Tissues Classification of the Cardiovascular System Using Texture Descriptors

Claudia Mazo[1(✉)], Enrique Alegre[2], Maria Trujillo[1], and Víctor González-Castro[2]

[1] Computer and Systems Engineering School, University of Valle, Cali, Colombia
claudia.mazo@correounivalle.edu.co
[2] Industrial and Informatics Engineering School, University of León, León, Spain

Abstract. In this paper, we present an approach to automatically classify tissues of the cardiovascular system using texture information. Additionally, this process makes possible to identify some cardiovascular organs, since some tissues belong to muscles associated to those, i.e. identifying the tissue makes possible to identify the organ. We have assessed rotation invariant Local Binary Patterns (LBPri) and Haralick features to describe the content of histological images. We also assessed Random Forest (RF) and Linear Discriminant Analysis (LDA) for the classification of these descriptors. The tissues were classified into four classes: (i) cardiac muscle of the heart, (ii) smooth muscle of the elastic artery, (iii) loose connective tissue, and (iv) smooth muscle of the large vein and the elastic artery. The experimental validation is conducted with a set of 2400 blocks of 100×100 pixels each. The classifier was assessed using a 10-fold cross-validation. The best AUCs (0.9875, 0.9994 and 0.9711 for the cardiac muscle of the heart, the smooth muscle of muscular artery, the smooth muscle of the large vein and the elastic artery classes, respectively) are achieved by LBPri and RF.

Keywords: Fundamental tissues · Histology images · Image processing · Organs of the cardiovascular system · Automatic classification

1 Introduction

The human body is composed of four basic types of tissues: epithelium, connective tissue, muscular tissue and nervous tissue, which vary in their composition and function. Currently, the recognition of healthy tissues and organs is carried out by histology experts and there are no automatic systems able to identify them. An automatic classification platform of histological images of the cardiovascular system would be helpful for educational purposes [1]. It would promote self-learning to on-campus students and also facilitate on-line learning to external or remote students [2] requiring low social and economic investment to train qualified professionals. Furthermore, it could be used for automatic labelling and sorting of large repositories of histology images that are available in hospitals or distributed

M. Valdés Hernández and V. González-Castro (Eds.): MIUA 2017, CCIS 723, pp. 123–132, 2017.
DOI: 10.1007/978-3-319-60964-5_11

through different storage devices of histologists. Thus, subjectivity, time, costs, difficulty and impracticality issues of the manual annotation of histology images may be solved by using automatic classification systems [3]. In this paper, we present an approach to automatically identify tissues and thereafter, indirectly, recognise organs of the cardiovascular system using rotation invariant Local Binary Pattern (LBPri) texture features and the Random Forest (RF) classifier.

Texture descriptors are used to describe the content of medical images [4] and histological images [5,6]. Some histological image studies focused on cell [7,8], tissue and organ recognition [9,10]. For instance, in [11] a statistical model for histological image classification is presented using texture features of Multichannel Gabor. The probability distribution of texture patterns of each category is approximated by a finite Gaussian mixture model in order to identify ten organs—adrenal gland, the heart, kidney, liver, lung, pancreas, spleen, testis, thyroid and uterus. The validation was performed using 778 histological images. This approach yields an accuracy between 44% and 93% which varies depending on the organ being identified. In particular, the approach yields an accuracy of 80% for the heart, which is the only organ of the cardiovascular system included in the study.

A similar proposal is found in [12], which aims to identify five organs of the gastrointestinal tract—oesophagus, stomach, small intestine, large intestine and anus—by means of a 2D stochastic method for the semantic analysis of the content of histological images. The proposed method is based on the classification of 64×64 pixels blocks. From each block a 25-dimensional feature vector is obtained by concatenating the total Gabor energy and the mean grey value. The model yields an accuracy between 59% and 82% depending on the organ being identified. However, this model has high computational complexity and the evaluation was conducted using a limited dataset of images, i.e. 40 histological images per region.

We have done preliminary advances in the segmentation and the classification of fundamental tissues [13]. In [14] coating epithelial tissue is recognised and classified into flat, cubic and cylindrical in histological images acquired using a 40x lens. We used the K-means algorithm, a sphericity measure and spatial information, obtaining a sensitivity of 0,85. In addition, in [10] loose connective and muscle tissues are recognised by clustering the Structure Tensor, the red and the green colour channels. Expert evaluations scored with an average of 4.85 the identification of loose connective tissue and an average of 4.82 the identification of the muscle tissue, out of 5. However, this approach does not differentiate among different muscle types.

The computer-aided recognition of fundamental tissues and organs presents several problems due to the hard boundary among fundamental tissues, the lack of a panoramic view of the entire sample and the low definition in image areas such as at the edges of the tissues and organs. Moreover, due to the importance and complexity of the appearances of fundamental tissues and organs, their automatic classification remains an open problem.

In this paper, we present an approach to automatically classify the fundamental tissues of the cardiovascular system. We used the LBPri [15] texture

descriptor to represent the image content. The RF [16] classifier is used to identify cardiac muscle of the heart, loose connective tissue (i.e., veins, arteries and the heart), smooth muscle of muscular artery, and smooth muscle of the large vein and the elastic artery. Additionally, we outperform our previous results [10,14] using more general and robust computer vision and machine learning methods, which allows us to recognise organs through the same process.

The rest of the paper is structured as follows: the problem statement is presented in Sect. 2. Section 3 explains the proposed approach to automatic classification of fundamental tissues and organs. We present and discuss the dataset, the experiments and the results in Sect. 4. Finally, in Sect. 5 some conclusions are presented.

2 Problem Statement

Different tissues have different roles and functions. For instance, epithelium forms the coverings of surfaces of the body and it has different purposes, e.g. protection, adsorption, excretion, secretion, filtration, and sensory reception. Connective tissue protects, supports, and binds together parts of the body. Muscular tissue produces movement. Nervous tissue receives stimuli and conduct impulses.

Automatic recognition of tissues is based on visual features. Although tissues have unique patterns that make them identifiable from others, their appearances may be different within the same organ, due to changes in the capture zone, cut or sample. In addition, colour is not a reliable feature in histological images since it depends on the applied stain. Nevertheless, the observation of spatial patterns suggests that texture may be a relevant characteristic for tissue recognition. Figure 1 illustrates tissue blocks of 100×100 pixels from different organs. It can

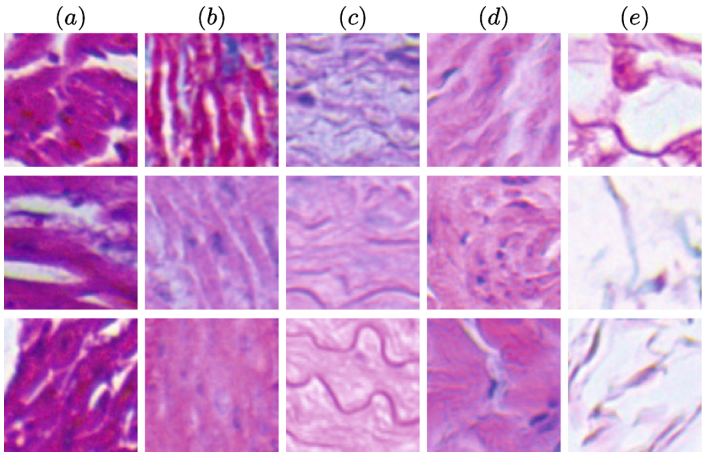

Fig. 1. Examples of blocks of the cardiac muscle tissue of the heart (a), smooth muscle tissue of the muscular artery (b), smooth muscle tissue of the elastic artery (c), smooth muscle tissue of the large vein (d) and loose connective tissue (e)

be observed that there are some intra-class similarity and inter-class difference among the blocks of the same tissue.

The classification of fundamental tissues is affected by the existing variability among blocks of the same tissue, which makes it a challenging and open problem.

3 Classification of Fundamental Tissues

The proposed approach is carried out in three steps: initially, an image is divided into blocks—an image capture protocol was designed to obtain a set of images with the same characteristics and to reduce errors in the automatic classification. Then, relevant texture features are extracted from the blocks and, finally, these descriptors are used to classify them. A general outline of our proposal is shown in Fig. 2. As it is depicted, first of all, blocks of 100 × 100 pixels are obtained from histological images. In this dataset, a block contains only one type of tissue. Afterwards, textural features are extracted from the blocks to obtain relevant information from them. After evaluating different texture descriptors, we propose to use LBPri to efficiently and robustly represent local micro-patterns. Thus, each block is represented by a vector of 36 features. Finally, Random Forest

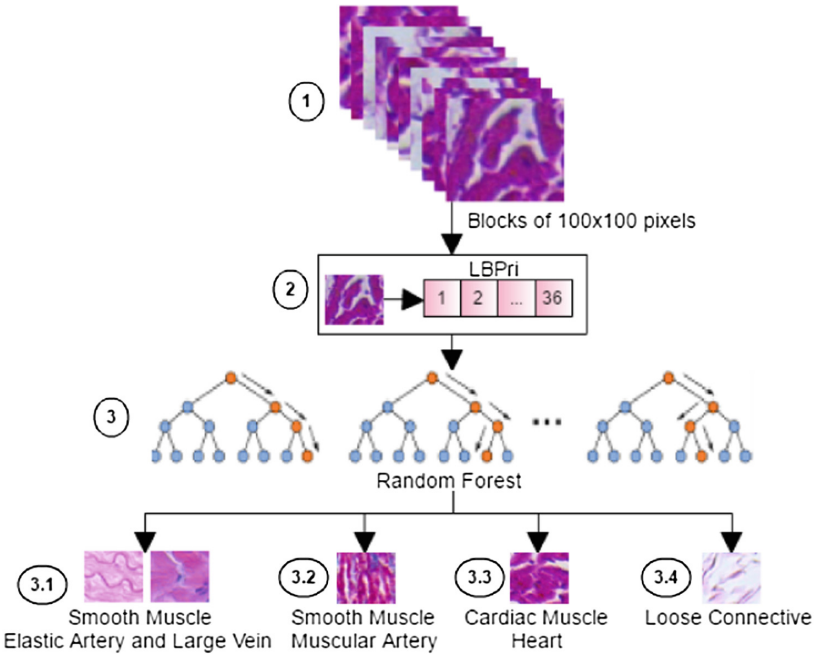

Fig. 2. Proposed approach for automatic classification of fundamental tissues associated with an organ: (1) image block of size 100 × 100 pixels. (2) LBPri histogram calculated per block. (3) Classification using RF. (3.1), (3.2), (3.3) and (3.4) blocks classified.

(RF) [16] is used as a classifier, after comparing its performance with that of the Linear Discriminant Analysis (LDA) [17]. RF is used to classify blocks as belonging to one of the four considered classes: (i) smooth muscle of the large vein and the elastic artery, (ii) smooth muscle of muscular artery, (iii) cardiac muscle of the heart and (iv) loose connective tissue (3.1, 3.2, 3.3 and 3.4 in Fig. 2, respectively). The smooth muscle of the large vein and the elastic artery were put in the same class taking into account similarities in the composition of the tunica media and in the muscle tissue, which main difference is the thickness of the layer. These similarities are reflected in texture features even when are recognised manually by an expert. The details of the complete process are presented in the rest of this section.

3.1 Tissues Description

The proposed approach uses texture descriptors to extract information about composition and characteristics of fundamental tissue morphology using a block-based recognition algorithm. These features are thereupon used for tissue recognition.

Histological images have particular characteristics depending on each fundamental tissue. Different texture descriptors were assessed for tissue identification and classification. LBP features and its variants have been commonly used as texture descriptors in recent years [18]. Haralick features have been also used in some biomedical applications [19,20]. In this paper, we assessed the rotation invariant Local Binary Patterns (LBPri) and some of the Haralick features (specifically contrast, homogeneity, angular second moment, correlation, entropy, and first and second correlation measures) The parameters *radius* and *number of neighbours* in the LBPri were set to 1 and 8, respectively – i.e. $LBP^{ri}_{1,8}$ (for simplicity, this term will be referred to as LBPri henceforth) – and Haralick features were calculated with distance equal to 3 using 4 directions – i.e. 0°, 45°, 90° and 135°. The feature vectors calculated with LBPri and Haralick have 36 and 48 values, respectively.

3.2 Tissues and Organs Classification

The proposed approach uses machine learning algorithms to recognise fundamental tissues and, in some cases, the organs are also identified. The classification is done for the following four classes: (i) cardiac muscle of the heart, (ii) loose connective tissue (i.e., vein, arteries and the heart), (iii) smooth muscle of the muscular artery, and (iv) smooth muscle of the large vein and the elastic artery. The RF using different parameters and the LDA were compared in order to select the classifier. The RF is an ensemble of decision trees, where the number of trees has a significant effect on the accuracy of the resulting model [16]. The LDA uses a linear combination of features that compute the directions, which represent the axes that maximise the separation of two or more classes [17].

4 Experiments and Results

In this section we analyze the results obtained using the descriptors evaluated with RF. Then, we used the best texture descriptor, according to the previous results, to compare performance with another classifier such as LDA. We use the Receiver Operating Characteristic (ROC) curves for each descriptor we use the area under such curves (AUC) as the quantitative measurement of the classifications. In the rest of the Section, "Muscular" represents the smooth muscle of the muscular artery, "Heart" represents the cardiac muscle of the heart, "Connective" represents the loose connective tissue (i.e. veins, arteries and the heart), and "Elastic-Vein" represents the smooth muscle of the elastic artery and large vein.

4.1 Experimental Setup

Tissue samples were stained with Hematoxylin and Eosin and Masson's trichrome using a laboratory protocol to control the process [21]. The image capture protocol was defined by taking into account some characteristics such as the microscope configuration, software configuration or sample manipulation in order to reduce errors in the automatic classification. We used a dataset of 600 blocks per class (i.e. 2400 blocks), which belong to images obtained from tissue samples of different organs and people and acquired using a $10x$ objective. The histological images were acquired with a Leica *DM750-M* microscope with a resolution of 2048×1536 pixels and stored in *PNG* format. We have made the dataset publicly available at http://biscar.univalle.edu.co/?page_id=1003. Algorithms were implemented in *C++* and *MATLAB*, using the *CImg* library in a *8-cores* CPU and *8*GB *RAM* computer. We used a 10-fold cross-validation in the classification experiments.

4.2 Texture Descriptors Assessment

We have assessed the descriptors LBPri and Haralick features (specifically contrast, angular second moment, homogeneity, correlation, entropy, first and second correlation measures) both on their own. We present the obtained results using RF classifier in Fig. 3, for each set of features per class (one vs all).

LBPri was the best descriptor with the AUCs (0.9986, 0.9952, 0.9453, 0.921) for smooth muscle of the muscular artery, cardiac muscle of the heart, loose connective tissue, and smooth muscle of the elastic artery and large vein, respectively. LBPri presents the highest True Positive Rates, indicating good classification results globally. The worst AUCs are achieved by Haralick features in all cases, which means LBPri is the most effective descriptor for classify blocks of histological images among both local pattern descriptors. This is because Haralick only considers four neighbourhood to distance equal to three while the LBPri descriptor considers eight neighbourhood with radius one, more local information. LBPri is simple, effective, and robust and is proving to be a powerful discriminator in many medical image classification problems.

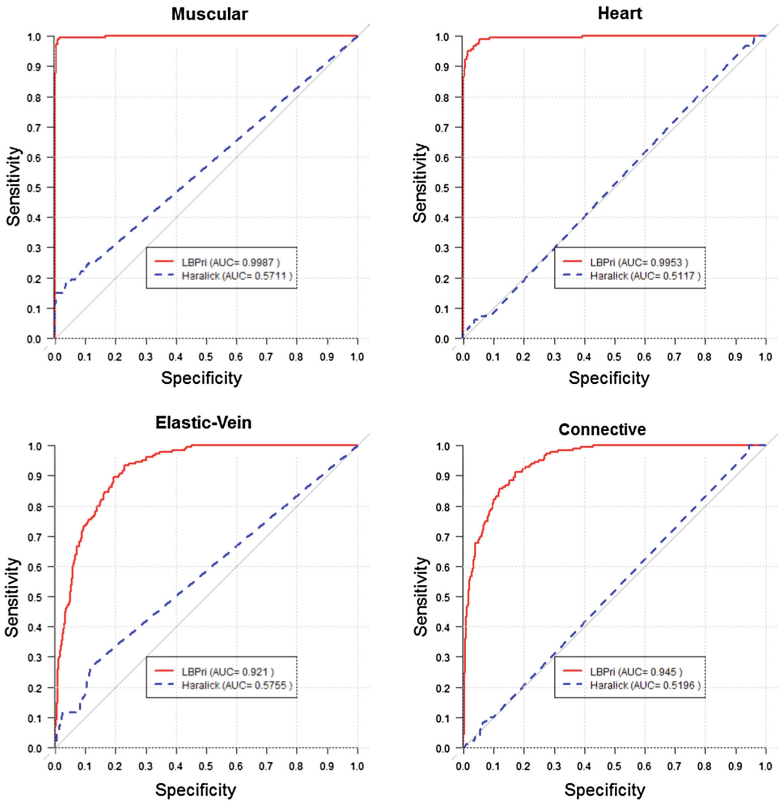

Fig. 3. ROC curves corresponding to the classification of each of the classes vs. the others using RF.

4.3 Evaluation of Different Classification Algorithms

We also compared the Random Forests (RF) and Linear Discriminant Analysis (LDA) using the texture descriptors that performed best in Sect. 4.2, i.e. LBPri. We assessed different parameters in the RF. Figure 4 depicts ROC curves and AUC values of testing performance comparisons in a multi-class classification.

As shown in Fig. 4, a direct comparison of the classification performances shows that the RF has the best classification results globally. The best results are achieved by RF for the cardiac muscle of the heart, the smooth muscle of muscular artery, the smooth muscle of the large vein and the elastic artery classes with AUCs of 0.9875, 0.9994 and 0.9711 respectively. LDA performed better in classifying loose connective tissue, achieving an AUC of 0.9897. Finally, we decided to use the RF with $Deep = 2$ and $MaxTrees = 100$ since they yielded the best performance.

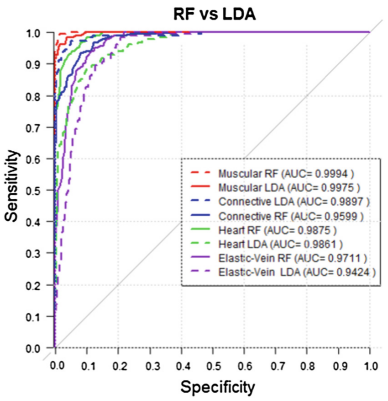

Fig. 4. Comparative evaluation using ROC curves classification of RF and LDA for each class.

5 Conclusions

In this paper, we presented a method that makes possible the recognition of the cardiac muscle of the heart, loose connective tissue (i.e. veins, arteries and the heart) and smooth muscle of the muscular artery, the large vein and the elastic artery with AUC greater than 0.90.

We have determined that, among those assessed, the best texture descriptor is the LBPri. In addition we have evaluated two different classifiers, being the best a RF with $Deep = 2$ and $MaxTrees = 100$ yielded AUC above 0.90.

Using the recognised tissues from a block, we were able to identify some organs. For instance, when cardiac muscle tissue is recognised, we know with no doubt that the specific image contains at least a partial view of tissue form the heart. In a similar way, the correct classification of smooth muscle tissue allows for the recognition of the presence in the image of muscular arteries, veins and large elastic arteries.

Taking into account these results, it is possible to recognise fundamental tissues from the cardiovascular system and also some of the organs associated to these tissues by extracting the appropriate texture feature with the selection of the proper classifier.

This is important for educational purposes: this method may be used in a system to promote self-learning for on-campus students and also facilitate on-line learning to external or remote students. As a result, it would make possible to train qualified professionals low social and economic investments.

In future works, we will extend this proposal by working in the following four lines: (i) to improve the classification method by exploring new classifiers such as Support Vector Machine (SVM), (ii) the classification of small blocks could be used to recognise every part of a whole histological image, (iii) to improve the recognition process by means of a histological ontology which would detect and

correct the wrong classifications of the blocks, and (iv) to add new cardiovascular tissue patterns with pathologies, and even from other organs.

Acknowledgements. This work has been supported by COLCIENCIAS and Asociación Universitaria Iberoamericana de Postgrado, AUIP. We thank Liliana Salazar, M.Sc., for providing insight and expertise that greatly assisted the research.

References

1. Izet, M.: E-learning as new method of medical education. Acta Informatica Medica **16**(2), 102–117 (2008). http://dx.doi.org/10.5455/aim.2008.16.102-117
2. Ruiz, J., Mintzer, M., Leipzig, R.: The impact of e-learning in medical education. Acad. Med. **81**(3), 207–212 (2006)
3. Hernadez, A.I., Porta, S.M., Miralles, M., Garca, B.F., Bolmar, F.: La cuanticacin de la variabilidad en las observaciones clnicas. Med. Clin. 424–429 (1990). http://www.ncbi.nlm.nih.gov/pubmed/2082114?dopt=Abstract
4. Nanni, L., Lumini, A., Brahnam, S.: Local binary patterns variants as texture descriptors for medical image analysis. Artif. Intell. Med. **49**(2), 117–125 (2010). doi:10.1016/j.artmed.2010.02.006
5. Herve, N., Servais, A., Thervet, E., Olivo-Marin, J.-C., Meas-Yedid, V.: Statistical color texture descriptors for histological images analysis. In: 2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, pp. 724–727 (2011). doi:10.1109/ISBI.2011.5872508
6. Ojansivu, V., Linder, N., Rahtu, E., Pietikinen, M., Lundin, M., Joen-Suu, H., Lundin, J.: Automated classification of breast cancer morphology in histopathological images. Diagn. Pathol. **8**(Suppl. 1), S29 (2013)
7. Mazo, C., Trujillo, M., Salazar, L.: An automatic segmentation approach of epithelial cells nuclei. In: Alvarez, L., Mejail, M., Gomez, L., Jacobo, J. (eds.) CIARP 2012. LNCS, vol. 7441, pp. 567–574. Springer, Heidelberg (2012). doi:10.1007/978-3-642-33275-3_70
8. Nanni, L., Paci, M., dos Santos, F.C., Skottman, H., Juuti-Uusitalo, K., Hyttinen, J.: Texture descriptors ensembles enable image-based classification of maturation of human stem cell-derived retinal pigmented epithelium. PLoS ONE **11**(2), e0149399 (2016). doi:10.1371/journal.pone.0149399
9. Diamond, J., Anderson, N.H., Bartels, P.H., Montironi, R., Hamilton, P.W.: The use of morphological characteristics and texture analysis in the identification of tissue composition in prostatic neoplasia. Hum. Pathol. **35**(9), 1121–1131 (2004)
10. Mazo, C., Trujillo, M., Salazar, L.: Identifying loose connective and muscle tissues on histology images. In: Ruiz-Shulcloper, J., Sanniti di Baja, G. (eds.) CIARP 2013. LNCS, vol. 8259, pp. 174–180. Springer, Heidelberg (2013). doi:10.1007/978-3-642-41827-3_26522
11. Zhao, D., Chen, Y., Correa, N.: Statistical categorization of human histological images. In: IEEE International Conference on Image Processing, ICIP 2005, vol. 3, pp. 628–631 (2005). doi:10.1109/ICIP.2005.1530470
12. Yu, F., Ip, H., Horace, H.S.: Semantic content analysis and annotation of histological images. Comput. Biol. Med. **38**(6), 635–649 (2008). doi:10.1016/j.compbiomed.2008.02.004
13. Boya, J.: Atlas de Histología y Organografía Microscópica. Editorial Medica Panamericana S.A., Madrid (2011)

14. Mazo, C., Trujillo, M., Salazar, L.: Automatic classication of coating epithelial tissue. In: Bayro-Corrochano, E., Hancock, E. (eds.) CIARP 2014. LNCS, vol. 8827, pp. 311–318. Springer, Cham (2014). doi:10.1007/978-3-319-12568-8_38
15. Pietikinen, M., Ojala, T., Xu, Z.: Rotation-invariant texture classication using feature distributions. Pattern Recogn. **33**, 43–52 (2000)
16. Bader-El-Den, M.: Self-adaptive heterogeneous random forest. In: 2014 IEEE/ACS 11th International Conference on Computer Systems and Applications (AICCSA), pp. 640–646 (2014). doi:10.1109/AICCSA.2014.7073259
17. Ghassabeh, Y.A., Rudzicz, F., Moghaddam, H.A.: Fast incremental LDA feature extraction. Pattern Recogn. **48**(6), 1999–2012 (2015). doi:10.1016/j.patcog.2014.12.012
18. Kylberg, G., Sintorn, I.-M.: Evaluation of noise robustness for local binary pattern descriptors in texture classification. EURASIP J. Image Video Process. **2013**, 17 (2013). http://dblp.uni-trier.de/db/journals/ejivp/ejivp2013.html#KylbergS13
19. Canada, B.A., Thomas, G.K., Cheng, K.C., Wang, J.Z., Liu, Y.: Towards efficient automated characterization of irregular histology images via transformation to frieze-like patterns. In: CIVR, pp. 581–590. ACM (2008)
20. Oliveira, D.L., Nascimento, M.Z., Neves, L.A., Batista, V.R., Godoy, M.F., Jacomini, R.S., Duarte, Y.A., Arruda, P.F., Neto, D.S.: Automatic classification of prostate stromal tissue in histological images using Haralick descriptors and local binary patterns. In: Journal of Physics: Conference Series, vol. 490, no. 1 (2013). http://stacks.iop.org/1742-6596/490/i=1/a=012151
21. Alturkistani, H.A., Tashkandi, F.M., Mohammedsaleh, Z.M.: Histological stains: a literature review and case study. Glob. J. Health Sci. **8**(3), 72–79 (2016). http://doi.org/10.5539/gjhs.v8n3p72