

# Liver Tissue Classification in Patients with Hepatocellular Carcinoma by Fusing Structured and Rotationally Invariant Context Representation

John Treilhard<sup>1</sup>(✉), Susanne Smolka<sup>3,4</sup>, Lawrence Staib<sup>1,2,3</sup>, Julius Chapiro<sup>3</sup>, MingDe Lin<sup>3,5</sup>, Georgy Shakirin<sup>6</sup>, and James S. Duncan<sup>1,2,3</sup>

<sup>1</sup> Department of Biomedical Engineering,  
Yale University, New Haven, CT 06520, USA  
j.treilhard@gmail.com

<sup>2</sup> Department of Electrical Engineering, Yale University, New Haven, CT 06520, USA

<sup>3</sup> Department of Radiology and Biomedical Imaging,  
Yale University, New Haven, CT 06520, USA

<sup>4</sup> Charité University Hospital, 10117 Berlin, Germany

<sup>5</sup> Philips Research North America, Cambridge, MA 02141, USA

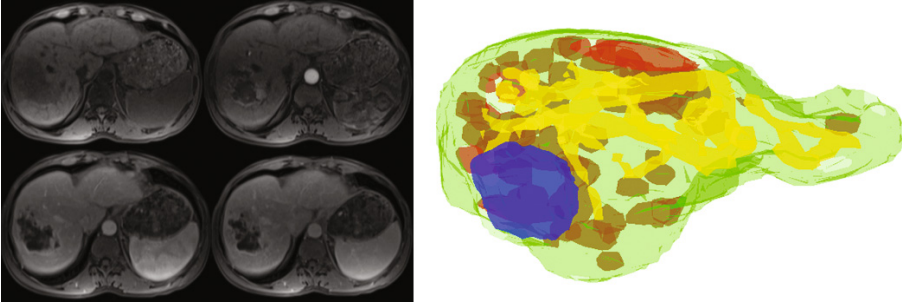
<sup>6</sup> Philips Research Aachen, 52074 Aachen, Germany

**Abstract.** This work addresses multi-class liver tissue classification from multi-parameter MRI in patients with hepatocellular carcinoma (HCC), and is among the first to do so. We propose a structured prediction framework to simultaneously classify parenchyma, blood vessels, viable tumor tissue, and necrosis, which overcomes limitations related to classifying these tissue classes individually and consecutively. A novel classification framework is introduced, based on the integration of multi-scale shape and appearance features to initiate the classification, which is iteratively refined by augmenting the feature space with both structured and rotationally invariant label context features. We study further the topic of rotationally invariant label context feature representations, and introduce a method for this purpose based on computing the energies of the spherical harmonic decompositions computed at different frequencies and radii. We test our method on full 3D multi-parameter MRI volumes from 47 patients with HCC and achieve promising results.

**Keywords:** Classification · Structured prediction · Rotationally invariant context features · Spherical harmonics · HCC · MRI

## 1 Introduction

Hepatocellular carcinoma (HCC) is the most common primary cancer of the liver, its worldwide incidence is increasing, and it's the second most common cause of cancer-related death [4]. Multi-parameter MRI is extremely useful for detecting and surveilling HCC, and in this work, we consider the problem of



**Fig. 1.** *Left:* DCE-MRI sequence, clockwise from top left: pre-contrast phase, arterial phase, portal venous phase, delayed phase. *Right:* 3D view of liver tissue classes from gold standard radiologist segmentation: green - whole liver, red - viable tumor, blue - necrosis, yellow - vasculature.

automated classification of pathological and functional liver tissue from multi-parameter MRI in patients with HCC. We present a structured prediction framework for liver tissue classification. The contributions of our work are: (1) to our knowledge, this is among the first works on fully automatic multi-class liver tissue classification from multi-parameter MRI in patients with HCC, despite this modality being a gold-standard for HCC diagnosis and surveillance, (2) the method is based on a novel integration of multi-scale shape (Frangi vesselness) and appearance (mean, median, standard deviation, gradient) features to implement a fine-grained, clinically relevant, and simultaneous delineation of four tissue classes: viable tumor tissue, necrotic tissue, vasculature, and parenchyma, (3) the introduction of a novel method for rotationally-invariant context feature representation via the energies of the spherical harmonic context representation computed at different frequencies and radii, and (4) the segmentation is refined by a novel iterative classification strategy which fuses structured and rotationally-invariant semantic context representation. Our method is applied to a clinical dataset consisting of full 3D multi-parameter MRI volumes from 47 patients with HCC (Fig. 1).

## 2 Methods

### 2.1 Shape and Appearance Features

Variations in patient physiology and HCC’s variable radiologic appearance necessitates the integration of multi-scale shape and appearance features in order to discriminate between tissue classes. Multi-scale shape and appearance features were extracted by computing Haralick textural features (contrast, energy, homogeneity, and correlation) and applying multi-scale median, mean, and standard deviation filters. The distance to the surface of the liver was associated as a feature to each voxel. Frangi filtering was applied to each temporal kinetic image to generate an image in which noise and background have been suppressed, and

tubular structures (namely vessels) are enhanced. To account for variable vessel width, the filter response is computed at multiple scales and the maximum vesselness found at any scale is taken as the final result.

## 2.2 Label Context Features

**Rotation-invariant context features.** Our method incorporates the “auto-context” structured prediction framework [6], which takes the approach of training a cascade of classifiers, and using the probabilistic label predictions output by the  $(n - 1)^{st}$  classifier as additional *label context* feature inputs for the  $n^{th}$  classifier. Adopting a structured classification framework can yield significantly improved performance; however, one critical question concerns the nature of the representation of label context features. One typical approach is to sparsely extract label context probabilities for each tissue class from a neighborhood of each voxel, and concatenate these orderly into a 1D vector [6]. This approach has a key advantage – namely, that it preserves the structural and semantic relationships between the extracted pixels: the process of orderly concatenation ensures that any two given positions in the label context feature representation share a consistent spatial orientation relationship and represent probabilities of consistent tissue classes. For these reasons, we call context features represented in this way *structured* context features.

On the other hand, structured context features are highly dependent on the arbitrary orientation of the label context patch from which they are drawn. However in the present setting, patches within the liver do not have a canonical frame of reference, and therefore, alternative *rotationally invariant* label context feature representations potentially have some advantages over structured context feature representations in terms of improving the generalization performance of a structured classifier. More specifically, an orientation-invariant representation  $T$  is a mapping  $T : \mathbb{R}^{d^3} \rightarrow \mathbb{R}^{d'}$  which maps a  $d \times d \times d$  patch into a 1D representation, such that

$$T(\mathbf{P}) = T(R(\mathbf{P})), \quad (1)$$

where  $R : \mathbb{R}^{d^3} \rightarrow \mathbb{R}^{d^3}$  denotes the operator rotating the patch in 3D. The patch  $\mathbf{P}$  in this setting refers to a patch of label probabilities corresponding to a particular tissue class. It is clear that the mapping  $T$  which simply reorders the elements of a 3D patch into a 1D vector does not satisfy Eq. (1).

Previous work on orientation invariant representation of contextual features [2] involves the decomposition of the 3D image patch into concentric spherical shells, and the computation of the distribution (via a histogram) of label context features on each of these shells, and finally the concatenation these histograms<sup>1</sup>. More precisely, define the spherical shells  $S$  about voxel  $i$  as:

$$S_r(i) = \{k \in \Omega : \|k - i\| = r\}. \quad (2)$$

<sup>1</sup> [2] actually proposes the use of “spin-context”, which involves computing “soft” histograms, but the principle is the same.

Furthermore, denote the histogram operator with  $m$  bins by  $H_m : \mathbb{R}^* \rightarrow \mathbb{N}^m$ , (where  $\mathbb{R}^*$  denotes a list of numbers of arbitrary length). It is clear that the mapping:

$$T(\mathbf{P}) = [H_m(\mathbf{P}|_{S_r})]_{r=1}^R \quad (3)$$

satisfies the rotation invariance property given in (1), where  $\mathbf{P}|_{S_r}$  denotes the restriction of  $\mathbf{P}$  to  $S_r$ . We note however that despite their orientation invariance, using the distributions of contextual features on concentric spherical shells involves the complete loss of structural information about the label context. Therefore, inspired by work in the field of shape descriptors [1], we propose an alternative representation of label context features by decomposing the label context features on spherical shells into their spherical harmonic representation, and using the norm of the energy of each frequency as features. We emphasize that the application of spherical harmonic decomposition for structure-preserving 3D rotation invariant label context representation has not appeared before in the literature. Following the notation in [1], let  $f(\theta, \phi)$  denote a function defined on the surface of a sphere; then  $f(\theta, \phi)$  admits a rotation invariant representation as:

$$SH(f) = \{\|f_l(\theta, \phi)\|_2\}_{l=0}^\infty, f_l(\theta, \phi) = \sum_{m=-l}^l a_{lm} Y_l^m(\theta, \phi) \quad (4)$$

where the  $\{Y_l^m(\theta, \phi)\}$  are a set of orthogonal basis functions for the space of  $L^2$  functions defined on the surface of a sphere, called spherical harmonics. In practice we limit the bandwidth of the representation to some small  $L > 0$ . Returning to our original motivation, we define the operator  $T$  acting on the label context patch  $\mathbf{P}$  by:

$$T(\mathbf{P}) = [SH(\mathbf{P}|_{S_r})]_{r=1}^R. \quad (5)$$

We note that this alternative representation has some advantages relative the representation in (3), particularly the preservation of structural properties regarding the distribution of context features, but their relative performance will ultimately be determined by the nature of the specific problem, data, and classification method under consideration. Note also that our notation suppresses that these rotationally invariant representations are computed from the probability maps generated for each tissue class separately, and subsequently concatenated - i.e. if  $\mathbf{P}_u$  denotes the patch of label probabilities corresponding to tissue class  $u$  (where there are  $U$  tissue classes total), then

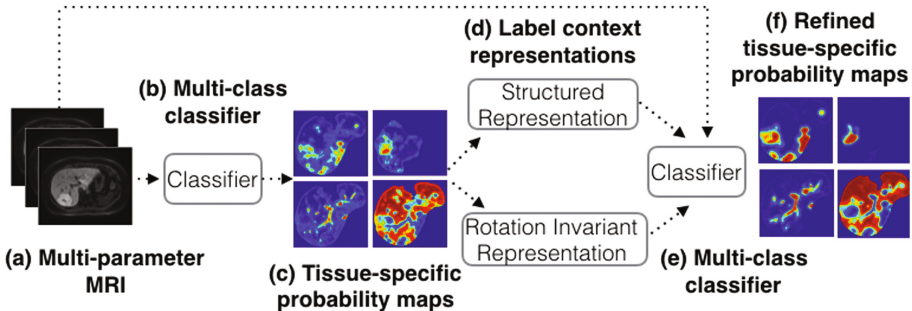
$$T(\mathbf{P}) = [SH(\mathbf{P}_u|_{S_r})]_{r=1..R, u=1..U}. \quad (6)$$

**Structured and rotationally-invariant context integration.** We note however that both rotation-invariant representations (3) and (5) share common disadvantages relative structured context feature representation: in particular, both fail to capture information *between* shells, and in addition they are computed separately for each tissue class and do not represent the relationship between probabilities of specific voxels belonging to different tissue classes. Since

both structured and rotationally-invariant context representations are endowed with advantages and disadvantages, we propose to integrate them into a single representation, which could exploit their respective strengths and minimize their weaknesses. Therefore we propose of a unified structured and rotationally-invariant context representation:

$$T(\mathbf{P}) = [T_1(\mathbf{P}), T_2(\mathbf{P})], \quad (7)$$

where  $T_1(\mathbf{P})$  is given by the orderly concatenation of the elements of  $\mathbf{P}$ , and  $T_2(\mathbf{P})$  is given by (3) or (5) (Fig. 2).

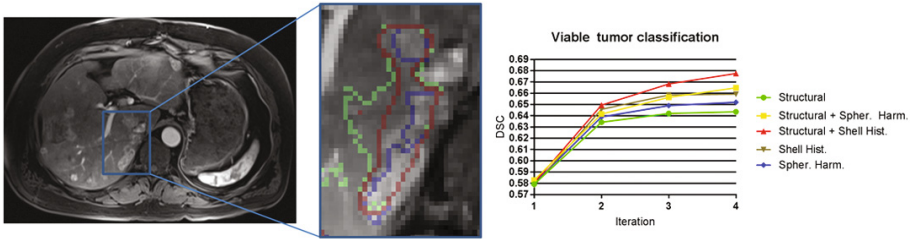


**Fig. 2.** Representation of the unique iterative classification method: (a) multi-parameter MRI input images, (b) multi-class classifier, in our case a random forest, (c) tissue-specific probability maps output by the classifier, (d) structured and rotationally-invariant label context representations are computed and used as input to another classifier, in addition to the original multi-parameter MRI imaging, (f) tissue-specific probability maps output by the classifier, and the entire process is iterated.

**Classification.** We introduce a structured classification framework where multi-scale shape and appearance features are extracted at each voxel from the multi-parameter MRI and used to train a random forest classifier with bagging and random node optimization. A cascade of random forests are trained, where the features used to train classifier  $n$  being augmented by the label-context of each voxel, inferred from the output of classifier  $(n - 1)$ , with the unique feature that these contextual features are represented both in *structured* and *rotationally-invariant* forms - see (7). We train a cascade of random forest classifiers, where at classification stage  $t$ , the training data is given by:

$$S_t = \{(y_{ji}, (X_j(N_i), T(\mathbf{P}_j^{(t-1)}(i))))\}, j = 1..m, i = 1..n, \quad (8)$$

where  $m$  denotes the number of images in the training set,  $n$  denotes the number of voxels in each training image,  $y_{ji}$  denotes the (tissue class) label of voxel  $i$  in image  $j$ ,  $X_j(N_i)$  denotes the concatenation of multi-scale shape and appearance features (described in Sect. 2.1) in a neighborhood  $N_i$  of voxel  $i$ , the mapping  $T$  is given by (5), and  $\mathbf{P}_j^{(t-1)}(i)$  denotes the patch of label probabilities surrounding voxel  $i$  in image  $j$  at classification iteration  $(t - 1)$  (Fig. 3).



**Fig. 3.** *Left:* Viable tumor tissue classification, where blue represents the gold-standard segmentation, green represents the classification based on structural context features alone, and red represents the segmentation achieved by our method integrating structured and rotationally invariant context features. *Right:* Dice similarity coefficient (DSC) for viable tumor tissue classification for each classification iteration (“Spher. Harm” = spherical harmonic contextual feature encoding, “Shell Hist.” = histograms of contextual features on shells).

## 3 Experiments

### 3.1 Data

We analyzed T1-w dynamic contrast enhanced (DCE) and T2-w MRI data sets from 47 patients with HCC. The DCE sequences consisted of 3 timepoints: the pre-contrast phase, the arterial phase (20 s after the injection of Gadolinium contrast agent), and the portal venous phase (70 s after the injection of Gadolinium contrast agent). Parenchyma, viable tumor, necrosis, and vasculature were segmented on the arterial phase image by a medical student and confirmed by an attending radiologist.

### 3.2 Numerical Results and Discussion

We evaluated our method on the dataset described in Sect. 3.1.

Regarding the results presented in Table 1, we observe first that classification based solely on bias field corrected pre-contrast, arterial, and portal venous phase intensities (row 1 of Table 1) provides a far inferior result to the classification computed using higher order features, for example multi-scale vessel filters and Haralick textural features. Furthermore, we observe a performance improvement by virtue of integrating structural and rotationally-invariant context feature representations; for the viable tumor tissue and necrosis tissue classes, the classification was significantly better (parametric paired two-tailed t-test,  $p < 0.05$ ) compared to using multi-scale shape and appearance features alone. We conclude that although both shell histogram and spherical harmonic context features add significant value following integration with structured context features, neither significantly outperformed the other in this problem. However, given a different problem setting and data, one might potentially outperform the other, and this deserves empirical evaluation with further data.

**Table 1.** Dice similarity coefficient evaluation of classification results.

Features	Viable tumor	Necrosis	Vasculature
(1) Intensity only	0.506	0.342	0.491
(2) Multi-scale shape + appearance	0.580	0.449	0.547
(3) = (2) + Structural Context	0.643	0.527	0.544
(4) = (2) + Shell Histogram Context	0.659	0.494	0.555
(5) = (2) + Spherical Harmonic Context	0.652	0.514	0.549
(6) = (3) + Spherical Harmonic Context	0.661	<b>0.544</b>	0.552
(7) = (3) + Shell Histogram Context	<b>0.678</b>	0.530	<b>0.557</b>

Regarding the overall performance, we note that the inter-reader variability for the Dice similarity coefficient for segmenting whole HCC lesions is reported as 0.7 [5]. In the context of this paper, we are assessing the accuracy of a much more difficult task, which is the segmentation of viable and necrotic tumor tissue segmentations, which could reasonably have a yet higher inter-reader variability. The inter-reader variability could be regarded as an upper-bound on how well an automated method could perform, and from this perspective, the results indicate the method is effective in the setting of the given problem. The MICCAI Multimodal Brain Tumor Image Segmentation Benchmark (BRATS - [3]) presents results related to the multi-class classification of brain tissue in patients with gliomas - they report slightly higher Dice similarity coefficient scores for the segmentations corresponding to the top performing methods, but we emphasize that we are considering different tissue classes in a different organ, and so the results are not directly comparable. Moreover, we are (to our knowledge) among the first to present multi-class tissue classification from MRI for a relatively large cohort of HCC patients.

## 4 Conclusion and Future Work

In this work we presented a novel method for addressing a previously unconsidered clinical problem: multi-class tissue classification from multi-parameter MRI in patients with HCC. The method integrated multi-scale shape and appearance features, along with structured and rotationally-invariant label context feature representations in a structured prediction framework. We also introduce a novel method for the rotationally-invariant encoding of context features and demonstrated that it delivered a significant performance improvement upon integration with structured context features, and our study considered full 3D clinical volumes in a substantial patient cohort. Furthermore, the methods we introduce here are independent of the specifics of our implementation, and can be translated to many other problems in the structured prediction setting.

**Acknowledgements.** This research was supported in part by NIH grant R01CA206180.

## References

1. Kazhdan, M., Funkhouser, T., Rusinkiewicz, S.: Rotation invariant spherical harmonic representation of 3D shape descriptors. In: Eurographics Symposium on Geometry Processing, pp. 156–165 (2003)
2. McKenna, S.J., Telmo, A., Akbar, S., Jordan, L., Thompson, A.: Immunohistochemical analysis of breast tissue microarray images using contextual classifiers. *J. Path. Inform.* **4**(2), 13 (2013)
3. Menze, B.H., Jakab, A., Bauer, S., et al.: The multimodal brain tumor segmentation benchmark (BRATS). *IEEE Trans. Med. Imag.* **34**, 1993–2024 (2015)
4. Park, J.-W., Chen, M., Colombo, M., et al.: Global patterns of hepatocellular carcinoma management from diagnosis to death: the BRIDGE Study. *Liver Int.* **35**(9), 2155–2166 (2015)
5. Tacher, V., Lin, M., Chao, M., et al.: Semiautomatic volumetric tumor segmentation for hepatocellular carcinoma. *Acad. Radiol.* **20**, 446–452 (2013)
6. Tu, Z., Bai, X.: Auto-context and its application to high-level vision tasks and 3D brain image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 1744–1757 (2010)