

# SCIENTIFIC REPORTS



OPEN

## Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGe): A new radiomics descriptor

Prateek Prasanna\*, Pallavi Tiwari\* & Anant Madabhushi

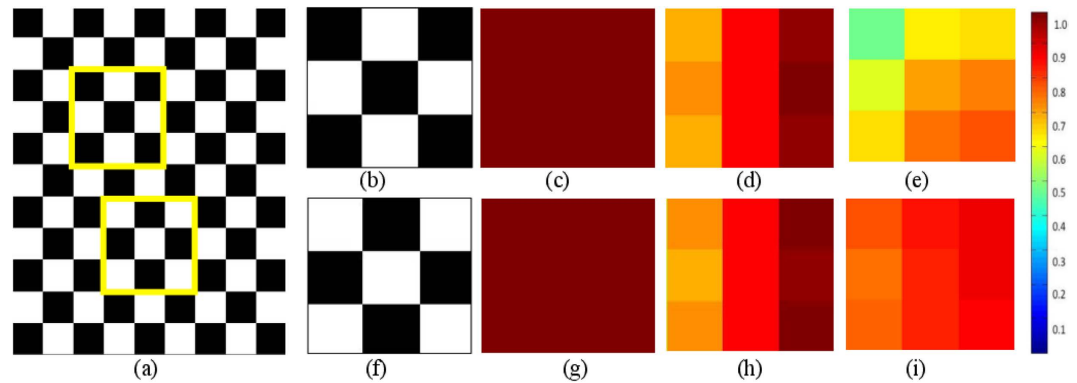
Received: 16 February 2016  
Accepted: 27 October 2016  
Published: 22 November 2016

In this paper, we introduce a new radiomic descriptor, Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGe) for capturing subtle differences between benign and pathologic phenotypes which may be visually indistinguishable on routine anatomic imaging. CoLIAGe seeks to capture and exploit local anisotropic differences in voxel-level gradient orientations to distinguish similar appearing phenotypes. CoLIAGe involves assigning every image voxel an entropy value associated with the co-occurrence matrix of gradient orientations computed around every voxel. The hypothesis behind CoLIAGe is that benign and pathologic phenotypes even though they may appear similar on anatomic imaging, will differ in their local entropy patterns, in turn reflecting subtle local differences in tissue microarchitecture. We demonstrate CoLIAGe's utility in three clinically challenging classification problems: distinguishing (1) radiation necrosis, a benign yet confounding effect of radiation treatment, from recurrent tumors on T1-w MRI in 42 brain tumor patients, (2) different molecular sub-types of breast cancer on DCE-MRI in 65 studies and (3) non-small cell lung cancer (adenocarcinomas) from benign fungal infection (granulomas) on 120 non-contrast CT studies. For each of these classification problems, CoLIAGe in conjunction with a random forest classifier outperformed state of the art radiomic descriptors (Haralick, Gabor, Histogram of Gradient Orientations).

There are several instances where benign and malignant pathologies might appear very similar on radiographic imaging. One such example is radiation necrosis (RN) (a relatively benign effect of radiation treatment) and recurrent brain tumors (rBT), which are visually almost indistinguishable on conventional MRI<sup>1</sup>; even though both RN and rBT have distinct cellular and architectural arrangements when examined on a pathology slide under a microscope. Another example is triple negative (TN) breast cancer (highly aggressive) and fibroadenomas (FA) (benign tumor) with similar morphologic appearances on MRI<sup>2</sup>. Similarly, fungal infections known as granulomas look strikingly similar to non-small cell lung cancers (adenocarcinomas) on routine non-contrast CT imaging. There is hence a need for identifying non-invasive markers that can reliably distinguish such similar appearing pathologies on routine imaging for early diagnosis as well as treatment evaluation. Identification of these imaging biomarkers could potentially obviate the need for unnecessary surgical interventions, as well as exposure to unnecessary radiation, for disease confirmation. "Radiomics"<sup>3-5</sup>, an emerging field in medical image analysis, refers to the quantitative extraction of shape, histogram, and/or texture-based features from radiographic images to distinguish disease phenotypes that are not visually appreciable on imaging. Two popular radiomic features are Haralick<sup>6</sup> and Gabor steerable filters<sup>7</sup>. Haralick features capture gray-level co-occurrence patterns<sup>8,9</sup>, where a matrix of co-occurring gray-level pairs in the image is constructed, from which second-order statistical texture features can be calculated. Haralick texture analysis is relatively popular in medical image analysis as it allows for capturing variations in gray-level image characteristics via second order intensity statistics (e.g. angular second moment, contrast, and difference entropy). However, Haralick features may fail to capture variations in subtly different sub-structures that may be morphologically different but may have identical co-occurring gray-level intensities. Figure 1(b) and (f) show one such example of two similar appearing texture patterns, where the corresponding Haralick energy feature (shown in Fig. 1(c) and (g)) for both the patterns was found to be identical.

Gabor filters<sup>7</sup> are modeled to mimic the way human visual system deciphers object appearances<sup>10</sup>. A Gabor filter can be defined as the modulation of a complex sinusoidal by a Gaussian function and is controlled by scale ( $t$ ) and

Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH 44120, USA. \*These authors contributed equally to this work. Correspondence and requests for materials should be addressed to P.P. (email: prateek.prasanna@gmail.com)



**Figure 1. Comparison of different feature representations for two similar-appearing toy images.** (a) A synthetic checkerboard image. Two  $3 \times 3$  patches ( $P_1, P_2$ ) obtained from the checkerboard image are shown in (b) and (f) respectively such that  $P_2 = \sim(P_1)$ . The corresponding feature representations for  $P_1$  and  $P_2$  are shown in (c) and (g) for Haralick entropy, (d) and (h) for Gabor ( $t = 4, \lambda = 22.5$ ), and (e) and (i) for CoLIAGE respectively.

orientation ( $\lambda$ ) parameters. Gabor features can be extracted as a response to convolution of an image with distinct Gabor filters obtained by varying each of the associated parameters ( $t, \lambda$ ) across the filter bank. However, Gabor filters capture a global response for an image at a specific value of ( $t, \lambda$ ), and may not capture local variations in orientations on a per-pixel basis within local neighborhoods, which may be an important attribute when computing differences across similar-appearing micro-textures. Figure 1(d) and (h) show a representative example of similar feature responses obtained for a Gabor filter at  $t = 2$ , and  $\lambda = \pi/4$  for the two texture patterns shown in Fig. 1(b) and (f) respectively.

Other popular radiomic features include histogram of gradient orientations (HOG)<sup>11</sup>. HOG yields a global patch-based signature by computing histogram distribution of orientations obtained from computing differences in image intensities in  $X$  and  $Y$  directions on a per pixel basis. A variant of HOG, called co-occurrence of histogram of gradient orientations (Co-HOG) was recently presented by Watanabe *et al.*<sup>12</sup> and Pang *et al.*<sup>13</sup> for pedestrian detection. Co-HOG is a multiple-gradient-orientation-based feature descriptor. Its computation yields a high dimensional feature vector that combines neighbor gradient orientations to quantify shape-based appearances in regions of interest. However, the Co-HOG approach in refs 12–14 (a) does not capture localized intensity-dependent variations across neighboring orientations, and (b) is susceptible to “curse of dimensionality” (due to a high dimensional feature space).

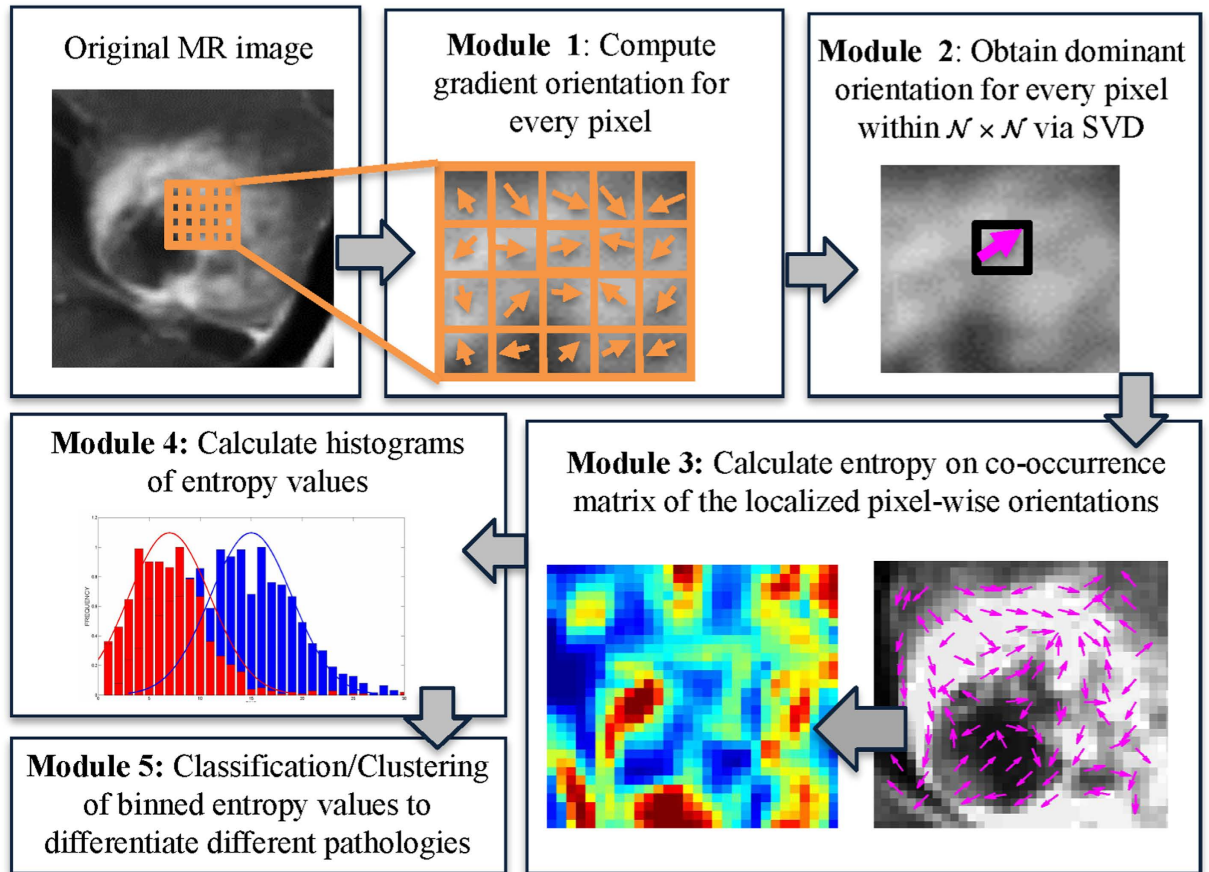
Recently, Lee *et al.*<sup>15</sup> developed a novel quantitative feature called cell orientation entropy (CoRE) to capture differences in orientation of nuclei with respect to the neighboring nuclei across different pathologies, and demonstrated differences in nuclear orientations across aggressive and benign conditions in the context of prostate cancer. If such pathologic differences are indeed reflected at the radiologic level (even though these differences may not be visually discernible), it begs the question whether a radiomic descriptor could be developed to capture the histologic anisotropy across different pathologies at the radiologic scale.

In this work, we present a new radiomic descriptor, Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGE), to capture anisotropic tensor gradient differences across similar appearing pathologies in an image. The rationale behind CoLIAGE, is that even though overall the global textural patterns or even the filter responses at a majority of pixel locations might be similar between two differing pathological conditions (e.g. RN versus rBT, FA versus TN breast cancer or adenocarcinomas versus granulomas), the organization and co-occurrences of local tensor gradients may differ across classes and will be relatively consistent within a class. CoLIAGE seeks to capture these local anisotropic differences in micro-structures by measuring entropy (a mathematical construct to measure disorder) of co-occurrences of pixel/voxel-level gradient orientations computed within a local neighborhood. The rationale being that the distribution of entropy of localized gradient field within a lesion will be high for aggressive disease conditions, potentially manifesting their inherent disorder and high heterogeneity appreciable at a cellular scale, as compared to benign pathologies which have a more coherent micro-architecture. An example of CoLIAGE is shown in Fig. 1(e) and (i) to distinguish two synthetic checkerboard images with similar-appearing patterns. CoLIAGE was found to capture localized variations in gradients on a per-pixel basis (reflected by high CoLIAGE values) across the two patterns, differences that were not appreciable on Haralick and Gabor feature representations.

The rest of the paper is organized as follows. The algorithm of CoLIAGE is detailed in the Methodology section, followed by the experimental setup to demonstrate the utility of CoLIAGE in the context of three problems involving brain tumors, breast cancer, and lung cancer. Subsequently we present and discuss the results followed by the concluding remarks.

## Methodology of CoLIAGE

In the following subsections, we describe the detailed mathematical formulation of our CoLIAGE strategy, both in 2-dimensions (2D) and in 3-dimensions (3D). A preliminary implementation of 2D CoLIAGE was previously presented in ref. 16. Figure 2 shows the workflow of CoLIAGE in 2D, while the 3D implementation is shown in Fig. 3.

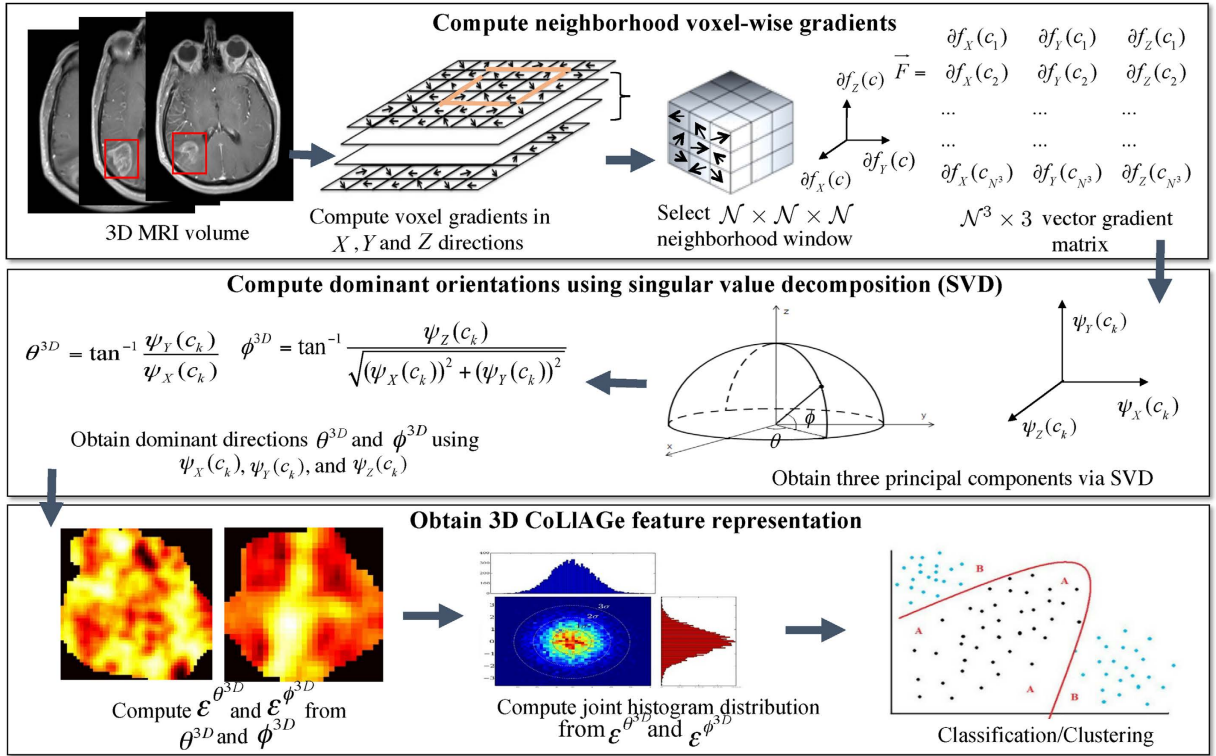


**Figure 2. Overview of 2-dimensional CoLIAGE and overall workflow.** The first module involves computing gradient orientations on a per-pixel basis within the lesion. In the second module, dominant direction for every pixel is computed within a  $\mathcal{N} \times \mathcal{N}$  neighborhood via singular value decomposition. A co-occurrence matrix of orientations and subsequently the associated entropy of each pixel is calculated in the third module. A histogram of entropy values is then aggregated for every pixel in the fourth module and used for classification in the fifth module to distinguish similar appearing pathological conditions.

**Notation.** An MRI image scene  $\mathcal{I}$  is defined as  $\mathcal{I} = (C, f)$  where  $\mathcal{I}$  is a spatial grid  $C$  of locations  $c \in C$ , in a 2-dimensional,  $\mathbb{R}^2$ , or a 3-dimensional space,  $\mathbb{R}^3$ . Each such spatial location,  $c \in C$  (in  $\mathbb{R}^2$  or  $\mathbb{R}^3$ ) is associated with an intensity value  $f(c)$ . A local neighborhood of  $\mathcal{N}$  pixels/voxels is defined within a window  $W$ , while the co-occurrence matrix computed from within  $W$  is denoted as  $\mathcal{M}$ . The entropy map is given as  $\mathcal{E}$ , and the final CoLIAGE feature vector is denoted as  $\mathbf{F}$ . For the sake of clarity, the notations used for computation of 2D CoLIAGE are denoted with superscript 2D, while the notations for computation of 3D CoLIAGE are denoted with superscript 3D. The common notations, operators and acronyms employed in this paper are listed in Table 1.

**Methodology.** Computation of 2D CoLIAGE (Algorithm 1) for every  $c \in C$  involves the following main steps,

- 1. Calculation of gradient magnitudes for every pixel:** For every  $c \in C$ , gradients along the X and Y directions are computed as,  $\nabla f(c) = \frac{\partial f(c)}{\partial x} \hat{i} + \frac{\partial f(c)}{\partial y} \hat{j}$ . Here,  $\frac{\partial f(c)}{\partial x}$  and  $\frac{\partial f(c)}{\partial y}$  are the gradient magnitudes along the X and the Y axes respectively, denoted by  $\partial f_X(c)$  and  $\partial f_Y(c)$ .
- 2. Computing local dominant orientations via singular value decomposition (SVD):** A  $\mathcal{N} \times \mathcal{N}$  window  $W$  centered around every  $c \in C$  is selected to compute the localized gradient field. We then compute  $\partial f_X(c_k)$  and  $\partial f_Y(c_k)$ ,  $k \in \{1, 2, \dots, \mathcal{N}^2\}$ . The vector gradient matrix  $\mathcal{F}$  associated with every  $c$  is given by  $\overline{\mathcal{F}} = [\partial f_X(c_k) \ \partial f_Y(c_k)]$ , where  $[\partial f_X(c_k) \ \partial f_Y(c_k)]$ ,  $k \in \{1, 2, \dots, \mathcal{N}^2\}$  is the matrix of gradient vectors in the X and Y directions for every  $c_k$  given by a  $\mathcal{N}^2 \times 2$  matrix,



**Figure 3. Workflow for extracting 3D CoLIAGe features.** Note that unlike 2D features, two dominant directions ( $\theta^{3D}, \phi^{3D}$ ) are computed for 3D CoLIAGe associated with each  $c \in C$ , which provide complimentary information about the degree of disorder of the principal gradient orientations in (X, Y) and (X, Y, Z) directions.

$$\begin{pmatrix} \partial f_X(c_1) & \partial f_Y(c_1) \\ \partial f_X(c_2) & \partial f_Y(c_2) \\ \vdots & \vdots \\ \partial f_X(c_{N^2}) & \partial f_Y(c_{N^2}) \end{pmatrix}$$

The most significant orientation for each pixel  $c_k$  within  $\mathcal{N} \times \mathcal{N}$  gradient field is obtained by performing singular value decomposition (SVD) of  $\bar{F}$ . The dominant principal components in X and Y directions are obtained from SVD as  $r_k^X$  and  $r_k^Y$  for every  $k \in \{1, 2, \dots, N^2\}$ . The most significant orientation for every  $c_k$  is then calculated as  $\theta^{2D}(c_k) = \tan^{-1} \left[ \frac{r_k^Y}{r_k^X} \right]$ .

- Calculation of second-order statistics for most significant orientations:** The objects of interest for calculating CoLIAGe features are the co-occurring directions given by discretization of the dominant orientation  $\bar{\theta}^{2D}(c_k)$  for every pixel  $c$ , such that  $\theta^{2D}(c_k) = \omega \times \text{ceil} \left[ \frac{\bar{\theta}^{2D}(c_k)}{\omega} \right]$ , where  $\omega$  is a discretization factor.

An  $N \times N$  co-occurrence matrix  $\mathcal{M}^{\theta^{2D}}$  subsequently captures pairs of orientations ( $p, q$ ) between pixels ( $c_p, c_k$ ) which co-occur in the neighborhood  $\mathcal{W}_p$ , such that,

$$\mathcal{M}_{\mathcal{W}_i}^{\theta^{2D}}(p, q) = \sum_{c_p} \sum_{c_k} \begin{cases} 1, & \text{if } \theta(c_j) = p, \theta(c_k) = q \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

where  $N = \frac{360^\circ}{\omega}$  is the number of discrete angular bins. Entropy measure,  $\mathcal{E}^{\theta^{2D}}(c)$  is then computed from every co-occurrence matrix on every  $c$  as,

$$\mathcal{E}^{\theta^{2D}}(c) = \sum_{p,q} -\mathcal{M}^{\theta^{2D}}(p, q) \log(\mathcal{M}^{\theta^{2D}}(p, q)). \quad (2)$$

- A **histogram of  $\mathcal{E}^{\theta^{2D}}$**  is computed by aggregating  $\mathcal{E}^{\theta^{2D}}(c_k), k \in \{1, \dots, |C|\}$ , where  $|\cdot|$  is the cardinality of set

Symbol	Description	Symbol	Description
$\mathcal{I}$	MRI scene	$\beta^{Acc}$	Average accuracy
$C$	Grid of MRI voxels	$c$	Spatial location of pixel/voxel of interest
$f(c)$	MRI signal intensity at $c$	$\nabla$	Spatial derivative
$\partial f_x(c)$	Gradient magnitude along X-axis	$\partial f_y(c)$	Gradient magnitude along Y-axis
$\partial f_z(c)$	Gradient magnitude along Z-axis	$W$	Local neighborhood around central location
$k$	Index of voxels/pixels of interest, $k \in 1, \dots,  C $	$\vec{\mathcal{F}}$	Vector gradient matrix
$\nu$	Bin size	$\omega$	Discretization factor
$r_k^a$	Principal component in $\mathbb{R}^2$ , $a \in X, Y$	$\psi_n(c_k)$	Principal component in $\mathbb{R}^3$ , $n \in X, Y, Z$
$\theta^{2D}$	Dominant orientation computed from $r_k^X, r_k^Y$	$\theta^{3D}$	Dominant orientation computed from $\psi_X, \psi_Y$
$\phi^{3D}$	Dominant orientation computed from $\psi_X, \psi_Y, \psi_Z$	$\mathcal{N}$	Window size dimensions (pixels/voxels)
$N$	Window size for co-occurrence computation	$\mathcal{W}$	Co-occurrence matrix neighborhood
$\mathcal{M}^{\theta^{2D}}$	Co-occurrence matrix computed from $\theta^{2D}$	$\mathcal{M}^{\theta^{3D}}$	Co-occurrence matrix computed from $\theta^{3D}$
$\mathcal{E}^{\theta^{2D}}$	Entropy map computed from $\mathcal{M}^{\theta^{2D}}$	$\mathcal{E}^{\theta^{3D}}$	Entropy map computed from $\mathcal{M}^{\theta^{3D}}$
$\mathcal{M}^{\phi^{3D}}$	Co-occurrence matrix computed from $\phi^{3D}$	$\mathcal{E}^{\phi^{3D}}$	Entropy map computed from $\mathcal{M}^{\phi^{3D}}$
$\mathbf{F}^{2D}$	2D CoLLAGe feature vector	$\mathbf{F}^{3D}$	3D CoLLAGe feature vector
UH	University Hospitals, Cleveland	CCF	Cleveland Clinic Foundation
UPenn	University of Pennsylvania	RN	Radiation Necrosis
rBT	Recurrent Brain Tumor	TN	Triple Negative
ER	Estrogen Receptor	HER2	Human Epidermal growth factor Receptor 2
FA	Fibroadenoma	RF	Random Forest
PET	Positron Emission Tomography	CT	Computed Tomography
FDG	Fluorodeoxyglucose	Gd	Gadolinium

**Table 1.** List of commonly used notations, symbols and acronyms in this paper.

C. The entropy histogram is divided into bin size  $\nu$ , optimized on the training set via grid search optimization.

A CoLLAGe feature vector,  $\mathbf{F}^{2D}$  can be obtained for every  $c \in C$  which consists of the binned histogram values in the form of  $\nu \times 1$  vectors.

---

**Data:** ROI volume  $C$

**Result:** 2D CoLLAGe features,  $\mathcal{F}$

---

**begin**

**for each pixel**  $c \in C$ , **do**

Obtain gradients  $\partial f_x(c)$  and  $\partial f_y(c)$  along X- and Y- axes;

**end**

**for each pixel**  $c \in C$ , **do**

Compute gradient vectors  $\vec{\partial f}_X(c_k)$  and  $\vec{\partial f}_Y(c_k)$  in  $\mathcal{N} \times \mathcal{N}$  neighborhood; Obtain localized gradient vector matrix  $\vec{\mathcal{F}} = [\vec{\partial f}_X(c_k) \vec{\partial f}_Y(c_k)]$ ;

Compute dominant orientation,  $\theta(c_k), k \in \{1, \dots, \mathcal{N}^2\}$  via SVD of  $\vec{\mathcal{F}}$ ;

**end**

Compute Co-occurrence matrix  $\mathcal{M}$  from  $\theta(x, y)$  using Equation 1;

Compute  $\mathcal{E}(c)$  from  $\mathcal{M}$  as described in Equation 2;

Obtain  $\nu \times 1$  dimensional feature vector  $\mathcal{F}$  from distribution of  $\mathcal{E}$ ;

**end**

---

### Algorithm 1: Computation of 2D CoLLAGe features

---

**Data:** ROI volume  $C$

**Result:** 3D CoLLAGe features,  $\hat{\mathcal{F}}$  and  $\tilde{\mathcal{F}}$

---

**begin**

**for each voxel**  $c \in C$ , **do**

Obtain gradients  $\partial f_i(c)$  along  $i$ -axes, where  $\partial f_i(c) = \frac{\partial f(c)}{\partial i}, i \in x, y, z$

Continued

end

Define  $\mathcal{N} \times \mathcal{N} \times \mathcal{N}$  neighborhood centered around  $c \in \mathcal{C}$

for each voxel  $c \in \mathcal{C}$ , do

  Compute gradient vectors in neighborhood,  $i \in (x, y, z)$ ,

$k \in \{1, \dots, \mathcal{N}^3\}$ , where  $\partial f_i(c) = [\partial f_i(c_1) \ \partial f_i(c_2) \ \dots \ \partial f_i(c_{\mathcal{N}^3})]^T$

  Obtain localized gradient vector matrix  $\vec{\mathcal{F}} = [\vec{\partial f}_X(c_k) \ \vec{\partial f}_Y(c_k) \ \vec{\partial f}_Z(c_k)]$ ;

  Compute dominant components  $\psi(c_{k11})$ ,  $\psi(c_{k12})$ , and  $\psi(c_{k13})$ , by SVD of;

  Obtain dominant directions  $\theta(c_k)$  and  $\phi(c_k)$ , using Equations 3 and 4 respectively;

end

  Compute co-occurrence matrix  $\hat{\mathcal{M}}$  from  $\theta(c_k)$  and  $\tilde{\mathcal{M}}$  from  $\phi(c_k)$  using Equation 1;

  Compute  $\hat{\mathcal{E}}(c_k)$  from  $\hat{\mathcal{M}}$  and  $\tilde{\mathcal{E}}(c_k)$  from  $\tilde{\mathcal{M}}$  as described in Equation 2;

  Obtain  $v \times 1$  dimensional feature vectors  $\hat{\mathcal{F}}$  and  $\tilde{\mathcal{F}}$  from distributions of  $\hat{\mathcal{E}}$  and  $\tilde{\mathcal{E}}$  respectively;

end

## Algorithm 2: Computation of 3D CoLIAGe features

**Extension of CoLIAGe to 3-Dimensions.** For 3D CoLIAGe (Algorithm 2), the local neighborhood around a voxel is first defined by a 3D window  $W$  of size  $\mathcal{N} \times \mathcal{N} \times \mathcal{N}$  along the X, Y, and Z-directions and gradient directions for every voxel are calculated from within  $W$ . We collate the gradient magnitudes along the three axes in  $W$  into a single gradient matrix  $\vec{\mathcal{F}}$  of size  $\mathcal{N}^3 \times 3$  given as:  $\vec{\mathcal{F}} = [\vec{\partial f}_X(c_k) \ \vec{\partial f}_Y(c_k) \ \vec{\partial f}_Z(c_k)]$ , where  $[\vec{\partial f}_X(c_k) \ \vec{\partial f}_Y(c_k) \ \vec{\partial f}_Z(c_k)]$ ,  $k \in \{1, 2, \dots, \mathcal{N}^3\}$  is the matrix of gradient vectors in the X, Y and Z directions respectively for every  $c_k$ . SVD of  $\vec{\mathcal{F}}$  for a voxel  $c_k$  yields three dominant principal components  $\psi_X(c_k)$ ,  $\psi_Y(c_k)$  and  $\psi_Z(c_k)$  in the X-, Y- and Z- directions respectively. Two dominant orientations  $\theta^{3D}(c_k)$  and  $\phi^{3D}(c_k)$  can then be obtained from the three principal components to capture variability in orientations across (X, Y), and (X, Y, Z), given by

$$\theta^{3D}(c_k) = \tan^{-1} \frac{\psi_Y(c_k)}{\psi_X(c_k)} \quad (3)$$

and

$$\phi^{3D}(c_k) = \tan^{-1} \frac{\psi_Z(c_k)}{\sqrt{\psi_Y^2(c_k) + \psi_X^2(c_k)}} \quad (4)$$

Two co-occurrence matrices  $\mathcal{M}^{\theta^{3D}}$  and  $\mathcal{M}^{\phi^{3D}}$  corresponding to  $\theta^{3D}(c_k)$  and  $\phi^{3D}(c_k)$  capture the orientation pairs between voxels in the neighborhood and are computed as given in Equation 1.  $\mathcal{E}^{\theta^{3D}}(c_k)$  from  $\mathcal{M}^{\theta^{3D}}(c_k)$  and  $\mathcal{E}^{\phi^{3D}}$  from  $\mathcal{M}^{\phi^{3D}}$  (computed using Equation 2), yield two distinct entropy representations in (X, Y) and (X, Y, Z) directions within the volume of interest. A joint histogram of CoLIAGe feature vector  $\mathbf{F}^{3D} = [\mathcal{E}^{\theta^{3D}}, \mathcal{E}^{\phi^{3D}}]$  can be obtained for further evaluation in a supervised or an unsupervised classification setting.

## Experimental Design

### Datasets and Preprocessing.

In this work, we employed three unique dataset cohorts obtained from different collaborating institutions to evaluate the efficacy of CoLIAGe on three extremely challenging clinical problems: (a) distinguishing radiation necrosis, a relatively benign effect of radiation, from tumor recurrence on T1-w MRI in brain tumors, (b) distinguishing different molecular sub-types of breast cancer on DCE-MRI, and (c) distinguishing adenocarcinomas from granulomas on non-contrast CT images. Details regarding inclusion criteria, pre-processing, and experimental design for each of the three datasets are provided below.

**Brain tumor dataset.** Imaging scans were acquired under an Institutional Review Board (IRB)-approved (IRB # CC00148) and HIPAA-compliant study at University Hospitals, Cleveland (UH). The patient cohorts were identified by performing a retrospective review of neuropathology in all brain tumor patients who underwent a surgery of a recurrent or progressive Gd T1w-enhancing lesion identified during follow-up post-9 months (or later) after the initial after brain radiation therapy. Follow-up MRI scans within 0–21 days prior to second resection or biopsy (for disease confirmation) were used for analysis. Written informed consent was obtained from all the subjects. Inclusion criteria were that the pathology specimen must have been obtained by resection (preferably) or by multiple biopsies (>2) via stereotactic guidance. Fewer than two biopsies were not allowed because of the potential for sampling error. Histology was re-reviewed by a neuropathologist, blinded to the original diagnosis and type of radiation, in order to quantify the percentage of radiation necrosis and recurrent tumor. In order to avoid any training errors due to “mixed” pathologies on the same lesion the presence of RN was strictly defined as >80% RN and of recurrent tumor as >80% recurrent tumor (other “mixed” cases with varying proportions of RN and tumor recurrence were excluded). We identified a total of 42 cases, from 2006 to 2014 that followed this strict inclusion criterion. Our retrospectively analyzed brain tumor dataset comprised 22 primary (10 RN, 12 rBT) and 20 metastatic (8 RN, 12 rBT) cases. Patient MRIs were acquired at 3 Tesla. Images have an in-plane resolution

	Primary Brain tumors	Metastatic Brain Tumors
Number of patients (samples)	22 (10 RN, 12 rBT)	20 (8 RN, 12 rBT)
Gender	12 F, 10 M	5 F, 15 M
Age (y, range)	53 (33–75)	50 (37–65)

**Table 2. Clinicopathologic characteristics of brain tumor studies.**

	Fibroadenoma	HER2+	ER+	Triple Negative
Number of patients (samples)	9	14	23	19
Number of lesions	12	18	25	21
Age (y, range)	46 (32–60)	50 (38–63)	45 (32–70)	51 (32–68)
Pre-menopausal (n)	5	5	15	9
Post-menopausal (n)	2	9	1	7
Peri-menopausal (n)	2	0	7	3
Lymph Node Positive Studies	0	1	8	3

**Table 3. Clinicopathologic characteristics of breast cancer studies.**

	Adenocarcinomas	Granulomas
Number of patients (samples)	65	55
Gender	39 F, 26 M	25 F, 30 M
Age (y, range)	67 (40–85)	58 (18–84)

**Table 4. Clinicopathologic characteristics of lung studies.**

of 0.8–0.93 mm/pixel. Slice thickness = 3–5 mm, TR = 400–750 ms, TE = 14–17 ms. The clinicopathologic characteristics of the brain tumor dataset have been summarized in Table 2.

**Breast cancer dataset.** Breast MRI data was prospectively collected in an Institutional Review Board-approved (IRB #02-13-42C), HIPAA-compliant study at the University of Pennsylvania (UPenn) between 2002 and 2007. Written informed consent was obtained from all subjects. Women without contraindication to MRI or gadolinium who presented with either a suspicious breast lesion or known malignancy prior to surgery were recruited to a larger single-institution study of MRI in the staging, diagnosis, and screening of breast cancer. Women who underwent neoadjuvant chemotherapy prior to surgery were excluded, as were women who had excisional biopsy prior to entry. From this data set we sub-selected women whose pathology revealed invasive cancer. Subjects whose images of the index lesion demonstrated substantial metallic artifact from prior biopsy were also excluded. This study examined MRI characteristics in 76 solid lesions from 65 patients for whom pathology results and, where applicable, ER, PR, and HER2 results were available. Reference standard diagnosis was made by histopathologic examination of tissue obtained by either core biopsy sampling or lumpectomy. Of the 76 lesions, 12 were benign fibroadenomas and 64 were invasive carcinomas. All of the carcinomas were immunohistochemically stained for hormone receptors and HER-2/neu. In cases in which staining for HER-2/neu was inconclusive, amplification was confirmed with fluorescence *in situ* hybridization. Of the 64 carcinomas, 21 were triple negative (ER–/PR–/HER2–) cancer, 18 were HER2+ (14 ER–/HER2+, 4 ER+/HER2+) cancer, and 25 were ER+ (ER+/HER2–) cancer. Patient MRIs were acquired at either 1.5 or 3 Tesla (Siemens Sonata or Trio, respectively, Malvern, PA). Imaging parameters for DCE-MRI varied over time and magnet type (in-plane resolution 0.20–0.70 mm/pixel, slice thickness 2–5 mm, TR = 7–26 ms, TE = 1.8–6.5 ms, flip angle 25–30 degrees. The clinicopathologic characteristics of the breast cancer dataset have been summarized in Table 3.

**Lung cancer dataset.** Two separate datasets of non-contrast Lung CT scans were prospectively collected in an Institutional Review Board-approved (IRB #02-13-42C), HIPAA-compliant study from two collaborating institutions: University Hospitals, Cleveland and Cleveland Clinic Foundation (CCF) in 2013 and 2014. Written informed consent was obtained for all the studies within the two cohorts. Histology was confirmed by anatomical pathologists at the respective institutions from the surgical specimen available for studies employed in both the cohorts. All patients underwent non-contrast CT scans prior to resection as part of routine care. Patients with multiple solitary nodules were excluded. Dataset 1 from University Hospitals, used as the training set, comprised 64 studies (31 adenocarcinomas and 33 granulomas). Dataset 2, from Cleveland Clinic, consisted of 56 cases (34 adenocarcinomas and 22 granulomas), and was used as an independent test set. The CTs were acquired using Siemens scanners with 2 mm slice thickness and 1 mm reconstruction. The tube voltage/current was 120 kV/150 mAs. The clinicopathologic characteristics of the lung cancer dataset have been summarized in Table 4.

Descriptor	#	Feature setting
Haralick	52	Window size ( $\mathcal{N} \times \mathcal{N}$ ); $\mathcal{N} \in \{3, 5, 7, 9, 11\}$ ; $\theta \in [0, 180^\circ]$
HOG	100	$v \in \{10, 15, 20, 25, 30\}$ ; $\kappa = 360^\circ/v$
Gabor	432	$t \in \{0, 2, 4, \dots, 32\}$ ; $\lambda \in \{0, 22.5, \dots, 180\}$ ; $\sigma_x^2 \in \{2, 4, 6\}$ , $\sigma_y^2 \in \{2, 4, 6\}$
CoLIAGe	500	Window size ( $\mathcal{N} \times \mathcal{N}$ ); $\mathcal{N} \in \{3, 5, 7, 9, 11\}$ ; $v \in \{10, 15, 20, 25, 30\}$

**Table 5. Summary of features and feature parameters used in this work.** A total of 26 parameters (across different features) were compared and evaluated in terms of their classification performance, for each of the three experiments in brain, breast and lung cancers.

**Pre-processing.** For the brain tumor cohort, the lesion ROIs were manually traced by an expert neuroradiologist, with over 25 years of experience, using the annotation tool in 3D Slicer after skull-stripping and intensity standardization. Intensity standardization<sup>17,18</sup> is an essential pre-requisite when comparing image intensities across different acquisitions, as it allows the gray scale MR intensities to have a fixed tissue-specific meaning within the same imaging protocol and the same body region, and within the same patient. For the brain MRI cohort, the segmented regions included enhancing and non-enhancing neoplastic tissue with the exclusion of edema. The neuroradiologist annotated all the 2D slices that had visible lesions.

For the breast cancer and lung cancer cohort, the best representative section, which was a central section of either the DCE-MRI volume or the non-contrast CT scan, was annotated by a breast radiologist and a thoracic radiologist respectively, who were both blinded to the pathologic diagnosis. Both for lung and breast cohorts, the lesion boundary was manually delineated on the basis of the image that demonstrated the greatest lesion conspicuity from neighboring tissues, which was then used for subsequent analysis.

**Comparison of CoLIAGe with other popular texture features.** While we qualitatively evaluated 3D CoLIAGe on a limited cohort of studies, quantitative analysis for all the three use cases in brain, breast, and lung, was restricted to 2D CoLIAGe, due to the variable slice thickness and anisotropic MRI volumes of our retrospective studies. We compared the performance of 2D CoLIAGe with other state-of-the-art texture descriptors (i.e. Haralick, Gabor, HOG). Apart from CoLIAGe, we extracted a total of 584 2D texture features for the ROIs on a per-pixel basis, including 52 Haralick features, 432 Gabor features and 100 HOG descriptors. The Gabor filter bank consisted of six different frequency-shift values  $t \in \{0, 2, 4, \dots, 32\}$ , eight orientation parameter values  $\lambda \in \{0, 22.5, 45, \dots, 180\}$ , and 9 different variance settings, generating a total of 432 different filters. Each filter yielded a real and imaginary response, which were used to calculate the total response magnitude for every region of interest. HOG features were computed via pixel-wise gradient orientations obtained from the grayscale intensity differences in the region of interest. These orientations were then equally binned into  $v$  bins,  $v \in \{10, 15, 20, 25, 30\}$ , with each bin encompassing  $36^\circ, 24^\circ, 18^\circ, 14.4^\circ$  or  $12^\circ$ . For Haralick and Gabor descriptors, the feature representations were obtained from every lesion by computing the median of feature values across all pixels within a lesion. Summary of parameters used for different feature sets is provided in Table 5.

**Comparison of CoLIAGe with expert diagnosis.** To further demonstrate the efficacy of CoLIAGe in distinguishing similar appearing pathologies on imaging, we performed a human machine comparison for our brain ( $N = 42$ ) and lung ( $N = 20$ ) cohort. For both the cohorts, collaborating expert readers (board certified attending radiologists and pulmonologists) independently provided a diagnosis, which was then compared with the analysis from the CoLIAGe classifier. In both the human-machine comparison experiments, the expert readers were kept blinded to the pathology reports. The experts assigned a score between  $[0, 1]$  to each lesion, with 0 referring to a high confidence that the nodule is “benign” (radiation necrosis or granuloma), and 1 being “malignant” (recurrent tumor or adenocarcinoma). Similarly, probability scores were assigned by the Random Forest classifier using CoLIAGe features. Using the assigned probabilities we computed the areas under the respective receiver operating characteristic curves (AUCs).

**Experimental Evaluation.** The feature sets were used to train a Random Forest (RF) classifier<sup>19</sup>, a bootstrapped aggregation of multiple decision tree classifiers, in conjunction with  $F^{2D}$  to distinguish between the categories of interest. Wilcoxon’s rank sum test<sup>20</sup> was employed to assess statistical significance and corrected for multiple comparisons for the experiments performed for the three use-cases. The sample size for the different experiments has been summarized in Table 6. In all our experiments, the RF classifier was used to assign every ROI into classes  $\{+1, -1\}$  based on the following classification tasks:

- **Experiment 1:** Distinguishing radiation necrosis from recurrent tumor on MRI,
- **Experiment 2:** Distinguishing triple negatives from other molecular subtypes of breast cancer (ER+, HER2+ and benign FA) on MRI,
- **Experiment 3:** Distinguishing adenocarcinoma from granuloma on non-contrast CT.

**Experiment  $E_j$ : Distinguishing Radiation Necrosis from Recurrent Tumor.** We computed  $F^{2D}$  for every slice with expert-annotated ROI on the primary and metastatic brain tumor cohorts and employed the feature set in a RF classifier setting to distinguish RN from rBT such that slices from the same patient are either used for training or for testing. A total of 50 trees were used for training the RF classifier. 3-fold randomized cross-validation was



Cancer Site	Cohort Size (Site)	Clinical challenge	Classification task	Number of samples
Brain Tumors	42 (UH)	RN versus rBT	Primary tumor cohort	10 RN, 12 rBT
			Metastatic tumor cohort	8 RN, 12 rBT
Breast Cancer	65 studies (UPenn)	Distinguish subtypes of breast cancer	TN-ER	21 TN, 25 ER+
			TN-HER2+	21 TN, 18 HER2+
			TN-FA	21 TN, 12 FA
Lung Cancer	120 studies (UH, CCF)	Distinguish granuloma from adenocarcinoma	Training Set (UH)	31 adenocarcinoma, 33 granuloma
			Test Set (CCF)	34 adenocarcinoma, 22 granuloma

**Table 6. List of studies employed in this work for three different clinical problems in brain, breast, and lung cancers.** The data acquisition sites include University Hospitals, Cleveland (UH), University of Pennsylvania (UPenn) and Cleveland Clinic (CCF).

used to train and evaluate classifier performance. This involved randomly splitting the entire dataset into 3 equally sized sets with 2 subsets used for classifier training and 1 subset used for independent evaluation. The diagnostic performance of each classifier trained with CoLLAGE and other comparative features was evaluated using average classification accuracy  $\beta^{Acc}$ , computed over 150 iterations of cross-validation runs. Quantitative results across CoLLAGE and the other texture features were compared by computing  $\beta^{Acc}$  for every feature set at the operating point. Additionally, for a subset of studies we computed  $F^{3D}$  for qualitative visualization. Qualitative results for both 2D and 3D CoLLAGE feature representation were visualized as standard heatmaps where high CoLLAGE values were shown in red while blue represented low CoLLAGE values.

*Experiment  $E_2$ : Distinguishing Molecular Subtypes of Breast Cancer.*  $F^{2D}$  was similarly computed for the slices with expert annotated ROI for the breast cancer cohort and was used to distinguish triple negative from the other breast cancer subtypes (ER+, HER2+) and benign fibroadenoma. A similar cross-validation technique was employed, along with a comprehensive comparison with other texture descriptors as given in  $E_1$ .

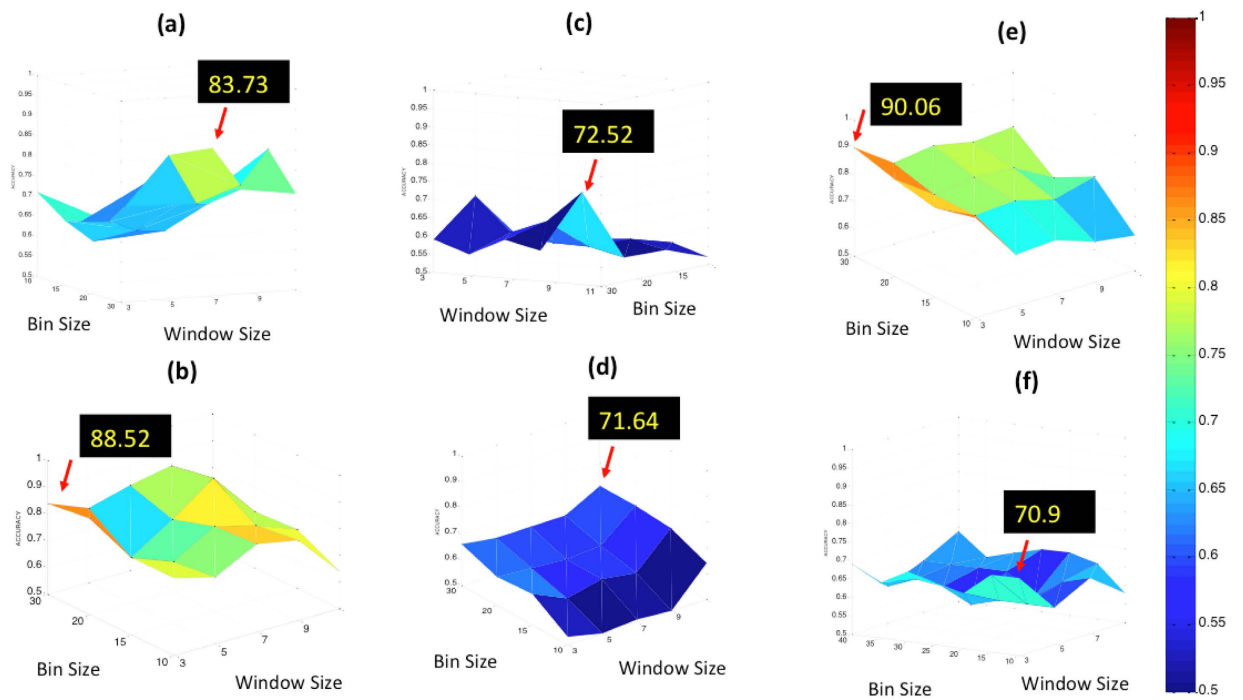
*Experiment  $E_3$ : Distinguishing Adenocarcinomas from Granulomas.* For every slice with the expert annotated ROI,  $F^{2D}$  was similarly computed for the lung cancer cohort and was used to distinguish adenocarcinomas from granulomas. Cross-validation, along with a comprehensive comparison with other texture descriptors as given in  $E_1$  was employed for the training set. This cross-validation strategy was employed for all the parameters as listed in Table 5. The model using the parameters that yielded the best results, for each category of feature descriptors, across the aggregated cross-validation runs was ‘locked down’, and then used to classify the cases in the independent test cohort.

**Parameter Sensitivity Analysis.** The sensitivity of CoLLAGE features was evaluated across its two key parameters, bin size of the entropy histogram ( $\nu$ ) and neighborhood ( $\mathcal{N}$ ) size for computing localized orientations. To account for smaller lesion sizes across the two cohorts, we restricted the  $\mathcal{N} \in \{3, 5, \dots, 11\}$ . Bin sizes  $\nu \in \{10, 15, \dots, 30\}$  were additionally considered at regular bin intervals of 5 to evaluate variation in  $\beta^{Acc}$ . We reported the variation in  $\beta^{Acc}$  as a function of  $\mathcal{N}$  and  $\nu$ .  $\nu > 20$ , and  $\mathcal{N} = 9$  (Fig. 4) were found to be optimal parameters for the brain tumor cohort. Similarly, for the breast cancer cohorts (results not shown), the following pairs of parameters were found to be optimal:  $\nu = 30$ , and  $\mathcal{N} = 5$  for TN versus ER+,  $\nu = 30$ , and  $\mathcal{N} = 11$  both for TN versus HER2+ and TN versus FA. For the lung cancer cohort, the best parameters were found to be  $\nu = 10$ , and  $\mathcal{N} = 3$ . The bounds of the different parameters were selected in a way that accounted for boundary effects that would arise in case of relatively small lesions ( $\mathcal{N} > 11$ ), and when  $\nu > 30$ . Figure 4 shows the parameter sensitivity for all the different classification experiments using CoLLAGE.

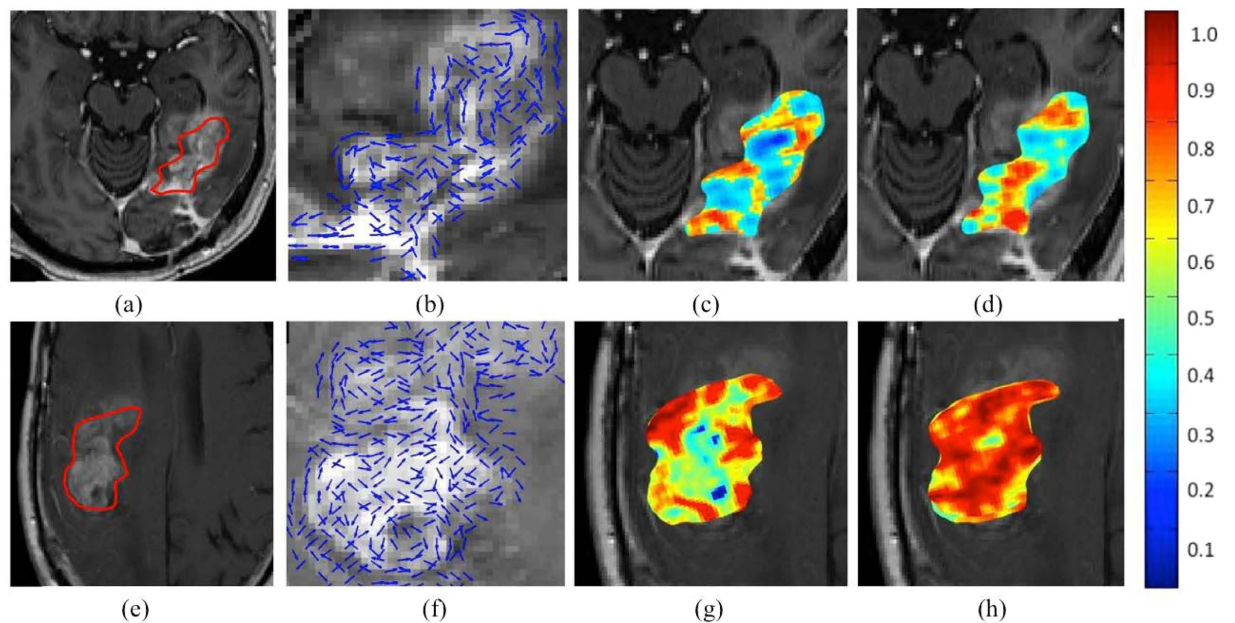
## Results and Discussion

**Experiment  $E_1$ : Distinguishing Radiation Necrosis from Recurrent tumor.** Figure 5 shows the qualitative feature maps of 3D CoLLAGE on a Gd-T1 MRI for radiation necrosis (a) and recurrent tumor (e) respectively. The localized gradient field,  $\theta^{3D}$ , for radiation necrosis and recurrent tumor is shown in (b) and (e), while the CoLLAGE entropy heatmaps,  $\mathcal{E}^{\theta^{3D}}$  and  $\mathcal{E}^{\phi^{3D}}$  are shown in (c), (d) for radiation necrosis and in (g), (h) for tumor recurrence respectively. High CoLLAGE values are reflected in red, while blue reflects under expression of CoLLAGE values. As may be evidenced from the CoLLAGE heatmaps, tumor recurrence has an over-expression of CoLLAGE both in (X, Y) as well as (X, Y, Z) directions compared to radiation necrosis. The over-expression of CoLLAGE values may be reflective of the higher structural heterogeneity of recurrent tumor, owing to the presence of more varied tissue types and hypercellularity, as compared to radiation necrosis.

The quantitative results including comparison of the classification performance of 2D CoLLAGE with Haralick, Gabor, and HOG, using the best parameter settings, are shown in Table 7. The best classification accuracy obtained for the popular-texture features was reported to be between 50% to 65%, while CoLLAGE was found to perform significantly better (over 20% improvement in classification accuracy ( $p$ -value  $< 0.001$ )) with an accuracy of  $83.79 \pm 5.43\%$  for primary cases, and  $88.52\% \pm 3.93$  for the metastatic brain tumor cohort. It has been previously shown that anomalies in brain tissue morphology are associated with directional patterns that can be



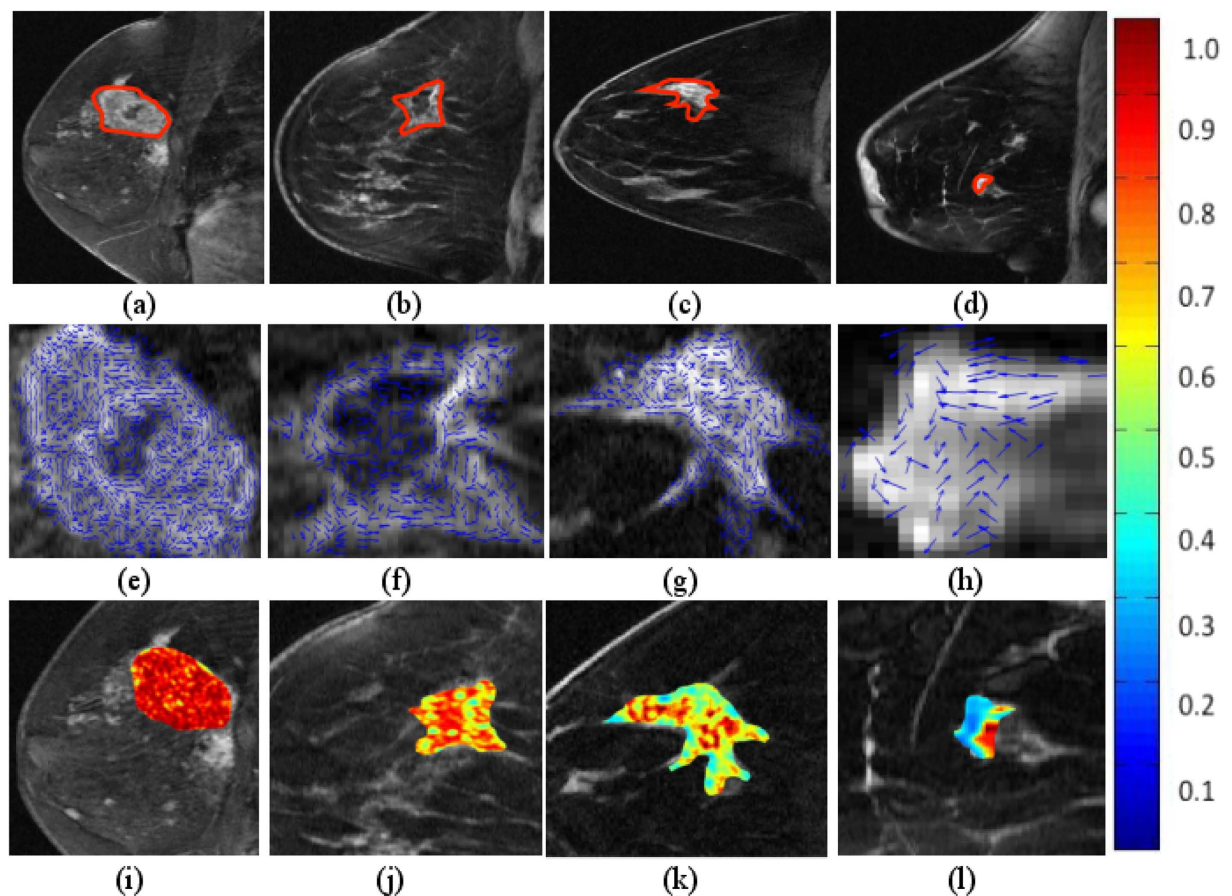
**Figure 4.** Parameter sensitivity analysis shown for CoLLAGE features for the six different classification experiments (a) RN vs rBT (primary brain tumors), (b) RN vs rBT (metastatic brain tumors), (c) TN vs ER+, (d) TN vs HER2+, (e) TN vs FA and (f) adenocarcinoma vs granuloma. The most optimal CoLLAGE parameters for each of the experiments were found to be as follows: (a)  $\mathcal{N} = 9, \nu = 30$  (b)  $\mathcal{N} = 9, \nu = 20$  (c)  $\mathcal{N} = 5, \nu = 30$  (d)  $\mathcal{N} = 11, \nu = 30$  (e)  $\mathcal{N} = 11, \nu = 30$  (f)  $\mathcal{N} = 3, \nu = 10$ , respectively.



**Figure 5.** CoLLAGE entropy maps on brain tumor cases. Qualitative feature maps of 3D CoLLAGE visualized on Gd-T1 MRI for radiation necrosis (a) and recurrent tumor (e) respectively. The localized gradient orientations are shown in (b) and (f) for radiation necrosis and tumor recurrence, while the CoLLAGE heatmaps,  $\mathcal{E}^{\theta^{3D}}$  and  $\mathcal{E}^{\phi^{3D}}$  are shown in (c), (d) for radiation necrosis and in (g), (h) for tumor recurrence respectively. As evident, tumor recurrence has a higher density of high entropy regions both in (X, Y) as well as (X, Y, Z) directions than radiation necrosis.

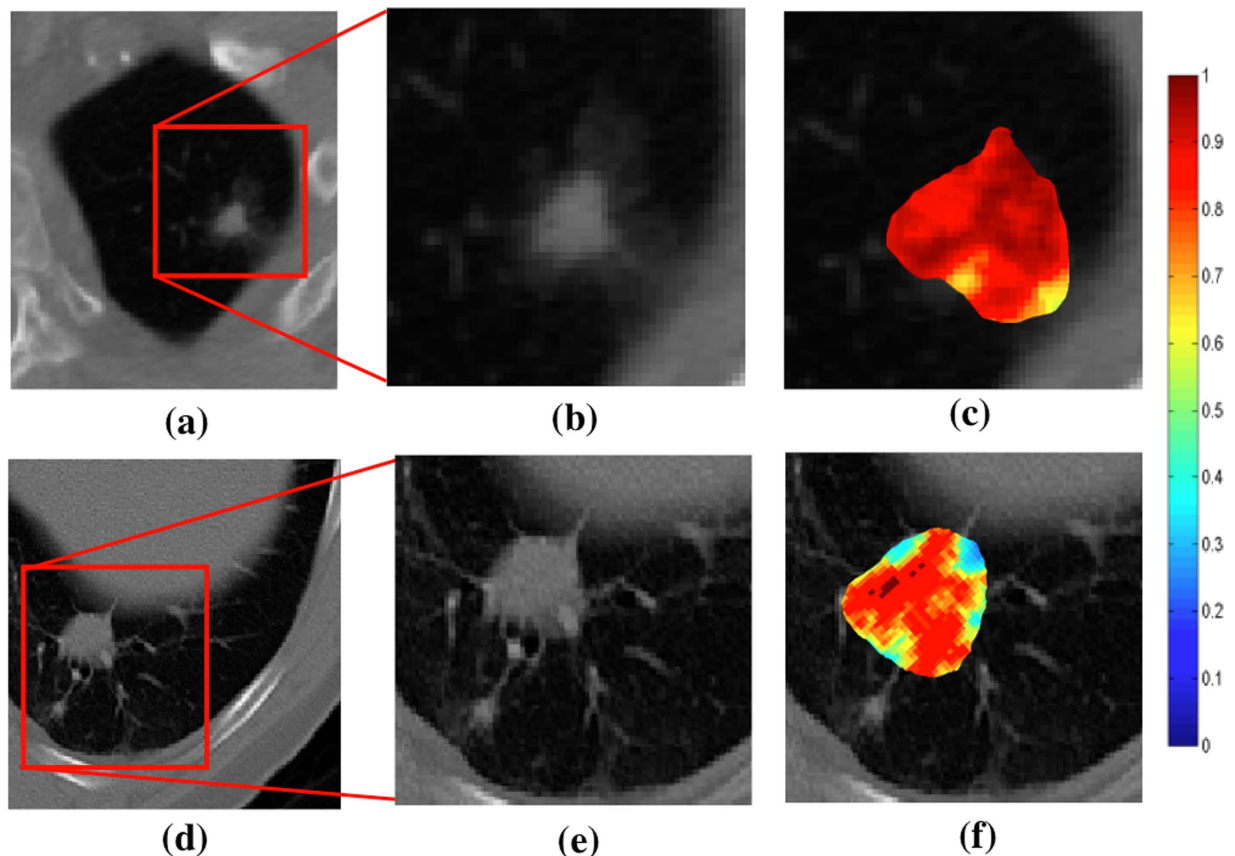
Classification Task	Haralick	HOG	Gabor	CoLIAGe
RN vs rBT (GBM)	62.19 ± 0.99	60.62 ± 3.21	59.68 ± 5.8	<b>83.73 ± 5.43</b>
RN vs rBT (Metastatic)	63.83 ± 2.42	72.99 ± 1.35	59.45 ± 1.73	<b>88.52 ± 3.93</b>
TN vs ER+	60.5 ± 4.49	68.49 ± 3.8	62.37 ± 4.29	<b>72.52 ± 5.19</b>
TN vs HER2+	54.69 ± 5.19	60.18 ± 4.08	62.67 ± 3.97	<b>71.64 ± 6.13</b>
TN vs FA	78.37 ± 3.75	72.7 ± 4.41	68.1 ± 1.85	<b>90.06 ± 4.38</b>
Adenocarcinoma vs Granuloma (Training)	75.5 ± 5	67.7 ± 5.4	66.2 ± 4.4	<b>70.9 ± 7.1</b>
Adenocarcinoma vs Granuloma (Testing)	69.6	67.3	62.4	<b>69.8</b>

**Table 7.**  $\beta^{Acc}$  for 2D CoLIAGe and the comparative strategies (Haralick, Gabor, and HOG) obtained across 150 iterations of 3-fold cross validation in a random forest classifier setting for the brain and breast cancer use-cases, as well as for independent training and test dataset for the lung cancer use-case.



**Figure 6.** CoLIAGe entropy maps on breast cancer cases. 2D DCE-MRI (1.5 Tesla) scan for TN+ (a), HER2+ (b), ER+ (c), and FA (d) with the lesion outlined in red. 4(e), (f), (g), (h) represent localized gradient orientations, while 4(i), (j), (k), (l) represent CoLIAGe heatmaps for the corresponding lesion on (a), (b), (c) and (d), where red represents higher while blue represents low CoLIAGe values.

captured by texture analysis. For example, gyrifications in gray matter create oriented spatial frequencies, that can be captured by wavelet-based features. Kovalev *et al.*<sup>21</sup> have analyzed gradient and anisotropy properties of 3D texture in the context of neurodegenerative diseases. According to Georgiadia *et al.*<sup>22</sup>, brain lesion texture is correlated with presence and type of cancer cells. A recent study<sup>23</sup> employed Haralick and wavelet texture features to distinguish radiation necrosis from metastatic brain tumor recurrence with a reported AUC of 94%. However, we believe that the results, reported on a per-slice basis, may have been affected by the classifier being contaminated by slices from the same patient being used both within the training as well as testing sets during classification. It is worth noting that the diagnostic accuracy of distinguishing radiation necrosis from tumor recurrence by an expert radiologist on visual inspection of MRI has been reported to be between 50–60%<sup>24</sup>.



**Figure 7. CoLIAGe entropy maps on lung cancer cases.** Qualitative feature maps of 2D CoLIAGe visualized on non-contrast lung CT for (a) adenocarcinoma and (d) granuloma shown in (c) and (f) respectively. Adenocarcinomas have higher CoLIAGe entropy values as compared to granulomas. Red represents higher while blue represents low CoLIAGe values.

**Experiment  $E_2$ : Distinguishing Molecular Sub-types of Breast Cancer.** Figure 6 shows the qualitative 2D CoLIAGe feature maps for each of the breast cancer sub-types, TN (i), HER2+ (j), ER+ (k), and FA (l). Higher CoLIAGe values are reflected in red, while blue reflects low CoLIAGe values. The corresponding localized orientations ( $\theta^{2D}$ ) for TN, HER2+, ER+ and FA are shown in Fig. 6(e), (f), (g) and (h) respectively. It is interesting to note that the gradient field was found to be more disordered across cancer sub-types, as compared to benign FA. Similar to Experiment  $E_1$ , a marked difference in CoLIAGe values was observed across different sub-types of breast cancer (TN, HER2+, ER+, and FA), suggesting that CoLIAGe may potentially be capturing local anisotropic differences in micro-structures on imaging that are otherwise not visually appreciable.

Table 7 shows the average  $\beta^{Acc}$  values obtained over 150 iterations of a 3 fold-cross validation via a RF classifier.  $\beta^{Acc}$  values obtained from CoLIAGe in the breast cancer cohort significantly outperformed ( $p$ -value  $< 0.001$ ) the other state-of-the-art texture descriptors (Haralick, Gabor, HOG), with an improvement of  $\approx 10\%$  for majority of the 3 classification tasks (TN versus ER+, TN versus HER2+, TN versus FA). Our results resonate with the findings reported in Agner *et al.*<sup>25</sup> that used a similar cohort of breast DCE-MRI studies to distinguish different sub-types of breast cancer using a novel texture kinetic approach. Similar to Agner *et al.*, the most prominent difference, both in qualitative and quantitative performance of CoLIAGe, is reported between FA, a benign condition, from TN, the most aggressive sub-type of breast cancer with  $\beta^{Acc}$  of  $90.06 \pm 4.38$ . It is worth noting that currently radiologists are unable to distinguish TNs from FAs on a routine MRI<sup>26–28</sup>.

**Experiment  $E_3$ : Distinguishing Adenocarcinomas from Granulomas.** Figure 7 shows the qualitative 2D CoLIAGe feature maps for a representative non-contrast CT image with pathologically proven adenocarcinoma (a) and granuloma (d), respectively. Higher CoLIAGe values are reflected in red, while blue reflects low CoLIAGe values. It may be observed that representative adenocarcinoma lesion has higher density of larger CoLIAGe entropy values as compared to the granuloma sample. Table 7 shows the average  $\beta^{Acc}$  values obtained over 150 iterations of a 3 fold-cross validation via a RF classifier. Using the locked-down classifier, CoLIAGe features showed the best classification results for the test set (69.8%).

Tumor heterogeneity has been previously shown to be associated with non-small cell lung carcinoma<sup>29</sup>. This heterogeneity can be attributed to the hypoxic microenvironment<sup>30</sup>. The subtle differences in the hypoxia-related heterogeneity as suggested in ref. 29 is perhaps manifested in the differential expression of CoLIAGe entropy.

Dennie *et al.*<sup>31</sup> have reported an AUC of 0.90 in distinguishing the two conditions. The texture analysis in ref. 31, using Haralick features, has yielded higher accuracy than FDG-PET/CT in distinguishing the two pathologies. However, the approach in ref. 31 has not been validated on an independent test set. Besides, there is no clear qualitative evidence of the Haralick features visibly distinguishing the two pathologies.

**Human-Machine Comparison Results.** On the hold-out lung cohort, the AUCs for the two experts were found to be 0.68, and 0.54 respectively. Using the CoLIAGE features, the associated AUC was computed as 0.78. For the brain tumor studies, AUCs for the two independent readers were 0.75 and 0.58 respectively, while the same for a CoLIAGE-based classifier was computed to be 0.80. To the best of our knowledge, no radiomics-based work recently has demonstrated such a rigorous and comprehensive human-machine reader comparison across multiple disease sites.

### Concluding Remarks

We presented a radiomic feature descriptor, Co-occurrence of Local Anisotropic Gradient Orientations (CoLIAGE), that captures higher order co-occurrence patterns of local gradient tensors at a voxel level to distinguish disease phenotypes that have similar morphologic appearances. We employed three clinically challenging datasets to evaluate the efficacy of CoLIAGE, distinguishing (a) radiation necrosis, a relatively benign effect of radiation, from tumor recurrence on T1-w MRI in brain tumors, (b) different molecular sub-types of breast cancer on DCE-MRI and (c) adenocarcinomas from granulomas on non-contrast CT. Additionally, we compared performance of CoLIAGE with other state-of-the-art texture descriptors (Haralick, Gabor, Histogram of Gradient orientations) as well as across two expert readers (for two use-cases), and demonstrated that CoLIAGE has significantly better classification accuracy than the other texture descriptors as well as the expert readers. Across the cross-validation and testing stages, CoLIAGE outperformed other texture features in 20 out of 21 comparative experiments (Table 7). Our results, on all three cohorts, seem to suggest that CoLIAGE has the potential to serve as a powerful radiomic descriptor in distinguishing similar appearing pathologies on imaging.

Nevertheless, there are a few limitations to our study. Firstly, CoLIAGE was only compared against three popularly used texture features (Haralick, Gabor, and HOG). Secondly, based on our parameter sensitivity analysis, it appears that parameter selection may be an important consideration when employing CoLIAGE for a specific problem. In this paper, histogram representation was used to collate CoLIAGE values to classify every image within a random forest classifier. However, the choice of feature representation and classification methods is flexible and can be modified depending on the specific application. Future work will focus on (1) rigorously evaluating efficacy of CoLIAGE across other texture features on a larger cohort of multi-institutional studies, and (2) identifying a domain-independent parameter selection strategy to evaluate robustness of CoLIAGE.

### References

1. Tiwari, P. *et al.* Texture descriptors to distinguish radiation necrosis from recurrent brain tumors on multi-parametric MRI. In *SPIE Medical Imaging*, 90352B–90352B (International Society for Optics and Photonics, 2014).
2. Agner, S. *et al.* Computerized image analysis for identifying triple-negative breast cancers and differentiating them from other molecular subtypes of breast cancer on DCE-MRI. *Radiology* **272**, 91–99 (2014).
3. Kumar, V. *et al.* Radiomics: the process and the challenges. *Magn Reson Imaging* **30**(9), 1234–48 (2012).
4. Lambin, P. *et al.* Radiomics: extracting more information from medical images using advanced feature analysis. *European Journal of Cancer* **48**, 441–446 (2012).
5. Gatenby, R., Grove, O. & Gillies, R. Quantitative imaging in cancer evolution and ecology. *Radiology* **269**, 8–14 (2013).
6. Haralick, R., Shanmugam, K. & Dinstein, I. H. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on* 610–621 (1973).
7. Jain, A. & Farrokhnia, F. Unsupervised texture segmentation using Gabor filters. In *Systems, Man and Cybernetics, 1990. Conference Proceedings., IEEE International Conference on*, 14–19 (IEEE, 1990).
8. Prasanna, P., Jain, S., Bhagat, N. & Madabhushi, A. “Decision support system for detection of diabetic retinopathy using smartphones,” 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, 176–179, 2013.
9. Prasanna, P., Patel, J., Partovi, S., Madabhushi, A. & Tiwari, P. “Radiomic features from the peritumoral brain parenchyma on treatment-naïve multi-parametric MR imaging predict long versus short-term survival in glioblastoma multiforme: Preliminary findings,” *European Radiology* 2016, doi:10.1007/s00330-016-4637-3.
10. Fogel, I. & Sagi, D. Gabor filters as texture discriminator. *Biological cybernetics* **61**, 103–113 (1989).
11. Dalal, N. & Triggs, B. Histograms of Oriented Gradients for human detection. In *CVPR 2005 vol. 1*, 886–893 (IEEE, 2005).
12. Watanabe, T., Ito, S. & Yokoi, K. Co-occurrence histograms of oriented gradients for pedestrian detection. In *Advances in Image and Video Technology* **5414**, 37–47 (2009).
13. Pang, Y., Yan, H., Yuan, Y. & Wang, K. Robust CoHOG feature extraction in human-centered image/video management system. *Systems, Man, and Cybernetics* **42**, 458–468 (2012).
14. Ito, S. & Kubota, S. Object classification using heterogeneous co-occurrence features. In *Computer Vision–ECCV 2010* **6315**, 701–714 (2010).
15. Lee, G. *et al.* Cell Orientation Entropy (COE): Predicting biochemical recurrence from prostate cancer tissue microarrays. In *MICCAI* (3), 396–403 (2013).
16. Prasanna, P., Tiwari, P. & Madabhushi, A. Co-occurrence of local anisotropic gradient orientations (collage): Distinguishing tumor confounders and molecular subtypes on MRI. In *MICCAI 2014*, 73–80 (Springer International Publishing, 2014).
17. Madabhushi, A. & Udupa, J. New methods of MR image intensity standardization via generalized scale. *Medical Physics* **33**, 3426–3434 (2006).
18. Lötjönen, Jyrki M. P. *et al.* Fast and robust multi-atlas segmentation of brain magnetic resonance images. *Neuroimage* **49**(3), 2352–2365 (2010).
19. Breiman, L. Random forests. *Machine learning* **45**, 5–32 (2001).
20. Wilcoxon, F. & Wilcoxon, R. A. *Some Rapid Approximate Statistical Procedures* (Lederle Laboratories, 1964).
21. Kovalev, V. A., Kruggel, F., Gertz, H.-J. & Von Cramon, D. Y. Three-dimensional texture analysis of mri brain datasets. *Medical Imaging, IEEE Transactions on* **20**, 424–433 (2001).
22. Georgiadis, P. *et al.* Enhancing the discrimination accuracy between metastases, gliomas and meningiomas on brain mri by volumetric textural features and ensemble pattern recognition methods. *Magnetic Resonance Imaging* **27**, 120–130 (2009).

23. Larroza, A. *et al.* Support vector machine classification of brain metastasis and radiation necrosis based on texture analysis in MRI. *Magnetic Resonance Imaging* **42**(5), 1362–1368 (2015).
24. Verma, N., Cowperthwaite, M., Burnett, M. & Markey, M. Differentiating tumor recurrence from treatment necrosis: a review of neuro-oncologic imaging strategies. *Neuro-oncology* **15**, 515–534 (2013).
25. Agner, S. *et al.* Textural kinetics: a novel dynamic contrast-enhanced (DCE)-MRI feature for breast lesion classification. *Journal of Digital Imaging* **24**, 446–463 (2011).
26. Schrading, S. & Christiane, K. K. Mammographic, US, and MR Imaging Phenotypes of Familial Breast Cancer 1. *Radiology* **15**, 58–70 (2008).
27. Uematsu, T., Kasami, M. & Yuen, S. Triple-Negative Breast Cancer: Correlation between MR Imaging and Pathologic Findings 1. *Radiology* **250**(3), 638–647 (2009).
28. Agner, S. *et al.* Computerized image analysis for identifying triple-negative breast cancers and differentiating them from other molecular subtypes of breast cancer on dynamic contrast-enhanced MR images: a feasibility study. *Radiology* **272**(1), 91–99 (2014).
29. Ganeshan, B. *et al.* Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *European radiology* **22**, 796–802 (2016).
30. Foster, J. G., Wong, S. C. & Sharp, T. V. The hypoxic tumor microenvironment: driving the tumorigenesis of non-small-cell lung cancer. *Future Oncology* **10**(16), 2659–2674 (2014).
31. Dennie, C. *et al.* Role of quantitative computed tomography texture analysis in the differentiation of primary lung cancer and granulomatous nodules. *Quantitative imaging in medicine and surgery* **6**(1) (2016).

## Acknowledgements

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award numbers R01CA136535-01, R01CA140772-01, R21CA167811-01, R21CA179327-01, R21CA195152-01 the National Institute of Diabetes and Digestive and Kidney Diseases under award number R01DK098503-02, the DOD Prostate Cancer Synergistic Idea Development Award (PC120857); the DOD Lung Cancer Idea Development New Investigator Award (LC130463); the Ohio Third Frontier Technology development Grant, the CTSC Coulter Annual Pilot Grant, the Case Comprehensive Cancer Center Pilot Grant, VelaSano Grant from the Cleveland Clinic, and the Wallace H. Coulter Foundation Program in the Department of Biomedical Engineering at Case Western Reserve University. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## Author Contributions

P.P. developed the method, P.P., P.T. and A.M. conceived the experiments, P.P. conducted the experiments, P.P. and P.T. analyzed the results. All authors reviewed the manuscript.

## Additional Information

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Prasanna, P. *et al.* Co-occurrence of Local Anisotropic Gradient Orientations (CoLLAGE): A new radiomics descriptor. *Sci. Rep.* **6**, 37241; doi: 10.1038/srep37241 (2016).

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

© The Author(s) 2016