



A comparison of methods for extracting information from the co-occurrence matrix for subcellular classification



Loris Nanni^{a,*}, Sheryl Brahnam^b, Stefano Ghidoni^a, Emanuele Menegatti^a, Tonya Barrier^b

^a Department of Information Engineering, University of Padua, Via Gradenigo, 6, 35131 Padova, Italy

^b Department of Computer Information Systems, Missouri State University, 901 S. National, Springfield, MO 65804, USA

ARTICLE INFO

Keywords:

Texture descriptors
Co-occurrence matrix
Subcellular localization
Support vector machine
Random subspace

ABSTRACT

In this paper we focus on cell phenotype image classification, a bioimaging problem that is concerned with finding the location of protein expressions within a cell. Protein localization is becoming increasingly critical in the diagnosis and prognosis of many diseases. In recent years several new approaches for describing a given image have been proposed. Some of the most significant developments have been based on binary encodings, such as local binary patterns and local phase quantization. In this paper we reexamine one of the oldest methods for representing an image that Haralick famously proposed in 1979 using the co-occurrence matrix for calculating a set of image statistics. Few methods have been proposed since that extract new features from the co-occurrence matrix. In this work we compare some recently proposed methods that are based on the co-occurrence matrix (CM) to classify cell phenotype images. We investigate the correlation among the different sets of features that can be extracted from the CM and then determine the best way to combine these different feature sets for optimizing system performance. Moreover, we combine our novel approach with state of the art descriptors to optimize performance. We validate our approach on various types of biological microscope images using five image databases for subcellular classification. We use these image features for training a stand-alone support vector machine and a random subspace of support vector machines to separate the classes in each dataset.

The Matlab code for some of the approaches tested in this paper will be available at http://www.dei.unipd.it/wdyn/?IDsezione=3314&IDgruppo_pass=124&preview=>.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

The development of new and improved tools for automatic analysis and classification has already proven beneficial in clinical practice and in medical and biological research (Hamilton, Wang, Kerr, & Teasdale, 2009; Murphy, 2006). In Karkanis, Iakovidis, Maroulis, Karras, and Tzivras (2003), for example, linear discriminant analysis of wavelet features, used in applications that range from traffic incident detection (Samant & Adeli, 2000) to face identification and verification (Shen, Bai, & Fairhurst, 2007), recently proved effective in the detection of tumors in endoscopic images. In Ameling, Wirth, Paulus, Lacey, and Vilarino (2009), image texture information successfully discriminated polyps in colonoscopy images. The Local Binary Pattern (LBP) operator (Ojala, Pietikainen, and Maenpaa (2002), which has distinguished itself from other image texture operators by its simplicity, effectiveness, and robustness,

has also proven very good at detecting a variety of tumors and masses. In Vécsei, Amann, Hegenbart, Liedlgruber, and Uhl (2011), for example, LBP was used to assign a Marsh-like score to endoscopic images of pediatric celiac diseases, thus providing concrete help for pathologists. In Oliver, Lladó, Freixenet, and Martí (2007) a Support Vector Machine (SVM) was coupled with the LBP operator to distinguish real masses from normal parenchyma in mammographic images, thus reducing the incidence of false positive samples. In Unay and Ekin (2008), for instance, LBP was used to explore brain magnetic resonance data, and in Nanni and Lumini (2008) the authors demonstrated how a combination of LBP with other texture descriptors is effective in classifying different cell phenotypes using SVM.

In this paper we focus on cell phenotype image classification, a bioimaging problem that is concerned with finding the location of protein expressions within a cell. Understanding the function of proteins at the cellular level is a major goal in biology (Chebira et al., 2007) since the knowledge of the subcellular locations of a protein is useful in understanding its specific function and in describing a cell's behavior under different conditions. Not only is protein localization an important topic in biology, but it is also

* Corresponding author. Tel.: +39 3493511673.

E-mail addresses: nanni@dei.unipd.it (L. Nanni), sbrahnam@missouristate.edu (S. Brahnam), ghidoni@dei.unipd.it (S. Ghidoni), emg@dei.unipd.it (E. Menegatti), TonyaBarrier@missouristate.edu (T. Barrier).

critical in the design of drug screening systems, drug discovery, and recently in the diagnosis and prognosis of many diseases. In fact, it has been suggested that the old paradigm of one protein, one biomarker, one clinical decision no longer holds and is being replaced by multiparametric analysis of genes and proteins, with protein patterns currently considered to offer better diagnostic possibilities than a single biomarker (Rosenblatt et al., 2008).

This change in view has been motivated in part by the abundance of recent research showing that aberrant subcellular localizations are associated with many diseases: cancer (Fernandes et al., 2009; Geerts et al., 2007; Knostman et al., 2007; Mezzananza & D., 0000; Perrone et al., 2007; Ralhan et al., 2010; Tang et al., 2010), heart disease (Bedard, Haaning, & Ware, 2011; Obrenovich et al., 2006), kidney disease (Nishibori et al., 2004), pulmonary fibrosis (Thomas et al., 2002), Alzheimer disease, etc. In Nishibori et al. (2004), for instance, subcellular localization of three missense mutants located in the proximal C-terminus part of the podocin protein (NPHS2) were found to be greatly altered in patients with steroid-resistant nephrotic syndrome. In Knostman et al. (2007) sodium/iodide symporter (NIS) expression, subcellular localization, and function were analyzed in MCF-7 human breast cancer cells. The authors found that NIS and intercellular localization are associated with pAkt expression in human breast cancer tissues, and in Perrone et al. (2007) COX-2 expression and its subcellular localization in lobular in situ neoplasia (LIN) of the breast was assessed as a candidate biomarker for breast cancer. In Ralhan et al. (2010) the subcellular localization of EpEx and Ep-ICD in the human colon adenocarcinoma cell line CX-1 was examined using immunofluorescence. Nuclear and cytoplasmic Ep-ICD expression was increased in cancers of the breast, prostate, head, neck, and esophagus compared to their corresponding normal tissues that showed membrane localization of the protein. In Bachmann, Strume, Puntervoll, Kalvenes, and Akslen (2005) alterations in the expression and subcellular localization of cell adhesion markers were found to be important in the development and progression of melanocytic tumors, and it is well-known that activated leukocyte cell adhesion molecule (ALCAM) is expressed at the cell surface of epithelial ovarian cancer cell lines (Piazza et al., 2005). Moreover, in Bedard et al. (2011) it was discovered that Mutations in the gene encoding zinc finger of the cerebellum protein 3 (ZIC3) can cause congenital heart defects. Cytosolic proteins, such as G protein-coupled receptor kinases (GRKs), like GRK2, are well-characterized in the heart, and G protein-couple receptor desensitization is emerging as a key feature in several cardiovascular diseases. GRKs are also found in cerebral tissues, and significant increase in GRK2 immunoreactivity endothelial cells has been found in patients with Alzheimer disease. In Obrenovich et al. (2006), the authors explored cellular and subcellular localization by immunoreactivity of GRK2 and demonstrated that the ultrastructural localization and overexpression of GRK2 occurs during the early stages of damage in aged humans and in Alzheimer disease cases.

As indicated above, protein subcellular localization is also becoming increasingly important as a prognostic indicator (Bachmann et al., 2005; Moreira et al., 2010; Surowiak et al., 2006). There is considerable evidence, for instance, that the loss of BLCAP expression is associated with tumor progression. In Moreira et al. (2010) the authors were able to classify urothelial carcinomas into four groups based on levels of expression and subcellular localization of BLCAP protein. Their findings suggested that BLCAP may have prognostic value in bladder cancer. Subcellular localization may even be useful in the prediction of chemotherapy response to ovarian cancer (Surowiak et al., 2006).

In the last two decades several research groups have investigated automated cell phenotype image classification (Boland, Murphy, & Murphy, 1998; Boland & Murphy, 2001; Conrad et al., 2004; Danckaert, Gonzalez-Couto, Bollondi, Thompson, & Hayes, 2002;

Lin et al., 2007; Nanni & Lumini, 2008; Perner, Perner, & Muller, 2002). Work that has focused on training generic classifiers with image descriptors include (Chen, Velliste, & Murphy, 2006; Glory & Murphy, 2007; Glory, Newberg, & Murphy, 2008). The most common image recognition methods used for this problem are Haralick texture measures (Huang & Murphy, 2004), Zernike moments and threshold adjacency statistics (Hamilton, Pantelic, Hanson, & Teasdale, 2007), Gabor filters, and a number of other ad hoc measures (Conrad et al., 2004). More recently, new methods have been developed that are based on fusion at the feature and score level. At the feature level, vectors are created by concatenating several descriptors (Chen et al., 2005; Hamilton et al., 2007; Huang & Murphy, 2004). A multi-resolution approach, proposed in Chebira et al. (2007), trained classifiers using descriptors extracted from different resolution spaces. Examples of fusion methods used at the score level include (Lin et al., 2007).

The aim of this work is to assess the discriminant power in this problem using a recent method for extracting features from the co-occurrence matrix (CM), where new features are extracted considering the co-occurrence matrix as a 3D shape (SHAPE). This set of features works rather poorly with respect to the standard Haralick feature set (HAR), but we have discovered that SHAPE can be coupled with HAR to boost performance to levels similar to that obtained by recent texture descriptor approaches (Nanni, Ghidoni, & Menegatti, 2013a; Nanni, Paci, Brahmam, Ghidoni, & Menegatti, 2013b). Moreover, when this new set of features is coupled with state-of-the-art global descriptors, further improvements are obtained. This is demonstrated by combining our proposed set of features based on the co-occurrence matrix (SHAPE) with descriptors proposed in Paci et al. (2013). Since the Q-statistic between HAR and SHAPE is quite low, thus confirming the low correlation between the information extracted by HAR and SHAPE, we were motivated to investigate their fusion. As expected, fusion of HAR and SHAPE produced better results. Our experiments were validated across five databases: (1) 2D HeLa dataset (Boland & Murphy, 2001), (2) locate endogenous mouse sub-cellular organelles, (3) locate transfected mouse subcellular organelles (Fink et al., 2006; Hamilton et al., 2007), (4) Chinese Hamster Ovary (Boland et al., 1998), and (5) the RNAi dataset (Zhang & Pham, 2011).

The rest of this paper is organized as follows. In Section 2 we describe our proposed approach using descriptors based on CM along with all the other descriptors combined with our system or used for comparison purposes. In Section 3 we describe the five benchmark databases used to validate our approach. In Section 4, we present our experimental results, and in Section 5 we conclude our paper with a few suggestions for further research.

2. Proposed approach

In this paper we compare recently proposed descriptors based on the co-occurrence matrix (CM). Our goal is to enhance the performance of the standard Haralick descriptor with a recent set of proposed features that can also be extracted from the co-occurrence matrix (Nanni et al., 2013a). Our tests using the Q-statistic show that the different approaches for extracting features from the co-occurrence matrix provide different information. This indicates that combining these descriptors should enhance performance. Moreover, by combining this novel set of descriptors from CM with state of the art descriptors, we predict even better gains in performance. In this section we describe our approach.

2.1. Descriptors extracted from the co-occurrence matrix

GLDM is a histogram evaluated over a two-dimensional domain of dimension $N_{GL} \times N_{GL}$, where N_{GL} is the number of gray levels in a

given image, usually 256. GLDM describes the gray level transitions between pixels. Each location on the histogram identifies a specific transition between two values: for example, the bin placed at location (20, 30) records the number of pixel couples whose first pixel has the value of 20 and whose second pixel has the value of 30. GLDM is constructed from the source image by considering each pixel couple and incrementing the corresponding bin. The way couples are determined depends on two parameters, d and θ , that respectively represent the distance of the two pixels and the direction they are aligned.

In this work, four directions are considered: 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). When choosing $d = 1$ and $\theta = 0$, for example, the pixels are grouped in couples that are adjacent to each other in the same row. Fig. 1 provides an illustration of this: the original image appears on the left-hand side, with the corresponding GLDM appearing on the right-hand side. The locations placed along the principal diagonal of the GLDM have coordinates with the same value for x and y and represent transitions between two pixels of the same color. Since it often happens that the two gray levels of the pixels composing the couple are the same, or at least similar, GLDM normally has a strong component on the diagonal and the neighboring region, as clearly visible in Fig. 1. Other smaller peaks can be distinguished towards the end of the diagonal, as well as at other locations, each representing neighborhoods of pixels where a number of couples have the same value.

Because GLDM evaluates pixel value transitions, it contains a large amount of information about image texture. However, this information in raw form is very complex to handle. As a result, it is common to extract a set of indicators from the GLDM that are evaluated according to a given algorithm. The use of indicators extracted from the GLDM was originally introduced together with the GLDM itself in Haralick (1979). The selection of a good set of indicators that provides optimal information from the matrix for a particular problem is crucial. The investigation of novel indicators is desirable because they can improve the performance of a texture analysis.

In this paper, we exploit some of the indicators that were introduced in the original paper by Haralick (1979) together with some novel features we have developed for enhancing system performance in this domain. Regarding the first group (HAR), we investigate the following 13 features introduced by Haralick:

- (1) Energy
- (2) Correlation

- (3) Inertia
- (4) Entropy
- (5) Inverse difference moment
- (6) Sum average
- (7) Sum variance
- (8) Sum entropy
- (9) Difference average
- (10) Difference variance
- (11) Difference entropy
- (12) Information measure of correlation 1
- (13) Information measure of correlation 2

Regarding the second group, we investigate some new indicators proposed originally for crowd detection (Ghidoni, Cielniak, & Menegatti, 2012) that explore the shape of the 3D function, as illustrated in Fig. 2. The main idea is to consider GLDM at selected peaks. For each peak, the level curve is obtained as the intersection between the GLDM itself and a horizontal plane at a given height (see Fig. 2(a)). The intersection (see Fig. 2(b)) defines a complex shape that can be made of one or more extracted blobs. The main blob is the one with the largest area and volume, and represents the majority of the pixel transitions. The main component is therefore the peak that takes into account the main textural component, even though information is present also in minor peaks. To keep our approach simple and processing time short, we decided to consider only one peak, and the main component is the one to be investigated in this case. To simplify the analysis, the main blob's shape is fitted to an ellipse (drawn in red in Fig. 2(c)). This approximation of the main blob's shape by an ellipse results in some information loss, but it offers the advantage of making the comparison among curves much easier.

Level curves are considered towards the base of the GLDM, starting at height 1 and going until height 19, with a distance of 2 between two consecutive planes. They are all at a rather low height because that region is very stable: there are no significant changes between two frames. The peaks of the GLDM, however, can undergo substantial changes (e.g., because of image noise) even between consecutive frames in a video flow framing a static scene. The conventional maximum height fixed to 19 has been experimentally determined, and represents a value that allows a detailed analysis of the GLDM shape and neglects noisy peaks. Higher values for the maximum height causes the analysis to be performed in regions of the GLDM that are not stable, which in turn cause noisier texture measurements. On the other hand, lower values would be more rough, since they would cause our analysis to exclude a great portion of the GLDM. We do not normalize GLDM

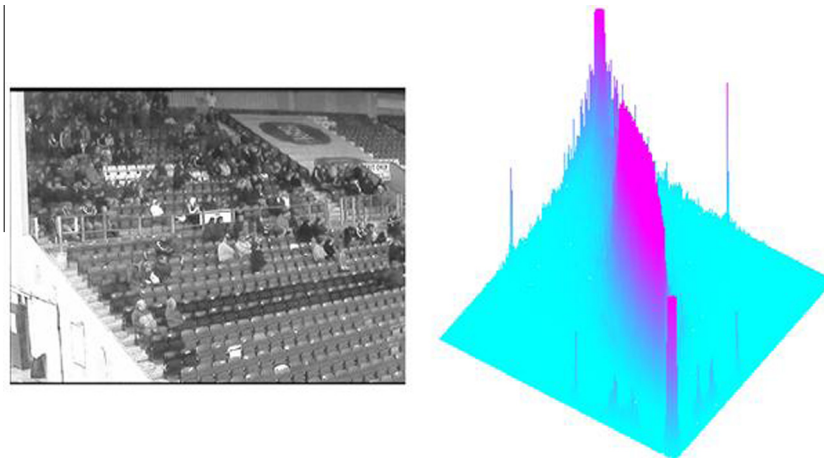


Fig. 1. An example of GLDM, shown as a 3D function (on the right), together with the image from which it has been calculated (on the left).

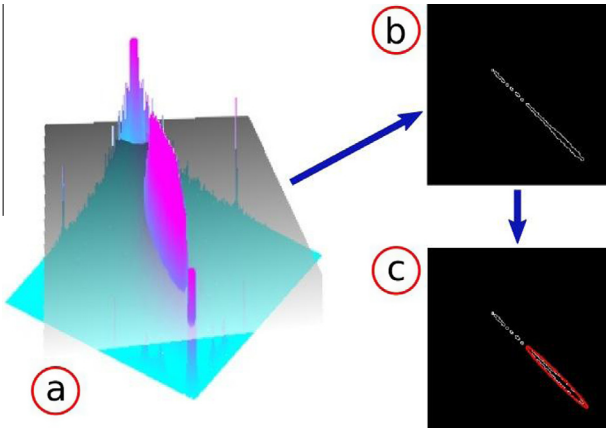


Fig. 2. The intersection with a plane provide a 2D image with multiple contours, an ellipse is fitted to the largest contour.

since normalization to the highest bin would introduce instabilities. Of course, another possible normalization could be performed with respect to the total volume of the GLDM, which depends mainly on the size of the original image. However, since all the images considered in this paper have the same size, this would have no effect on results.

For each level, a set of descriptors regarding the ellipses derived from the GLDM are evaluated. The features describing all levels are then jointly analyzed for deriving yet another set of features that describe the evolution of the level curves. In this fashion the features describing the shape of the 3D function are obtained.

We obtain a number of features through a joint analysis of the ellipses. The first feature describes how evenly (EV) the minor axis of the ellipses decreases as the height of the level curve is increased. Considering the length of the minor axis of the ellipse as a function of the height of the level curve, $w(h)$, EV is evaluated as:

$$EV = \sqrt{\frac{1}{N} \left(\int_{h_{min}}^{h_{max}} \left(dw(h) - \frac{w_{max} - w_{min}}{h_{max} - h_{min}} \right)^2 \right)}$$

where w_{max} and w_{min} are the maximum and minimum values for w , respectively, and h_{max} and h_{min} have the same meaning for h . Considering the discrete nature of the GLDM histogram, the actual formula for evaluating EV is:

$$EV = \sqrt{\frac{1}{N} \frac{\sum_{i=1}^{N-1} [(w_i - w_{i+1}) - \bar{\Delta w}]^2}{d}}$$

where w_i represents the minor axis (width) of the i th layer, N the total number of layers, $\bar{\Delta w}$ is the average difference between ellipses at two consecutive layers, and d is the height distance between two consecutive level curves. The sum is divided by i since, in the general case, the distance between two consecutive layers cannot be assumed to be constant.

A second descriptor evaluates the Minor Axis Spread (MAS), which provides information about the GLDM steepness. MAS is evaluated as the difference between the maximum and minimum observed for the length of the ellipse minor axis:

$$MAS = w_{max} - w_{min}$$

A third descriptor represents the minimum value (MV) for the minor axis. This provides information about the upper part of the GLDM: the larger MV , the smoother the upper part of the function:

$$MV = \min_i w_i$$

A fourth descriptor is the mean of the width/height ratio (WHR) of all ellipses. This measures the relationship between the contrast of the image (the higher the contrast, the longer the major axis) and the presence of strong variations between pixels (which causes a long minor axis). This mean is obtained as follows:

$$WHR = \frac{1}{N} \sum_{i=1}^N \frac{w_i}{h_i}$$

A fifth descriptor is the total volume (TV) under the analyzed level curves, which can be calculated as follows:

$$TV = w_1 h_1 l_1 + \sum_{i=2}^N w_i h_i (l_i - l_{i-1})$$

where w_i and h_i are the minor and major axes, respectively, of the ellipse found on the i th layer at height l_i .

Finally, descriptors can be based on the area of the smallest ellipse, or minimum area (MA), as well as on the ratio between the smallest and the largest ellipses, or the maximum–minimum area Ratio ($MMAR$):

$$MA = \min_i w_i h_i = w_N h_N$$

$$MMAR = \frac{w_1 h_1}{w_N h_N}$$

It should be noted that, by construction, the N th layer is the one with the smallest ellipse, while the largest one is placed on the first layer considered.

All the descriptors mentioned above measure the characteristics of the ellipses approximating the main components, or blobs, of the level curves. Two other features measure other characteristics of the co-occurrence matrix. The first is based on the volume of the peak (PV), or the portion of the GLDM that is above the highest level curve considered:

$$PV = \sum_{i=1}^M (H_i - l_N)$$

where l_N is the height of the highest layer, and M is the number of bins with height $H_i > l_N$.

The second is based on the number of blank locations (BL), i.e., the number of bins with zero occurrences, defined as:

$$BL = \sum_{i=1}^{N_{cl}} \sum_{j=1}^{N_{cl}} \begin{cases} 1 & \text{iff GLDM}(i,j) = 0 \\ 0 & \text{else} \end{cases}$$

where $GLDM(i,j)$ is the value of the GLDM at location (i,j) .

For both approaches (HAR and SHAPE) features are obtained in our experiments using $d = 1$ and $d = 3$ with an angle of $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$.

2.2. Local Quinary Pattern (LQP)

In the short space of a decade, the LBP operator (Ojala et al., 2002) has become one of the most appreciated and useful texture operators in machine learning. This popularity is mainly due to the high descriptive power offered by LBP and its invariance to local grayscale variations. LBP makes the assumption that an image is constituted by micropatterns. In its original form, the LBP operator is defined for each pixel in the image on the basis of a circular neighborhood of P pixels of grey level q_p and radius R around the central pixel of grey value q_c , as follows:

$$LBP(P, R) = \sum_{p=0}^{P-1} s(q_p - q_c) 2^p$$

where,

$$s(x) = \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

The extracted binary codes are represented as a histogram. In Ojala et al. (2002) changes to the original formulation were proposed that succeeded in achieving both rotation invariance and a reduction in the size of the histogram.

Due to the importance and simplicity of this operator, various modifications and improvements have been proposed. One important variant of LBP is the Local Ternary Pattern (LTP) Tan and Triggs (2007), which attempts to reduce LBP's sensitivity to noise, especially in near-uniform image regions.

LTP introduces the threshold τ to the canonical LBP $s(x)$ function, which becomes:

$$s(x) = \begin{cases} 1, & x \geq \tau \\ 0, & |x| \leq \tau \\ -1, & x \leq -\tau \end{cases}$$

Due to the length of the ternary coding, the LTP code is split into a positive and a negative LBP code.

More complex encodings based on LBP include the Local Quinary Pattern (LQP) operator, first proposed in Nanni, Lumini, and Brahnam (2010), which exploits two thresholds τ_1 and τ_2 , thus allowing the $s(x)$ function to assume the values $[-2; -1; 0; 1; 2]$:

$$s(x) = \begin{cases} 2, & x \geq \tau_2 \\ 1, & \tau_1 \leq x < \tau_2 \\ 0, & -\tau_1 \leq x < \tau_1 \\ -1, & -\tau_2 \leq x < -\tau_1 \\ -2, & x < -\tau_2 \end{cases}$$

In LQP the quinary pattern is split into four binary patterns, and the histograms that are computed from these patterns are then concatenated to form the descriptor.

In LTP and its variants, threshold selection is critical, as it reduces the sensitivity of these new operators to noise. Thresholds are usually set manually to obtain the best performance for specific problems, but some automatic adaptive procedures have been proposed in Vécsei et al. (2011), Akhloufi and Bendada (2010) that exploit such local statistics as the mean value and the standard deviation inside each neighborhood.

A multi-threshold approach (Paci et al., 2013) is used in our experiments that simply combines the scores of all the classifiers trained with an LTP-based descriptor having a different value for τ . For example, in the case of LTP, if we test $\tau = 1$, $\tau = 3$ and $\tau = 5$, then the multi-threshold LTP is given by the fusion among three SVM, one trained with LTP $\tau = 1$, another with LTP $\tau = 3$, and yet another with LTP $\tau = 5$.

The multi-threshold approach we use for LQP (MLQP) combines by sum rule all the LQP with rotation uniform bins, that is by concatenating histograms obtained with $(P = 8, R = 1)$ and $(P = 16, R = 2)$, with $\tau_1 = \{1, 3, 5, 7, 9\}$ and $\tau_2 = \{\tau_1 + 2, \tau_1 + 4, \dots, \tau_1 + 11\}$ (as in Paci et al. (2013)).

2.3. Local phase quantization (LPQ)

LPQ is another important operator that analyses textural information based on the blur invariance of the Fourier Transform Phase (Ojansivu & Heikkilä, 2008). Blur insensitive information is extracted by locally computing the phase of the 2D Short Term Fourier Transform (STFT) for each pixel position of the image over a rectangular neighborhood.

The spatially invariant blurring of an original image $f(\mathbf{x})$ results in an observed blurred image, $g(\mathbf{x})$, expressed by

$$g(\mathbf{x}) = f(\mathbf{x}) * h(\mathbf{x})$$

where $\mathbf{x} = [x, y]^T$ represents the spatial coordinate vector and $h(\mathbf{x})$ the point spread function. Thus, in the Fourier space, this means

$$G(\mathbf{u}) = F(\mathbf{u}) \cdot H(\mathbf{u})$$

where $G(\mathbf{u})$, $F(\mathbf{u})$, and $H(\mathbf{u})$ are the Discrete Fourier transforms (DFT) of the blurred $g(\mathbf{x})$, $f(\mathbf{x})$ and $h(\mathbf{x})$, respectively, and $\mathbf{u} = [u, v]^T$ is the frequency coordinate vector.

The magnitude and phase aspects can be separated resulting in

$$|G(\mathbf{u})| = |F(\mathbf{u})||H(\mathbf{u})|$$

and

$$\angle G(\mathbf{u}) = \angle F(\mathbf{u}) + \angle H(\mathbf{u}).$$

If the blur $h(\mathbf{x})$ is centrally symmetric, the Fourier transform is always real-valued, and its phase is a two-valued function given by

$$\angle H(\mathbf{u}) = \begin{cases} 0, & H(\mathbf{u}) \geq 0 \\ \pi, & H(\mathbf{u}) < 0 \end{cases}$$

This means that the original and the blurred images have the same phase at those frequencies which make $H(\mathbf{u})$ positive. However, due to the finite size of the observed image, in practice the blurring invariance cannot be strictly achieved: when the extent of blur is comparable to the image size, the loss of information due to the border effect becomes too strong.

LPQ is based on the above properties of blur invariance. It uses the local phase information extracted using STFT computed over a rectangular neighborhood N_x of size M by M at each pixel position \mathbf{x} of the image $f(\mathbf{x})$:

$$F(\mathbf{u}, \mathbf{x}) = \sum_{\mathbf{y} \in N_x} f(\mathbf{x} - \mathbf{y}) e^{-2\pi i \mathbf{u}^T \mathbf{y}} = \mathbf{w}_u^T \mathbf{f}_x$$

where \mathbf{w}_u is the basis vector of the 2-D DFT at frequency \mathbf{u} , and \mathbf{f}_x is a vector containing all M^2 elements from N_x .

One of the main difficulties when analyzing texture is that results have a strong dependence on resolution and scale, an effect that is much stronger with respect to e.g. edge-based approaches. Using rectangular neighborhood of different size permits to extract more reliable features, this is already well shown in the literature, e.g. (Chan, Tahir, Kittler, & Pietikäinen, 2013).

The LPQ operator considers only the following four frequency vectors: $\mathbf{u}_1 = [a, 0]^T$, $\mathbf{u}_2 = [0, a]^T$, $\mathbf{u}_3 = [a, a]^T$, and $\mathbf{u}_4 = [a, -a]^T$, where a is small enough to reside below the first zero crossing of $H(\mathbf{u})$ that satisfies

$$\angle G(\mathbf{u}) = \angle F(\mathbf{u}), \text{ for all } \angle H(\mathbf{u}) \geq 0.$$

Now, if we let

$$\mathbf{F}_x^c = [F(\mathbf{u}_1, \mathbf{x}), F(\mathbf{u}_2, \mathbf{x}), F(\mathbf{u}_3, \mathbf{x}), F(\mathbf{u}_4, \mathbf{x})]$$

and

$$\mathbf{F}_x = [\text{Re}\{\mathbf{F}_x^c\}, \text{Im}\{\mathbf{F}_x^c\}]^T,$$

then the corresponding 8 by M^2 transform matrix is

$$\mathbf{W} = [\text{Re}\{\mathbf{w}_{u1}, \mathbf{w}_{u2}, \mathbf{w}_{u3}, \mathbf{w}_{u4}\}, \text{Im}\{\mathbf{w}_{u1}, \mathbf{w}_{u2}, \mathbf{w}_{u3}, \mathbf{w}_{u4}\}]^T$$

And, thus,

$$\mathbf{F}_x = \mathbf{W} \mathbf{f}_x.$$

The coefficients need to be decorrelated before quantization to maximally preserve the information. Assuming a Gaussian distribution, a whitening transform can achieve independence

$$\mathbf{G}_x = \mathbf{V}^T \mathbf{F}_x$$

where \mathbf{V} is an orthonormal matrix derived from the singular value decomposition of the covariance matrix of the transform coefficient vector \mathbf{F}_x .

After computing \mathbf{G}_x for all the image position, the resulting vectors are quantized using the following scalar quantizer:

$$q_j = \begin{cases} 1, & g_j \geq 0 \\ 0, & g_j < 0 \end{cases}$$

where g_j represent the j th component of \mathbf{G}_x . These quantized coefficients are represented as integers between 0 and 255 using the binary coding:

$$b = \sum_{j=1}^8 q_j 2^{j-1}$$

These integer values are then organized in a feature vector that is useful for classification tasks.

As with LBP, a ternary coding for the LPQ operator is proposed and formulated as follows:

$$q_j = \begin{cases} 1, & g_j \geq \rho \cdot \tau_j \\ 0, & -\rho \cdot \tau_j \leq g_j \leq \rho \cdot \tau_j \\ -1, & g_j \leq -\rho \cdot \tau_j \end{cases}$$

where τ_j is set to half of the standard deviation of the j th component of \mathbf{G}_x and ρ is a given weight.

In this paper we combine by sum rule all the LPQ¹ with ternary coding as follows: $\rho = \{0.2, 0.4, 0.6, 0.8, 1.0, 1.2, 1.4, 1.6, 1.8\}$ (as in Paci et al. (2013)). We named this fusion MLP.

2.4. Support vector machine

The basic SVM algorithm is the non-linear generalization of the General Portrait algorithm (Vapnik & Chervonenkis, 1964; Vapnik & Lerner, 1963). It is a binary classifier that learns the boundary between samples of two different populations by projecting the samples onto a multidimensional feature space and drawing a separating hyperplane. Some advantages using SVM include the ability of managing extremely large feature vectors, improbable overfitting due to the maximized distance between the hyperplane and the closest samples, and the flexibility offered by various kernel functions. In this study, both linear and radial basis function kernels are tested. For each dataset, the best kernel and the best set of parameters are chosen by 5-fold cross validation using the training data.

2.5. Random subspace (RS)

RS was introduced in for decision trees and works well when the features are highly correlated (Ho, 1998). The method consists in training N_c classifiers with randomly built subsets, T_i , of the original training set T . Considering l samples ($1 \leq l \leq L$) and c classes ($1 \leq j \leq c$), from the original d -dimensional training set

$$T = \{(\mathbf{x}_l, t_l) | 1 \leq l \leq L\}, \quad \mathbf{x}_l \in \mathbb{R}^d, \quad t_l \in C = \{1, \dots, c\},$$

the $d \times k$ -dimensional subsets

$$T_i = \{(\mathbf{P}_i(\mathbf{x}_l), t_l) | 1 \leq l \leq L, 1 \leq i \leq N_c\}$$

are generated, where $\mathbf{P}_i(\mathbf{x}_l)$ represents the random selection of $k \times d$ features ($0 \leq k \leq 1$) from the d features of the sample l . The classification results of the N_c classifiers are then combined according to a fusion rule (e.g., the voting and sum rules) (Wang & Tang, 2006). In this study $N_c = 50$ and $k = 0.5$.

3. Datasets

In this section we describe the datasets used in our experimental section. Each of the datasets includes subgroups of either sub-cellular structures, such as organelles, or cell classes, such as different cell lines, that were used as classes to both train and test the classifiers.

3.1. 2D HeLa dataset

The 2D HeLa² dataset (Boland & Murphy, 2001) contains 862 single cell images from fluorescence microscope acquisitions on HeLa cells (see Fig. 3 for some examples). The images are 16 bit greyscale of size 512 by 382 pixels. There are 10 different classes in that dataset that correspond to 10 different stains (see Fig. 4):

- (1) Actin filaments (98 images);
- (2) Endosomes (84 images);
- (3) Endoplasmic reticulum (86 images);
- (4) Golgi giantin (87 images);
- (5) Golgi GPP130 (85 images);
- (6) Lysosome (91 images);
- (7) Microtubules (91 images);
- (8) Mitochondria (73 images);
- (9) Nucleolus (80 images);
- (10) Nucleus (87 images).

3.2. LOCATE ENDOGENOUS mouse sub-cellular organelles (LE)

This dataset³ (Fink et al., 2006; Hamilton et al., 2007) contains images of endogenous proteins and features of specific organelles that were stained with fluorescent antibodies, as well as some other probes from mouse cells (see Fig. 3 for some examples). Images are 8 bit greyscale and of size 768 by 512 pixels. The classes correspond to 10 located organelles:

- (1) Actin;
- (2) Endoplasmic Reticulum;
- (3) Endosome;
- (4) Golgi;
- (5) Lysosome;
- (6) Microtubules.
- (7) Mitochondria;
- (8) Nucleus;
- (9) Peroxisome;
- (10) Plasma Membrane;

For each organelle/class, about 50 images were acquired, for approximately 502 images, each containing up to 13 cells each.

3.3. Locate transfected mouse subcellular organelles (LT)

This dataset³ Hamilton et al., 2007; Fink et al., 2006 contains images of epitope-tagged or fluorescence-tagged proteins transiently expressed in the specific organelle and subsequently detected. Images are 8 bit greyscale and of size 768 by 512 pixels (see Fig. 5 for some examples). In addition to the 10 classes contained in the LOCATE ENDOGENOUS dataset, the supplementary Cytoplasm class is also included in this dataset. For each organelle/class, about 50 images were acquired for a total of 553 images containing up to 13 cells each (see Fig. 6).

² Download from <http://ome.grc.nia.nih.gov/iicbu2008/hela/index.html>.

³ Download from <http://locate.imb.uq.edu.au/>.

¹ With radius $x = 3$ or $x = 5$.

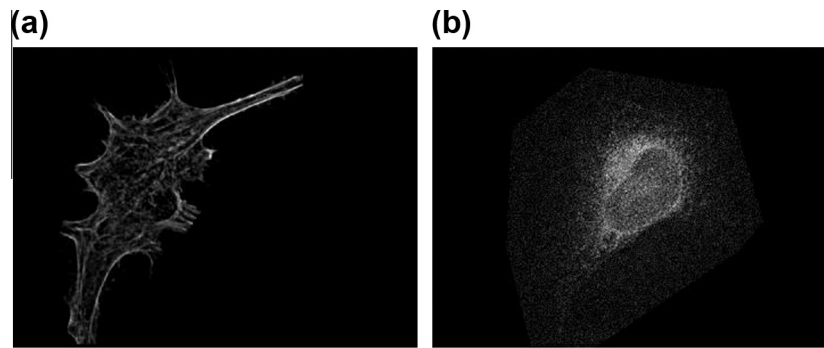


Fig. 3. 2D HeLa dataset samples: (a) Actin filament and (b) Endoplasmic reticulum.

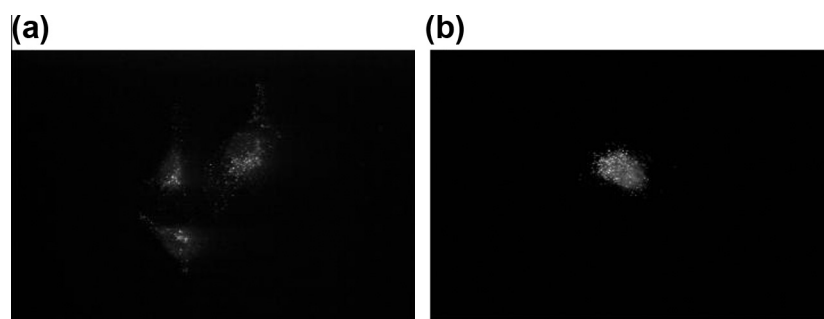


Fig. 4. Locate endogenous dataset samples: (a) Endosomes and (b) Peroxisomes.

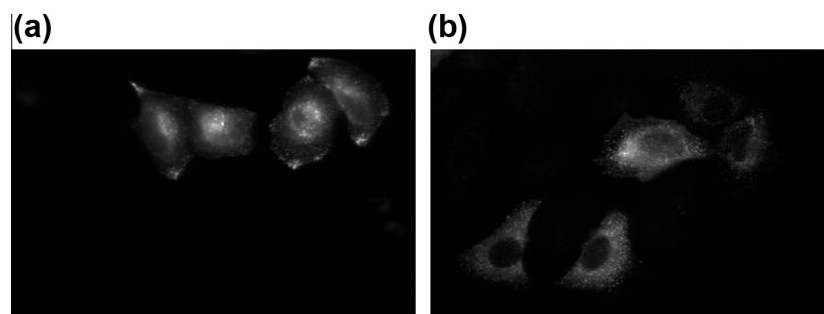


Fig. 5. Locate transfected dataset samples: (a) Endosomes, (b) Peroxisomes.

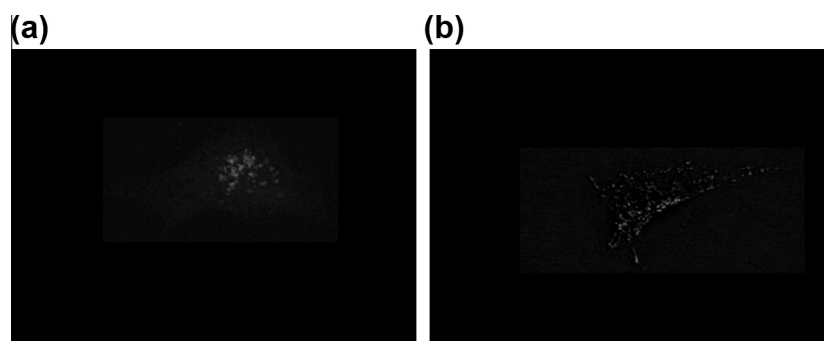


Fig. 6. CHO dataset samples: (a) Giantin and (b) Lamp2.

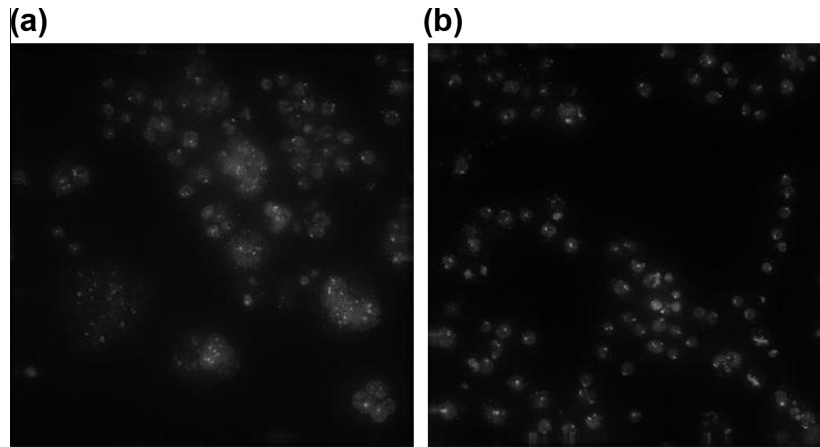


Fig. 7. RNAi dataset samples: (a) CG01258 and (b) CG07922.

Table 1

Performance of the descriptors HAR, SHAPE, and FUS.

	HAR	SHAPE	SH + PAT	SH + PAT + OV	FUS	FUS+
HeLa	83.84 (86.63)	69.53 (72.44)	70.70 (74.30)	70.93 (74.19)	86.05 (86.63)	85.58 (86.74)
LE	90.80 (92.00)	79.00 (82.60)	80.00 (81.80)	80.80 (83.00)	91.40 (92.60)	92.00 (93.40)
LT	92.00 (92.55)	65.27 (70.18)	68.73 (71.27)	64.73 (71.45)	92.18 (92.36)	92.36 (92.00)
CHO	94.46 (95.69)	97.85 (99.38)	98.15 (99.08)	98.15 (98.77)	98.77 (99.08)	98.77 (99.08)
RNAi	85.00 (87.00)	77.00 (76.50)	76.50 (77.50)	76.50 (78.00)	88.50 (88.50)	88.50 (90.00)
AVG	89.22 (90.77)	77.73 (80.22)	78.81 (80.79)	78.22 (81.08)	91.38 (91.83)	91.44 (92.24)

3.4. CHO

This dataset⁴ (Boland et al., 1998) contains 327 fluorescent microscopy images from Chinese Hamster Ovary cells. Images are 16 bit grayscale of size 512 by 382 pixels (see Fig. 7 for some examples). The following 5 classes are included in this dataset:

- (1) Giantin (77 images);
- (2) Hoechst (69 images);
- (3) Lamp2 (97 images);
- (4) Nop4 (33 images);
- (5) Tubulin (51 images).

3.5. RNAi

This dataset⁵ (Zhang & Pham, 2011) contains 200 fluorescence microscopy images of fly cells subjected to a set of gene-knockdowns using RNAi. The cells are stained with DAPI to visualize their nuclei. Images are 16 bit grayscale of size 1024 by 1024 (see Fig. 7 for some examples). Ten genes were selected to represent the classes, and their Gene IDs are used for the class labels. Each class contains 20 images.

4. Experimental results

For the evaluation protocol, we use 5-fold cross-validation for testing each of the texture descriptors. The performance indicator

is accuracy since that is the indicator widely used in papers published using these datasets.

The first experiment was aimed at establishing which version of the co-occurrence matrix works best to classify the images in the five datasets. Table 1 reports different approaches (see Section 2.1 for a description of these features):

- HAR: the standard Haralick method.
- SHAPE: the shape based features extracted from the entire co-occurrence matrix.
- SH + PAT: three shape based descriptors (each used for training a different SVM, then the set of SVMs is combined by sum rule). The first descriptor is extracted from the whole co-occurrence matrix; the second and third from two subwindows of the GLDM defined from the coordinates (0, 0) to (127, 127) and (128, 128) to (255, 255), respectively.
- SH + PAT + OV: where five descriptors are combined: the three used in SH + PAT and two others extracted from two subwindows of the GLDM defined by coordinates (0, 0) to (191, 191) and (64, 64) to (255, 255), respectively.
- FUS: the fusion by sum rule of HAR and SHAPE.
- FUS+: the fusion by sum rule of HAR and SH + PAT + OV.

In each cell of Table 1, two values are reported: the first is the accuracy obtained by a stand-alone SVM and the second is the accuracy obtained by an RS of 50 SVMs (RSS). The last row of the table, AVG, reports the average accuracy on the five datasets.

It is clear that FUS+ is the best approach and that the different approaches obtain higher performance when RSS is used as classifier. This is in contrast with the histogram-based descriptors reported in Paci et al. (2013), where RSS and stand-alone SVMs obtained similar performance using the datasets tested in this paper. This is probably due to correlation in the features extracted by HAR and SHAPE, a problem which is partially solved by RSS.

Table 2

Combining HAR and SHAPE.

	FUS (W = 1)	FUS (W = 2)	FUS (W = 3)	CONC
HeLa	86.63	86.98	86.98	86.74
LE	92.60	93.20	92.80	90.60
LT	92.36	94.00	94.55	93.27
CHO	99.08	98.77	97.54	99.38
RNAi	88.50	90.00	89.00	84.00
AVG	91.83	92.59	92.18	90.79

⁴ Download from <http://ome.grc.nia.nih.gov/iicbu2008/hela/index.html#cho>.

⁵ Download from <http://ome.grc.nia.nih.gov/iicbu2008/rnai/index.html>.

Table 3

Combining HAR and SH + PAT + OV.

	FUS+(W = 1)	FUS+(W = 2)	FUS+(W = 3)
HeLa	86.74	86.74	86.63
LE	93.40	93.80	93.60
LT	92.00	93.64	93.82
CHO	99.08	98.15	97.23
RNAi	90.00	89.50	89.00
AVG	92.24	92.36	92.05

Table 4

Q-statistic between HAR and SHAPE.

HeLa	LE	LT	CHO	RNAi
0.63	0.61	0.36	0.98	0.71

Table 5

Accuracy obtained using sum rule to combine the classifiers.

	NewHARA	MLQP	MLP	FUS	FUS + NH	FUS + 0.5 × NH
HeLa	86.98	92.56	93.72	93.95	94.42	94.42
LE	93.20	98.20	98.80	98.80	98.80	99.20
LT	94.00	98.00	98.36	98.00	98.18	98.36
CHO	98.77	99.38	99.08	99.69	99.38	99.38
RNAi	90.00	90.63	74.38	90.00	92.00	92.00
AVG	92.59	95.75	92.87	96.08	96.56	96.67

The novel approach based on SHAPE (viz., SH + PAT + OV) outperforms the original SHAPE method. Moreover, the simple sum rule with HAR (i.e., FUS+) outperforms HAR in most of the tested datasets.

In Table 2 we attempt to improve FUS by combining HAR and SHAPE with a weighted sum rule, where the weight of SHAPE fixed to 1 while the weight of HAR is W . The label CONC in Table 2 refers to an RSS trained by concatenating the features HAR and SHAPE.

It is clear that at least in the datasets tested in this paper the best performance is obtained by the weighted sum rule, as seen by the fact that HAR, which varies in weight, outperforms SHAPE.

We perform the same tests, see Table 3, also using FUS+. In these tests FUS+ and FUS perform equally well.

As an additional experiment, we investigate the relationship between HAR and SHAPE by evaluating the error independence. Table 4 reports the average Yule's Q-statistic (Kuncheva & Whitaker, 2003) in the three tested datasets for each couple of feature vectors. For two classifiers, G_i and G_j , the Q-statistic is a *a posteriori* measure that is defined as:

$$Q_{ij} = \frac{N^{11}N^{00} - N^{01}N^{10}}{N^{11}N^{00} + N^{01}N^{10}}$$

where N^{ab} is the number of instances in the testing set classified as correctly ($a = 1$) or incorrectly ($a = 0$) by classifier G_i and correctly ($b = 1$) or incorrectly ($b = 0$) by the classifier G_j . Q varies between -1 and 1 , and $Q_{ij} = 0$ for statistically independent classifiers. Classifiers that tend to recognize the same patterns correctly will have $Q > 0$, and those which commit errors on different patterns will have $Q < 0$. In this problem the Q-statistic values are low enough (see Sundaram, Ramar, Arumugam, and Prabin (2011) to validate the idea of concatenating the vectors.

In Table 5, we compare the following approaches:

- NH: the weighted sum rule “FUS ($W = 2$),” as reported in Table 2;
- MLQP: the descriptor described in Section 2.2;

- MLP: the descriptor described in Section 2.3;
- FUS⁶: fusion by sum rule of MLQP and MLP;
- FUS + NH₆, fusion by sum rule of MLQP, MLP, and NH;
- FUS + $0.5 \times$ NH₆: fusion by weighted sum rule between MLQP, MLP, and NH, where the weight of MLQP and MLP is 1 and the weight of NH is 0.5.

We want to stress that in this set of experiments we have not reported the performance obtained by standard approaches, such as LBP, LTP, and LPQ⁷, since it was shown in Paci et al. (2013) that these methods were outperformed by MLQP and MLP using the same datasets.

The results reported in Table 5 show that the fusions of MLQP, MLP, and NH outperform FUS. The main drawback of fusion methods is the high computation time, which makes such methods unsuitable for real-time applications. In the HeLa dataset, for instance, the computation time needed to extract all the descriptors that built MLQP was the following:

- 46 s on an Intel Duo P8600 processor (2 cores) running standard MATLAB code;
- 38 s on an Intel Duo P8600 processor (2 cores) running the MATLAB parallel toolbox to exploit the multi-core architecture of the chip;
- 6.5 s on an Intel i5-2500 processor (4 cores) using Matlab parallel toolbox to exploit the multi-core architecture.

The difference in the computation time between the Duo P8600 and the i5-2500 is quite remarkable; in a few years, it may be possible for newer multi-core processors to performed these fusions almost in real-time.

In Table 6, we offer a brief comparison with other descriptors and classifiers representative of the state of art, focusing in particular on those works using the same datasets investigated in this paper. In Zhang and Pham (2011), the HeLa, CHO, and RNAi datasets were used to assess the performances of a multi-layer perceptron (using stand-alone and random subspace methods) trained with the Grey Level Co-occurrence Matrix second-order statistics and curvlet-based features. In Nanni, Brahnam, and Lumini (2010) a set of four global and two local descriptors obtained in the HeLa, LE, and LT datasets obtained the following accuracies: 95.8%, 99.5%, and 97.0%, respectively. But these results were obtained via concatenation of both global and local descriptors. The best results obtained using a single descriptor were 90.8%, 95.7% and 91.8%, respectively.

Other recent results include the following:

- In Chebira et al. (2007) a multi-resolution system obtains an accuracy of 95.4% in the HeLa dataset;
- In Hamilton et al. (2007) an ensemble of neural networks trained using the concatenation between the Haralick texture features and the Threshold Adjacency Statistics obtained the following accuracies in the LE and LT datasets: 98.2% and 93.2%, respectively.
- In Nanni et al. (2010) the same descriptors were applied to HeLa, LE, and LT datasets and used to feed a RS of Levenberg–Marquardt neural networks with five hidden nodes; this approach resulted in the following accuracies: 88.9%; 94.7% and 93.3%, respectively.

⁶ Before the fusion the scores of each classifier are normalized to mean 0 and standard deviation 1.

⁷ The value of the parameters were selected using an internal 10-fold cross validation using only the training data (test set is blind).

Table 6

Brief comparison between our results and other state of art descriptors and classifiers.

	New	Zhang ^a	Nanni ^b		Chebira ^c	Hamilton ^d	Lin ^f
			Concatenated descriptors	Single descriptor			
HeLa	94.42	91.20	95.8	90.8	95.4	– 88.9 ^e	93.6
CHO	99.38	98.86	–	–	–	–	–
RNAi	92.00	91.03	–	–	–	–	–
LE	99.2	–	99.5	95.7	–	98.2 94.7 ^e	–
LT	98.36	–	97.0	91.8	–	93.2 93.3 ^e	–

^a Zhang and Pham (2011).^b Nanni et al. (2010).^c Chebira et al. (2007).^d From the original paper of Hamilton et al. (2007).^e Nanni et al. (2010) same descriptors used in Hamilton et al. (2007), classification with a random subspace of Levenberg–Marquardt neural networks.^f Lin et al. (2007).

- In Lin et al. (2007) a novel AdaBoost method trained using global and local features obtained an accuracy of 93.6% in the HeLa dataset.

It is clear that our proposed system gets very good results; we expect to improve the performance further by combining several texture descriptors with our ensemble method.

5. Discussion and conclusion

In this paper we focused on the problem of cell phenotype image classification by studying a range of features that can be extracted from the co-occurrence matrix (CM). Starting from the analysis of the best methods and descriptors proposed in the literature, we compared a number of different approaches and demonstrated the power of combining standard Haralick features with a set of novel features derived from the CM using five different benchmark datasets.

We believe the descriptors described in this paper will prove valuable for diagnostic and prognostic purposes using many types of medical imagery. In the near future, we plan to test the proposed fusion between HAR and SHAPE on more biological and medical image classification problems to confirm the value of combining different sets of features from the CM. We also plan on studying more methods for extracting features from CM. Yet another future study might focus on the development of heterogeneous ensembles of classifiers that combine results using the same set of descriptors.

Acknowledgment

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007–2013) under grant agreement No. 284607.

References

- Akhlofi, M., & Bendada, A. (2010). Locally adaptive texture features for multispectral face recognition. In *IEEE international conference on systems man and cybernetics (SMC)*, 2010 (pp. 3308–3314).
- Ameling, S., Wirth, S., Paulus, D., Lacey, G., & Vilarino, F. (2009). Texture-based polyp detection in colonoscopy. In Hans-Peter Meinzer, Thomas Martin Deserno, & Heinz Handels, et al. (Eds.), *Bildverarbeitung für die Medizin* (pp. 346–350). Berlin and Heidelberg, Germany: Springer.
- Bachmann, I. M., Straume, O., Puntervoll, H. E., Kalvenes, M. B., & Akslen, L. A. (2005). Importance of p-cadherin, b-catenin, and wnt5a/frizzled for progression of melanocytic tumors and prognosis in cutaneous melanoma. *Clinical Cancer Research*, 11, 606–614.

- Bedard, J. E., Haaning, A. M., & Ware, S. M. (2011). Identification of a novel ZIC3 isoform and mutation screening in patients with heterotaxy and congenital heart disease. *Plos One*, 6(8), e23755.
- Boland, M. V., Markey, M. K., & Murphy, R. F. (1998). Automated recognition of patterns characteristic of subcellular structures in fluorescence microscopy images. *Cytopathology*, 33(3), 366–375.
- Boland, M. V., & Murphy, R. F. (2001). A neural network classifier capable of recognizing the patterns of all major subcellular structures in fluorescence microscope images of HeLa cells. *Bioinformatics*, 17(12), 1213–1223.
- Chan, C.-H., Tahir, M. A., Kittler, J., & Pietikäinen, M. (2013). Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(5), 1164–1177.
- Chebira, A., Barbotin, Y., Jackson, C., Merryman, T., Srinivasa, G., Murphy, R. F., et al. (2007). A multiresolution approach to automated classification of protein subcellular location images. *BMC Bioinformatics*, 8, 210.
- Chen, X., Velliste, M., & Murphy, R. (2006). Automated interpretation of subcellular patterns in fluorescence microscope images for location proteomics. *Cytometry*, 69(7), 631–640.
- Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., et al. (2005). VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Research*, 33, D325–D328.
- Conrad, C., Erfle, H., Warnat, P., Daigle, N., Lorch, T., Ellenberg, J., et al. (2004). Automatic identification of subcellular phenotypes on human cell arrays. *Genome Research*, 14(6), 1130–1136.
- Danckaert, A., Gonzalez-Couto, E., Bollondi, L., Thompson, N., & Hayes, B. (2002). Automated recognition of intracellular organelles in confocal microscope images. *Traffic*, 3, 66–73.
- Fernandes, A. P., Capitanio, A., Selenius, M., Brodin, O., Rundlöf, A.-K., & Björnstedt, M. (2009). Expression profiles of thioredoxin family proteins in human lung cancer tissue: Correlation with proliferation and differentiation. *Histopathology*, 54, 313–320.
- Fink, J. L., Aturaliya, R. N., Davis, M. J., Zhang, F., Hanson, K., Teasdale, M. S., et al. (2006). LOCATE: A protein subcellular localization database. *Nucleic Acids Research*, 34.
- Geerts, D., Wallick, C. J., Koomoa, D.-L. T., Koster, J., Versteeg, R., Go, R. C. V., et al. (2007). Expression of prenylated rab acceptor 1 domain family, member 2 (praf2) in neuroblastoma: Correlation with clinical features, cellular localization, and cerulenin-mediated apoptosis regulation. *Human Cancer Biology*, 13, 6312–6319.
- Ghidoni, S., Cielniak, G., & Menegatti, E. (2012). Texture-based crowd detection and localisation. In *The 12th International Conference on Intelligent Autonomous Systems*, 2012, Jeju Island, Korea.
- Glory, E., & Murphy, R. (2007). Automated subcellular location determination and high throughput microscopy. *Developmental Cell*, 12, 7–16.
- Glory, E., Newberg, J., & Murphy, R. F. (2008). Automated comparison of protein subcellular location patterns between images of normal and cancerous tissues. *PMC*, PMC2901901, 304–307.
- Hamilton, N., Pantelic, R., Hanson, K., & Teasdale, R. D. (2007). Fast automated cell phenotype classification. *BMC Bioinformatics*, 8–110.
- Hamilton, N. A., Wang, J. T. H., Kerr, M. C., & Teasdale, R. D. (2009). Statistical and visual differentiation of subcellular imaging. *BMC Bioinformatics*, 10(94).
- Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67(5), 786–804.
- Ho, T. K. (1998). The random subspace method for constructing decision forests. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(8), 832–844.
- Huang, K., & Murphy, R. (2004). Boosting accuracy of automated classification of fluorescence microscope images for location proteomics. *BMC Bioinformatics*, 5(78).
- Karkanis, S. A., Iakovidis, D. K., Maroulis, D. E., Karras, D. A., & Tzivras, M. (2003). Computer-aided tumor detection in endoscopic video using color wavelet

- features. *IEEE Transaction on Information Technology in Biomedicine*, 7(3), 141–152.
- Knostman, K. A. B., McCubrey, J. A., Morrison, C. D., Zhang, Z., Capen, C. C., & Jhiang, S. M. (2007). PI3K activation is associated with intracellular sodium/iodide symporter protein expression in breast cancer. *BMC Cancer*, 7, 137.
- Kuncheva, L. I., & Whitaker, C. J. (2003). Measures of diversity in classifier ensembles and their relationship with the ensemble accuracy. *Machine Learning*, 51(2), 181–207.
- Lin, C. C., Tsai, Y. S., Lin, Y. S., Chiu, T. Y., Hsiung, C. C., Lee, M. I., et al. (2007). Boosting multiclass learning with repeating codes and weak detectors for protein subcellular localization. *Bioinformatics*, 23(24), 3374–3381.
- Mezzanzanica, D., Fabbì, M., Bagnoli, M., Staurengo, S., Losa, M., Balladore, E., Alberti, P., Lusa, L., Ditto, A., Ferrini, S., Pierotti, M. A., Barbareschi, M., Pilotti, S., & Canevari, S. Subcellular localization of activated leukocyte cell adhesion molecule is a molecular predictor of survival in ovarian carcinoma patients. *Clinical Cancer Research*, 14, pp. 726–733.
- Moreira, J. M. A., Ohlsson, G., Gromov, P., Simov, R., Sauter, G., Celis, J. E., et al. (2010). Bladder cancer-associated protein, a potential prognostic biomarker in human bladder cancer. *Molecular & Cellular Proteomics*, 9, 161–177.
- Murphy, R. F. (2006). Putting proteins on the map. *Nature Biotechnology*, 24(10), 1223–1224.
- Nanni, L., Brahnam, S., & Lumini, A. (2010). Novel features for automated cell phenotype image classification. In H. R. Arabnia (Ed.), *Advances in computational biology: Advances in experimental medicine and biology (AEMB)* (pp. 207–213). New York: Springer.
- Nanni, L., Ghidoni, S., & Menegatti, E. (2013). A comparison of multi-scale approaches for extracting image descriptors from the co-occurrence matrix (in press).
- Nanni, L., & Lumini, A. (2008). A reliable method for cell phenotype image classification. *Artificial Intelligence in Medicine*, 43(2), 87–97.
- Nanni, L., Lumini, A., & Brahnam, S. (2010). Local binary patterns variants as texture descriptors for medical image analysis. *Artificial Intelligence in Medicine*, 49(2), 117–125.
- Nanni, L., Paci, M., Brahnam, S., Ghidoni, S., & Menegatti, E. (2013b). Virus image classification using different texture descriptors. *IPCV*.
- Nishibori, Y., Liu, L., Hosoyamada, M., Endou, H., Kudo, A., Takenaka, H., et al. (2004). Disease-causing missense mutations in NPHS2 gene alter normal nephrin trafficking to the plasma membrane. *Kidney International*, 66(5), 1755–1765.
- Obrenovich, M. E., Smith, M. A., Siedlak, S. L., Chen, S. G., de la Torre, J. C., Perry, G., et al. (2006). Overexpression of GRK2 in Alzheimer disease and in a chronic hypoperfusion rat model is an early marker of brain mitochondrial lesions. *Neurotoxicity Research*, 10(1), 43–56.
- Ojala, T., Pietikainen, M., & Maenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7), 971–987.
- Ojansivu, V., & Heikkilä, J. (2008). Blur insensitive texture classification using local phase quantization. In *ICISP*, 2008 (pp. 236–243).
- Oliver, A., Lladó, X., Freixenet, J., & Martí, J. (2007). False positive reduction in mammographic mass detection using local binary patterns. In *Medical Image Computing and Computer-Assisted Intervention (MICCAI). Lecture Notes in Computer Science* (pp. 286–293). Berlin: Springer. 4791.
- Paci, M., Nanni, L., Lathi, A., Aalto-Setälä, K., Hyttinen, J., & Severi, S. (2013). Non-binary coding for texture descriptors in sub-cellular and stem cell image classification. *Current Bioinformatics*, 8(2), 208–219.
- Perner, P., Perner, H., & Muller, B. (2002). Mining knowledge for hep-2 cell image classification. *Artificial Intelligence in Medicine*, 26(1–2), 161–173.
- Perrone, G., Zagami, M., Santini, D., Vincenzi, B., Gullotta, G., Morini, S., et al. (2007). COX-2 expression in lobular in situ neoplasia of the breast: Correlation with histopathological grading system according to the Tavassoli classification. *Histopathology*, 51(1), 33–39.
- Piazza, T., Cha, E. I. B., Polito, L., Bargellesi, A., Sassi, F., Ferrini, S., et al. (2005). Internalization and recycling of ALCAM/CD166 detected by a fully human single-chain recombinant antibody. *Journal of Cell Science*, 118(7), 1515–1525.
- Ralhan, R., He, H. C.-H., So, A. K.-C., Tripathi, S. C., Kumar, M., Hasan, R., et al. (2010). Nuclear and cytoplasmic accumulation of Ep-ICD is frequently detected in human epithelial cancers. *Plos One*, 5(11), 1–15.
- Rosenblatt, R., Jonmarker, S., lewensohn, R., Egevad, L., Sherif, A., Kalkner, K. M., et al. (2008). Current status of prognostic immunohistochemical markers for urothelial bladder cancer. *Tumor Biology*, 29, 311–322.
- Samant, A., & Adeli, H. (2000). Feature extraction for traffic incident detection using wavelet transform and linear discriminant analysis. *Computer-Aided Civil and Infrastructure Engineering*, 15(4), 241–250.
- Shen, L., Bai, L., & Fairhurst, M. (2007). Gabor wavelets and general discriminant analysis for face identification and verification. *Image and Vision Computing*, 25, 553–563.
- Sundaram, M., Ramar, K., Arumugam, N., & Prabin, G. (2011). Histogram modified local contrast enhancement for mammogram images. *Applications of Soft Computing*, 11(8), 5809–5816.
- Surowiak, P., Materna, V., Kaplenko, I., Spaczynski, M., Dolinska-Krajewska, B., Gebarowska, E., et al. (2006). Abcc2 (Mrp2, cMOAT) can be localized in the nuclear membrane of ovarian carcinomas and correlates with resistance to cisplatin and clinical outcome. *Clinical Cancer Research*, 12, 7149–7158.
- Tan, X., & Triggs, B. (2007). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *Analysis and Modelling of Faces and Gestures, LNCS 4778*, 168–182.
- Tang, J. P., Tan, C. P., Li, J., Siddique, M. M., Guo, K., Chan, S. W., et al. (2010). VHZ is a novel centrosomal phosphatase associated with cell growth and human primary cancers. *Molecular Cancer*, 9, 128–136.
- Thomas, A. Q., Lane, K., III, J. P., Prince, M., Markin, C., Speer, M., Schwartz, D. A., Gaddipati, R., Marney, A., Johnson, J., Roberts, R., Haines, J., Stahlman, M., & Loyd, J. E. (2002). Heterozygosity for a surfactant protein c gene mutation associated with usual interstitial pneumonitis and cellular nonspecific interstitial pneumonitis in one kindred. *American Journal of Respiratory and Critical Care Medicine*, 165(9), 1322–1328.
- Unay, D., & Ekin, A. (2008). Intensity versus texture for medical image search and retrieval. In *Fifth IEEE international symposium on biomedical imaging: From Nano to Macro*, pp. 241–244.
- Vapnik, V., & Chervonenkis, A. (1964). A note on one class of perceptrons. *Automation and Remote Control*, 25.
- Vapnik, V., & Lerner, A. (1963). Pattern recognition using generalized portrait method. *Automation and Remote Control*, 24, 774–780.
- Vécsei, A., Amann, G., Hegenbart, S., Liedlgruber, M., & Uhl, A. (2011). Automated marsh-like classification of celiac disease in children using local texture operators. *Computers in Biology and Medicine*, 41(6), 313–325.
- Wang, X., & Tang, X. (2006). Random sampling for subspace face recognition. *Journal of Computer Vision*, 70(1), 91–104.
- Zhang, B., & Pham, T. D. (2011). Phenotype recognition with combined features and random subspace classifier ensemble. *BMC Bioinformatics*, 12(1), 128.