Co-occurrence of Local Anisotropic Gradient Orientations (CoLlAGe): Distinguishing Tumor Confounders and Molecular Subtypes on MRI

Prateek Prasanna*, Pallavi Tiwari*, and Anant Madabhushi

Department of Biomedical Engineering, Case Western Reserve University, USA
pallavi.tiwari@case.edu

Abstract. We introduce a novel *biologically inspired* feature descriptor, Co-occurrence of Local Anisotropic Gradient Orientations (CoLlAGe), that captures higher order co-occurrence patterns of local gradient tensors at a pixel level to distinguish disease phenotypes that have similar morphologic appearances. A number of pathologies (e.g. subtypes of breast cancer) have different histologic phenotypes but similar radiographic appearances. While texture features have been previously employed for distinguishing subtly different pathologies, they attempt to capture differences in global intensity patterns. In this paper we attempt to model CoLlAGe to identify higher order co-occurrence patterns of gradient tensors at a pixel level. The assumption behind this new feature is that different pathologies, even though they may have very similar overall texture and appearance on imaging, at a local scale, will have different co-occurring patterns with respect to gradient orientations. We demonstrate the utility of CoLlAGe in distinguishing two subtly different types of pathologies on MRI in the context of brain tumors and breast cancer. In the first problem, we look at CoLlAGe for distinguishing radiation effects from recurrent brain tumors over a cohort of 40 studies, and in the second, discriminating different molecular subtypes of breast cancer over a cohort of 73 studies. For both these challenging cohorts, CoLlAGe was found to have significantly improved classification performance, as compared to the traditional texture features such as Haralick, Gabor, local binary patterns, and histogram of gradients.

1 Introduction

A number of pathologies have different histologic phenotypes but similar radiographic appearances. One such instance is the problem of discriminating fibroadenoma (FA), a benign breast tumor from triple negative (TN), an aggressive breast cancer [1]. Both FA and TN have distinct cellular and architectural

* P. Prasanna and P. Tiwari are joint first authors. Research was supported by R01CA136535-01, R01CA140772-01, R21CA167811-01, R01DK098503-02, PC120857, the QED award, the Ohio Third Frontier Technology development Grant, and the Coulter Translational Award. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

P. Golland et al. (Eds.): MICCAI 2014, Part III, LNCS 8675, pp. 73–80, 2014.

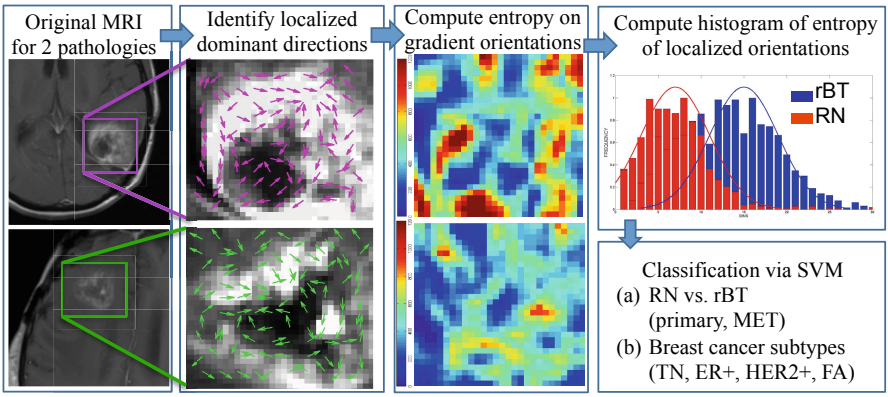


Fig. 1. Overview of CoLlAGe and overall workflow. The 1st module involves computing localized gradient orientations, while in the 2nd module entropy of localized gradient orientations is computed for every pixel. A histogram of entropy values is then aggregated for every pixel and is subsequently used for classification.

arrangements when examined on a pathology slide under a microscope. However they have very similar morphologic appearances on MRI. In this paper, we present a new *biologically inspired* feature descriptor, Co-occurrence of Local Anisotropic Gradient Orientations (CoLlAGe), that captures higher order co-occurrence patterns of local gradient tensors at a pixel level to distinguish disease phenotypes that have similar morphologic appearances. While texture features have emerged as a popular way of characterizing and distinguishing subtly differing pathologies, these operators typically tend to capture global textural patterns. One such class of texture features are grey-level co-occurrence matrix (GLCM) [2] and Gabor steerable filters. These texture descriptors involve computing global relationships between pixels by averaging responses to various filter operators within the neighborhood of a single global descriptor. While some texture features can provide pixel-level responses (e.g. local binary patterns (LBP)) [3], these filters are often employed to provide pixel level or patch based classification. LBP, unlike GLCM, provides a signature for every pixel by capturing localized intensity variations across the pixel. However, LBP, is highly dependent on the radius parameter, which is critical in extracting local patterns. Additionally, both global and per-pixel texture representations are based on intensity variations and are domain agnostic. When examined on a histopathology slide under a microscope at a high magnification, the differences between subtly different classes may be manifested in differently oriented nuclei, lymphocytes, and/or glands. The differences in histologic architecture, which are no doubt reflected on the imaging, hence need a new class of features to capture subtle differences in image patterns on a local scale.

The rationale behind our approach, CoLlAGe, 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 pathologies (e.g. FA versus TN), the organization of local gradients may differ across classes and will be relatively consistent within a class.

2 Previous Work and Novel Contributions

A popular texture descriptor that captures orientation variations is histogram of gradient orientations (HoG) [4]. HoG computes a global patch based signature by computing histogram distribution of orientations computed on a per pixel basis. However HoG, similar to the other texture descriptors, is domain agnostic and is not designed to capture localized per voxel texture characteristics depicted on imaging. A variant of HoG, called co-occurrence of histogram of gradient orientations (Co-HoG) was recently presented by Watanabe et al. [5] where a high-dimensional feature vector was computed for every pixel by accumulating values from co-occurrence matrix computed on gradient orientations for pedestrian detection on a per-pixel basis. However, the approach in [5], (a) did not capture localized variations across neighboring orientations, (b) is susceptible to "curse of dimensionality" (due to a high dimensional feature space), and (c) similar to its counterpart HoG, is domain agnostic.

Recently, deep learning (DL) has emerged as a powerful tool for learning alternative representations of data for improved classification [6]. DL approaches train multiple convolution layers on a large annotated dataset to learn abstract but useful patterns between classes. Although DL has shown tremendous promise in identifying complex differentiating patterns across diseases, the identified features are not intuitive and cannot be used to understand the underlying disease characteristics. Additionally, DL strategies require large annotated training dataset to obtain meaningful results.

CoLlAGe, on the other hand, is designed to be "biologically intuitive". Firstly, CoLlAGe captures neighborhood orientation variation via a localized gradient tensor field that may reflect the underlying cellular arrangement of the phenotype on imaging. Secondly, CoLlAGe computes co-occurrence matrix on localized gradient tensors to capture co-occurring patterns of orientation disorder in a localized fashion. While co-occurrence matrices are commonly used to describe image texture, to our knowledge, this is the first attempt at employing co-occurrences on localized gradient orientations to capture underlying orientation variations on imaging.

We demonstrate the utility of CoLlAGe in the context of two problems involving brain tumors and breast cancer. The first application involves evaluating radiation therapy response for distinguishing radiation necrosis (RN), a radiation induced effect, from recurrent brain tumors (rBT) for primary and metastatic (MET) brain tumors [7]. The second application involves identifying phenotypic imaging signatures of molecular sub-types of breast cancer: triple negative (TN), estrogen receptor positive (ER+), human epidermal growth factor receptor positive (HER2+), and benign fibroadenoma (FA) on dynamic contrast enhanced (DCE)-MRI [1].

3 Co-occurrence of Localized Gradient Orientations

A region of interest (ROI) on an MRI volume is defined as, $\mathcal{C} = (C, f)$, where $f(c)$ is the associated intensity at every pixel c on a 3D grid C. Computation of CoLlAGe for every $c \in C$ involves following steps,

1. **Calculation of gradient orientations 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 $\partial f_Y(c)$. The gradient orientation θ of every $c \in C$ is then calculated as $\theta(c) = \tan^{-1}\frac{\partial f_Y(c)}{\partial f_X(c)}$.
2. **Computing local dominant orientations via principal component analysis (PCA):** A $\mathcal{N}\times\mathcal{N}$ window centered around every $c \in C$ is selected. We then compute $\partial f_X(c_k)$ and $\partial f_Y(c_k)$, $k \in \{1, 2, \ldots, \mathcal{N}^2\}$. The vector gradient matrix $\overrightarrow{\mathbf{F}}$ associated with every c is given by $\overrightarrow{\mathbf{F}} = [\overrightarrow{\partial f}_X(c_k)\ \ \overrightarrow{\partial f}_Y(c_k)]$, where $[\overrightarrow{\partial f}_X(c_k)\ \ \overrightarrow{\partial f}_Y(c_k)]$, $k \in \{1, 2, \ldots, \mathcal{N}^2\}$ is the matrix of gradient vectors in the X and Y directions for every c_k. The most significant orientation for each pixel c_k within $\mathcal{N} \times \mathcal{N}$ is obtained by performing principal component analysis on $\overrightarrow{\mathbf{F}}$. The dominant principal components in X and Y directions are obtained as r_X^k and r_Y^k for every $k \in \{1, 2, \ldots, \mathcal{N}^2\}$. The most significant orientation for every c_k is then calculated as $\phi(c) = \tan^{-1}\frac{r_Y^k}{r_X^k}$.
3. **Calculation of second-order statistics for most significant orientations:** The objects of interest for calculating CoLlAGe features are the co-occurring directions given by discretization of the dominant orientation $\bar{\phi}(c_k)$ for every pixel c, such that $\phi(c_k) = \omega \times \text{ceil}(\frac{\bar{\phi}(c_k)}{\omega})$, where ω is a discretization factor.
   An $N \times N$ co-occurrence matrix $\mathcal{M}$ subsequently captures orientation pairs between pixels which co-occur in the neighborhood $\mathcal{W}_i$, such that,

$$\mathcal{M}_{\mathcal{W}_i}(p, q) = \sum_{c_j, c_k}^{\mathcal{W}_i} \sum_{p,q=1}^{N} \begin{cases} 1, & \text{if } \phi(c_j)=p \text{ and } \phi(c_k)=q \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

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

$$\mathcal{E}(c) = \sum_{p,q} -\mathcal{M}(p, q) log(\mathcal{M}(p, q)). \tag{2}$$

4. A **histogram of** $\mathcal{E}$ is computed by aggregating $\mathcal{E}(c_k)$, $k \in \{1, \ldots, |C|\}$, where $|.|$ is the cardinality of set C. The entropy histogram is divided into bin size v, optimized on the training set via grid search optimization.
   A CoLlAGe feature vector, $\mathcal{F}$ is then obtained for every $\mathcal{C}$ which consists of the binned histogram values in the form of v×1 vectors. $\mathcal{F}$ is then employed within a classifier for classification purposes.

Data: ROI volume $\mathcal{C}$
Result: *CoLlAGe* features, $\mathcal{F}$

begin
  for each pixel $c \in \mathcal{C}$, **do**
    Obtain gradients $\partial f_X(c)$ and $\partial f_Y(c)$ along X- and Y- axes ;
    Obtain gradient orientation $\theta(c) = \tan^{-1} \frac{\partial f_Y(c)}{\partial f_X(c)}$;
  end
  for each pixel $c \in \mathcal{C}$, **do**
    Compute gradient vectors $\overrightarrow{\partial f_X}(c_k)$ and $\overrightarrow{\partial f_Y}(c_k)$ in $\mathcal{N} \times \mathcal{N}$ neighborhood;
    Obtain localized gradient vector matrix $\overrightarrow{\mathcal{F}} = [\overrightarrow{\partial f_X}(c_k) \ \overrightarrow{\partial f_Y}(c_k)]$;
    Compute dominant orientation, $\phi(c_k)$, $k \in \{1, \ldots, \mathcal{N}^2\}$ via PCA;
  end
  Compute Co-occurrence matrix $\mathcal{M}$ from $\phi(c)$;
  Compute $\mathcal{E}(c)$ from $\mathcal{M}$;
  Obtain v×1 dimensional feature vector $\mathcal{F}$ from distribution of $\mathcal{E}$;
end

Algorithm 1. Computation of *CoLlAGe* features

4 Experimental Results and Description

4.1 Data Description

Dataset 1 comprised two cohorts of 20 primary (10 RN, and 10 rBT) and 20 MET (12 RN and 8 rBT) patient studies respectively. All the studies were retrospectively acquired with 3 Tesla Gadolinium-constrast (Gd-C) T1-w MRI), and were histologically confirmed on biopsy samples by an expert pathologist. Dataset 2 comprised DCE-MRI studies from 65 women with 73 breast lesions for whom pathology results and 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 73 lesions, 9 were benign FA, 21 were TN, 18 were HER2+, and 25 were ER+.

4.2 Implementation Details

Figure 1 shows the work-flow of CoLlAGe, and its implementation in the context of clinical problems in brain tumors and breast cancer. We compared CoLlAGe against GLCM, Gabor, HoG, and LBP features, and evaluated their performance using a support vector machine (SVM) classifier [8] with a radial basis function (RBF) kernel. A 3-fold cross-validation strategy was employed and the performance of each of the texture descriptors was compared over different window, $\mathcal{N} \in \{3, 5, 7, 9\}$, and bin sizes, $v \in \{10, 20, 30, 40, 50\}$. $v = 30$, and $\mathcal{N} = 7$ were found to be optimal parameters across different descriptors and was employed for further evaluation. The average accuracy values were reported over 100 runs of 3-fold cross validation for both the cohorts. Wilcoxon's rank sum test [9] was performed to report statistical significance and corrected for multiple comparisons for the experiments performed for the two use-cases.

Table 1. Summary of features and feature parameters used in this work

Descriptor	#	Feature setting	Description
Haralick	26	$\mathcal{N} \in \{3, 5, 7, 9\}$	gray-level co-occurrence such as angular second moment, contrast and entropy
LBP	59	radius $R = 8$	Histogram of intensity variations within R
HoG	20	bin-size=18°	Histogram of gradient orientations
CoLlAGe	30	$\omega = 20$; $\mathcal{N} \in \{3, 5, 7, 9\}$; $v \in \{10, 20, 30, 40\}$	Entropy of localized gradients for $\mathcal{N} \times \mathcal{N}$

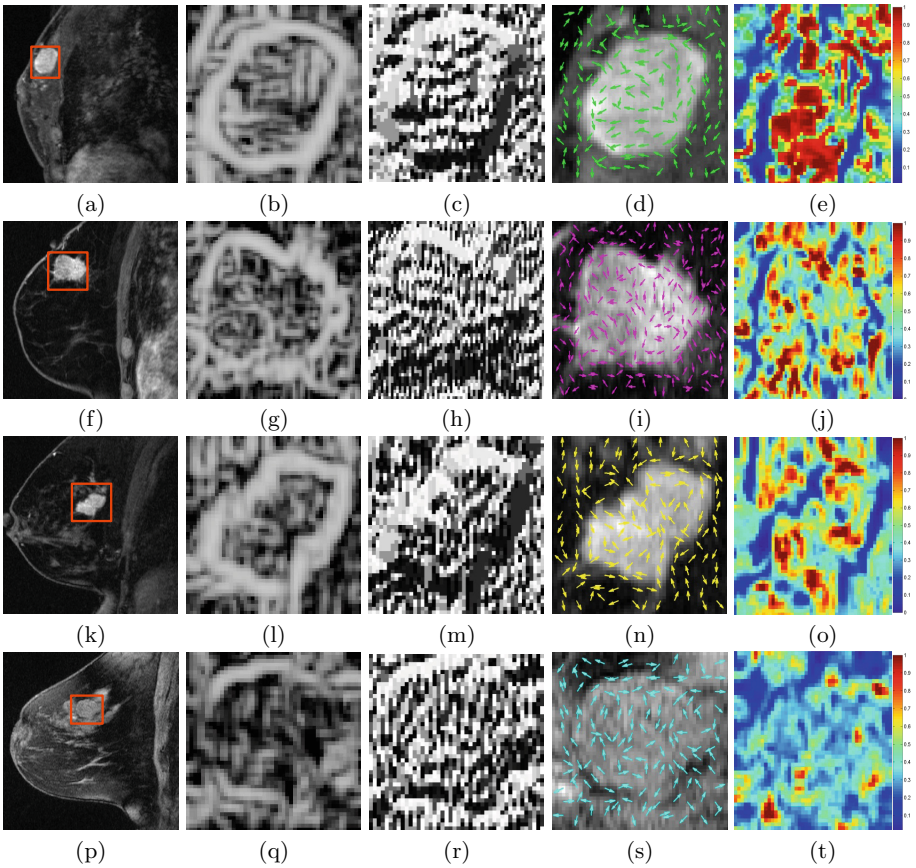


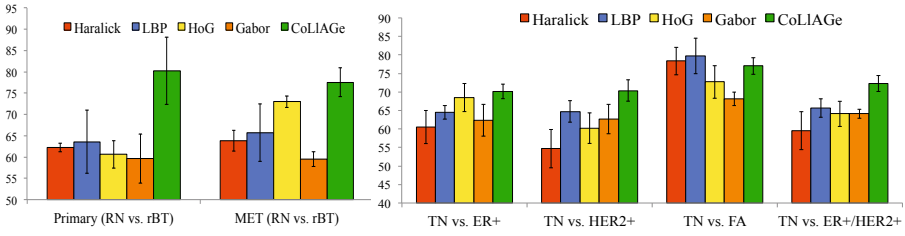
Fig. 2. Original DCE-MRI for ER+ (a), TN (f), HER2+ (k), and FA (p) with the lesion outlined in red. 2(b), (g), (l), (q) correspond to Haralick, while 2(c), (h), (m), (r) correspond to LBP representations of the lesion on ER+, TN, HER2+, and FA images. 2(d), (i), (n), (s) represent localized gradient orientations, while 2(e), (j), (o), (t) represent entropy heatmaps for the corresponding lesion on (a), (f), (k) and (p), where red represents higher while blue represents low entropy values.

4.3 Distinguishing RN from rBT for Primary and MET Patients

Figure 1 shows a representative RN, and rBT patient for primary brain tumor cohort. The orientations shown in magenta correspond to rBT, while the ones in green correspond to RN. The heatmaps represent entropy values obtained from the localized orientations on a per-pixel basis, where higher values of entropy are shown in red while lower values are shown in blue. It is interesting to note that the entropy values for rBT on a per pixel basis are substantially higher than those of RN suggesting orientation disorder in recurrent tumor. The histogram plots shown in red (RN), and blue (rBT) seem to suggest a clear separation between entropy distributions across the two morphologies. Figure 3 (a) demonstrates the quantitative results obtained for both primary as well as MET patients in distinguishing RN from rBT. For both the cohorts, CoLlAGe (80.25 $\pm$ 7.89% for primary cases, 77.55% $\pm$ 3.35 for MET) was found to significantly outperform Haralick (62.19% $\pm$ 0.99% for primary, 63.83% $\pm$ 2.42% for MET), Gabor (59.68% $\pm$ 5.8% for primary, 59.45% $\pm$ 1.73% for MET), LBP (63.63 $\pm$ 3.21% for primary, 65.75% $\pm$ 6.76% for MET), and HoG features (60.62 $\pm$ 3.21% for primary, 72.99% $\pm$ 1.35 for MET).

4.4 Distinguishing TN from Other Breast Cancer Subtypes

Figure 2 illustrates the qualitative comparison of CoLlAGe with the other texture descriptors, Haralick, LBP, and HoG in differentiating molecular subtypes, ER+ (2(a)), TN (2(f)), HER2+ (2(k)), and FA (2(p)) on DCE-MRI. Note the apparent differences in entropy heatmaps across 2(e), (j), (o), and (t) corresponding to ER+, TN, HER2+, and FA respectively. The most prominent difference is reported between FA, a benign condition, from ER+, a subtype of breast cancer. The results suggest that the orientations of cancer subtypes are more disordered than the benign condition. Similarly, ER+ reported overall higher entropy values than TN, and HER2+ cancer subtypes. The accuracy values averaged over 100 runs of 3 fold cross-validation for different feature descriptors are shown in Figure 3(b). It is interesting to note that although CoLlAGe significantly outperforms the other texture descriptors (Haralick, Gabor, LBP, and HoG) for



(a) (b)

Fig. 3. Mean accuracy values obtained for different feature descriptors (Haralick, LBP, HoG, Gabor, CoLlAGe) for (a) brain tumor, and (b) breast tumor datasets respectively.

distinguishing the more difficult disease subtypes (TN vs. HER2+, TN vs. ER+ and TN vs. ER+/HER2+), the results of CoLlAGe were comparable to Haralick descriptors for distinguishing TN vs. FA.

5 Concluding Remarks

In this work, we presented a new feature descriptor, Co-occurrence of Local Anisotropic Gradient Orientations (CoLlAGe), that captures higher order co-occurrence patterns of local gradient tensors at a pixel level to distinguish disease phenotypes that have similar morphologic appearances. We demonstrated the utility of CoLlAGe in identifying MRI phenotypes for clinically challenging problems in the context of breast and brain tumors for (a) distinguishing radiation necrosis, a treatment related effect from recurrent brain tumors over a cohort of 40 MRI studies, and (b) distinguishing breast cancer subtypes (ER+, HER2+, TN, and FA) over a cohort of 73 DCE-MRI studies. CoLlAGe was found to significantly outperform traditional texture descriptors such as Haralick, Gabor, local binary patterns, and histogram of gradients for the two use-cases, except for distinguishing FA from TN, where the results were found to be comparable to Haralick texture descriptors. In future work, we will seek to understand the correlation of CoLlAGE with pathologic correlates and evaluate its applicability across other disease sites such as prostate cancer.

References

1. 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 (pre-print, 2014)
2. Haralick, R.M., et al.: Textural Features for Image Classification. Systems, Man and Cybernetics (6), 610–621 (1973)
3. Ojala, T., et al.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. PAMI 24(7), 971–987 (2002)
4. Dalal, N., Triggs, B.: Histograms of Oriented Gradients forHuman Detection. In: CVPR 2005, vol. 1, pp. 886–893. IEEE (2005)
5. Pang, Y., et al.: Robust Cohog Feature Extraction in Human-Centered Image/Video Management System. Systems, Man, and Cybernetics 42(2), 458–468 (2012)
6. Bengio, Y., Courville, A.C., Vincent, P.: Unsupervised feature learning and deep learning: A review and new perspectives. CoRR (2012)
7. Tiwari, P., et al.: Texture descriptors to distinguish radiation necrosis from recurrent brain tumors on multi-parametric mri. In: SPIE, pp. 90352B–90352B (2014)
8. Furey, T.S., et al.: Svm Classification and Validation of Cancer Tissue Samples using Microarray Expression Data. Bioinformatics 16(10), 906–914 (2000)
9. Wilcoxon, F., Wilcox, R.A.: Some Rapid Approximate Statistical Procedures. Lederle Laboratories (1964)