Uncertainty, Robustness, and Harmonization in Radiomics Studies of Lung Cancer

Dennis Mackin

AAPM 2017

Research group and collaborators

Our group
- Laurence Court
- Joy Zhang
- Jinzhong Yang
- Ryan Williamson
- Joonsang Lee
- Rachel Ger
- Melinda Chi

Past members
- David Fried
- Xenia Fave

Imaging Physics
- A. Kyle Jones
- Rick Layman

Center for Radiation Oncology Research
- Alma Faust
- Sunil Krishnan

Radiation Oncology and Radiology
- Daniel Gomez
- Chaan Ng
- Joe Chang
- Dave Fuller
- Mona Jomaa
- Hesham Elhalawani

Statistics
- Shouhao Zhou
- Susan Tucker
- Francesco Stingo
- Arvind Rao

Medical Center
- Cristina Dodge (Texas Children's Hospital)
- Brian Taylor (Baylor C. of Medicine, VA)
- Edgardo Rodriguez-Rivera (Methodist Hospital)
- Ramiro Pino (Methodist Hospital)

2

Acknowledgements and Disclosures

NCI award R03CA178495 has provided salary support and funded some of this research

I have no financial conflicts of interests to disclose.

Radiomics workflow

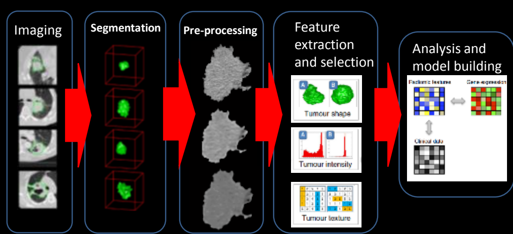


Figure adapted from Aerts et al., Nature Communications 2015, courtesy of L. Court

Features

What is a robust quantitative imaging feature?

- **Repeatability** is the variability in features extracted from images acquired under the same conditions
  - Same subject, imaging system, and image acquisition parameters
- **Reproducibilty** is the variability in features extracted from images acquired under different conditions
  - Same subject, but may have different scanner, kernel, image FOV, slice thickness, etc

Properties of an ideal radiomics feature

Property	Methodology for evaluation
Repeatable	Test-retest data
Reproducible	Compare metrics through different analysis pipelines
Low redundancy with other features	Quantify and rank statistical correlations between features
Predictive or prognostic of a clinical endpoint	Improved models

Table based on Hatt et al, Eur J Nucl Mol Imaging, Published online June 2016

7

Radiomics focuses on improvements of image analysis, using an automated high-throughput extraction of large amounts (200+) of quantitative features of medical images.

Lambin *et al.*, 'Radiomics: extracting more information from medical images using advanced feature analysis', Euro. J. of Cancer, 48(4), 2012

extraction of large amounts (200+) of quantitative features

Image Intensity histogram
(first order features)

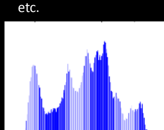
- Mean, standard deviation, skewness, kurtosis, entropy, etc.

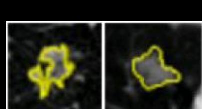
Shape features

- Volume, surface area, convex hull, roundness, sphericity, etc.

Texture features

- Gray level run length
- Gray level co-occurrence
- Neighborhood gray tone difference matrix





Hua, K.-L.,et al., *OncoTargets and ther.*, (2015)

Courtesy of Laurence Court

Feature Categories

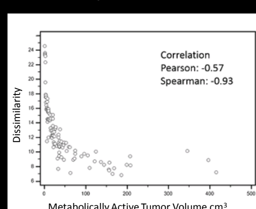
Shape	Intensity Histogram	Gray Level Run-Length Matrix (GLRL)	Gray Level Co-occurrence Matrix (GLCM)	Neighborhood Intensity Difference Matrix
NumberOfVoxel	Max	ShortRunEmphasis	Contrast	Busyness
Volume	Min	LongRunEmphasis	Correlation	Coarseness
Mass	Median	GrayLevelNonuniformity	Energy	Complexity
Roundness	Mean	RunLengthNonuniformity	Homogeneity	Contrast
Orientation	Range	RunPercentage	Entropy	TextureStrength
Convex	Kurtosis	LowGrayLevelRunEmphasis	AutoCorrelation	
NumberOfObjects	Skewness	HighGrayLevelRunEmphasis	Dissimilarity	
ConvexHullVolume3D	StandardDeviation	ShortRunLowGrayLevelEmphasis	ClusterShade	
ConvexHullVolume	MeanAbsoluteDeviation	ShortRunHighGrayLevelEmphasis	ClusterProminence	
SurfaceArea	MedianAbsoluteDeviation	LongRunLowGrayLevelEmphasis	ClusterTendency	
SurfaceAreaDensity	InterQuartileRange	LongRunHighGrayLevelEmphasis	MaxProbability	
MeanBreadth	Percentile		DifferenceEntropy	
VoxelSize	Quantile		Homogeneity2	
Compactness1	Entropy		InformationMeasureCorr1	
Compactness2	Uniformity		InformationMeasureCorr2	
Max3DDiameter			InverseDiffMomentNorm	
SphericalDisproportion			InverseDiffNorm	
Sphericity			InverseVariance	
			SumAverage	
			SumEntropy	
			SumVariance	
			Variance	

Wavelet and Laplacian of Gaussian versions of these features bring numbers into the hundreds.

Also popular are gray level size zone [GLSZ, Thibault et al., *Int. J. of Pat. Rec. and AI* (2013)] and fractal features

Feature Redundancy

- Features designed for 2D aerial photos not 3D medical images
  - Many correlated with volume
- Many features from a few matrices
  - Gray level co-occurrence matrix
  - Gray level run-length matrix
  - high degree of correlation (i.e. redundant)
- Spearman rank correlation often more effective than Pearson correlation which measure the linearity of the relationship.



Hatt, M., et al., *Eur. J. of nuc. med. and mol. imaging,*(2017)

Reducing Volume Dependence

Feature	Original algorithm		Corrected algorithm	
Busyness (13)	$\frac{\left[\sum_{i=0}^{G_h} p_i s(i)\right]}{\left(\sum_{i=0}^{G_h}\sum_{j=0}^{G_h} ip_i - jp_i\right)}$	[1]	$\frac{\left[\sum_{i=0}^{G_h} p_i \frac{s(i)}{N(i)}\right]}{\left(\sum_{i=0}^{G_h}\sum_{j=0}^{G_h} ip_i - jp_i\right)}$	[2]
Coarseness (13)	$\left[\epsilon + \sum_{i=0}^{G_h} p_i s(i)\right]^{-1}$	[3]	$\left[\epsilon + \sum_{i=0}^{G_h} p_i \frac{s(i)}{N(i)}\right]^{-1}$	[4]
Gray-level nonuniformity (14)	$\frac{\sum_{i=1}^{N_g}\left[\sum_{j=1}^{N_r} p(i,j)\right]^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j)}$	[5]	$\frac{1}{N_v}\frac{\sum_{i=1}^{N_g}\left[\sum_{j=1}^{N_r} p(i,j)\right]^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j)}$	[6]
Run-length nonuniformity (14)	$\frac{\sum_{j=1}^{N_r}\left[\sum_{i=1}^{N_g} p(i,j)\right]^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j)}$	[7]	$\frac{1}{N_v}\frac{\sum_{j=1}^{N_r}\left[\sum_{i=1}^{N_g} p(i,j)\right]^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j)}$	[8]
Energy	$\sum_{i}^{N_v} p(i)^2$	[9]	$\frac{1}{N_v}\sum_{i}^{N_v} p(i)^2$	[10]

Spearman correlation for these features went from [0.95 - 1.00] to [0.0 – 0.77] for CT images from 107 NSCLC patients..

Fave, X., et al., *Trans.Cancer Research,* (2016)

Reducing Voxel Size Dependence

$$f_m(P,T) = f(P,T) * V(P,T) \quad (1)$$

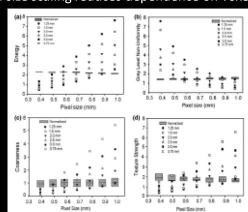
$$f_m(P,T) = \frac{f(P,T)}{V(P,T)} \quad (2)$$

$$f_m(P,T) = \frac{f(P.T)}{log[n(P,T)]} \quad (3)$$

V(P,T) = voxels size = $\frac{\text{volume}}{\text{number voxels}}$

P = pixel spacing
T = slice thickness
f = uncorrected feature value
n = number of voxels

Voxel size scaling reduces dependence on voxel size



Shafiq-ul-Hassan, et al, *Medical physics,* 44(3), 2017

Standardization – Radiomics Software

Software package	Imaging modality and format	ROI definition	Features and image pre-processing	Website
IBEX (free open source)	CT, PET, MR	DICOM-RT	109: intensity, texture, shape	http://bit.ly/IBEX_MDAnderson
	DICOM, Pinnacle native format	Editing and free drawing	Smoothing, resampling, enhancement	
CGITA (free open source)	Designed for PET; CT, MR tested-DICOM	DICOM-RT, PMOD	72: intensity, texture	http://code.google.com/p/cgita
		Region growing and FCM	No pre-processing	
MaZda (free open source)	Designed for MR	Thresholding, deformable surface	279: intensity, texture, shape, wavelet	http://www.eletel.p.lodz.pl/programy/mazda/
	DICOM		Resampling, discretization, normalization	
RADIOMICS™ (OncoRadiomics, Maastricht, Netherlands) (commercial)	CT, PET, MR	DICOM-RT	543: intensity, texture, shape, wavelet	http://www.oncoradiomics.com
	DICOM	Plug-in for several TPS	Laplacian of Gaussian	
			Resampling, discretization	
TexRAD™ (Feedback plc, Cambridge, UK) (commercial)	CT, PET, MR	DICOM-RT	~30: texture and filtering (Laplacian of Gaussian)	http://www.texrad.com
	DICOM	Editing, thresholding		

Larue, R. T., et. al., *British j. of rad.,* 20160665. (2017)

14

Image Biomarker Standardization Initiative

Aim: To achieve consensus on and provide:
- feature nomenclature
- common image feature definitions
- feature and image processing benchmarks
- reporting guidelines

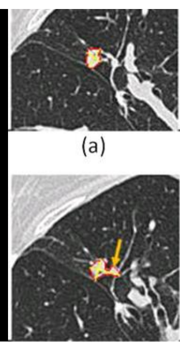
Current status:		current	09-10-16
• no agreement	(< 3 institutions or < 50% identical)	52	192
• agreement	(> 50% identical)	59	85
• standardised	(> 80% identical)	240	25

Contact: Dr. Alex Zwanenburg
https://arxiv.org/abs/1612.07003v4

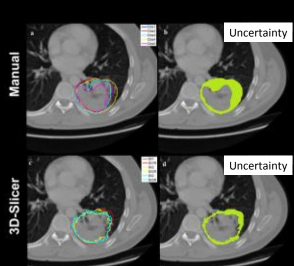
Segmentation

Segmentation

- Inter-observer variability
- Intra-observer bias (and even variability)
- Manual contouring is time intensive – radiomics studies have hundreds or thousands of contours
- Yet – manual delineation remains the gold standard



Semi-automated Segmentation in NSCLC



Study compared contours in 20 stage IB – IIIB non-small cell lung cancer

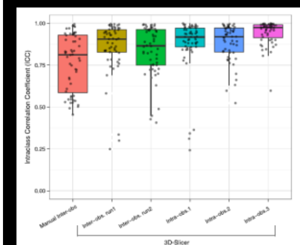
Methods
- 3D Slicer Competitive region growing algorithm
- Compared 3 semi-auto contours by 3 users to 5 manual contours (PET + CT)

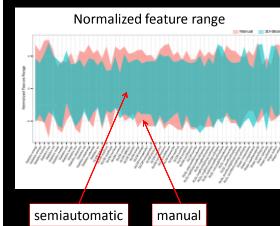
Results
- Semi-automated showed less uncertainty
- Both manual and semi-automated diameters were correlated to pathology results (Spearman r = 0.92 and 0.89 respectively)

Conclusion
- Semi-automated contouring is accurate and more stable than manual contouring.

Velazquez *et al. Scientific Reports*, Article number: 3529 (2013) doi:10.1038/srep03529

Segmentation – effect on feature robustness





Parmar et al, PlosONE, 9(7), e102107, 2014

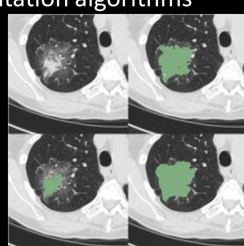
19

Lung nodule segmentation algorithms

- 3 lung nodule segmentation algorithms
- 52 tumors in 41 CT image sets
- Nodules range from 0.03 to 66 cc
- Compared multiple runs of each algorithm

- Results
  – Intra-algorithm results were less variable than inter-algorithm results
  – Least biased was also least repeatable

- Conclusion
  – Large differences between algorithms underscored need to use the same segmentation algorithm throughout a study



Kalpathy-Cramer, *et al. Journal of digital imaging,* 29(4), 476-48 (2016)

Features values depend on segmentation

PET feature	50% SUVmax Value	40% SUVmax Value	40% SUVmax Change	FLAB Value	FLAB Change
Volume	26.59	34.61	+30%	40.36	+51%
Entropy	3.64	3.89	+7%	3.98	+10%
Total energy	$4.72 \cdot 10^4$	$5.28 \cdot 10^4$	+12%	$5.57 \cdot 10^4$	+18%
Homogeneity	0.29	0.25	-12%	0.24	-17%
Dissimilarity	2.97	3.60	+21%	3.95	+33%
Zone percentage	0.35	0.39	+9%	0.39	+10%
Size-zone variability	0.36	0.38	+6%	0.38	+4%
Intensity variability	0.09	0.08	-15%	0.07	-18%

FLAB: Fuzzy locally adaptive Bayesian delineation

Larue, R. T., et. al., *British j. of rad.,* 20160665. (2017)

CT

Test/retest Studies from 4D-CT

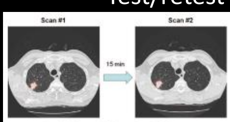


FIG. 1. Test-retest CT images taken 15 min apart.

Concordance Correlation Coeffient

Inter-scan differences

$$CCC \equiv 1 - \frac{\langle (x-y)^2 \rangle}{\sigma_x^2 + \sigma_y^2 + (\mu_x - \mu_y)^2}$$

Inter-patient differences

Reproducible → CCC > 0.9

Repeatability for 328 features on 25 image pairs
- Average scans: 93% had CCC > 0.90
- T50 (end-of-exhilation): 73% had CCC > 0.90
- Breath-hold : 61% had CCC > 0.90

Reproducibility for 328 features
- Avgerage scans: 86% had CCC > 0.90
- T50 (end-of-exhilation): 52% had CCC > 0.90
- Breath-hold : 42% had CCC > 0.90

Reproducibility Contrast enhanced CT
- 23% have CCC > 0.90 but . .
- Some initial scans taken at referring clinic
- Avg. time between scans was 38 days (range, 17-72)

Hunter et al, Med Phys 12,12916, 2013
Fried et al, Int. J. Rad. Onc. Bio. Phys. 90(4), 834-843, 2014

23

CT Scan Acquisition Parameters

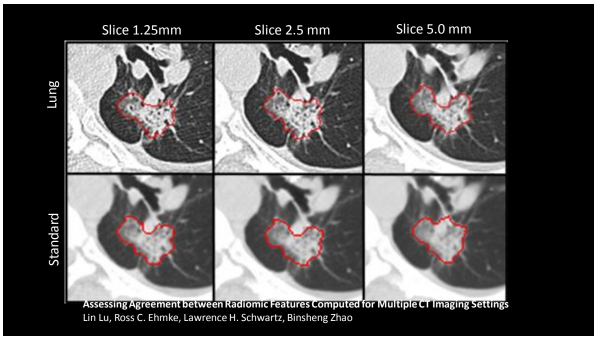
- Reconstruction FOV (Pixel size)
- Data Collection Diameter
- Slice thickness (mm)
- Reconstruction FOV (Pixel size)
- Data Collection Diameter
- Algorithm

- Effective mAs
- Kilovolt Peak (kVp)
- Scan type (axial, helical)
- Pitch factor

- Scanner Model

8/1/2017

24


Slice 1.25mm
Slice 2.5 mm
Slice 5.0 mm
Lung
Standard
Assessing Agreement between Radiomic Features Computed for Multiple CT Imaging Settings
Lin Lu, Ross C. Ehmke, Lawrence H. Schwartz, Binsheng Zhao

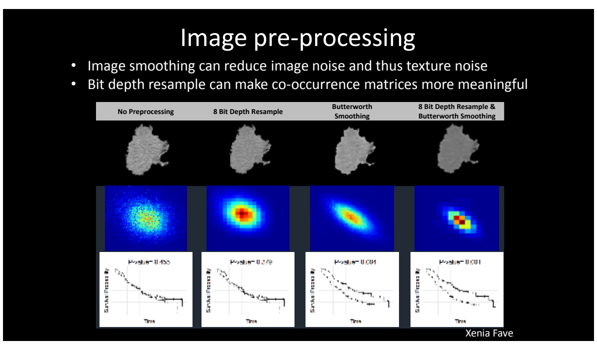
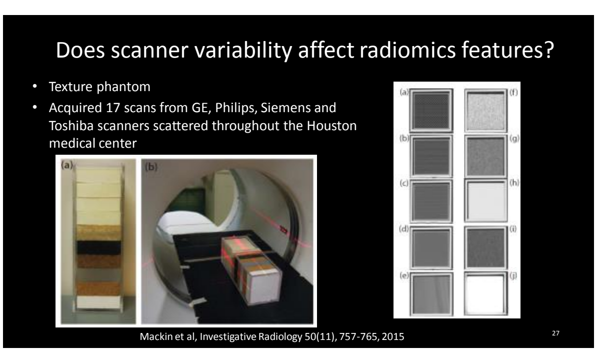
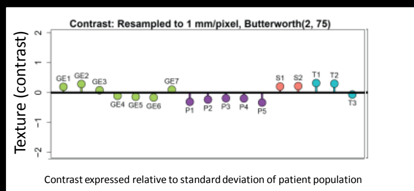


Image pre-processing
• Image smoothing can reduce image noise and thus texture noise
• Bit depth resample can make co-occurrence matrices more meaningful
No Preprocessing
8 Bit Depth Resample
Butterworth Smoothing
8 Bit Depth Resample & Butterworth Smoothing
Xenia Fave

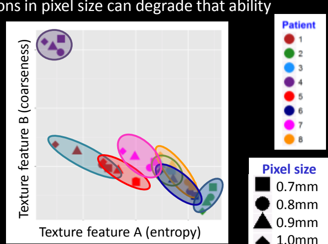

Does scanner variability affect radiomics features?
• Texture phantom
• Acquired 17 scans from GE, Philips, Siemens and Toshiba scanners scattered throughout the Houston medical center
(a)
(b)
(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)
(j)
Mackin et al, Investigative Radiology 50(11), 757-765, 2015
27

Texture is scanner dependent



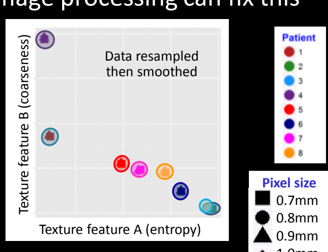
Pixel size

- Radiomics models allow us to cluster patients into groups
- Variations in pixel size can degrade that ability



Dennis Mackin, Chaan Ng

Image processing can fix this



Reducing Voxel Size Dependence (revisit)

$$f_m(P,T) = f(P,T) * V(P,T) \quad (1)$$

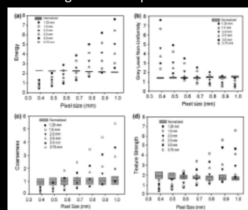
$$f_m(P,T) = \frac{f(P,T)}{V(P,T)} \quad (2)$$

$$f_m(P,T) = \frac{f(P.T)}{log[n(P,T)]} \quad (3)$$

V(P,T) = voxels size = $\frac{\text{volume}}{\text{number voxels}}$

P = pixel spacing
T = slice thickness
f = uncorrected feature value
n = number of voxels

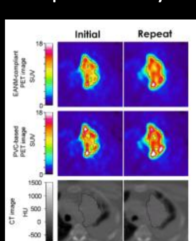
Voxel size scaling reduces dependence on voxel size



Shafiq-ul-Hassan, et al, *Medical physics,* 44(3), 2017

PET

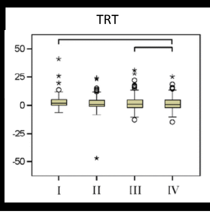
Repeatability – FDG PET

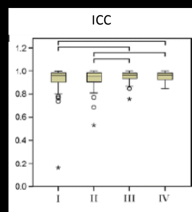


- Test-retest for FDG-PET of NSCLC patients
- 19 ROIs
- 4mm voxels
- volume > 10 cc
- 105 Features
  - Intensity histogram (n=27)
  - Texture (GLRL and GLCM; n=69)
  - Shape (n=9)
- Compared semi-auto delineations on PET to manual contours in CT using 2 PET reconstruction approaches
  - point spread function (PSF)
  - European Association for Nuclear Medicine (EANM)

van Velden, et al., *Mol. Img. and Bio.,* 18(5), 2016

Repeatability – FDG PET





$$\mathrm{TRT}_r = \frac{1}{n} \times \sum_{i=1}^{n} \frac{\text{test-retest}}{\text{mean(test, retest)}} \times 100\%$$

ICC > 0.9 for 63% of features

Features more sensitive segmentation than PET reconstruction

van Velden, et al., *Mol. Img. and Bio.*, 18(5), 2016

Reproducibility - PET

- Internal vs. outside scans – esophageal cancer patients (Van Rossum et al)
  - 7 patients
  - 31d between scans (11-42d)
  - Contoured using MiM.
  - In most cases, high ICC
- Internal vs. outside scans – NSCLC (Fried)
  - 53 patients

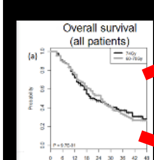
Feature category	Example	Median ICC
Conventional PET	SUVmax	0.87
Geometry	Roundness	0.92
First-order	Skewness	0.86
Texture	busyness	0.69 - 0.83

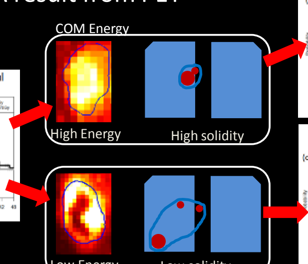
Van Rossum, J Nucl. Med. 57,691-700 2016

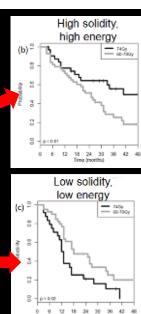
	All	Reconstruction Type		Time Difference (days)			Low Volume Change
QIF	All (N = 53)	2D - 3D (n = 40)	3D - 3D (n = 10)	0-30 (n = 16)	31-60 (n = 25)	61+ (n = 12)	<25% (n=24)
Average (all metrics)	0.78	0.72	0.93	0.81	0.82	0.62	0.74

Fried, UT GSBS Dissertation, 2015, slide used with permission from Laurence Court 35

A result from PET



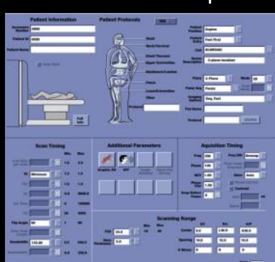




Fried et al, IJROBP 94, 368-376, 2015

MRI

MRI Acquisition Parameters



GE MRI Scanner Operator Interface

- Field of view
- Slice thickness
- Acquisition matrix

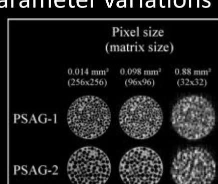
- Magnetic field strength
- Echo time
- Repetition time
- Flip angle
- Bandwidth (Hz)
- Scan duration

- Pulse sequences
- Diffusion-weighted imaging
- Dynamic contrast-enhanced imaging

Effects of MRI acquisition parameter variations

Phantom: Polystyrene/Agar gel
- Insert 1: 1.25 – 2.00 mm spheres
- Insert 2: 2.00 – 3.15 mm spheres

Varied Parameters
- number of acquisitions (NA)
- repetition time (TR)
- echo time (TE)
- sampling bandwidth (SBW)



Results
- Texture features were sensitive to variations in any of the parameters
- Reducing the imaging resolution reduced the sensitivity

Mayerhoefer, et al., *Medical physics,* 36(4), 1236-1243.

Normalizing the Intensities in MRI

Literature search (2004 – 2017)*

- Scopus.com
- Key words: MRI and (Texture or Radiomics)
- Excluded conference papers
- 746 papers
- Only 4% (32 papers) mentioned Normalization
- 19 Papers cited Collewet *et al.*

Method	Number of Papers
$\mu \pm 3\sigma$	20
Histogram normalization	4
Non-uniformity correction	2
Other methods	6

*Courtesy of Joon Sang Lee

Influence of normalization on texture

- 16 old cheeses (43 days)
- 16 new cheeses (18 days)
- T2 weighted MRI
- Used textures
  1. Co-occurrence matrix features
  2. Run-length matrix
  3. Gradient matrix
  4. Harr wavelet energy

- Normalization methods:
  1. Intensity rescaled to 64 bins
  2. Rescaled by $\frac{2000}{\max(\mathrm{ROI})}$; 64 bins
  3. Rescaled by $\frac{\mathrm{mean(all\ ROIs)}}{\mathrm{mean(ROI)}}$; 64 bins
  4. $\mu \pm 3\sigma$; 64 bins

Collewet *et al.*, Mag. Res. Img. 22, 81-91, 2003

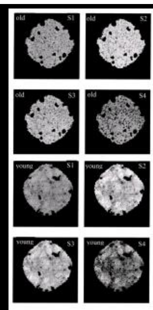
4 Normalizations for T2 weighted cheese

Normalization scheme	Classification error
S1	4 [12.50%]
S2	1 [3.13%]
S3	4 [12.50%]
S4	0 [0.00%]

Results for 1 nearest neighbor classification

Intensity normalization scheme:
1. Intensity rescaled to 64 bins
2. Rescaled by $\frac{2000}{\max(\mathrm{ROI})}$; 64 bins
3. Rescaled by $\frac{\mathrm{mean(all\ ROIs)}}{\mathrm{mean(ROI)}}$; 64 bins
4. $\mu \pm 3\sigma$; 64 bins

Different scaling for each image



camembert cheese

Collewet *et al.*, Mag. Res. Img. 22, 81-91, 2003

Summary

Radiomics workflow (Summary)

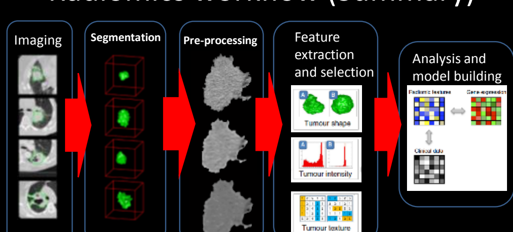
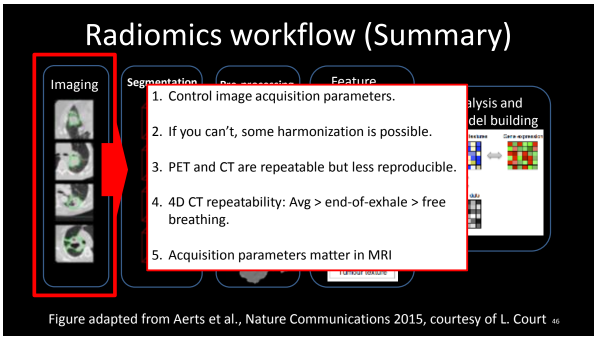
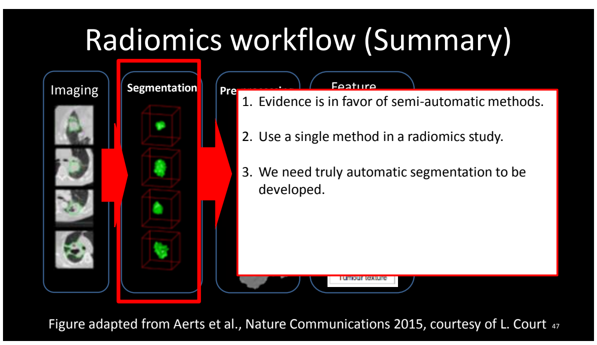
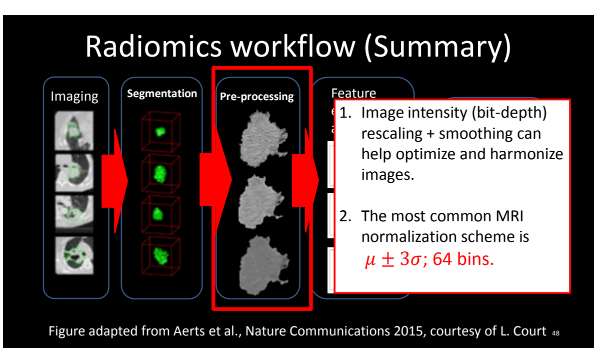


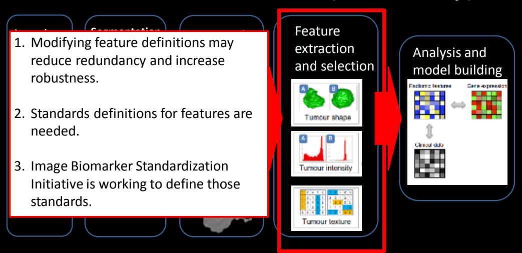
Figure adapted from Aerts et al., Nature Communications 2015, courtesy of L. Court 45


Radiomics workflow (Summary)
Imaging
1. Control image acquisition parameters.
2. If you can't, some harmonization is possible.
3. PET and CT are repeatable but less reproducible.
4. 4D CT repeatability: Avg > end-of-exhale > free breathing.
5. Acquisition parameters matter in MRI
Figure adapted from Aerts et al., Nature Communications 2015, courtesy of L. Court 46


Radiomics workflow (Summary)
Imaging
Segmentation
1. Evidence is in favor of semi-automatic methods.
2. Use a single method in a radiomics study.
3. We need truly automatic segmentation to be developed.
Figure adapted from Aerts et al., Nature Communications 2015, courtesy of L. Court 47


Radiomics workflow (Summary)
Imaging
Segmentation
Pre-processing
Feature
1. Image intensity (bit-depth) rescaling + smoothing can help optimize and harmonize images.
2. The most common MRI normalization scheme is $\mu \pm 3\sigma$; 64 bins.
Figure adapted from Aerts et al., Nature Communications 2015, courtesy of L. Court 48

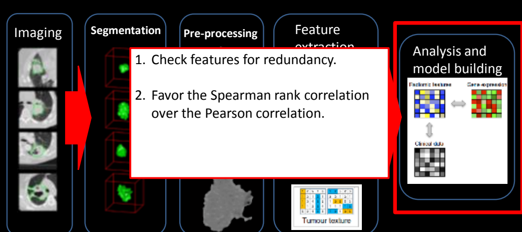
Radiomics workflow (Summary)



1. Modifying feature definitions may reduce redundancy and increase robustness.
2. Standards definitions for features are needed.
3. Image Biomarker Standardization Initiative is working to define those standards.

Figure adapted from Aerts et al., Nature Communications 2015, courtesy of L. Court 49

Radiomics workflow (Summary)



1. Check features for redundancy.
2. Favor the Spearman rank correlation over the Pearson correlation.

Figure adapted from Aerts et al., Nature Communications 2015, courtesy of L. Court 50

Summary

- Segmentation – For now, semi-automated approaches have the most support in the literature, and the 3D slicer implementations are probably the most studied.
- For 4D CT repeatability: Average > end-of-exhalation > free breathing.
- Some harmonization is possible with image intensity rescaling (bit-depth resampling) and smoothing
- $\mu \pm 3\sigma$ with 64 intensity bins is the most commonly used normalization for MRI

In radiomics, nothing has been settled. To my knowledge, none of these approaches have been verified to be the best.

References (1)

- Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M. M., Leemans, C. R., Dekker, A., Quackenbush, J., Gillies, R. J. and Lambin, P. (2014) 'Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach', Nature Communications, 5, 1-8.
- Collewet, G., Strzelecki, M. and Mariette, F. (2004) 'Influence of MRI acquisition protocols and image intensity normalization methods on texture classification', Magnetic Resonance Imaging, 22(1), 81-91.
- Fried, D. V., Tucker, S. L., Zhou, S., Liao, Z., Mawlawi, O., Ibbott, G. and Court, L. E. (2014) 'Prognostic value and reproducibility of pretreatment CT texture features in stage III non-small cell lung cancer', International Journal of Radiation Oncology* Biology* Physics, 90(4), 834-842.
- Hatt, M., Tixier, F., Pierce, L., Kinahan, P. E., Le Rest, C. C. and Visvikis, D. (2017) 'Characterization of PET/CT images using texture analysis: the past, the present… any future?', European journal of nuclear medicine and molecular imaging, 44(1), 151-165.
- Hua, K.-L., Hsu, C.-H., Hidayati, S. C., Cheng, W.-H. and Chen, Y.-J. (2015) 'Computer-aided classification of lung nodules on computed tomography images via deep learning technique', OncoTargets and therapy, 8.
- Hunter, L. A., Krafft, S., Stingo, F., Choi, H., Martel, M. K. and Kry, S. F. (2013) 'High quality machine-robust image features: Identification in nonsmall cell lung cancer computed tomography images', Medical physics, 40(12), 121916.
- Kalpathy-Cramer, J., Zhao, B., Goldgof, D., Gu, Y., Wang, X., Yang, H., Tan, Y., Gillies, R. and Napel, S. (2016) 'A comparison of lung nodule segmentation algorithms: methods and results from a multi-institutional study', Journal of digital imaging, 29(4), 476-487.
- Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G., Granton, P., Zegers, C. M., Gillies, R., Boellard, R. and Dekker, A. (2012) 'Radiomics: extracting more information from medical images using advanced feature analysis', European Journal of Cancer, 48(4), 441-446.
- Larue, R. T., Defraene, G., De Ruysscher, D., Lambin, P. and 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.
- Lu, L., Ehmke, R. C., Schwartz, L. H. and Zhao, B. (2016) 'Assessing agreement between radiomic features computed for multiple CT imaging settings', PloS one, 11(12), e0166550.
- Lu, L., Lv, W., Jiang, J., Ma, J., Feng, Q., Rahmim, A. and Chen, W. (2016) 'Robustness of radiomic features in [11C] Choline and [18F] FDG PET/CT imaging of nasopharyngeal carcinoma: impact of segmentation and discretization', Molecular Imaging and Biology, 18(6), 935-945.

References (2)

- Mackin, D., Fave, X., Zhang, L., Fried, D., Yang, J., Taylor, B., Rodriguez-Rivera, E., Dodge, C. and Jones, A. K. (2015) 'Measuring Computed Tomography Scanner Variability of Radiomics Features', Investigative radiology, 50(11), 757-765.
- Mayerhoefer, M. E., Szomolanyi, P., Jirak, D., Materka, A. and Trattnig, S. (2009) 'Effects of MRI acquisition parameter variations and protocol heterogeneity on the results of texture analysis and pattern discrimination: an application-oriented study', Medical physics, 36(4), 1236-1243.
- Parmar, C., Velazquez, E. R., Leijenaar, R., Jermoumi, M., Carvalho, S., Mak, R. H., Mitra, S., Shankar, B. U., Kikinis, R. and Haibe-Kains, B. (2014) 'Robust radiomics feature quantification using semiautomatic volumetric segmentation', PloS one, 9(7), e102107.
- Saha, A., Yu, X., Sahoo, D. and Mazurowski, M. A. (2017) 'Effects of MRI scanner parameters on breast cancer radiomics', Expert Systems with Applications, 87, 384-391.
- Shafiq-ul-Hassan, M., Zhang, G. G., Latifi, K., Ullah, G., Hunt, D. C., Balagurunathan, Y., Abdalah, M. A., Schabath, M. B., Goldgof, D. G. and Mackin, D. (2017) 'Intrinsic dependencies of CT radiomic features on voxel size and number of gray levels', Medical physics, 44(3), 1050-1062.
- Thibault, G., Fertil, B., Navarro, C., Pereira, S., Cau, P., Levy, N., Sequeira, J. and Mari, J.-L. (2013) 'Shape and texture indexes application to cell nuclei classification', International Journal of Pattern Recognition and Artificial Intelligence, 27(01), 1357002.
- van Velden, F. H., Kramer, G. M., Frings, V., Nissen, I. A., Mulder, E. R., de Langen, A. J., Hoekstra, O. S., Smit, E. F. and Boellaard, R. (2016) 'Repeatability of radiomic features in non-small-cell lung cancer [18F] FDG-PET/CT studies: impact of reconstruction and delineation', Molecular Imaging and Biology, 18(5), 788-795.
- Velazquez, E. R., Parmar, C., Jermoumi, M., Mak, R. H., Van Baardwijk, A., Fennessy, F. M., Lewis, J. H., De Ruysscher, D., Kikinis, R. and Lambin, P. (2013) 'Volumetric CT-based segmentation of NSCLC using 3D-Slicer', Scientific reports, 3, 3529.
- Zhao, B., Tan, Y., Tsai, W.-Y., Qi, J., Xie, C., Lu, L. and Schwartz, L. H. (2016) 'Reproducibility of radiomics for deciphering tumor phenotype with imaging', Scientific reports, 6.

Thank you & Questions