OPEN ACCESS

UNIVERSITY OF DUNDEE

University of Dundee

TUMOR LOCALIZATION IN TISSUE MICROARRAYS USING ROTATION INVARIANT SUPERPIXEL PYRAMIDS

Shazia Akbar*, Lee Jordan†, Alastair M. Thompson‡, Stephen J. McKenna*

*School of Computing, University of Dundee, Dundee, UK
†Pathology, Ninewells Hospital, Dundee, UK
‡Department of Surgical Oncology, MD Anderson Cancer Center, Houston, United States

ABSTRACT

Tumor localization is an important component of histopathology image analysis; it has yet to be reliably automated for breast cancer histopathology. This paper investigates the use of superpixel classification to localize tumor regions. A superpixel representation retains information about visual structures such as cellular compartments, connective tissue, lumen and fatty tissue without having to commit to semantic segmentation at this level. In order to localize tumor in large images, a rotation invariant spatial pyramid representation is proposed using bags-of-superpixels. The method is evaluated on expert-annotated oestrogen-receptor stained TMA spots and compared to other superpixel classification techniques. Results demonstrate that it performs favorably.

Index Terms— tumor classification, superpixels, spatial bag-of-words, rotation invariant spatial pyramid

1. INTRODUCTION

Tissue microarray (TMA) spots such as shown in Fig. 1(a) enable relatively high-throughput inspection of various cancers for multiple molecular markers [1]. They are used for validation of diagnostic markers in annotated clinical samples, testing new antibodies and probes, or determining optimal staining conditions. The growth of tissue banks has lead to an increased workload for pathology experts, exceeding the manual skills available. Furthermore, expert pathology review suffers from inter- and intra-observer variations.

Tumor localization is an important component of histopathology image analysis such as immunohistochemical (IHC) scoring and has not been reliably automated in breast TMAs. It is challenging because breast tissue has complex visual structure; currently no detailed guidelines exist on how to visually identify tumor regions.

One approach to automate tumor analysis is to first explicitly detect and segment cells or subcellular components and then build a feature representation based on these detected objects [3, 4]. However, in the images considered here, cellular compartment boundaries other than those of nuclei are not often apparent. Nuclei detection is non-trivial because of

(a)   (b)
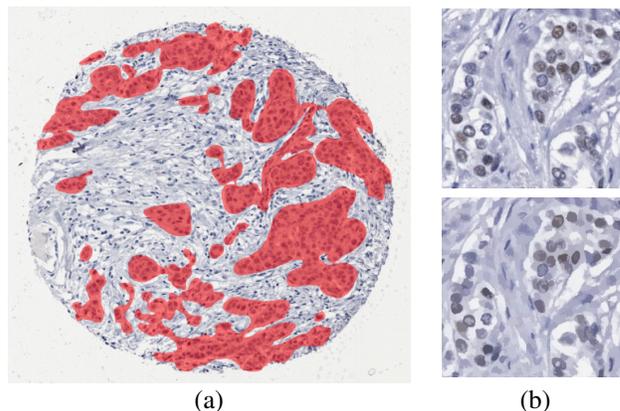
Fig. 1. (a) A tissue microarray spot with annotated tumor regions. (b) An image patch (top) and a SLIC [2] superpixel image with each superpixel rendered as the average RGB value of the pixels contained in it (bottom).

the different cell types, varied appearance, and often heavily clustered nuclei. An alternative approach is to treat the problem as one of texture segmentation, without explicit detection of objects [5,6]. However, statistical texture descriptors computed directly from the pixels may fail to capture structural aspects of tissue.

We investigate features based on superpixels to localize tumor regions. Fig. 1(b) shows a superpixel representation. The intuition is that a superpixel-based representation can retain information about visual structures such as cellular compartments, connective tissue, lumen and fatty tissue without having to commit to semantic segmentation at this level. Superpixels have previously been applied in histopathology for cell nuclei segmentation [7] and tumor classification [8]. Here, we use features from superpixels in a Bag-of-Words (BoW) framework, called Bags-of-Superpixels (BoS).

We make the following contributions in this paper. (i) Incorporation of spatial information in BoS (iii) extension of the spatial pyramid [9] to build a Rotation Invariant Superpixel Pyramid (RISP) (iv) an adaptation of the autocorrelogram incorporating superpixels. Tumor localization is formulated as

To appear in the Proceedings of 2015 IEEE International Symposium on Biomedical Imaging (ISBI 2015).

Appearance	Mean RGB, greyscale mean and variance, 13 Haralick texture features ($f_1 - f_{13}$).
Geometry	Compactness, eccentricity, area, perimeter.
Neighbors	Number of immediate superpixel neighbors, variance of superpixel perimeter shared with neighboring superpixels.

Table 1. Features extracted from each superpixel.

a superpixel classification problem and is evaluated on expert-annotated oestrogen-receptor stained TMA spots.

2. METHODS

2.1. Superpixel Features

A superpixel image representation is an over-segmentation in which similar pixels are grouped into perceptually consistent units, regularized to be compact and of similar size. We use SLIC [2] to construct superpixels as it is computationally efficient and retains tissue structure. The number of superpixels is assigned such that the area of a single superpixel rarely exceeds the area of a cell nucleus. Many cell nuclei are assigned two or more superpixels. An example SLIC representation of a breast TMA patch is shown in Fig. 1(b) in which the average superpixel area is 223 pixels ($\sigma = 33$). An expert pathologist considered such superpixel images to retain the important structure of the tissue; tumor regions are still identifiable as are structures such as ducts and lobules.

For each superpixel, we obtain a set of features to describe its geometric and photometric properties as well as its adjacency relationship to neighboring superpixels (Table 1). These features are normalized and concatenated to form a descriptor per superpixel.

2.2. Bag-of-Superpixels (BoS)

To perform superpixel classification, one approach is to use the features directly. However this produces poor results due to lack of context. To incorporate context, we use superpixel features in the BoW framework.

Extracted superpixel descriptors are quantized using a *K*-means dictionary to give visual words. A circular window with radius r is positioned at the centre point of the superpixel to be classified. Visual words of superpixels within the circular window are histogrammed, resulting in a Bag-of-Superpixels (BoS). To provide complementary local features, the superpixel descriptor for the centre superpixel is concatenated with BoS.

2.3. Spatial Bag-of-Superpixels (S-BoS)

Whilst BoS is a simple yet powerful representation, its main drawback is lack of spatial information. Here, we model the

Input: Image I, radius r, no. of rings Q, codebook C, no. of superpixels M ;
Run SLIC on I to generate superpixels, $\mathbf{s} = [s_1...s_M]$;
Extract superpixel features, $\mathbf{F} = \{\mathbf{f}_1...\mathbf{f}_M\}$;
for *each superpixel, s_i, in* $\mathbf{s}$ **do**
    Identify superpixels $\mathbf{t}$ within circular window with radius, r, centered at s_i;
    Initialize spatial BoS histogram, H_i;
    for *each superpixel, t_j, in* $\mathbf{t}$ **do**
        Lookup codeword, v_j for $\mathbf{f}_{t_j}$ in C;
        Compute $d = ||c(t_j) - c(s_i)||$, c returns the centre point of a superpixel;
        Increment $H_i(v_j, \lfloor \frac{Qd}{r} \rfloor)$;
    end
    Normalize H_i;
end
Algorithm 1: Spatial Bag-of-Superpixels (S-BoS)

spatial distribution of each visual word in a spatial Bag-of-Superpixel (S-BoS) histogram. Spatial information is captured in the form of equally spaced rings within the circular window in BoS (Algorithm 1). This approach has some similarities to [8], however we analyze neighboring superpixels from their visual words instead of RGB values. The resulting histogram captures the BoS representation within Q rings.

2.4. Rotation Invariant Superpixel Pyramid

Spatial pyramids partition an image repeatedly to compute a BoW histogram per cell or sub-region. In [9], sub-regions consist of square grids and the representation is not rotation invariant. Instead we use a Rotation Invariant Superpixel Pyramid (RISP) in which S-BoS histograms are computed per level (Fig. 2). The number of rings grows exponentially with p in each level. By utilizing circular rings, histograms computed from each ring are rotationally invariant. Previous work [10] has used a technique similar to this for cell classification. However, here we use superpixel visual words and describe a generalisable multi-level pyramid applicable to other computer vision tasks.

S-BoS histograms are concatenated for each level to form a multiscale representation, RISP. We evaluate RISP and the S-BoS histogram, varying Q.

2.5. Superpixel autocorrelogram

In previous work on scene categorization, spatial information between interest regions has been modeled in the form of a codebook correlogram [11]. To improve computational costs by reducing the dimensionality of the correlogram, Huang *et al* [12] proposed the color autocorrelogram. We propose a superpixel autocorrelogram which captures spatial information between pairs of superpixels. Similarly to [12], only identi-
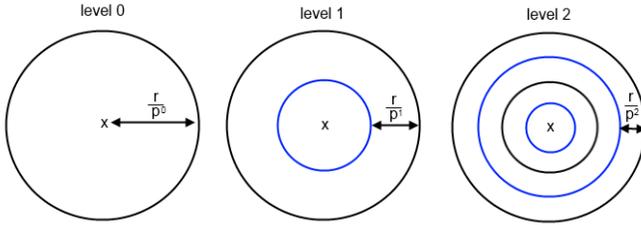
Fig. 2. Levels 0, 1 and 2 of RISP. The BoW representation is as level 0 after which partitions are applied iteratively according to p. In the above, $p = 2$. Therefore for each ring in level l, two more are created in $l + 1$.

Fig. 3. Precision-recall curves for RISP, BoS, S-BoS, superpixel features (SF), method as described in [8] (Gorelick) and the superpixel autocorrelogram (Corr) with 200 codewords.

		Number of Codewords			
		25	**50**	**100**	**200**
BoS		0.623	0.644	0.670	0.676
S-BoS	$Q = 2$	0.638	0.657	0.680	0.688
	$Q = 4$	**0.656**	0.672	0.671	0.687
RISP	2 levels $\{0, 1\}$	**0.656**	**0.676**	0.685	0.690
	3 levels $\{0, 1, 2\}$	0.645	**0.676**	**0.686**	**0.692**

Table 2. F1 measures for BoS, S-BoS and RISP.

cal pairs of codewords are counted. However we opt to use circular windows instead of a regular grid to retain rotation invariance. The distance distribution between superpixel centre points is captured resulting in a two-dimensional autocorrelogram. In reported experiments, 5 spatial bins were used.

3. EVALUATION

A data set of 32 TMA tumor spots subjected to nuclear staining for estrogen receptor (ER) was used. Spots were 0.6mm (3000 pixels) in diameter. Tumor regions were annotated by a highly experienced pathologist using a Bamboo tablet and stylus at 40x objective (Fig. 4).

Tumor localization was evaluated using 8-fold cross-validation. A balanced training set was used with identical number of positive and negative samples. A linear support vector machine (SVM) classifier was implemented in the LI-BLINEAR framework. Grid search was performed to obtain an optimal cost parameter. 50,000 SLIC superpixels were extracted from each TMA spot image with the compactness parameter set to 5. For a 3600x3600 pixel image, 50,000 SLIC superpixels can be computed in 29 seconds on an Intel i5-2410M 2.3GHz processor. A circular window with a radius $r = 100$ pixels was used. TMA spots were segmented manually to reduce background interference.

Figure 3 shows precision-recall curves for superpixel features with no context (SF), Bags-of-Superpixels (BoS), spatial Bags-of-Superpixels (S-BoS), Rotation Invariant Superpixel Pyramids (RISP), superpixel autocorrelograms (Corr) and an implementation of a method described by Gorelick *et al* [8] which encapsulates RGB values in rings (Gorelick). Results are shown for 2 and 4 S-BoS rings and a 3-level RISP. Features from the central superpixel were concatenated to each reported method to complement contextual information extracted from the circular window. There is a clear improvement in classification performance between BoS and S-BoS with 2 rings, showing the benefits of incorporating spatial information. 3-level RISP which incorporates 7 (1+2+4) rings also shows a small improvement compared to S-BoS; how-

ever increasing the number of rings in S-BoS results in minor improvement. Both RISP and S-BoS show significant improvement compared to the superpixel autocorrelogram and the method described in [8].

Table 2 shows F1 measures for RISP for various numbers of codewords. As the number of codewords increases and as more levels are incorporated in RISP, the accuracy continues to increase; differences between 100 and 200 codewords are marginal. Compared to BoS (i.e. level 0), RISP shows improvement for all dictionary sizes. For comparison, our previous method [13] using pixel-level features including differential invariants up to 2nd order and intensity spin features, results in an F1 measure of 0.569 with a linear SVM. Compared to 3-level RISP this is a decrease of 0.123. Examples of results achieved with RISP are shown in Fig. 4.

A closer look at the difference images in Fig. 4 suggests most disagreements occur near the annotation boundary. To test this theory, a distance transform was applied to binary manual annotations. Tumor probabilities obtained using RISP were thresholded at 0.5. The distribution of distances to the closest boundary for pixels in disagreement with the pathologist's annotation is shown in Fig. 5. A large proportion of disagreements lies within a short distance of nearby tumor regions; 41% of disagreements are within 20 pixels (approximately one cell diameter). This confirms most disagreements between the pathologist's and automated annotations occur around the tumor boundary.

Image   Annotation   RISP   Difference
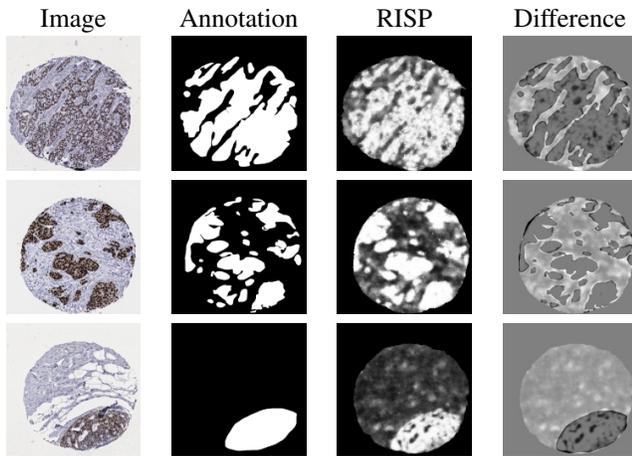
Fig. 4. TMA images are shown on the left alongside manual annotations. The third column shows tumor probabilities (3-level RISP, 200 codewords). The last column contains difference images where bright pixels are false positive and dark pixels are false negative errors.
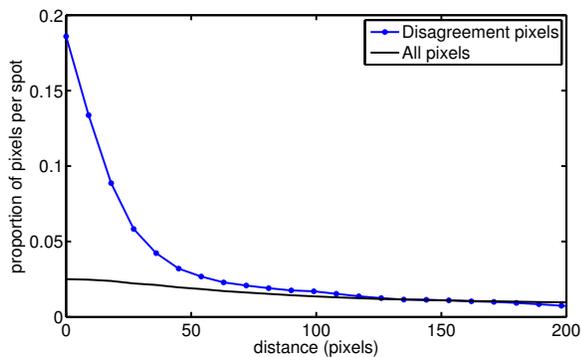
Fig. 5. Distribution of distances to the closest boundary for pixels within disagreement regions. The black line shows the distribution of distances for all pixels.

4. CONCLUSION

Superpixel classification algorithms were proposed for the problem of tumor localization in breast tissue. The bag-of-words framework was adapted to incorporate essential spatial information. Results show spatial information in the form of circular rings improves classification performance compared to Bags-of-Superpixels (BoS). When compared to Gorelick [8], the proposed Rotational Invariant Superpixel Pyramid (RISP) performs favourably.

Most disagreements between the pathologist and automated segmentation occur near the tumor boundary. It is unlikely that hand-drawn manual annotations are accurate to the pixel-level, therefore future work will consider how to deal with annotation uncertainty.

5. REFERENCES

[1] J. Kononen, L. Bubendorf, A. Kallionimeni, et al., "Tissue microarrays for high-throughput molecular profiling of tumor specimens," *Nature Medicine*, vol. 4, no. 7, pp. 844–847, 1998.

[2] R. Achanta, A. Shaji, K. Smith, et al., "SLIC superpixels compared to state-of-the-art superpixel methods," *PAMI*, vol. 34, pp. 2274–2282, 2012.

[3] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *TBME*, vol. 57, no. 4, pp. 841–852, 2010.

[4] M.N. Gurcan, T. Pan, H. Shimada, and J. Saltz, "Image analysis for neuroblastoma classification: Segmentation of cell nuclei," in *EMBS*, 2006, pp. 4844–4847.

[5] C. Wang, "Robust automated tumour segmentation on histological and immunohistochemical tissue images," *PLoS One*, vol. 6, 2011.

[6] A. Tabesh and M. Teverovskiy, "Tumor classification in histological images of prostate using color texture," in *ACSSC*, 2006, pp. 841–845.

[7] A.H. Beck, A.R. Sangoi, S.L. Leung, et al., "Systematic analysis of breast cancer morphology uncovers stromal features associated with survival," *Science Translational Medicine*, vol. 3, 2011.

[8] L. Gorelick, O. Veksler, M. Gaed, et al., "Prostate histopathology: Learning tissue component histograms for cancer detection and classification," *TMI*, vol. 32, pp. 1804–1818, 2013.

[9] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *CVPR*, 2006, vol. 2, pp. 2169–2178.

[10] A. Wiliem, C. Sanderson, Y. Wong, et al., "Automatic classication of human epithelial type 2 cell indirect immunouorescence images using cell pyramid matching," *Pattern Recognition*, 2013.

[11] Y. Zheng, H. Lu, C. Jin, and X. Xue, "Incorporating spatial correlogram into bag-of-features model for scene categorization," *ACCV 2009*, pp. 333–342, 2010.

[12] J. Huang, S.R. Kumar, M. Mitra, et al., "Image indexing using color correlograms," in *CVPR*, 1997, pp. 762–768.

[13] S. J. McKenna, T. Amaral, S. Akbar, L. Jordan, and A. Thompson, "Immunohistochemical analysis of breast tissue microarray images using contextual classifiers," *JPI*, 2013.

To appear in the Proceedings of 2015 IEEE International Symposium on Biomedical Imaging (ISBI 2015).