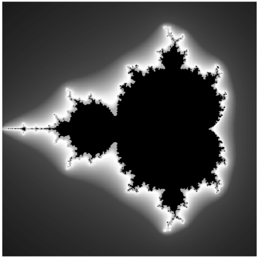
RADIOMICS


Feature Extraction
Acquisition
Segmentation
Shape
Intensity
Texture
Wavelet

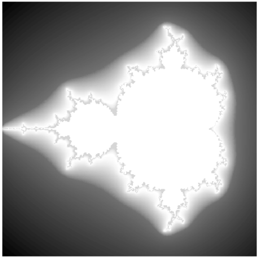
FILTERING



Original

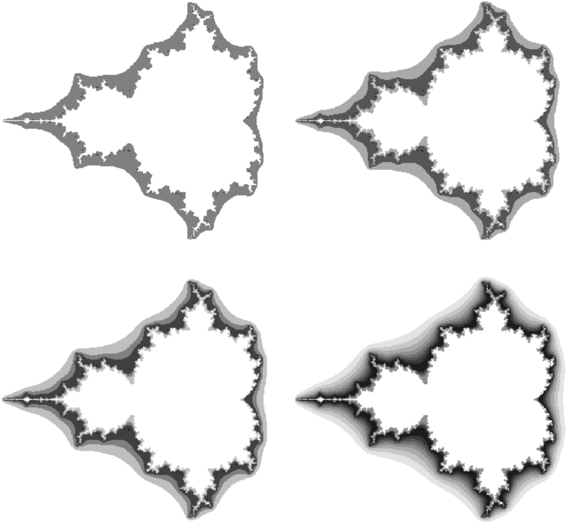


Squareroot



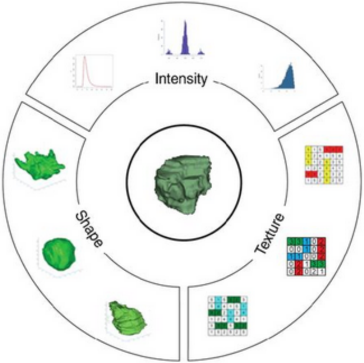
Logarithm

Discretization



Radiomics

- Shape-based (16)
- First-order statistics (19)
- Texture (75)



[6] Zhang, Ouyang, & Gu et al (2017). Advanced nasopharyngeal carcinoma

Radiomics

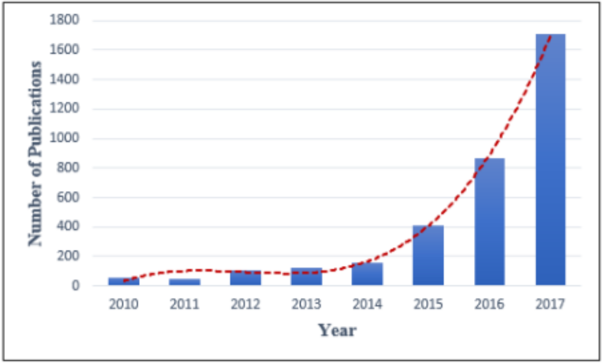
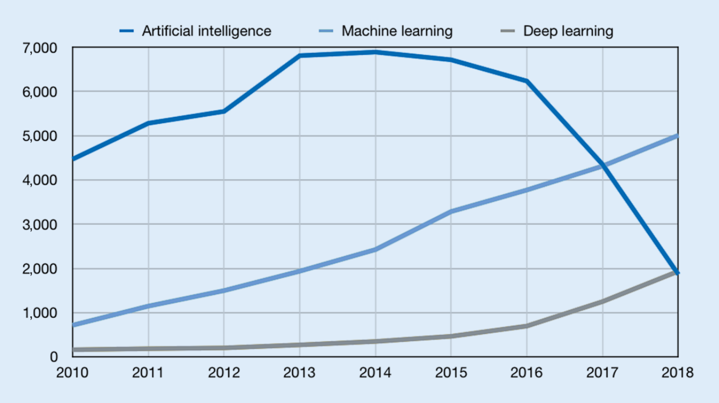


Fig. 1: Increasing interest in Radiomics based on data from Google Scholar. [7]

[7] Mohammadi et al. (2018). From Hand-Crafted to Deep Learning-based Cancer Radiomics: Challenges and Opportunities

RADIOMICS + AI

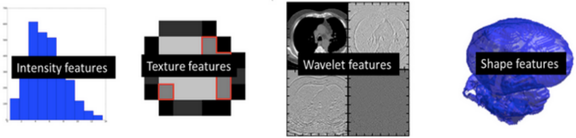


AI in PubMed publications by year. [8]

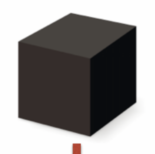
[8] Langs, Röhrich, Hofmanninger et al (2018). Radiologe

Radiomics + AI

Radiomic Features



Machine Learning

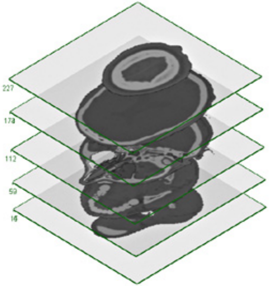


Treatment outcome

CASE

Goal: Classify patients by radiotherapy response
Data set: 198 head and neck cancer patients

- Disease-free survival/locoregional relapse
- Clinical variables (sex, age, smoking habits, etc.)
- PET + CT stacks



Case

Base study:

- Graylevel discretization (PET = 16 bins, CT = 128 bins)
- Square root transformation
- 4-fold nested cross-validation (CV)
  - PLSR, logistic regression (L2), LDA, QDA
  - Area Under Curve (TPR / FPR) of 0.66

Parmar & **Grossmann et al.**: [9]

- 464 lung cancer patients (CT)
- Validation set approach
- Adopted hyperparameters
- AUC of 0.68 (RF + WLCX)

[9] Parmar, Grossmann et al. 2015. *Machine learning methods for quantitative radiomic biomarkers*

The .632+ bootstrap method

- K-fold CV: nearly unbiased, but can be
  - highly variable
  - pessimistic
- Based on .632 estimator by Bradley Efron

$$\varepsilon_{.632} = 0.368\varepsilon_T + 0.632\varepsilon_V$$

Issue: Cannot handle overfitting

The .632+ bootstrap method

- Improved: Efron & Tibhirani

$$\varepsilon_{.632+} = \omega\varepsilon_T + (1-\omega)\varepsilon_V$$

- variable weighting

$$\omega = \frac{0.632}{1 - 0.368R(\gamma)}$$

R: relative overfitting rate (degree of overfitting)
γ: no-information error rate ($\geq$ majority class)

.632+ ESTIMATOR

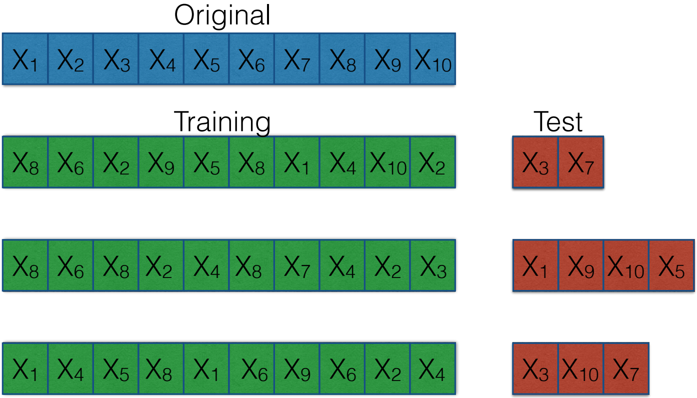
Alternative:

$$\varepsilon_{.632+} = \varepsilon_{.632} + (\overline{\varepsilon_T} - \varepsilon_V)\frac{0.367 \cdot 0.632\overline{R}}{1 - 0.367\overline{R}}$$

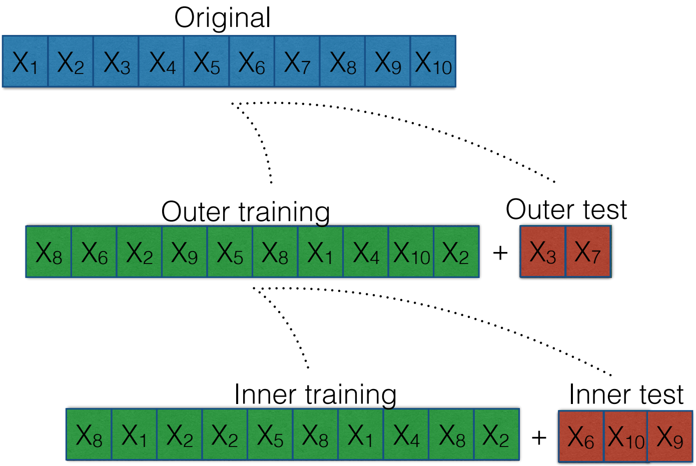
where

- $\overline{\varepsilon_T} = \min\{\varepsilon_T, \gamma\}$
- $\overline{R} = \begin{cases} \varepsilon_T & \text{if } \gamma > \varepsilon_V \\ 0 & \text{if } \gamma \leq \varepsilon_V \end{cases}$

Bootstrap Out-of-Bag



Nested bootstrap Out-of-Bag



EXPERIMENTS

Preprocessing:

- Graylevel discretization (PET = 16 bins, CT = 128 bins)
- Square root transformation

Feature extraction:

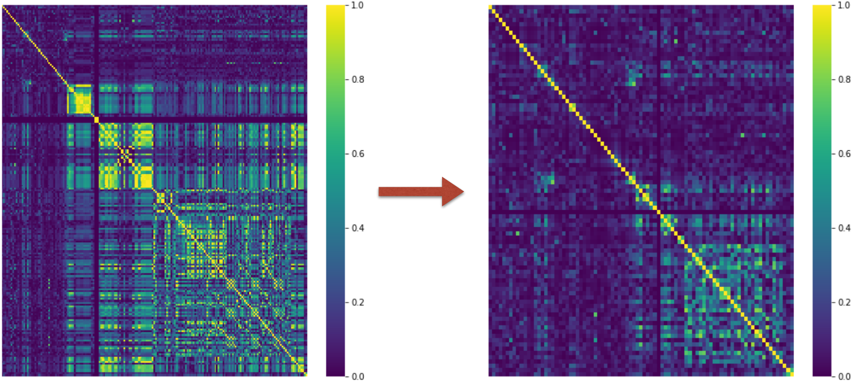
- 188 PyRadiomics image features
- 49 clinical variables (encoded)

EXPERIMENTS

Postprocessing: Remove unattainable + correlated features

- 37 PET textural features failed extraction (flat image regions)
- In total 87 features to analysis

Correlations (≥ 0.85)



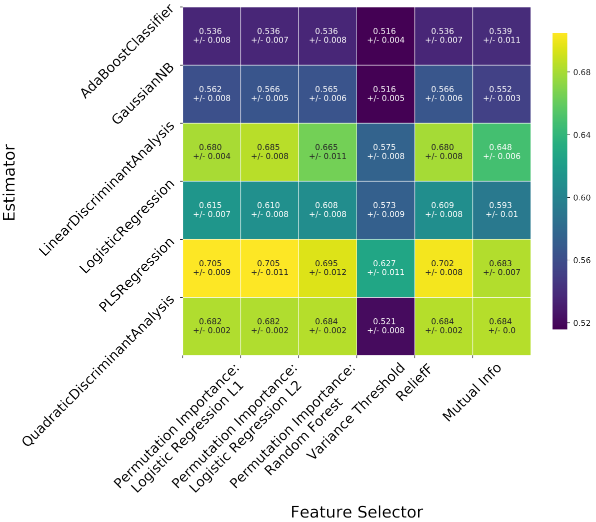
Model comparisons

- 10 experimental parallels (DFS/LRR)
- 50 out-of-bag nested iterations
- AUC (TPR / FPR) based .632+ score

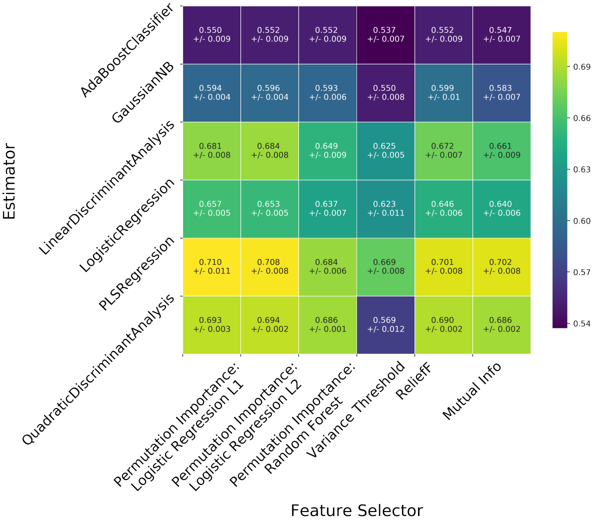
Note: majority of events

- Disease free survival: 0.67
- Locoregional relapse: 0.75

Locoregional relapse



Disease free survival



Observations

- PLSR (best performing):
  - Disease free survival: $0.67 < 0.710$)
  - Locoregional relapse: $0.75 > 0.705$)
- Improved performance compared to base study (0.66)

Limitations

- Alternative image transformations
- Feature engineering/Deep feature synthesis
- Increase model complexity (boosting/DL)

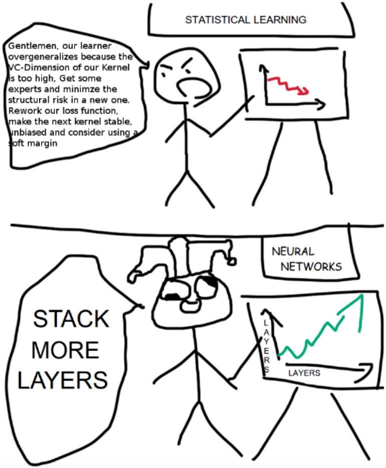
Limitations

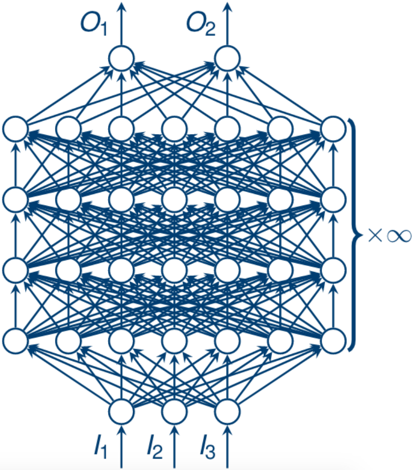
TABLE I: Popular data sets for performing Radiomics.

Data Set	Reference	Year	Imaging Modality	Application
LIDC-IDRI	[15]	2015	CT	Lung Tumor
NSCLC-Radiomics	[2]	2014	CT	Lung Tumor
NSCLC-Radiomics-Genomics	[2]	2014	CT, Gene expression	Lung tumor
LGG-1p19qDeletion	[45]	2017	MRI	Brain Tumor
Head-Neck-PET-CT	[46]	2017	PET, CT	Head-and-Neck Cancer
BRATS2015	[50]	2015	MRI	Brain Tumor

[10]

[10] Mohammadi et al. (2018). From Hand-Crafted to Deep Learning-based Cancer Radiomics: Challenges and Opportunities

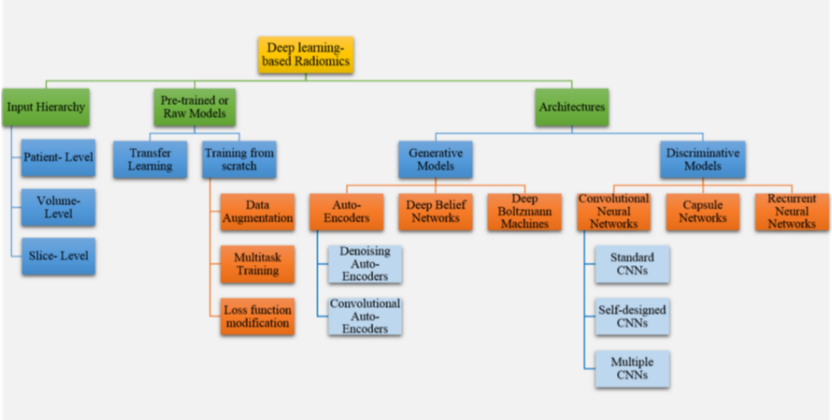




[11]

[11] https://www.reddit.com/r/ProgrammerHumor/comments/5si1f0/machine_learning_approaches/

Deep learning in radiomics



[12]

[12] Afshar, Parnian, et al. "From Hand-Crafted to Deep Learning-based Cancer Radiomics. (2018).

Feature reduction techniques

Category	Description	Methods
Supervised	Considers the relation of features with the class labels and features are selected mostly based on their contribution to distinguish classes.	
• Filtering (Univariate)	Test the relation between the features and the class label one by one.	Fisher score (FSCR); Wilcoxon rank sum test; Gini index (GINI); Mutual information feature selection (MIFS); Minimum redundancy maximum relevance (MRMR), and; Student t-test [44].
• Wrapper (Multivariate)	Considers both relevancy and redundancy.	Greedy forward selection, and Greedy backward elimination.
Unsupervised	Does not consider the class labels and its objective is to remove redundant features.	
• Linear	Features have linear correlations.	Principle Component Analysis (PCA), and; Multidimensional scaling (MDS)
• Nonlinear	Features are not assumed to be lied on a linear space.	Isometric mapping (Isomap), and; Locally linear embedding (LLE).

[13]

[13] Afshar, Parnian, et al. "From Hand-Crafted to Deep Learning-based Cancer Radiomics. (2018).