Automatic Classification of Colorectal and Prostatic Histologic tumor Images using Multiscale Multispectral Local Binary Pattern Texture Features and Stacked Generalization

Rémy Peyret[1], Ahmed Bouridane[1], Fouad Khelifi[1], Muhammad Atif Tahir[2], Somaya Al-Maadeed[3]

1 Northumbria University, Newcastle Upon Tyne, UK

2 National University of Computer and Emerging Sciences (NUCES-FAST), Karachi Campus, Pakistan

3 Qatar Univerity, Doha, Qatar

Abstract

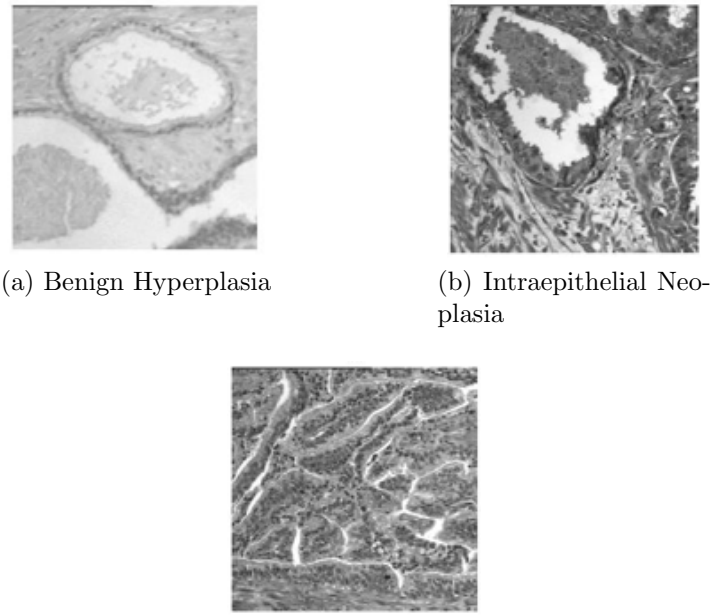
This paper proposes a new multispectral multiscale local binary pattern feature extraction method for automatic classification of colorectal and prostatic tumor biopsies samples. A multilevel stacked generalization classification technique is also proposed and the key idea of the paper considers a grade diagnostic problem rather than a simple malignant versus tumorous tissue problem using the concept of multispectral imagery in both the visible and near infrared spectra. To validate the proposed algorithm performances, a comparative study against related works using multispectral imagery is conducted including an evaluation on three different multiclass datasets of multispectral histology images: two representing images of colorectal biopsies - one dataset was acquired in the visible spectrum while the second captures near-infrared spectra. The proposed algorithm achieves an accuracy of 99.6% on the different datasets. The results obtained demonstrate the advantages of infrared wavelengths to capture more efficiently the most discriminative information. The results obtained show that our proposed algorithm outperforms other similar methods.

Keywords: Multiscale Multispectral Local Binary Pattern, Stacked generalization, histology, colorectal cancer, prostate cancer, automatic diagnosis

1. Introduction

The World Health Organization has declared that the cancer burden is a worldwide health problem. According to their 2014's report, 14 million new cases were diagnosed in 2012 and 8 million people died from it in the same period [1]. Colorectal cancer is the third most incident globally and prostate is in second position amongst men representing respectively 9.7% and 7.9% of all cancers for both sexes [1]. Both colorectal and prostatic tissues are glandular thus having a similar histological appearance. They also are both subject to the same tumor types; adenocarcinoma being the most commonly diagnosed cancerous tumor type in these organs. The European Association of Urology's guidelines [2] advise to perform a biopsy and a histological analysis on the sample for prostate cancer diagnosis. This method is also the most widely used for colorectal cancer diagnosis [3]. However, this process is very time-consuming for pathologists as they have to manually analyze every sample to spot the particular features characterizing the type of tumor and the various cancer stages. This process results in a high intra- and inter-observer variability [4], [5] thus affecting the diagnostic reliability. This paper aims to propose an algorithm that will automatically classify the samples into different categories of the cancer hence assisting the pathologist to make the appropriate diagnostics. This will in turn help to reduce the diagnosis time and act as a second opinion for the pathologist to reduce the intra- and inter-observer diagnostic variability.

Fig. 1 shows microscopic images of three different colorectal tumor biopsies. The first one is a Benign Hyperplasia, a benign tumor with little risk of evolving to cancer while the second represents an Intraepithelial Neoplasia, a tumor with high risk of evolving to cancer. The last image shows a carcinoma, which is a cancerous tumor. One of the features taken into account by the pathologists for the diagnosis is the general structure and organization of the tissue. In a normal structure, the epithelial cells are organized around the lumen which is separate from the plasma cells; whereas in the case of a carcinoma the normal structure becomes completely random and can be chaotic. Therefore, the proposed algorithm uses image texture features so as to capture and quantify its structure in order to classify the samples into the various types of malignancy. The sole pixel intensity can be insufficient to characterize the type of cell or sub-cellular components and so can have



(a) Benign Hyperplasia

(b) Intraepithelial Neoplasia

(c) Carcinoma

Figure 1: Example of images from the different classes of colorectal tissue.

negative effect on the feature extraction. Consequently, using the spectral response of each point of the sample to describe the tissue is adopted in this paper to improve the classification performances; hence the use of multispectral images of the biopsies. This paper also aims to investigate the advantages of using the pixels response from a wider electromagnetic spectrum ranging from the visible light to the infrared (IR) in comparison to other methods that can be used in biopsy image analysis as shown in Fig. 2.

The main contributions of this paper are three-fold. First, our work considers a grade diagnostic problem rather than a simple malignant versus benign tumor problem in the context of multispectral imaging and this is almost absent in the literature, especially for colorectal cancer [6]. Second, it introduces a new Multispectral Multiscale Local Binary Pattern (MMLBP) texture feature which is an adjusted LBP to multispectral data taking into account the third dimension (spectrum) of the data. The MMLBP differs from the traditional LBP in that it considers the joint information across spatial and spectral directions of the image. In addition, a stacked generalization technique is devised in order to fuse the different scales of the MMLBP and GLCM features at the score level. Finally, a new dataset is also introduced in

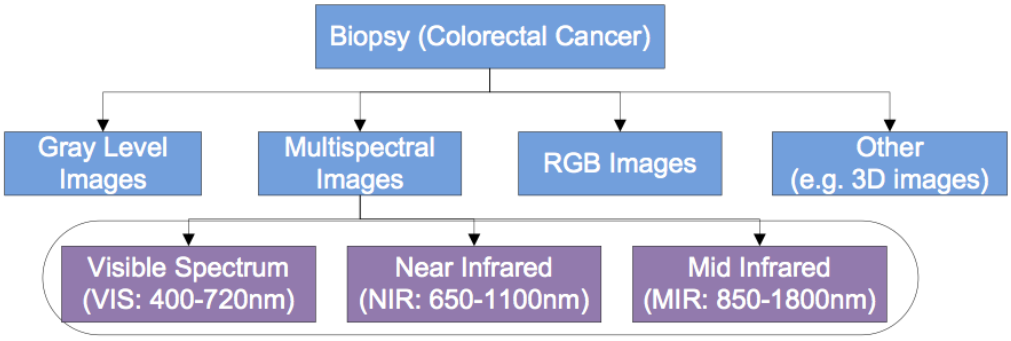


Figure 2: Imaging technologies for the classification of colorectal cancer.

the paper. This dataset is composed of multispectral images with a spectrum extending to the infrared (IR).

This paper is organized as follows: Section 2 gives a briefly a review of existing systems including a discussion on prostate and colorectal cancer tissue analysis . Section 3 reports some related feature extractors using LBP approach and some of its variants. Section 4 describes the proposed methodology including the contributions made and the proposed system. Section 5 explains the implementations and experimental analysis. Section 6 evaluates the results and a comparative study against existing techniques. It also shows the advantages of using IR images to improve the classification accuracy. Section 7 concludes the paper.

2. Related work

2.1. Previous work on prostate and colorectal cancer tissue analysis

Several techniques available in the literature extract textural features using panchromatic images [6], [7]. Esgiar *et al.* [8] computed a gray-level co-occurrence matrix (GLCM) on each colorectal histological image and extracted some of the GLCM features [9]. They then proposed a malignant versus benign classification using an SVM classifier. In [10], Kalkan at al. combined the same features with structural ones before computing a feature selection and a four-class classification, and achieved a 75.15 % accuracy. In [11], the authors used a 8-class dataset of 5000 images. The classes involved where the following: tumor epithelium, simple stroma, complex stroma, immune cells, debris, normal mucosal glands adipose tissue and background (no

tissue). The authors compared several texture descriptors such as GLCM, LBP, perception feature - mimicking the human perception at an abstract level - and Gabor filters. Their results range from 74 to 97 %. In the case of prostate cancer, several authors used the GLCM features for the same task of carcinoma detection [12]. The authors of [13], [14] used fractal analysis for prostate cancer grading or carcinoma detection.

Multispectral images have been used for texture feature extraction. In [15], Masood *et al.* applied GLCM features after segmenting the image data through a pre-processing phase. the approach consists of using the spectral dimensions to segment the image into four clusters representing four different tissue types: nuclei, cytoplasm, glands and stroma. Chaddad *et al.* proposed an improved version of the snake algorithm for the segmentation and extraction of GLCM texture features of multispectral-segmented images [16]. In [17], the authors proposed a method for characterizing the continuum of colorectal cancer using several texture features after segmentation. As features extraction, the GLCM features, the Laplacian of Gaussian and discrete wavelets were used. A few other studies used wavelet transforms [18] [19] and Laplacian of Gaussian [20] [10]. In [21], Roula *et al.* worked on prostate histological images and extracted GLCM features from each spectral band and combined them with morphological features for the discrimination phase using a quadratic discriminant analysis. They showed that multispectral analysis significantly improved the classification scores. The authors of [22] also demonstrated that the use of texture features in multispectral images improved the results when using texture features of panchromatic images on a colorectal histology dataset. They compared the performance of different texture features on multispectral images, namely the GLCM features and the multiscale LBP and used PCA for dimensional reduction followed by a SVM classifier. In [23], Tahir *et al.* first extracted statistical and structural features as well as the GLCM features. They then used a Round-Robin Tabu Search for dimensional reduction of the multispectral data before classification. They achieved a classification accuracy between 98% and 100%.

None of the previously mentioned authors used a multispectral texture feature detector that uses the spectral dimension directly. They either combined several results of 2-dimensional texture detector run on each spectral band, or used the dimensional reduction to create a 2D image on which the texture was to be detected. Khelifi *et al.* [24] developed a multi-band texture detection extending the GLCM. For this purpose, they used a spatial and spectral gray level dependence method (SSGLDM) assuming a joint infor-

mation between spectral bands exists. They applied this technique to the prostate cancer case.

However, only few studies use LBP texture features in this field [22, 25, 26] and none of them uses the joint information of spatial and spectral dimensions. For example, the authors of [25] select a single band from which the texture extraction was conducted. In [26], the LBP histogram is built on all three color channels of the image.

2.2. Previous work on multispectral texture analysis

Some methods for other applications, such as image segmentation, used a 3D histogram as a mean to fuse information from three color channels of a colour image [27]. Hassan El Maia *et al.* [28] proposed a method for multispectral image classification using the mutual information of GLCM features. In [29], the authors used a method developed in [30] for automatic face recognition. This algorithm was a modified LBP that computed a LBP on each color band of the spectrum separately and added opponent features to capture the spacial correlation between the bands. Radu-Mihai Coliban *et al.* [31] proposed a pseudo-morphology based on the Euclidean distance in $\mathbb{R}^n$. Using the proposed pseudo-morphology, the authors introduced a pseudo-granulometry and a morphological covariance to characterize the image texture. In [32], the authors use a neural network structure to classify multispectral texture information extracted from the images.

2.3. Previous work on IR texture analysis

In the field of facial recognition, the IR spectrum has been used and has proven to increase the recognition rates in many cases. Abdelhakim Bendada *et al.* [33] introduced a differential LTP descriptor and extend their method to the IR spectrum. They showed that a high recognition rate was achieved with the IR spectrum. The authors [34] developed a method for synthesising the visible and near IR face images in order to take advantage of both the illumination invariance of IR images and the detailed texture information provided by the face images captured in the visible range of the electromagnetic spectrum. The authors compared their method to the conventional LBP applied separately to the near IR and the visible images and showed that the combined use of the IR and the visible spectra increased the identification rate by 8.76pp (from 88.83% to 97.59%).

Thematic mapping imagery uses the infrared spectrum to acquire information not captured by the visible spectrum. Yun Zang [35] used an algorithm of conditional variance detection on multispectral images captured on a visible and IR spectrum for classification of urban treed areas.

3. Feature extraction using LBP approach: a review

In this section, the conventional LBP and its rotation invariant and three-dimensional variants are discussed.

3.1. Conventional LBP

Ojala *et al.* described LBP texture features as a local characterisation of a pixel's neighborhood at a radius R sampled into a set of P neighbors on a circle centered around the central pixel and of radius R. Let g_0 be the intensity of the central pixel x and g_p the intensity of its pth neighbor. The LBP is defined as follows [36]:

$$LBP_{P,R}(x) = \sum_{p=1}^{P} s(g_0 - g_p)2^{p-1} \tag{1}$$

where,

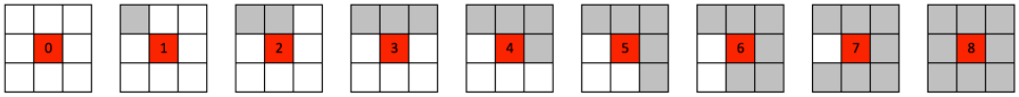
$$s(x) = \begin{cases} 0 \text{ if x} \leq 0 \\ 1 \text{ if x} > 0 \end{cases}$$

LBP is computed for the whole image, before it is pooled into a LBP histogram of size 256. The resulting LBP histogram, which is invariant to intensity changes, is then used as a texture feature descriptor to characterize the image.

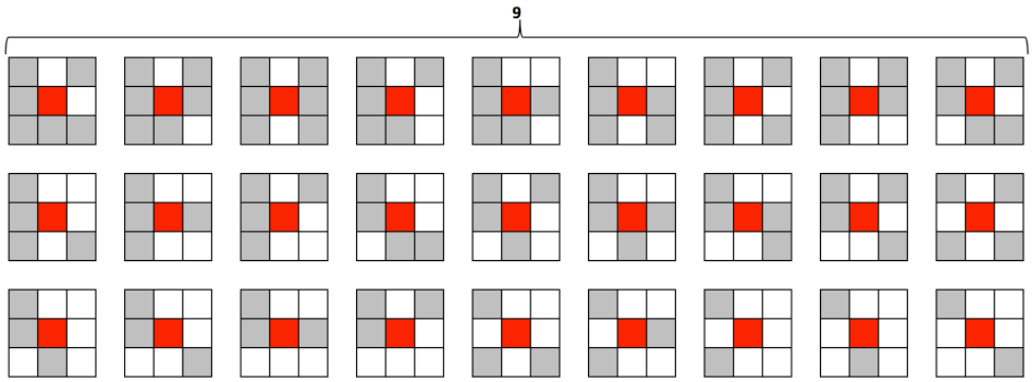
3.2. Rotation Invariant Uniform LBP

A rotation invariant LBP, referred to as LBP^{riu2}, using uniform patterns has also been proposed as illustrated in Fig. 3a. They operate as templates for microstructures such as bright spot (0), flat area or dark spot (8) and edges of varying positive or negative curvature (1-7) [36]. These structures define a uniformity measure U corresponding to the number of transition in the pattern as follows:

$$U(LBP_{P,R}) = |s(g_{P-1} - g_c) - s(g_0 - g_c)| + \sum_{p=1}^{P-1} |s(g_p - g_c) - s(g_{p-1} - g_c)| \tag{2}$$



(a) Uniform LBP patterns and their corresponding labels



(b) Non-uniform LBP patterns

Figure 3: The 36 unique possibilities for a circular symetric set of LBP patterns and their corresponding labels for rotation invariant, uniform LBP. The red squares correspond to the central pixel, the white and grey squares represent the 0 and 1 bits in the 8-bits output of the operator. The numbers are the unique $LBP_{P,R}(x)$ labels.

Fig. 3a shows the 9 patterns with a U measure of at most 2 when the 27 other patterns shown of Fig. 3b have a uniformity measure of at least 4. Therefore, patterns having $U(LBP_{P,R}) \leq 2$ are said to be uniform. The following operator defines a gray-scale and rotation invariant texture description [36]:

$$LBP_{P,R}^{riu2}(x) = \begin{cases} \sum_{p=1}^{P} s(g_p - g_c) \; if \; x \leq 2 \\ P+1 \qquad otherwise \end{cases} \tag{3}$$

In this way, $P+1$ uniform patterns are assigned to a unique label corresponding to the number of 1 bits in the pattern while the non-uniform patterns are grouped under the same category. The final texture feature used is a histogram of $P+2$ bins generating all the $LBP_{P,R}^{riu2}$ outputs accumulated over the image.

This form of LBP seems more adapted to the problem at hand because of the rotation invariance it provides. Indeed, in the case of histopathology, sample orientation and cells direction are not relevant criteria to consider for classification because they vary independently to the sample's class. A

second advantage of this $LBP_{P,R}^{riu2}$ over a conventional $LBP_{P,R}$ is its smaller size thus making it faster to process in a classification step.

3.3. Three-dimensional LBP

Since multispectral images are three-dimensional data the conventional LBP concept needs to be modified to deal with this datatype. In the literature, two methods are usually described when dealing with 3D images for applications such as video processing and face recognition [37]. The proposed method is inspired from Volume LBP and LBP-TOP (for Local Binary Pattern-Three Orthogonal Plan) [37]. Here, we briefly discuss VLBP and LBP-TOP before explaining the proposed extension of LBP to multispectral LBP. To extend LBP to Dynamic Texture analysis, Zhao *et al.* define a neighborhood as the joint distribution of $3P + 3$ image pixels where P is the number of neighbors on one frame as shown on [37]. A similar technique to the conventional LBP can be applied and a Volume Local Binary Pattern (VLBP) is defined as follows:

$$VLBP_{P,R}(x) = \sum_{p=1}^{3P+2} s(g_0 - g_p)2^{p-1} \tag{4}$$

The VLBP local features are pooled into a histogram of size 2^{3P+2}. This histogram's size increases very quickly when the number of neighbors P grows and may become very computationally intensive. On the other hand, using a small P may lead to a loss of some critical information for diagnosis. To address this issue, a LBP-TOP feature is proposed by considering three orthogonal planes intersecting on a central pixel as shown in [37]. The technique computes a two-dimensional LBP on each of these plans and concatenates the output histograms which will be of size $3 * 2^P$ instead of 2^{3P+2} in the previous case. Here, the circles are considered in the time dimension because this LBP-TOP is meant to be applied on video processing so the motion direction of texture is unknown.

4. The Proposed MMLBP System for Cancer Classification

As illustrated in Fig. 4, the proposed system is composed of two main stages. First, MMLBP features are extracted and, then, an Independent Component Analysis is performed to reduce the dimensionality of the feature space. Then, a stacked generalization employing the Support Vector Machine classifier is used at the matching stage.

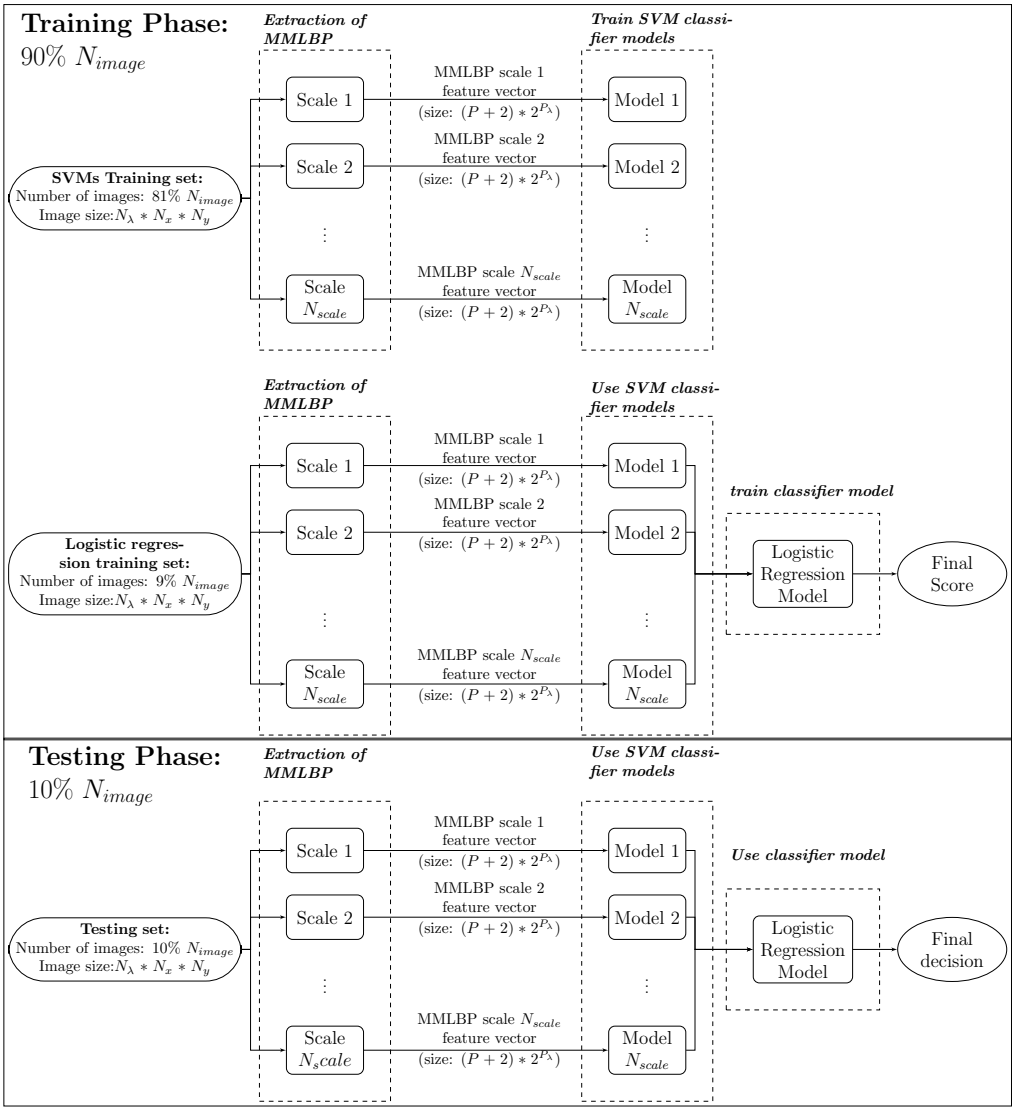


Figure 4: Block diagram of stacking training and testing with MMLBP texture features. N_{image} represents the number of images in the dataset; N_x, N_y are the number of lines and columns in each image, respectively, and N_λ is the number of spectral bands.

4.1. Proposed Multispectral LBP Texture Feature

In the proposed technique the third dimension is spectral (not temporal) thus no texture motion is considered. Therefore, unlike in the aforementioned 3D-LBP variants, a neighborhood of only P points in the spatial plan and P_λ

on a straight line in the spectral dimension intersecting the spatial plan at the central pixel was considered as shown in Fig. 5 with $P_\lambda = 2$. As explained above, this technique is adopted to make the LBP rotation invariant in the spatial dimensions while still using the same U measure described in (3) in the XY plan. The idea is to assign the $LBP^{riu2}_{P,R}$ patterns to different

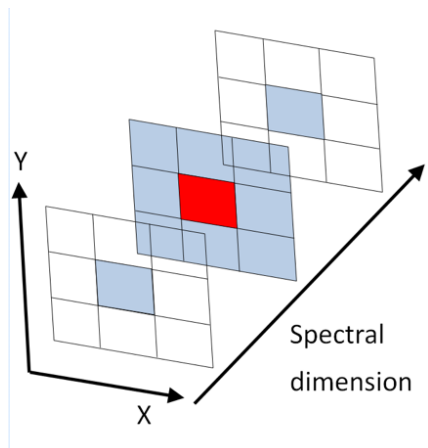


Figure 5: Multispectral LBP descriptor: the neighborhood considered for multispectral LBP. X and Y being the spatial dimensions. Each tile represents a pixel. The red tile is the central pixel considered, and the blue tiles are the pixel considered in the neighborhood.

categories depending on the P_λ pixels in the neighboring plans. On top of the $LBP^{riu2}_{P,R}$ computed using equation 3, $LBP^{\lambda}_{P_\lambda,R}$ is calculated using the following equation:

$$LBP^{\lambda}_{P_\lambda,R}(x) = \sum_{q=1}^{P_\lambda - 1} s(g'_q - g_c)2^q \tag{5}$$

where, g'_q is the pixel value in the pixel of plan q aligned to the central pixel. The $MMLBP_{P,P_\lambda,R}$ is defined as follows:

$$MMLBP_{P,P_\lambda,R} = LBP^{riu2}_{P,R} + (P+1)LBP^{\lambda}_{P_\lambda,R} \tag{6}$$

The $MMLBP_{P,P_\lambda,R}$ outputs are then pooled into a histogram of size $(P+2) * 2^{P_\lambda}$. It is worth noting that the scale is controlled by $R \in [1..N_{scale}]$ As a result, the histograms built from each scale are concatenated to form the MMLBP and each scale is considered as a separate feature which is fed to the stacked classifier as shown in Fig. 6.

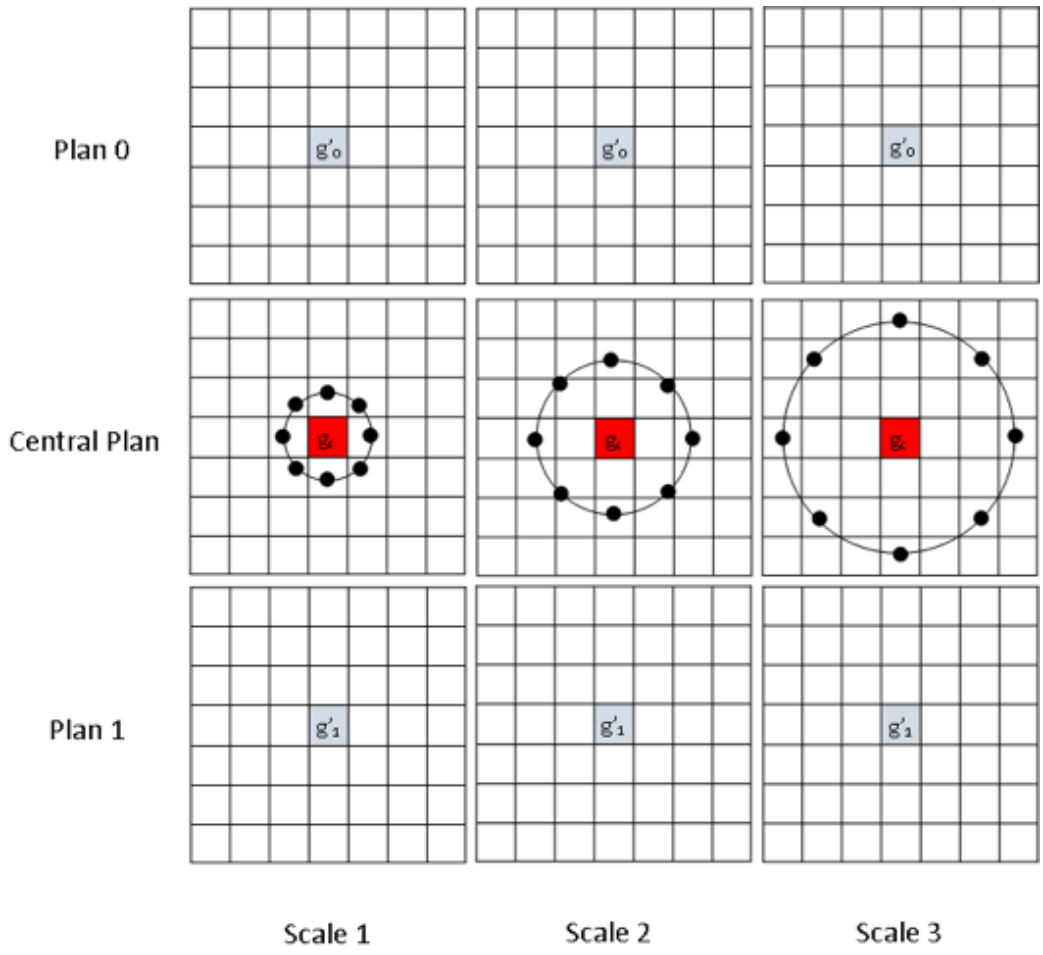


Figure 6: Multiscale neighborhood for MMLBP

4.2. Dimensionality Reduction using Independent Component Analysis and Classification using Support Vector Machine

In order to address the curse of dimensionality problem and hence reduce the learning cost, the Independent Component Analysis (ICA) is applied before classification. In contrast to more widely used Principal Component Analysis (PCA), this technique presents the advantage of being able to decorrelate the signal and reduce statistical dependencies between the features as much as possible [38]. In fact, it could be seen as a version of PCA that defines orthogonal directions. The ICA transformed data are computed using only the training data of the SVM classifier. The testing data are projected

to the new basis before classification. The number of components used for classification will be optimized as described in a further section.

The classification step consists of a multiclass Support Vector Machine (SVM) classifier with a Gaussian kernel. A SVM constructs a hyper-plan separating the classes; it tries to find the maximum distance to the nearest training data point of each class and can be described as follows. Given a training vector $x_i \in \mathbf{R}^P$, $I \in [|1, n|]$, in two classes and a vector $y \in \{1, -1\}^n$, the SVM solves the following optimization problem:

$$\min_{w,b,\zeta} (1/2\ w^T w + C \sum_{i=1}^{n} \zeta_i) \text{ subject to: } y_i(w^T \phi(x_i) + b) \geq 1 - \zeta_i, \zeta_i \geq 0,\ i \in [|0, n|] \tag{7}$$

Its dual form is:

$$\min_{\alpha} (1/2\ \alpha^T Q \alpha - e^T \alpha) \text{ subject to: } y^T \alpha = 0, 0 \leq \alpha_i \leq C,\ i = [|0, n|] \tag{8}$$

Where, $C > 0$ is the upper bound, Q is an n by n positive semi definite matrix $Q_{i,j} = y_i y_j K(x_i, y_j)$ where $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is the kernel. Here training vectors are implicitly mapped into a higher dimensional space by the function ϕ. The decision function is:

$$f_{decision} = sgn(\sum_{i=1}^{n} y_i \alpha_i K(x_i, x) + \rho) \tag{9}$$

The kernel function used here is the radial basis function or Gaussian kernel:

$$K(x) = e^{-\gamma |x - x'|^2} \tag{10}$$

Where γ is a positive parameter. The kernel parameters are optimized using a grid search method which will be detailed in Section 5.1. In order to find the appropriate compromise between the sizes of training and testing datasets and hence avoid over-fitting that might be caused by a leave-one-out technique; a 10-fold cross-validation is used. The one-versus-all technique is used to build the multiclass classifier.

4.3. Logistic Regression for Stacked Generalization

Stacked generalization (or stacking) is an ensemble method for classification [39]. It uses the output of a first layer of classifiers as inputs to another

classifier - called meta-classifier - for the final decision. In this paper, this system is used to fuse the different scales of multispectral LBP texture feature at score level.

Fig. 4 shows the two steps of training and testing for the stacking algorithm. A logistic regression model is used as a meta-classifier for its relatively low computing cost. The first layer of classifiers is composed by SVM classifiers with a Gaussian kernel as described in Section 4.2. In addition to a 10-fold cross-validation carried out at the meta-classifier level, an internal cross-validation of the training data is implemented in order to prevent bias and improve stability of the different classifiers.

5. Experiment and Setup

5.1. Datasets

To evaluate the performance of the proposed technique, three different datasets are used in the experimentation process.

5.1.1. Dataset 1: Colorectal tumor Tissue from Texas

The first one, described in [17], is composed of colorectal biopsy images acquired using multispectral imagery at low magnification power (x40). The database consists of 29 three-dimensional images having a spatial resolution of 512*512 pixels and 16 spectral bands corresponding to wavelengths between 500 and 650 nm. The images are divided into 3 classes of tumor tissue types: Carcinoma (Ca), the class containing the cancerous samples, Benign Hyperplasia (BH), a class with benign tumors, and Intraepithelial Neoplasia (IN), containing images with tissues at a precancerous stage.

5.1.2. Dataset 2: Prostatic tumor Tissue

The second dataset [23], with some samples shown in Fig. 7, consists of multispectral images taken at 16 spectral channels (from 500 to 650 nm) and at x40 magnification power. 592 different samples (multispectral images) of size 128*128 have been used to carry out the analysis. The samples are evaluated by two highly experienced independent pathologists and labeled into four classes: 165 cases of Stroma (Str), which is normal muscular tissue, 106 cases of Benign Prostatic Hyperplasia (BPH), a benign condition, 144 cases of Prostatic Intraepithelial Neoplasia (PIN), a pre-cancerous stage, and 177 cases of Prostatic Carcinoma (PCa), an abnormal tissue development corresponding to cancer.

5.1.3. Dataset 3: Colorectal Vis-IR

Colorectal Vis-IR: The third dataset is also composed of multispectral colorectal histology data with a (x40) magnification power. This dataset was developed by University of Qatar with the collaboration of the Al-Ahli Hospital, Doha and will be made available for public use[1]. It is split into 4 classes, each of them composed of 10 images. The images are acquired on a wider spectrum than in the first dataset as it is spread on the visible (Vis) and infrared (IR) ranges of the electromagnetic spectrum - shown in Figure 2 - with an interval of 23 nm between each wavelength. That is to say, in the visible range, the wavelength interval is 23 nm starting from 465 nm to 695 nm and in the IR range, the wavelength interval is also 23 nm and ranges from 900 nm to 1590 nm. The 4 classes are: Carcinoma (Ca), containing the images of cancerous colon biopsies, Tubular Adenoma (TA), a pre-cancerous stage, Hyperplastic Polyp (HP), a benign polyp and No Remarkable Pathology (NRP).

5.2. Experiments

The first two datasets are used for the first sets of experiments and the 3rd dataset is used in the last experiment to show how the performance can be improved by using the IR imagery.

The proposed system is first compared with the results given by the algorithm described in [22] by using a conventional LBP extracted from a panchromatic image that is generated by averaging the spectral bands of the multispectral image. It is also compared to another variant of LBP adapted to multispectal images. It consisting in extracting LBP histograms from each band and then concatenating them to from a final descriptor. This method is referred to as the concatenated LBP. It is worth mentioning that these LBP variants were used with an SVM classifier for a fair comparison. For the same reason, they were also applied using the same number of scales N_{scale}. Many authors use GLCM texture features - see Appendix: GLCM texture feature. The results obtained using the proposed system - that is the stacked classification of the GLCM feature model combined with the MMLBP as shown in Figure 8, this is denoted as stacked MMLBP + GLCM - were also compared to the ones given by MMLBP alone. In order to assess our algorithm's robustness, the two first datasets presented in Section 5.1 are used.

[1] it is expected that the first release will take place in January 2018

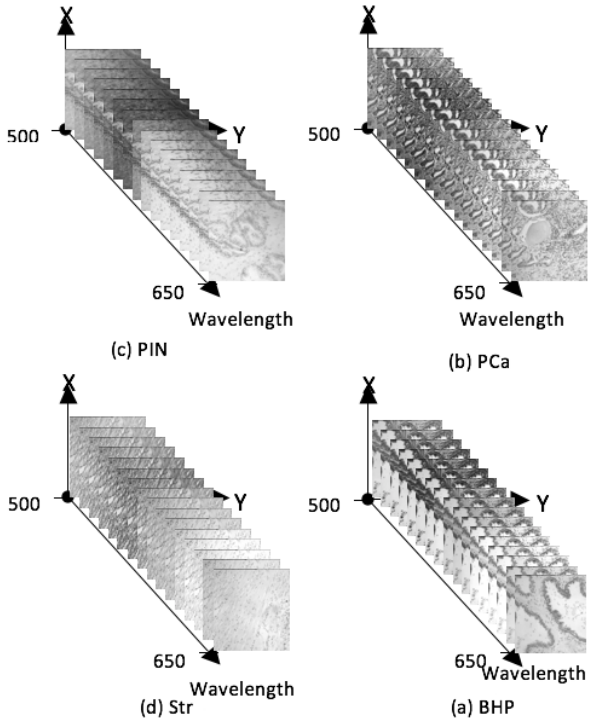


Figure 7: Sample of multispectral images from Dataset 2

In the second set of experiments, the impact of spatial resolution variations in performance is addressed .

The algorithm is also compared against different algorithms from the literature. An adapted version of Masood *et al.*'s algorithm [15] to the multiclass problem is implemented. In this method they use the GLCM features after segmentation of the image to train an SVM classifier. The results given by the algorithm described by the authors of [16] are used for comparison. It consists of using a snake algorithm for image segmentation and uses the GLCM feature as well. Our method is also compared against Khelifi *et al.*'s results [24]. They define a multispectral form of the GLCM before extracting the GLCM features. Finally the results shown in [23] are used for comparison purposes. In that paper, Tahir *et al.* describe a Round-Robin Tabu search algorithm for prostatic tumor classification.

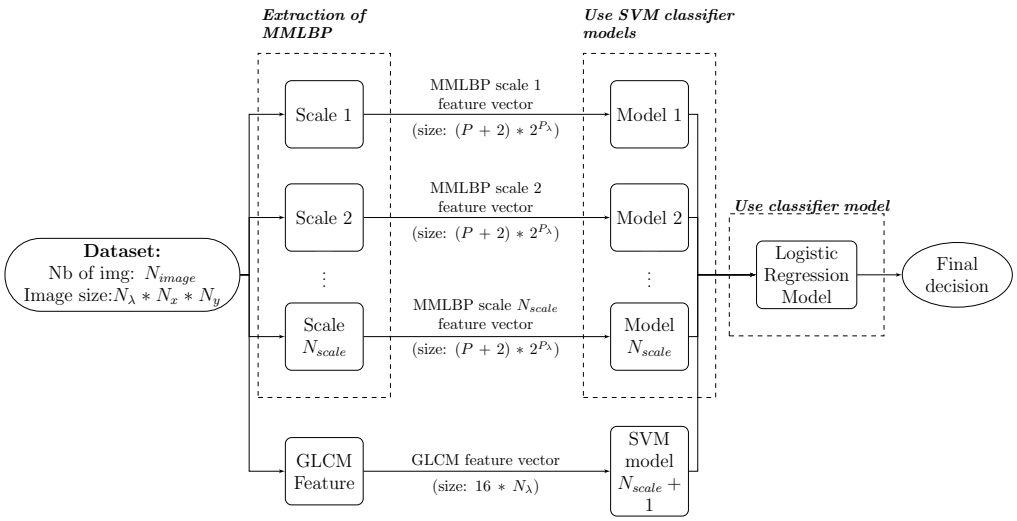


Figure 8: Block Diagram of the proposed algorithm.

5.3. Evaluation Measures

In order to avoid accuracy variations, the cross-validation is run ten times and the accuracy is averaged. The standard deviation is calculated on the mean accuracies of each cross-validation.

In addition of the accuracy and the standard deviation, the ROC curve and the Area Under Curve and the confusion matrix are also computed and used to assess the performances of the proposed algorithm. These performance measures are useful metrics to allow for a better understanding of what each class captures before one-versus-all combination to obtain the overall accuracy.

6. Results and Discussion

6.1. Training procedure

As illustrated by Fig. 4, the double 10-fold cross-validation run on the datasets means that, for each experiment, 90% of the dataset is used for training the logistic regression classifier model and the remaining 10% are used for the testing phase. 90% of this training set (or 81% of the total dataset) is used for training the SVM models and in the remaining 10% of the training set (or 9% of the whole dataset), the trained SVM models are

used to train the logistic regression model. Table 1 displays the SVMs and logistic regression training sets and the testing set sizes for each dataset.

Table 1: Number of images used in each phase for each the tested dataset

Data-set	SVMs training set	Logistic regression training set	testing set	dataset size
1	23	3	3	29
2	480	53	59	592
3	29	3	4	36

6.2. Parameters Tuning

As discussed previously, a total of 3 parameters need to be optimized for each SVM classification: the number of components selected in the ICA, and the C and γ parameters of the SVM kernel from Eq. (8), (10). A three-dimensional grid search was performed with the following parameters, with a step equals to 1:

$$C = 10^i, \text{ with } \; i = [|-3:3|],$$
$$\gamma = 10^j, \text{ with } \; j = [|-3:3|],$$
$$N_{comp} = 10 * k, \text{ with } \; k = [|1:50|].$$

For each combination of the parameters in these intervals, the accuracy is calculated and averaged with a 10-fold cross-validation. The parameters giving the maximum average accuracy are then chosen as the model parameters.

6.3. Proposed Algorithm Discussion

Table 2 shows a comparison of the classification accuracies obtained using different features and classification methods. First a conventional LBP followed by a SVM classification is performed and an accuracy of 88.3 % is achieved on dataset 1 and 77.4 % on dataset 2. This shows this option is not robust to the data. When using a concatenated version of multispectral LBP followed by an SVM classification, the results are improved and an accuracy of 95.8 % is achieved on dataset 1. However, only 89 % accuracy is obtained on dataset 2 hence indicating the instability of the method. When using stacked generalization with MMLPB texture feature, the results are improved again and an accuracy of 99.0 % and 99.2 % on dataset 1 and dataset

2, respectively, thus demonstrating the robustness of the proposed algorithm. This can be explained because the stacking method selects the best features for classification and discards the features that drop the accuracy and this is independent to the data. When GLCM texture features are combined to the MMLBP texture features the results are improved by 0.3 - 0.6 percentage points (pp). It can also be seen that the multispectral information brings significant improvement over the conventional LBP as illustrated by the performance of the concatenated multispectral LBP method. Furthermore, the stacking classification process enhances the performance further as demonstrated by the results of the stacked LBP compared to the concatenated LBP.

Table 2: Accuracy comparison of different feature extraction and classification methods

Dataset	Conventional LBP (%)	Concatenated LBP (%)	Stacked MMLBP (%)	Proposed algorithm: Stacked MMLBP + GLCM (%)
1	88.3 ± 2.7	95.8 ± 0.5	99.0 ± 0.3	99.6 ± 0.4
2	77.4 ± 4.0	89.0 ± 0.9	99.2 ± 0.3	99.5 ± 0.3

Figure 9 and 10 displays the ROC curves and shows the Area Under Curve (AUC) for the different classes in a binary classification following the one versus all scheme. This is done to assess the positive and negative false alarm rates for each class. Table 3 and Table 4 show the confusion matrices obtained when using different datasets.

Table 3: Confusion Matrix for dataset 1

	Class BH	Class Ca	Class IN
Class BH	144	0	0
Class Ca	0	144	0
Class IN	1	4	139

As can be seen from Fig. 9, the system performs better on classes BH and Ca than it does on class IN. Fig. 6.3 displays some examples of correctly classified and misclassified images from dataset 1. Fig. 11a shows a sample of class IN which has been misclassified as class Ca by the proposed algorithm. Fig. 11b and 11c show correctly classified samples from class IN and Ca, respectively. As one can see the contrast on Fig. 11a is not as pronounced as

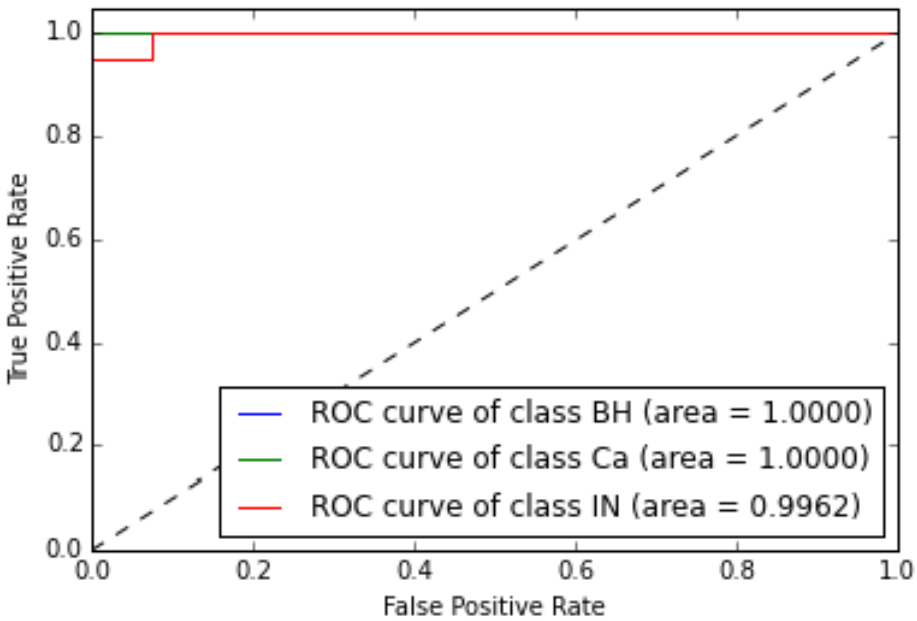


Figure 9: ROC for the proposed algorithm for Dataset 1

Table 4: Confusion Matrix for dataset 2

	Class BPH	Class PCa	Class PIN	Class Str
Class BPH	128	0	0	0
Class PCa	0	173	3	0
Class PIN	0	0	144	0
Class Str	0	0	0	144

what can be observed on Fig. 11b. This is especially true for the epithelial cells: in Fig. 11a the outer border of the cytoplasm of the cell is not as visible as it is on Fig. 11b. On the other hand, Fig. 11c presents hyperchromatism meaning the nuclei of the cells are well contrasted with the rest of the tissue and the border of the cytoplasm is not very clear. These features described on Fig. 11c is similar to what can be observed on Fig. 11a. Also both Fig. 11a and 11c show an area with stroma tissue respectively at the bottom right and at the top left of the images. The combination of both these features can explain the misclassification of Fig. 11a.

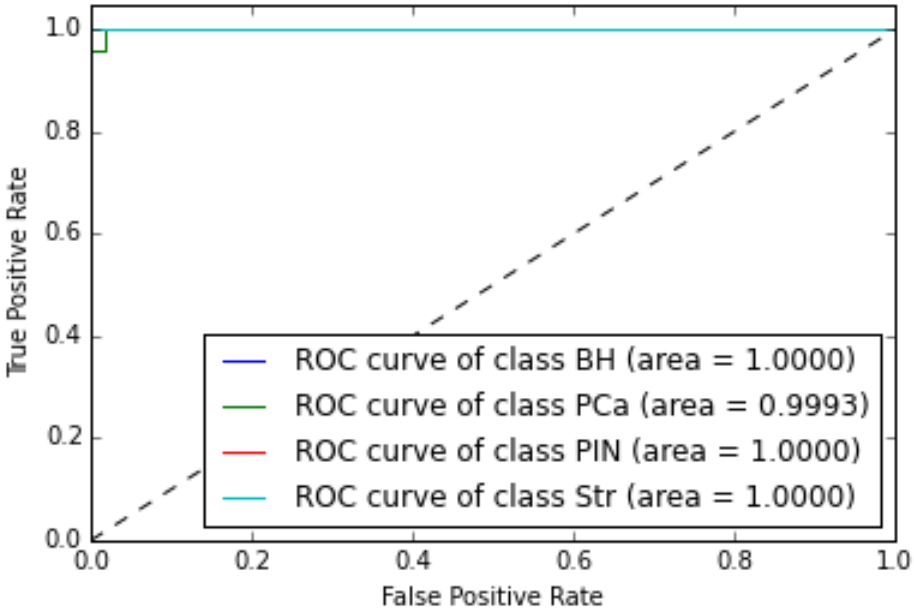


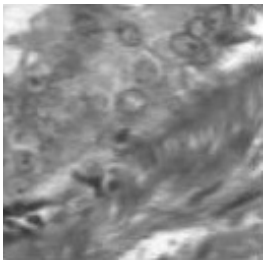
Figure 10: ROC for the proposed algorithm for Dataset 2

6.4. Impact of the Spatial resolution

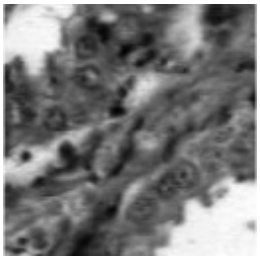
Table 5 shows the impact of image spatial resolution on the results. As can be seen, the accuracy is marginally influenced by the change of resolution. It varies from 99.6% ± 0.4 for the full resolution to 98.7% ± 0.4 for a spatial resolution of 25% the original one for Dataset 1. For a resolution of 10%, the accuracy drops to 96.0%. The same consistency can be seen on Dataset 2 until 50% of the original resolution then a drop by 2 points in accuracy is noticed for 25% of the original resolution. The drop further continues with a resolution of 10% the original one. This shows the robustness of the MMLBP algorithm presented in this paper to spatial resolution reduction until a certain percentage depending on the dataset.

Table 5: Accuracy comparison of different spatial resolution

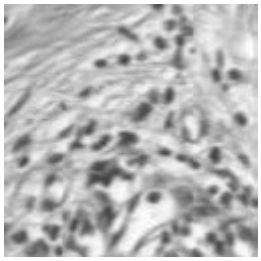
Data-set	Resolution 100%	Resolution 75%	Resolution 50%	Resolution 25%	Resolution 10%
1	99.6 ± 0.4	98.8 ± 0.4	99.4 ± 0.4	98.7 ± 0.4	96.3 ± 0.4
2	99.5 ± 0.3	99.8 ± 0.3	99.5 ± 0.3	97.6 ± 0.3	96.0 ± 0.4



(a) Image of class IN classified as Ca



(b) Image of class IN classified as IN



(c) Image of class Ca classified as Ca

Figure 11: Example of correctly classified and misclassified samples from dataset 1.

6.5. Comparison to Existing Algorithms

Table 6 depicts the performance accuracy obtained when comparing the proposed algorithm against some existing methods in the literature. [24]'s algorithm is tested on both the Texas and the Prostate datasets. The authors of [16] report a 98.9 % accuracy on the Texas dataset. Masood *et al.*'s algorithm is evaluated using the Texas and Prostate datasets using a multi-class classifier instead of the authors' binary classifier [15]. As can be seen in Table 6, the proposed method outperforms these three other algorithms in terms of accuracy. Tahir *et al.*'s algorithm is evaluated using the prostate dataset as reported by the authors who achieved a 98.9 % accuracy. The proposed algorithm is implemented on the same dataset and the results of 99.5 % accuracy clearly show that the proposed technique outperforms [23]'s algorithm.

6.6. Extension to the Infrared Spectrum

The algorithm is first evaluated on the visible and near infrared ranges separately on Dataset 3. Once this done, it is evaluated on a combined

Table 6: Accuracy comparison to literature methods

Dataset	Khelifi *et al.* [24] (%)	Tahir *et al.* [23] (%)	Chaddad *et al.* [16] (%)	Masood *et al.* [15] (%)	Proposed algorithm: Stacked MMLBP + GLCM (%)
1	89.9	n/a	98.9 ± 0.1	86.3 ± 0.3	99.6 ± 0.4
2	75.6	98.9	n/a	85.1 ± 2.0	99.5 ± 0.3

dataset including both the Vis and IR data by fusing the accuracy results at a score level using the stacking technique discussed in Section 3.2. Table 7 proves that using both the visible and infrared ranges of the light spectrum improves slightly the results. On the Qatar dataset, the proposed algorithm scores 99.2 % when using only the bands representing the wavelengths in the visible spectrum; this same algorithm scores 99.5 % when using the wavelengths from the infrared as well as the visible range. One can notice that the IR alone does not perform as well as the Vis spectrum with this algorithm but it adds different information and helps improving the accuracy when combined.

Table 7: Accuracy of proposed algorithm on Qatar dataset

Dataset	Accuracy
Dataset 3 Vis	99.2 ± 0.1
Dataset 3 IR	96.2 ± 0.5
Dataset 3 Vis+IR	99.5 ± 0.1

7. Conclusion

Multispectral texture features form an attractive method for extracting information from histologic images of colorectal or prostate tumor tissue for classification purposes. This paper proposed a MMLBP feature combined with GLCM using a stacked generalization for feature fusion at the score level for classification. The proposed method showed that this technique gives better results than similar and existing ones available in the literature attaining a classification accuracy above 99 % on all the datasets tested. This

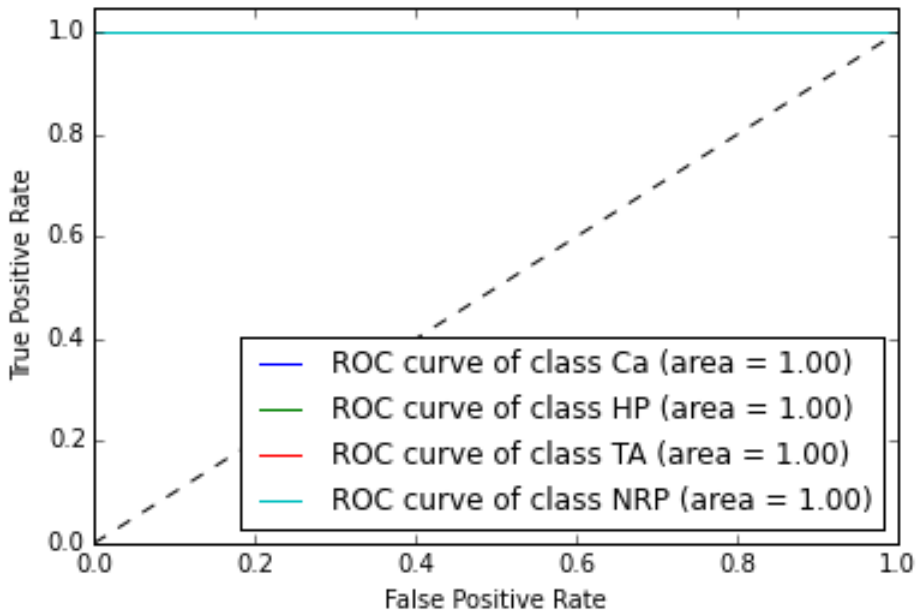


Figure 12: ROC for the proposed algorithm for dataset 3

study also showed that results can be improved when combing both infrared and visible information extracted from tissue samples.

Future work will focus on investigating the use of morphological features in order to improve the result. They will be easily combined to the texture features at decision level for classification thanks to the stacked generalization technique used here.

Appendix A. GLCM Texture Features

The GLCM texture features [9] are calculated from the GLCM extracted from the different layers of the multispectral [4] image where each layer represents the tissue response to a different wavelength. This GLCM matrix reflects how often a pixel with the intensity value I occurs in a specific spatial relationship (r, θ) to a pixel j. Four different spatial relationships are computed: $r = 1$ and $\theta = 0, 45, 90, 135$.

$$GLCM_{i,j,\lambda} = \sum_{p=1}^{n} \sum_{q=1}^{m} \begin{cases} 1, \text{ if } \begin{cases} I(p,q,\lambda) = i \\ I(p+\Delta x, q+\Delta y, \lambda) = j \end{cases} \\ 0, \text{ otherwise} \end{cases}$$

The following GLCM features are computed from the normalized GLCM matrices $p_{r,\theta}(i,j,\lambda)$ of the image:

- Energy:
$$\sum_{i,j} p(i,j,\lambda)^2$$
- Contrast:
$$\sum_{i,j} |i-j|^2 p(i,j,\lambda)$$
- Homogeneity:
$$\sum_{i,j} \frac{p(i,j,\lambda)}{1-|i-j|}$$
- Correlation:
$$\sum_{i,j} \frac{(\mu_i - i)(\mu_j - j)}{\sigma_i \sigma_j} p(i,j,\lambda)$$

For each multispectral image, the GLCM features are calculated on each GLCM from each layer and concatenated into a large vector of size $4 * number\ of\ multispectral\ layers$. The features are then rescaled and normalized to fit in the interval $[0,1]$ using the following equation:

$$x = \frac{x - min(x)}{max(x) - min(x)} \tag{A.1}$$

Where x is a vector representing the feature to be normalized, x' is the normalized feature, $max(x)$ and $min(x)$ are respectively the maximum and minimum values of x.

Acknowledgment

This publication was made possible using a grant from the Qatar National Research Fund through National Priority Research Program (NPRP) No. 6-249-1-053.

References

[1] J. Ferlay, M. Ervik, R. Dikshit, S. Eser, C. Mathers, M. Rebelo, D. Parkin, D. Forman, F. Bray, Globocan 2012 v.1.1, cancer incidence and mortality worldwide: Iarc cancerbase no. 11, 2012.

[2] A. Heidenreich, EAU Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Treatment of Clinically Localised Disease., Technical Report, European Urology, 2011.

[3] S. Kunhoth, S. A. Maadeed, A. Bouridane, R. A. Saady, A. Bouridane, Medical and computing insights into colorectal tumors, International Journal of Life Sciences Biotechnology and Pharma Research 4 (2015) 122–126.

[4] P. A. Humphrey, A. S. for Clinical Pathology, Prostate Pathology, American Society for Clinical Pathology, 2003.

[5] G. D. Thomas, M. F. Dixon, N. C. Smeeton, N. S. Williams, Observer variation in the histological grading of rectal carcinoma., Journal of Clinical Pathology 36 (1983) 385–391.

[6] S. Rathore, M. Hussain, A. Ali, A. Khan, A recent survey on colon cancer detection techniques, IEEE/ACM Transactions on Computational Biology and Bioinformatics 10 (2013) 545–563.

[7] C. Mosquera-Lopez, S. Agaian, A. Velez-Hoyos, I. Thompson, Computer-aided prostate cancer diagnosis from digitized histopathology: A review on texture-based systems, IEEE Reviews in Biomedical Engineering 8 (2015) 98–113.

[8] A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, A. Murray, Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa, IEEE Transactions on Information Technology in Biomedicine 2 (1998) 197–203.

[9] R. M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification, IEEE Transactions on Systems, Man, and Cybernetics SMC-3 (1973) 610–621.

[10] H. Kalkan, M. Nap, R. P. W. Duin, M. Loog, Automated classification of local patches in colon histopathology, Pattern Recognition (ICPR), 2012 21st International Conference on (2012) 61–64.

[11] J. N. Kather, C.-A. Weis, F. Bianconi, S. M. Melchers, L. R. Schad, T. Gaiser, A. Marx, F. G. Zöllner, Multi-class texture analysis in colorectal cancer histology, Scientific Reports 6 (2016) 27988 EP –.

[12] S. Bouatmane, M. A. Roula, A. Bouridane, S. Al-Maadeed, Round-robin sequential forward selection algorithm for prostate cancer classification and diagnosis using multispectral imagery, Machine Vision and Applications 22 (2011) 865–878.

[13] A. N. Esgiar, R. N. G. Naguib, B. S. Sharif, M. K. Bennett, A. Murray, Fractal analysis in the detection of colonic cancer images, IEEE Transactions on Information Technology in Biomedicine 6 (2002) 54–58.

[14] A. Tabesh, M. Teverovskiy, H. Y. Pang, V. P. Kumar, D. Verbel, A. Kotsianti, O. Saidi, Multifeature prostate cancer diagnosis and gleason grading of histological images, IEEE Transactions on Medical Imaging 26 (2007) 1366–1378.

[15] K. Masood, N. Rajpoot, H. Qureshi, K. Rajpoot, Co-occurrence and morphological analysis for colon tissue biopsy classification., Proc. 4th International Workshop on Frontiers of Information Technology. (2006) 211–216.

[16] A. Chaddad, C. Desrosiers, A. Bouridane, M. Toews, L. Hassan, C. Tanougast, Multi texture analysis of colorectal cancer continuum using multispectral imagery, PLoS ONE 11 (2016) 1–17.

[17] A. Chaddad, C. Tanougast, A. Dandache, A. A. Houseini, A. Bouridane, Improving of colon cancer cells detection based on haralick's features on segmented histopathological images, Computer Applications and Industrial Electronics (ICCAIE), 2011 IEEE International Conference on (2011) 87–90.

[18] S. Hilado, L. Gan Lim, R. Naguib, E. Dadios, J. Avila, Implementation of wavelets and artificial neural networks in colonic histopathological classification., Journal of Advanced Computational Intelligence and Intelligent Informatics 18 (2014) 792–797.

[19] N. K, J. AK, S. B., Prostate cancer detection: Fusion of cytological and textural features, J Pathol Inform (2011).

[20] S.-X. Rao, D. M. Lambregts, R. S. Schnerr, W. van Ommen, T. J. van Nijnatten, M. H. Martens, L. A. Heijnen, W. H. Backes, C. Verhoef, M.-S. Zeng, G. L. Beets, R. G. Beets-Tan, Whole-liver ct texture analysis in colorectal cancer: Does the presence of liver metastases affect the texture of the remaining liver?, United European Gastroenterology Journal 2 (2014) 530–538.

[21] M. Roula, A. Bouridane, P. Miller, A quadratic classifier based on multispectral texture features for prostate cancer diagnosis., Proceedings of the Seventh International Symposium on Signal Processing Applications (2003) 37–40.

[22] R. Peyret, A. Bouridane, S. A. Al-Maadeed, S. Kunhoth, F. Khelifi, Texture analysis for colorectal tumour biopsies using multispectral imagery, 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (2015) 7218–7221.

[23] M. A. Tahir, A. Bouridane, Novel round-robin tabu search algorithm for prostate cancer classification and diagnosis using multispectral imagery, IEEE Transactions on Information Technology in Biomedicine 10 (2006) 782–793.

[24] R. Khelifi, M. Adel, S. Bourennane, Multispectral texture characterization: application to computer aided diagnosis on prostatic tissue images, EURASIP Journal on Advances in Signal Processing 2012 (2012) 1–13.

[25] K. Masood, N. Rajpoot, Texture based classification of hyperspectral colon biopsy samples using clbp, 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro (2009) 1011–1014.

[26] A. Greenblatt, C. Mosquera-Lopez, S. Agaian, Quaternion neural networks applied to prostate cancer gleason grading, 2013 IEEE International Conference on Systems, Man, and Cybernetics (2013) 1144–1149.

[27] M. Mirmehdi, M. Petrou, Segmentation of color textures, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (2000) 142–159.

[28] H. E. Maia, A. Hammouch, D. Aboutajdine, Color-texture analysis by mutual information for multispectral image classification, in: 2009 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, pp. 359–364.

[29] C. H. Chan, J. Kittler, K. Messer, Multispectral local binary pattern histogram for component-based color face verification, in: 2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp. 1–7.

[30] T. Maenpaa, M. Pietikainen, J. Viertola, Separating color and pattern information for color texture discrimination, in: Object recognition supported by user interaction for service robots, volume 1, pp. 668–671 vol.1.

[31] R. M. Coliban, M. Ivanovici, Color and multispectral texture characterization using pseudo-morphological tools, in: 2014 IEEE International Conference on Image Processing (ICIP), pp. 630–634.

[32] P. R. Chowdhury, B. Deshmukh, A. K. Goswami, S. S. Prasad, Neural network based dunal landform mapping from multispectral images using texture features, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing 4 (2011) 171–184.

[33] A. Bendada, M. A. Akhloufi, Multispectral face recognition in texture space, in: 2010 Canadian Conference on Computer and Robot Vision, pp. 101–106.

[34] H. I. Kim, S. H. Lee, Y. M. Ro, Multispectral texture features from visible and near-infrared synthetic face images for face recognition, in: 2015 IEEE International Symposium on Multimedia (ISM), pp. 593–596.

[35] Y. Zhang, Texture-integrated classification of urban treed areas in high-resolution color-infrared imagery, in: 2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems, pp. 1359–1365 Vol. 67, No. 12.

[36] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 24 (2002) 971–987.

[37] G. Zhao, M. Pietikainen, Dynamic texture recognition using local binary patterns with an application to facial expressions, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (2007) 915–928.

[38] X. Shi, Independent Component Analysis, Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 60–83.

[39] D. H. Wolpert, Stacked generalization, Neural Networks 5 (1992) 241–259.