MNK Publication

International Journal of Latest Research in Science and Technology **ISSN (Online):2278-5299**
Volume 3, Issue 6: Page No.192-200, November-December 2014
https://www.mnkpublication.com/journal/ijlrst/index.php

REGION-BASED APPROACHES AND DESCRIPTORS EXTRACTED FROM THE CO-OCCURRENCE MATRIX

[1]Loris Nanni, [2]Sheryl Brahnam, [3] Stefano Ghidoni, [4] Emanuele Menegatti
[1] Department of Information Engineering at the University of Padua, Padua, Italy
[2]Computer Information Systems, Missouri State University, USA
[3]Department of Information Engineering at the University of Padua, Padua, Italy
[4] Department of Information Engineering at the University of Padua, Padua, Italy

Abstract-* Recently proposed texture descriptors extracted from the co-occurrence matrix across several datasets is surveyed and validated in this paper; moreover, two new methods for extracting features from the Gray Level Co-occurrence Matrix (GLCM) are proposed. The descriptors are extracted not only from the entire GLCM but also from subwindows. These texture descriptors are used to train a support vector machine. We also explore region-based approaches, which use different methods to divide each image into two different regions; different descriptors are extracted from each region. In this work methods based on saliency detection, edge detection, and wavelets are compared, and some of their fusions are reported as well. Region-based approaches are combined with different methods for extracting features from the GLCM and with three state-of-the-art descriptors: local ternary patterns, local phase quantization, and rotation invariant co-occurrence among adjacent local binary patterns. Experimental results show that the tested approaches improve performance of standard methods. The generality of the proposed descriptors is demonstrated on 15 datasets, and different statistical comparisons based on the Wilcoxon signed rank test are reported that confirm the goodness of the proposed approaches. Experiments show that the new methods for extracting features from the GLCM greatly improve the standard features that are typically extracted, and that the region-based approach boosts the performance of texture descriptors extracted from the whole image. The MATLAB source code of all the proposed approaches will be made available to the public at https://www.dei.unipd.it/node/2357.

Keywords - Co-occurrence matrix; texture descriptors; support vector machine; ensemble; region-based

I. INTRODUCTION

Although multi-objective texture analysis is frequently involved in image classification, a satisfactory, overarching definition of texture is still lacking. An interesting catalogue of texture definitions can be found in [1] that illustrates the many different ways this term can be characterized. Texture, for instance, can be viewed globally as a pattern composed of repeated local subpatterns [2] or, alternatively, as a region where a set of local properties remain constant, vary gradually, or are approximately periodic [3]. Very different methods for managing texture have been developed based on how texture is defined. Some of the best methods reported in the literature include the Scale-Invariant Feature Transform (SIFT) [4], Speeded Up Robust Features (SURF) [5], Histogram Of Oriented Gradients (HOG) [6], Gradient Location and Orientation Histogram (GLOH) [7], Region Covariance Matrix (RCM) [8], edgelet [9], Gray Level Co-occurrence Matrix (GLCM) [10], and Local Binary Patterns (LBP) [11].

One of the earliest, most studied, and extensively used methods for analyzing texture is the GLCM, originally proposed by Haralick [12] in 1979 for analyzing satellite images. GLCM is a set of features, or descriptors, that are evaluated starting from a histogram. Within the last decade, GLCM has become the focus of several research groups developing new methods for increasing the discriminability of GLCM descriptors. In [13], for instance, different values of the distance parameter that influences the GLCM are examined, and in [14] features are extracted from areas presenting high discrimination by weighted summation of the GLCM elements. In [15] descriptors are derived from the GLCM, which is formed by calculating the gradient value of each pixel in the neighborhood of interest points. In [16] the edge orientation co-occurrence matrix of superior order is combined with GLCM, in this way taking into consideration both the gray levels of the image pixels and the local features as edges. Several studies have performed multiscale analyses with GLCM. In [17] and in [18], for instance, GLCM descriptors are extracted by varying window sizes. In [19] the image is rescaled multiple times, and the co-occurrence descriptors are extracted from each rescaling.

In the last couple of years, region-based approaches have been proposed that significantly improve texture descriptor performance (see, [20] and [21]). Region-based approaches separate the texture image into two regions, or maps, using powerful preprocessing methods. In [20], for instance, a given image is split into two maps using the Difference of Gaussians (DoG) filter: one map corresponds to the "positive" and the other map the "negative" sides of the image edges. Textural information is then extracted from the two maps using several descriptors. A similar approach is taken in [21], where Sobel filtering splits the texture into an

Publication History

Manuscript Received	:	*15 December 2014*
Manuscript Accepted	:	*22 December 2014*
Revision Received	:	*26 December 2014*
Manuscript Published	:	*31 December 2014*

edge and a non-edge region. Different texture descriptors (e.g., LBP and LTP) are extracted from the original image and then combined with each of the two maps. This technique is particularly intriguing since it can be combined with many state-of-art texture descriptors, an opportunity exploited in [22] and [23], in which several descriptors are coupled with a variant of this technique.

The primary goal of this work is to report a comparison for extracting features given the co-occurrence matrix using region-based approaches ([22] and [24]); the comparison spans across a large set of very different computer vision problems.

This work expands some other reported multiscale comparisons ([25] and [26]) for extracting features from the co-occurrence matrix, where a smaller set of datasets are used for assessing the performance. Specifically, five different extraction methods are investigated (detailed in section 2), with two (numbers iv and v) tested in this paper for the first time:

i. Extracting descriptors using the standard approach proposed by Haralick [10];
ii. Extracting gray-level run-length features [26];
iii. Extracting descriptors from different 2D shapes by considering the GLCM as a 3D shape, where the 2D shapes are obtained by intersecting the co-occurrence matrix with a set of horizontal planes at given heights [26];
iv. Extracting fractal dimensions decomposing the GLCM into a set of binary images. The decomposition of the input image is achieved employing the Two-Threshold Binary Decomposition (TTBD) algorithm [27];
v. Extracting curvature based shape features [28], where 2D shapes are obtained by intersecting the co-occurrence matrix with a set of horizontal planes at given heights. For each shape a set of features is extracted.

In addition, four methods are applied for region-based approaches:

i. Separation of the texture image into two different regions according to edge and non-edge pixels, as in [21];
ii. Separation of the texture image into different regions according to Daubechies (db4) wavelet decomposition [29]. Each image is divided in two regions: pixels with value higher than mean value (mv) and pixels with value lower than mv;
iii. Separation of the texture image in different regions according to saliency detection [30]. For each image we extract two regions given some threshold: the first contains the pixels with higher saliency (those above the threshold), and the second contains the pixels with lower saliency (those below the threshold). Two different thresholds are tested.
iv. Separation of the texture image in two different regions by filtering the image using DoG, as in [20]. DoG is used to compute two maps corresponding to the "positive" and the "negative" sides of the image edges (i.e., the two different regions).

The region-based approaches are combined with two methods for extracting features from the co-occurrence matrix and with three state-of-the-art texture descriptors: the local ternary pattern (LTP) [31], local phase quantization (LPQ) [32], and the Rotation Invariant Co-occurrence among adjacent Local Binary Patterns (RICLBP) [33]. In other words, these texture descriptors are extracted from each of the two regions produced by the region-based approaches. Histograms are extracted, and for each histogram a specific Support Vector Machine (SVM) [34] is trained. Finally, the partial scores obtained by the different SVMs are combined by sum rule. Several fusions are also performed that investigate the best set of features.

Since Haralick-based features, LPQ, LTP, and RICLBP are widely used in the literature, this work has much practical value for other researchers. The approaches presented in this paper are evaluated across several datasets, described in section III, representing very different image classification problems. The results presented in section IV clearly confirm that region-based approaches outperform the texture descriptors extracted from the entire image. Moreover, the experimental results show that the proposed set of features discovered in the fusion experiments performs much better than the widely used standard set of features proposed by Haralick [10].

II. PROPOSED SYSTEM

Every approach evaluated in the experimental section is presented in this section, with the different extraction methods from the co-occurrence matrix detailed first, followed by a description of the region-based approaches. The base classifier used in all experiments is an SVM with a radial basis function kernel. To reduce computation time, the following set of parameters were used in all experiments: gamma=0.1 and cost= 1000.

A. *GLDM Co-Occurrence Matrix*

The GLDM [12] is a specific co-occurrence matrix that is obtained as the histogram on a 2D domain of dimension NGL×NGL, where NGL is the number of gray levels in the image (normally 256). The co-occurrence matrix counts the number of gray level transitions between two given pixel values such that the bin of the histogram whose coordinates are equal to the value is incremented. The two pixels selected for comparison depend on two parameters: d, the distance between the two pixels, and θ, the direction in which they are aligned. To illustrate, if d=1 and θ=0 then the two pixels would be adjacent to each other and lie on the same row. Four directions are considered in this work: the horizontal (H), the vertical (V), the diagonal top left-bottom right, or right-down (RD), and the top right-bottom left, or left-down (LD).

B. *Standard Haralick Statistics*

The idea of using statistical indicators of GLDM to describe texture in an image was originally proposed by Haralick in [12]. In the experimental section, this approach is labelled HAR, and the following indicators are evaluated (for details, see [26]):

- Energy
- Correlation
- Inertia
- Entropy
- Inverse difference moment

- Sum average
- Sum variance
- Sum entropy
- Difference average
- Difference variance
- Difference entropy
- Information measure of correlation 1
- Information measure of correlation 2

A set of 13 descriptors is calculated from each co-occurrence matrix, which is evaluated at θ ={0°, 45°, 90°, 135°} and *d*={1, 3}. The HAR descriptor is the concatenation of the features extracted for each distance and orientation value.

Reported in the experimental section are the performances of the following:

- HR: where features are extracted from the entire image only. Several comparisons of different parameters settings for HR are reported in [25]; accordingly, only the best configuration reported in [25] is tested in this paper.
- HRsub: where a feature set is extracted from the entire co-occurrence matrix as well as from each subwindow. In this paper four subwindows are extracted, and each feature set trains a separate SVM. All 5 SVMs (the four subwindows plus the entire co-occurrence matrix) are then combined by weighted sum rule (weight of 4 for the SVM trained on the whole matrix and weight of 1 for the other four SVMs). The four subwindows are labelled SW1-4, and the coordinates defining each subwindow are the following: SW1: (0, 0) to (127, 127), SW2: (128, 128) to (255, 255), SW3: (128, 0) to (255, 128), and SW4: (0, 128) to (128, 255).

C. *Shape*

The approach called SHAPE is an exploration of the co-occurrence matrix as a 3D function and has been explored in some detail in other papers such as [26] and [25]. The basic idea is to intersect the GLDM with a set of horizontal planes at given heights. A set of features is extracted from the contours of the intersection, which defines a complex shape made up of one or more extractable blobs. In SHAPE features are extracted from the main blob, i.e., the blob with the largest area. The main blob is fitted to an ellipse to simplify analysis. Using an ellipse makes the comparison among curves much easier and offsets the resultant loss of information.

Level curves are considered towards the base of the co-occurrence matrix, starting at height 1 and then going until height 19, with a distance of 2 between two consecutive planes. Level curves are all at a relatively low height because that region is very stable. The upper part of the co-occurrence matrix is much more unstable because of image noise. For this reason, the co-occurrence matrix is not normalized since normalization to the highest bin would introduce instabilities. Other types of normalization could be performed with respect to the total volume of the co-occurrence matrix, but results would depend on the size of the original image (which would be constant in most cases, thereby making the normalization irrelevant).

A set of descriptors extracted from the ellipses is calculated from the co-occurrence matrix for each level curve. A feature set describing all levels are then jointly analyzed, from which a final set of nine features is selected that describe the evolution of the level curves (see [25] for details).

These features provide a characterization of the input image that can be used as input for a classifier, which is the same idea exploited in the HAR approach. SHAPE features are evaluated on the entire co-occurrence matrix and on 12 subwindows of the GLDM whose coordinates are defined as follows: #1: (0, 0) to (127, 127); #2: (128, 128) to (255, 255); #3: (0, 0) to (191, 191); #4: (64, 64) to (255, 255); #5: (0, 0) to (95, 95); #6: (31, 31) to (95, 95); #7: (63, 63) to (127, 127); #8: (95, 95) to (159, 159); #9: (127, 127) to (191, 191); #10: (159, 159) to (223, 223); #11: (191, 191) to (255, 255); and #12: (63, 63) to (191, 191). Several experiments using the entire GLDM along with these same subwindows are reported in [26] and [25].

For each of these 13 windows (the 12 subwindows and the entire GLDM) a different feature vector is extracted, and each feature vector trains a separate SVM. Each of the 13 vectors is derived from co-occurrence matrices evaluated at θ ={0°, 45°, 90°, 135°} and *d* = {1, 3}, and the feature vector is obtained by concatenating the features extracted for each value of the distance. All 13 SVMs are combined by weighted sum rule, where a weight of 1 is assigned to the first five SVMs (i.e., to the SVMs based on the entire GLDM and subwindows #1, #2, #3 and #4), and a weight of 0.5 is assigned to the remainder (as in [26]). In the experimental section, SH refers to the case where features are extracted from the entire co-occurrence matrix only, while SHsub is the method based on all 13 windows.

D. *Gray-Level Run-Length Features (GL)*

GL [35] features are extracted from a run-length matrix that is derived from the characteristics of the gray level runs within an image. A gray level run is a set of consecutive pixels having the same value. A run length is the size of the set. Each location p(*i*,*j*) of the run-length matrix counts the number of runs of length *j* given gray level *i*. Several descriptors can be derived from the run-length matrix, as described in [35]. The following descriptors are used in the GL experiments, which comprise the same set used in [26]:

- Short Run Emphasis (SRE)
- Long Run Emphasis (LRE)
- Gray Level Nonuniformity (GLN)
- Run Length Nonniformity (RLN)
- Run Percentage (RP)
- Low Grey-Level Run Emphasis (LGRE)
- High Grey-Level Run Emphasis (HGRE)
- Short Run Low Grey Level Emphasis (SRLGE)
- Short Run High Grey Level Emphasis (SRHGE)
- Long Run Low Grey Level Emphasis (LRLGE)
- Long Run High Grey Level Emphasis (LRHGE)

The GL approach has its own orientation: all values considered in [26] are evaluated in our system, namely θ_{GL}={0°, 45°, 90°, 135°}.

In this paper, the descriptors described above are calculated from a run-length matrix that is evaluated on the

GLDM, using θ ={0°, 45°, 90°, 135°} and d={1, 3}. The final GL descriptor is obtained by concatenating all the features for all values of θ and d.

The performance of the following GL variants is also reported:

- GR: as in HR (described in section II.B) but using the GL descriptors;
- GRsub: as in HRsub (also described in section II.B) but using the GL descriptors.

E. *Shape Fractal Analysis (SFTA)*

The Segmentation-based Fractal Texture Analysis, or SFTA, method proposed in [27] is a feature extraction algorithm that decomposes a given image into a set of binary images through the application of what the authors call the Two Threshold Binary Decomposition (TTBD). For each resulting binary image, fractal dimensions of its region boundaries are calculated that describe the texture patterns.

TTBD takes an input grayscale image and returns a set of binary images by first computing a set of T threshold values from the gray level distribution information in input image. This is accomplished by recursively applying to each image region the multilevel Otsu algorithm [36], an algorithm that quickly finds the threshold that minimizes the input image intra-class variance until the desired number of thresholds is obtained. The input image is decomposed into a set of binary images by selecting pairs of thresholds from T and applying a two-threshold segmentation (see [27] for details).

The SFTA extraction algorithm extracts a feature vector from the resulting binary images' size, mean gray level, and the boundaries' fractal dimension. Fractal measurements are used to describe the boundary complexity of objects, with each regions' boundaries of a binary image represented as a border image. The fractal dimension is computed from each boarder image using a box counting algorithm (see [27] for details).

In this paper, first level curves are obtained as described in section II.C then each binary image is described by the SFTA extraction algorithm. As in SHAPE, level curves are considered towards the base of the co-occurrence matrix, starting at height 1 and then going until height 19, with a distance of 2 between two consecutive planes. For each level curve, a binary image is created; the binary image is then described by SFTA. The descriptors of the different level curves are concatenated to describe a given co-occurrence matrix.

In the experimental section the performance of the following SFTA variants are also reported:

- SF: as in HR (described in section II.B) but using the SFTA features;
- SFsub: as in HRsub (also described in section II.B) but using the SFTA features;

F. *Shape Curvature Histogram (SCH)*

SCH, proposed in [28], is a feature that describes the curvature of shapes within an image as a compact histogram. An advantage of SCH is that it does not require shapes to be within closed boundaries, unlike other shape-based feature extraction methods.

To extract SCH features, first an edge map is generated, using the Canny edge detector [37]. The gradient direction is then computed on the pixels in the edge map. Once the orientation of each pixel is obtained, the curvature for an edge pixel is calculated as the difference between the maximum and minimum gradient angle of all pixels defining a given neighborhood (see [28] for details). Finally, a histogram is created based on the curvature values for the edge pixels.

In the experiments reported in this paper, different binary images are built as in SFTA. Then each binary image is described by the SCH descriptor. The following SCH variants are also reported:

- SC: as in HR (described in section II.B) but using the SCH features;
- SCsub: as in HRsub (also described in section II.B) but using the SCH features.

G. *Region-Based Descriptors*

Inspired by the edge-based LBP variant (Edge), proposed in [21], region-based descriptors are based on the evidence that when an observer needs to fix attention on a particular object, the most likely perceived locations are those that present the highest spatial frequency edge information [38]. The Edge descriptor is computed as follows:

- Obtain the LBP image (LBPI) (see, [39]);
- Apply Sobel to detect the edges in the original image. Two binary maps are created from the edge information: the edge map (E), where edge pixels are set to 1 and non-edge pixels to 0, and the non-edge map (NE), where edge pixels are set to 0 and non-edge pixels to 1;
- Combine LBPI with the E and NE masks to obtain two histograms: HE, for edge pixels, and HNE, for non-edge pixels, (see [21] for details);
- Form the final histogram using weighted concatenation, such that $H=[w_E \times H_E,\ w_{NE} \times H_{NE}]$, $w_E > w_{NE}$], where wE and wNE represent the empirically determined weights that express the greater relevance of edge regions in capturing the viewer's visual attention.

A slightly different method is employed when using SVM as the classifier:

- Extract a set of descriptors (e.g., LBP, LTP, or LPQ) to obtain the labeled image DescI (each DescI is extracted in a similar way to LBPI), where Desc represents the name of the descriptor;
- Compute two binary maps: Map+ and Map-, using one of the following methods: Sobel (as in [21]), saliency, wavelet or Difference of Gaussians (the last three methods are described in the next three subsections);
- Compute the two histograms, H+ and H-, by combining DescI with Map+ and Map-, respectively;
- Train two different SVMs on H+ and H- and combine results by sum rule.

In the experimental section, the method based on Sobel for finding Map+ and Map- is labeled Ed.

In using this approach to build the co-occurrence matrix, first Map+ and Map- are extracted. Then two different co-occurrence matrices are calculated: the first considering the pixels that belong to Map+ and the second considering the pixels that belong to Map-.

H. *Saliency (SAL)*

The method proposed in [30] is used for extracting saliency maps from the image. In this method two regions are extracted from each original image: Map+, which contains the pixels with highest saliency (i.e., those pixels above a given threshold), and Map-, which contains the pixels with the lowest saliency (i.e., those pixels below the given threshold). In the experiments reported in this paper, two different threshold values were evaluated: 0.7 and 0.5. Thus, for each image, a total of two saliency maps and four histograms were extracted.

I. *Wavelet (WA)*

Wa is a method for computing the two binary maps of an image that is based on the wavelet decomposition [29] by Daubechies with four wavelets (db4), where the horizontal (CH), vertical (CV), and diagonal (CD) coefficients matrices are considered. Each of these matrices are resized to the size of the original image. The mean value (mv_CH, mv_CV, and mv_CD) of each of these images is then calculated. For each of the mv values, the original image is divided into two regions: Map+, which contains the pixels of {CH, CV, CD} with a value higher than {mv_CH, mv_CV, mv_CD}, and Map-, which contains the pixels of {CH, CV, CD} with a value lower than {mv_CH, mv_CV, mv_CD}. In this way six histograms are extracted from each image.

J. *Difference of Gaussians (Do)*

In [20], the Difference of Gaussians filter was used to compute the two maps from a given image: Map+ and Map-, with Map+ corresponding to the "positive" and Map- corresponding to the "negative" sides of the image edges. Textural information is then extracted from these two maps. A Gaussian lowpass filter with size 5 and sigma equal to 1 and 4 are used in the Do experiments reported in this paper.

III. DATASETS

The approaches proposed in this paper were tested across several datasets to assess their generalizability. These datasets represent different computer vision problems:

- PS: the Pap Smear dataset [40], containing images representing cells that are used in the diagnosis of cervical cancer.
- VI: the dataset in [26], containing images of viruses extracted using negative stain transmission electron microscopy. The 10-fold validation division of images used in [26] is used in the experiments reported in this paper. However, the masks for subtracting image backgrounds were not utilized. Instead, features were extracted from the entire image since this produced better results.
- CH: the dataset in [41], containing fluorescent microscopy images taken from Chinese Hamster Ovary cells and belonging to five different classes.
- SM: the dataset in [42], containing images extract from video-based smoke detection surveillance systems. The same division of the dataset into training/testing sets in [42] is used in all experiments on SM reported in this paper.
- HI: the Histopatology dataset [43], containing images from different organs representative of the four fundamental tissues.
- BR: the dataset in [44], containing 273 malignant and 311 benign breast cancer images.
- PR: a dataset containing 118 DNA-binding Proteins and 231 Non-DNA-binding proteins. Texture descriptors are extracted from the 2D distance matrix, which represents each protein, and this matrix is obtained from the 3D tertiary structure of a given protein (considering only atoms that belong to the protein backbone, see [45] for details).
- HE: the 2D HeLa dataset [41], containing single cell images, divided into 10 staining classes, from fluorescence microscope acquisitions on HeLa cells.
- LO: the Locate endogenous mouse sub-cellular organelles dataset [46], containing 502 images unevenly distributed among 10 classes of endogenous proteins or features of specific organelles.
- TR: the Locate transfected mouse sub-cellular organelles dataset [46], containing 553 images unevenly distributed in 11 classes of fluorescence-tagged or epitope-tagged proteins transiently expressed in specific organelles.
- PI: the dataset in [47] containing pictures extracted from digitalized pages of the Holy Bible of Borso d'Este, duke of Ferrara (Italy) from 1450 A.D. to 1471 A.D. PI is composed of 13 classes, characterized by a clear semantic meaning and significant search relevance.
- RN: a dataset containing 200 fluorescence microscopy images evenly distributed among 10 classes of fly cells subjected to a set of gene-knockdowns using RNAi. The cells were stained with DAPI to visualize their nuclei.
- HP: the full HEp-2 dataset, containing cell images, extracted from the specimen images of the positive sera of 419 patients. A public set of 13596 cell images of 6 classes is available [48].
- PA: the dataset in [49], containing 2338 paintings by 50 painters representative of 13 different painting styles: abstract expressionism, baroque, constructivism, cubism, impressionism, neo-classical, pop art, post impressionism, realism, renaissance, romanticism, surrealism, and symbolism. A split training/testing set is provided by the authors of [49]; those sets were used in all the experiments on PA reported in this paper.
- LE: a dataset containing images of several species of Brazilian flora [50]. A total of 400 samples, divided into 20 classes (20 samples per class), were collected. Three windows (128×128 pixels) were extracted manually from each sample making a total of 1200 textures. The protocol used in the experiments reported in this paper was a fivefold cross validation technique with the constraint that all windows extracted from a given leaf had to belong either to the training set or to the testing set, not both.

A descriptive summary of each dataset along with the URL where each dataset can be downloaded is reported in table I. If a dataset contains RGB images, these were converted to gray level images before the feature extraction step.

The testing protocol was the fivefold cross validation method, except for the VI, SM, and PA datasets, where the specific protocols and testing/training sets were used (the protocols were obtained from the creators of each of these datasets).

TABLE I DESCRIPTIVE SUMMARY OF THE DATASET

Dataset	#Classes	#Samples	Sample Size	URL for Download
PS	2	917	Various	http://labs.fme.aegean.gr/decision/downloads
VI	15	1500	41×41	http://www.cb.uu.se/~gustaf/virustexture
CH	5	327	512×382	http://ome.grc.nia.nih.gov/iicbu2008/hela/index.html#cho
SM	2	2868	100×100	http://staff.ustc.edu.cn/~yfn/vsd.html
HI	4	2828	various	http://www.informed.unal.edu.co/histologyDS
BR	2	584	various	upon request to Geraldo Braz Junior [ge.braz@gmail.com]
PR	2	349	various	upon request to Loris Nanni [nanni@dei.unipd.it]
HE	10	862	512×382	http://ome.grc.nia.nih.gov/iicbu2008/hela/index.html
LO	10	502	768×512	http://locate.imb.uq.edu.au/downloads.shtml
TR	11	553	768×512	http://locate.imb.uq.edu.au/downloads.shtml
PI	13	903	various	http://imagelab.ing.unimo.it/files/bible_dataset.zip
RN	10	200	1024×1024	http://ome.grc.nia.nih.gov/iicbu2008/rnai/index.html
HP	6	13596	various	http://mivia.unisa.it/datasets/biomedical-image-datasets/hep2-image-dataset/
PA	13	2338	various	http://www.cat.uab.cat/~joost/painting91.html
LE	20	1200	128×128	Upon request to bruno@ifsc.usp.br

IV. EXPERIMENTAL RESULTS

The performance indicator used in all experiments is the area under the ROC curve (AUC) because it provides a better overview of classification results [51]. In the multiclass problem, AUC is calculated using the one-versus-all approach, where a given class is considered as "positive" and all the other classes are considered as "negative," and the average AUC is reported in all tables. The last row labelled *Av* in all the tables included in this section reports the average performance on all 15 datasets.

The aim of the first experiment reported in table II is to establish the usefulness of extracting features not only from the entire co-occurrence matrix but also from different subwindows. By examining table II, it is clear that all methods improve when features are extracted from GLDM subwindows. Even the standard HR improves when coupled with subwindow extraction. To statistically validate these experiments, the Wilcoxon signed rank test [52] was used for all methods. The HR version based on subwindows (HRsub) outperforms with p-value 0.05 the version based on the entire co-occurrence matrix. Thus, this experiment validates previous results [26] that were obtained using a smaller set of datasets, lending weight to the superiority of the co-occurrence matrix subwindow approach.

The aim of the second experiment reported in table III is to show the performance gain that is possible by fusing different descriptors extracted from the co-occurrence matrix. The descriptors chosen for this experiment were the following:

- W2: the weighted sum rule between HRsub (weight 2) and GRsub (weight 1);
- W3: the weighted sum rule of HRsub (weight 2), GRsub (weight 1), and SHsub (weight 0.5);
- W5: the weighted sum rule of HRsub (weight 2), GRsub (weight 1), SHsub (weight 0.5), SFsub (weight 0.5), and SCsub (weight 0.5).

TABLE III USEFULNESS OF EXTRACTING FEATURES FROM SUBWINDOWS OF THE CO-OCCURRENCE MATRIX

Dataset	HR	HRsub	GR	GRsub	SH	SHsub
PS	89.3	92.1	80.2	84.2	82.5	86.6
VI	95.9	96.8	88.8	93.1	84.6	89.9
CH	99.7	99.8	99.2	98.9	98.7	98.8
SM	99.2	99.3	99.0	99.2	98.2	99.0
HI	88.8	89.9	83.3	87.6	82.3	85.3
BR	92.7	93.5	84.8	90.5	88.7	91.9
PR	90.6	91.1	84.3	89.8	81.2	82.6
HE	97.0	97.3	92.2	94.2	94.1	94.7
LO	99.1	99.5	98.4	99.0	95.5	96.9
TR	98.9	99.2	98.0	98.6	90.7	90.8
PI	87.8	90.4	81.1	85.7	81.0	85.0
RN	95.1	95.0	89.2	90.8	95.2	95.7
HP	88.9	89.9	86.3	86.5	80.8	81.8
PA	84.1	87.5	79.8	84.1	78.1	83.4
LE	97.2	97.4	90.4	92.7	89.4	93.9
Av	93.6	94.6	89.0	91.7	88.1	90.4

Dataset	SF	SFsub	SC	SCsub
PS	78.9	80.0	79.3	80.2
VI	74.7	84.2	76.4	82.3
CH	98.2	98.8	97.4	93.5
SM	97.4	98.1	97.7	98.1
HI	79.6	81.7	74.7	76.7
BR	76.6	83.9	86.8	88.6
PR	83.3	87.0	81.6	84.8
HE	94.1	95.1	84.5	88.9
LO	98.4	98.5	97.3	97.6
TR	97.2	97.7	95.9	96.8
PI	79.0	81.8	76.5	80.6
RN	87.4	88.2	84.4	85.0
HP	84.4	85.3	86.7	87.1
PA	76.9	81.1	76.6	80.2
LE	90.6	91.9	85.2	87.7
Av	86.4	88.9	85.4	87.2

TABLE IIIII AUC OBTAINED USING FUSION APPROACHES

Dataset	HR	HRsub	W2	W3	W5
PS	89.3	**92.1**	91.4	91.6	91.4
VI	95.9	**96.8**	96.7	96.7	96.6
CH	99.7	99.8	**99.9**	**99.9**	**99.9**
SM	99.2	99.3	**99.5**	**99.5**	**99.5**
HI	88.8	89.9	90.8	90.8	**90.9**
BR	92.7	93.5	94.3	94.7	**95.1**
PR	90.6	91.1	92.4	91.9	**92.7**
HE	97.0	97.3	97.2	97.3	**97.5**
LO	99.1	99.5	99.5	**99.6**	**99.6**
TR	98.9	99.2	99.3	**99.4**	**99.4**
PI	87.8	90.4	90.9	**91.0**	90.7
RN	95.1	95.0	96.2	96.9	**97.2**
HP	88.9	89.9	90.8	**90.9**	90.7
PA	84.1	87.5	88.0	88.5	**88.8**
LE	97.2	97.4	**97.5**	97.4	97.4
Av	93.6	94.6	95.0	95.1	**95.2**

Table III shows that the best result was obtained by W5, which confirms that all the subwindow methods extract different information from the co-occurrence matrix.

Moreover, all the ensembles (W2, W3, and W5) outperform HRsub with p-value 0.05 using Wilcoxon signed rank test.

TABLE IVV AUC OBTAINED USING THE REGION-BASED APPROACHES WITH TEXTURE DESCRIPTORS

Dataset	Preprocessing Method						
	O	Ed	Sal	Wa	Do	All	All+O
LTP							
PS	91.4	87.8	89.3	87.7	86.6	89.0	**91.7**
VI	93.5	94.0	93.4	93.5	**94.9**	94.4	94.1
CH	99.9	99.9	**100**	99.9	99.9	99.9	99.9
HI	91.6	92.7	90.6	92.4	92.6	**92.8**	92.3
BR	96.9	96.4	96.1	96.2	95.9	96.6	**97.6**
PR	89.7	87.0	90.3	87.8	85.1	90.0	**93.2**
HE	98.6	98.6	98.6	98.5	98.6	**98.9**	98.7
LE	99.5	99.5	**99.7**	99.6	99.3	**99.7**	99.6
LT	99.3	99.3	**99.6**	99.4	99.4	**99.6**	99.5
SM	99.7	99.7	99.5	99.7	**99.8**	99.7	99.7
PIC	92.9	90.2	**93.3**	91.6	89.3	92.5	92.9
RN	97.0	97.0	**97.4**	97.0	96.6	97.2	97.3
HP	88.8	90.3	91.0	91.0	91.2	91.6	**91.7**
PA	89.0	90.2	89.7	89.5	90.2	90.6	**90.7**
LE	**97.9**	**97.9**	97.2	97.8	97.8	**97.9**	**97.9**
Av	95.0	94.7	95.0	94.7	94.5	95.4	**95.8**
RICLBP							
PS	91.8	92.0	92.5	91.9	93.3	**92.7**	92.4
VI	97.6	97.7	97.4	97.5	**97.8**	97.7	**97.8**
CH	99.2	99.8	**99.9**	99.7	99.8	99.8	99.7
HI	92.8	93.7	93.4	93.6	93.8	**94.0**	93.5
BR	92.8	93.8	94.0	94.6	93.9	**95.0**	93.9
PR	88.6	89.3	89.2	88.5	88.0	**89.6**	89.4
HE	97.3	98.4	**98.7**	98.2	98.8	98.6	98.2
LE	99.0	99.5	**99.6**	99.3	**99.6**	99.5	99.4
LT	98.7	99.1	**99.4**	98.5	99.3	99.2	99.0
SM	99.8	99.8	99.8	99.8	**99.9**	99.8	**99.9**
PIC	90.1	91.5	**94.7**	91.3	92.7	93.4	94.3
RN	96.6	96.8	**97.4**	96.8	97.2	97.1	96.9
HP	92.5	93.9	93.9	94.1	94.1	94.4	**94.7**
PA	85.9	88.6	88.8	88.3	88.5	89.4	**90.5**
LE	97.4	97.7	97.6	97.6	98.0	97.8	**98.1**
Av	94.6	95.4	95.8	95.3	95.6	**95.9**	95.8
LPQ							
PS	90.2	89.3	90.8	90.4	90.3	**90.9**	90.7
VI	94.9	94.4	94.5	94.7	94.1	**95.2**	**95.2**
CH	99.2	99.6	**99.8**	99.6	99.6	**99.8**	99.6
HI	92.0	92.9	92.7	92.8	92.5	**93.2**	92.7
BR	95.7	97.3	96.5	96.2	96.1	**96.8**	96.3
PR	86.2	88.7	**90.5**	88.9	86.6	90.2	88.7
HE	97.2	98.0	**98.5**	98.0	98.2	98.4	98.0
LE	97.6	98.2	**99.4**	98.1	98.7	98.7	98.2
LT	97.7	98.4	**99.2**	97.6	98.8	98.8	98.3
SM	99.8	99.8	**99.9**	99.8	99.8	**99.9**	**99.9**
PIC	90.7	91.7	**95.3**	92.1	91.6	94.0	94.3
RN	95.2	94.9	**95.7**	94.5	95.5	95.2	95.3
HP	91.0	91.4	92.0	91.5	90.9	92.5	**93.2**
PA	88.3	89.1	89.4	89.6	89.4	90.1	**91.0**
LE	99.0	98.6	98.8	98.8	98.4	98.8	**99.0**
Av	94.3	94.8	**95.5**	94.8	94.7	**95.5**	95.3

The aim of the third experiment reported in tables IV and V is to show the performance gain that can be achieved using the region-based approaches combined with state of the art texture descriptors (LTP, LPQ, and RICLBP). Also reported in table IV is the performance obtained by the following:

- O: the specific texture descriptor being evaluated (LTP, LPQ, and RICLBP) applied alone to the original image;
- All: the fusion by sum rule of Sal, Ed, and Wa (note: before fusion, the scores of each method were normalized to mean 0 and standard deviation 1. The method labeled Do is not included because it did not enhance fusion performance);
- All + O: the fusion by sum rule of O, Sal, Ed, and Wa.

In table V the performance obtained combining HR and GR with the region-based approaches is reported. In addition to All and All+O described above, the performance obtained by the following methods is reported:

- All + S: the fusion by sum rule of All and Xsub, with X∈{HR, GR};
- All + 2 × S: the fusion by weighted sum rule of Xsub, with weight 2, and All with weight 1.

The results reported in IV and V are interesting in this regard: the region-based approaches outperform with a p-value of 0.05 all the standard texture descriptors. It is also interesting to note that the ensemble of descriptors extracted from the co-occurrence matrix obtains a performance that is comparable with recent state-of-the-art descriptors (i.e., LTP, LPQ, and RICLBP).

TABLE V AUC OBTAINED USING THE REGION-BASED APPROACHES COMBINING HR AND GR

Dataset	Preprocessing Method							
	Ed	Sal	Wa	Do	All	All+ O	All+ S	All+ 2×S
HR								
PS	91.6	90.7	89.9	91.3	91.8	91.7	92.1	**92.3**
VI	96.7	96.5	96.6	96.5	97.1	97.1	**97.2**	**97.2**
CH	**99.9**	**99.9**	99.8	**99.9**	**99.9**	**99.9**	**99.9**	**99.9**
HI	**99.6**	99.5	99.4	99.5	**99.6**	**99.6**	**99.6**	**99.6**
BR	89.7	89.6	89.6	90.5	90.1	90.0	**90.3**	**90.3**
PR	95.9	94.3	95.4	96.1	95.9	95.6	95.8	**98.8**
HE	89.3	91.6	91.3	89.8	91.7	91.6	91.8	**91.9**
LE	98.0	97.9	97.2	97.9	**98.1**	**98.1**	**98.1**	**98.1**
LT	99.1	**99.5**	98.7	99.3	99.3	99.3	99.4	99.4
SM	99.2	99.2	99.1	99.3	**99.5**	**99.5**	**99.5**	**99.5**
PIC	89.1	**91.2**	88.7	90.8	90.7	90.4	90.8	90.9
RN	95.4	**96.3**	95.3	95.4	95.7	95.6	95.8	95.8
HP	90.2	90.3	89.8	**91.1**	90.5	90.3	90.5	90.5
PA	85.7	85.4	84.8	87.0	86.3	86.2	86.8	87.2
LE	97.2	97.0	97.2	97.8	97.3	97.3	97.4	**97.5**
Av	94.4	94.6	94.2	94.8	94.9	94.8	95.0	**95.3**
GR								
PS	84.6	85.5	84.0	85.8	86.6	86.5	**86.9**	**86.9**
VI	92.3	91.4	92.1	92.5	94.0	93.9	94.3	**94.5**
CH	99.1	**99.9**	99.1	99.2	99.6	99.6	99.5	99.5
HI	99.2	98.9	99.2	99.4	99.4	99.4	99.4	**99.5**
BR	84.2	85.0	86.5	86.6	87.2	87.3	87.8	**88.2**
PR	88.5	90.1	91.6	86.9	92.4	92.1	92.7	**92.8**
HE	90.1	88.9	91.4	90.0	92.0	91.4	92.0	**92.1**
LE	92.8	93.8	92.9	93.8	93.8	93.7	94.0	**94.2**
LT	99.0	98.4	99.1	**99.2**	**99.2**	**99.2**	**99.2**	**99.2**
SM	98.5	98.3	98.3	98.5	98.7	98.7	98.8	**98.9**
PI	84.0	84.8	84.6	85.3	86.4	86.2	86.7	**86.8**
RN	93.0	89.1	**93.6**	91.5	**93.6**	93.4	**93.6**	**93.6**
HP	87.7	87.2	87.5	87.6	88.3	88.3	**88.4**	88.3
PA	81.3	81.3	82.2	83.0	83.7	83.8	84.3	**84.6**
LE	91.4	91.6	91.5	92.0	92.8	92.9	93.2	**93.5**
Av	91.0	90.9	91.5	91.4	92.5	92.4	92.7	**92.8**

V. CONCLUSION

The goal of this study was to extend recent work on texture analysis techniques based on the co-occurrence matrix and region-based approaches. Different strategies for extending the texture descriptors extracted from the co-occurrence matrix are compared and combined. These methods were improved by extracting features not only from the entire co-occurrence matrix but also from subwindows. Two new methods for extracting features from the co-occurrence matrix are proposed. Moreover, the ensemble approach proposed in this paper is shown to improve the performance of SHAPE (as reported in [26] and [25]) and standard Haralick-based features [12].

In this work different region-based methods (specifically those based on saliency detection, edge detection, and wavelets) are compared and some of their fusions are reported. The region-based approaches are combined with the highest performing methods for extracting features from the co-occurrence matrix and with three state-of-the-art descriptors: local ternary pattern, local phase quantization, and rotation invariant co-occurrence among adjacent local binary pattern. For all experiments SVM was used as the base classifier.

The generality of the proposed approach was validated across 15 different datasets, representing very different image classification problems. Results in the experimental section were also compared with some state-of-the-art descriptors.

REFERENCES

[1] J. M. Coggins: 'A framework for texture analysis based on spatial filtering'. Ph.D. Thesis, Michigan State University, 1982.

[2] S. W. Zucker: "Towards a model of texture", Computer Graphics Image Processing, vol. 5, pp. 190-202, 1976.

[3] J. Sklansky: "Image segmentation and feature extraction", IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-8, pp. 237-247, 1978.

[4] D. G. Lowe: "Distinctive image features from scale-invariant keypoints", International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[5] H. Bay, T. Tuytelaars, and L. V. Gool: "SURF: Speeded up robust features", European Conference on Computer Vision, vol. 1, pp. 404-417, 2006.

[6] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," in Proc. 9th European Conference on Computer Vision, San Diego, CA, 2005,

[7] K. Mikolajczyk, and C. Schmid: "A performance evaluation of local descriptors", IEEE Trans Pattern Analysis Mach Intell, vol. 29, no. 10, pp. 1615-1630, 2005.

[8] O. Tuzel, F. Porikli, and P. Meer, "Region covariance: A fast descriptor for detection and classification," in Proc. 9th European Conference on Computer Vision, 2006, pp. 589-600.

[9] B. Wu, and R. Nevatia, "Detection of multiple, partially occluded humans in a single image by bayesian combination of edgelet part detectors," in Proc. IEEE International Conference on Computer Vision, 2005, pp. 90-97.

[10] R. M. Haralick, K. Shanmugam, and I. Dinstein: "Textural features for image classification", IEEE Transactions on Systems, Man, and Cybernetics, vol. 3, no. 6, pp. 610-621, 1973.

[11] T. Ojala, M. Pietikäinen, and D. Harwood: "A comparative study of texture measures with classification based on featured distribution", Pattern Recognit Lett, vol. 29, no. 1, pp. 51-59, 1996.

[12] R. M. Haralick: "Statistical and structural approaches to texture", Proceedings of the IEEE, vol. 67, no. 5, pp. 786-804, 1979.

[13] A. Gelzinis, A. Verikas, and M. Bacauskiene: "Increasing the discrimination power of the co-occurrence matrix-based features", Pattern Recognit, vol. 40, no. 9, pp. 2367-2372. , 2007.

[14] R. Walker, P. Jackway, and D. Longstaff: "Genetic algorithm optimization of adaptive multi-scale GLCM features", International Journal of Pattern Recognition and Artificial Intelligence, vol. 17, no. 1, pp. 17-39, 2003.

[15] S. Chen, W. Chengdong, D. Chen, and W. Tan, "Scene classification based on gray level-gradient co-occurrence matrix in the neighborhood of interest points," in Proc. IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS 2009) 2009, pp. 482-485.

[16] D. Mitrea, P. Mitrea, S. Nedevschi, R. Badea, M. Lupsor, M. Socaciu, A. Golea, C. Hagiu, and L. Ciobanu: "Abdominal tumor characterization and recognition using superior-order cooccurrence matrices, based on ultrasound images", Computational and Mathematical Methods in Medicine, vol., pp. 2012.

[17] Y. Hu, "Unsupervised texture classification by combining multi-scale features and k-means classifier," in Proc. Chinese Conference on Pattern Recognition, 2009, pp. 1-5.

[18] F. Pacifici, and M. Chini, "Urban land-use multi-scale textural analysis," in Proc. IEEE International Geoscience and Remote Sensing Symposium, 2008, pp. 342-345

[19] P. Rakwatin, N. Longepe, O. Isoguchi, M. Shimada, and Y. Uryu, "Mapping tropical forest using ALOS PALSAR 50m resolution data with multiscale GLCM analysis," in Proc. IEEE International Geoscience and Remote Sensing Symposium, 2010, pp. 1234-1237

[20] N.-S. Vu, T. Nguyen, and C. Garcia: "Improving texture categorization with biologically-inspired filtering", Image and Vision Computing, vol. 32, no. 6-7, pp. 424-436, 2014.

[21] A. Abdesselam: "Improving local binary patterns techniques by using edge information", Lecture Notes on Software Engineering, vol. 1, no. 2, pp. 360-363, 2013.

[22] F. L. C. dos Santosa, M. Paci, L. Nanni, S. Brahnam, and J. Hyttinen: "Computer vision for virus image classification", Biosystems Engineering, vol., in press.

[23] L. Nanni, M. Paci, F. L. C. D. Santos, S. Brahnam, and J. Hyttinen, "Analysis of virus textures in transmission electron microscopy images," in Proc. KES Innovation in Medicine and Healthcare (InMed-14), San Sebastián, Spain, 2014,

[24] L. Nanni, M. Paci, F. L. C. dos Santosa, H. Skottoman, K. Juuti-Uusitalo, and J. Hyttinen: "image-based classification of maturation of human stem cell-derived retinal pigmented epithelium", Expert Systems with Applications, vol., submitted.

[25] L. Nanni, S. Ghidoni, and E. Menegatti: "A comparison of multi-scale approaches for extracting image descriptors from the co-occurrence matrix", Computer Communication & Collaboration, vol., submitted 2013.

[26] L. Nanni, S. Brahnam, S. Ghidoni, E. Menegatti, and T. Barrier: "Different approaches for extracting information from the co-occurrence matrix", PLoS ONE, vol. 8, no. 12, pp. 1-9, 2013.

[27] A. F. Costa, G. E. Humpire-Mamani, and A. J. M. Traina, "An Efficient Algorithm for Fractal Analysis of Textures," in Proc. IBGRAPI 2012 (XXV Conference on Graphics, Patterns and Images), Ouro Preto, Brazil, 2012, pp. 39-46

[28] M. Gadermayr, M. Liedlgruber, A. Uhl, and A. Vécsei, "Shape Curvature Histogram: A Shape Feature for Celiac Disease Diagnosis," in Proc. 3rd International MICCAI - MCV Workshop 2014, pp. 175-184.

[29] S. Mallat: "A theory for multiresolution signal decomposition", IEEE Trans Pattern Analysis Mach Intell, vol. 11, pp. 674-693, 1989.

[30] X. Hou, J. Harel, and C. Koch: "Image signature: highlighting sparse salient regions", IEEE Trans Pattern Analysis Mach Intell, vol. 34, no. 1, pp. 194-201, 2012.

[31] X. Tan, and B. Triggs: "Enhanced local texture feature sets for face recognition under difficult lighting conditions", Analysis and Modelling of Faces and Gestures, vol. LNCS 4778, pp. 168-182, 2007.

[32] V. Ojansivu, and J. Heikkila, "Blur insensitive texture classification using local phase quantization," in Proc. ICISP, 2008, pp. 236-243.

[33] R. Nosaka, and K. Fukui: "HEp-2 cell classification using rotation invariant co-occurrence among local binary patterns.", Pattern Recognition in Bioinformatics, vol. 47, no. 7, pp. 2428-2436, 2014.

[34] V. N. Vapnik: The Nature of Statistical Learning Theory (Springer-Verlag, 1995. 1995).

[35] X. Tang, "Texture information in run-length matrices," in Proc. IEEE Transactions On Image Processing, 1998, pp. 1602-1609.

[36] P. Liao, T. Chen, and P. Chung: "A fast algorithm for multilevel thresholding", Journal of Information Science and Engineering, vol. 17, no. 5, pp. 713-727, 2001.

[37] J. Canny: "A computational approach to edge detection", IEEE Transactions on Pattern Recognition and Machine Intelligence, vol. 8, no. 6, pp. 679-698, 1986.

[38] R. J. Baddeley, and B. W. Tatler: "High frequency edges (but not contrast) predict where we fixate: A Bayesian system identification analysis", Vision research, vol. 46, no. 18, pp. 2824-2833, 2006.

[39] G. Heusch, Y. Rodriguez, and S. Marcel, "Local binary patterns as an image preprocessing for face authentication," in Proc. 7th International Conference on Automatic Face and Gesture Recognition (FGR 2006), Southampton 2006,

[40] J. Jantzen, J. Norup, G. Dounias, and B. Bjerregaard, "Pap-smear benchmark data for pattern classification," in Proc. Nature inspired Smart Information Systems (NiSIS), Albufeira, Portugal, 2005, pp. 1-9.

[41] M. V. Boland, and R. F. Murphy: "A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells", BioInformatics, vol. 17, no. 12, pp. 1213-1223, 2001.

[42] F. Yuan: "Video-based smoke detection with histogram sequence of LBP and LBPV pyramids", Fire Safety Journal, vol. 46, no. 3, pp. 132-139, 2011.

[43] A. Cruz-Roa, J. C. Caicedo, and F. A. González: "Visual pattern mining in histology image collections using bag of features", Artificial Intelligence in Medicine, vol., no. 52, pp. 91-106, 2011.

[44] G. B. Junior, A. Cardoso de Paiva, A. C. Silva, and A. C. Muniz de Oliveira: "Classification of breast tissues using Moran's index and Geary's coefficient as texture signatures and SVM", Computers in Biology and Medicine, vol. 39, no. 12, pp. 1063-1072, 2009.

[45] L. Nanni, J.-Y. Shi, S. Brahnam, and A. Lumini: "Protein classification using texture descriptors extracted from the protein backbone image", Journal of Theoretical Biology, vol. 3, no. 7, pp. 1024-1032, 2010.

[46] N. Hamilton, R. Pantelic, K. Hanson, and R. D. Teasdale: "Fast automated cell phenotype classification", BMC Bioinformatics, vol., pp. 8-110, 2007.

[47] D. Borghesani, C. Grana, and R. Cucchiara: "Miniature illustrations retrieval and innovative interaction for digital illuminated manuscripts in multimedia systems", Multimedia Systems, vol. 20, pp. 65-79, 2014.

[48] P. Foggia, G. Percannella, A. Saggese, and M. Vento: "Pattern recognition in stained HEp-2 cells: Where are we now?", Pattern Recognition, vol. 4, no. 7, pp. 2305-2314, 2014.

[49] F. Khan, S. Beigpour, J. van de Weijer, and M. Felsberg: "Painting-91: a large scale database for computational painting categorization", Machine Vision and Applications, vol. 25, no. 6, pp. 1385-1397, 2014.

[50] D. Casanova, J. Joaci de Mesquita, and O. M. Bruno: "Plant leaf identification using gabor wavelets," International Journal of Imaging Systems and Technology, vol. 19, no. 3, pp. 236 - 243, 2009.

[51] T. Fawcett: "ROC graphs: Notes and practical considerations for researchers", in Editor (Ed.)^(Eds.): 'Book ROC graphs: Notes and practical considerations for researchers' (HP Laboratories, 2004, edn.), pp.

[52] J. Demšar: "Statistical comparisons of classifiers over multiple data sets", Journal of Machine Learning Research, vol. 7 pp. 1-30, 2006.