Accurate HEP-2 cell classification based on Sparse Coding of Superpixels

Shahab Ensafi a,b,∗, Shijian Lu b, Ashraf A. Kassim a, Chew Lim Tan c

a Electrical and Computer Engineering, Faculty of Engineering, National University of Singapore, 4 Engineering Drive 3, Block E4, 5-45, Singapore 117583, Singapore
b Institute for Infocomm Research, Agency for Science, Technology and Research (A*STAR), 1 Fusionopolis Way, 21-01 Connexis (South Tower), Singapore 138632, Singapore
c Department of Computer Science, School of Computing, National University of Singapore, Computing 1, 13 Computing Drive, Singapore 117417, Singapore

A R T I C L E   I N F O
Article history:
Available online 26 February 2016

Keywords:
Autoimmune diseases
HEP-2 cells
Computer-aided diagnosis
Cell classification
Superpixel
Sparse coding

A B S T R A C T
Autoimmune Diseases (AD) are among the top 10 leading causes of death in female children and women in all age groups up to 64 years. They are widely diagnosed by various antibody tests that typically apply the Indirect Immunofluorescence (IIF) to the Human Epithelial Type-2 (HEP-2) cells. Automated classification of HEP-2 cells has attracted much research interest in recent years, and many of these approaches employ patch-based models and the Bag of Words (BoW) scheme, but often face several typical constraints such as the need to process a huge number of overlapped image patches, tuning of various parameters and etc. We propose a superpixel based HEP-2 cell classification technique by calculating the sparse codes of image patches which are prepared in a more intelligent way. In particular, the superpixel approach guides the determination of the right image patches while aggregating the neighboring pixels of similar patterns. In addition, we introduce “extended superpixels” which is able to capture the most discriminative gradient information across the boundary of the HEP-2 cell images. The proposed technique has been evaluated over two public datasets (ICPR2012 and ICIP2013) and experiments show superior performance in both classification accuracy and speed of model training and cell classification.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

The diagnosis of Autoimmune Diseases (AD) commences with the imaging of affected organ through Indirect Immunofluorescence (IIF). Specifically, antibodies are first stained in a tissue and then bound to a fluorescent chemical compound. In case of Antinuclear Antibodies (ANAs), the antibodies bound to the nucleus of the Human Epithelial type 2 (HEP-2) demonstrate different visual patterns that can be captured and visualized within microscope images [1]. Categorizing the patterns of the HEP-2 cell images distinguishes the phase and severeness of the disease.

Automatic classification of the HEP-2 cells has been attracting much research interest in recent years, due to its importance to the diagnosis of AD. This problem originated from the increasing demand of AD tests and the lack of certified physicians to facilitate the test. Additionally, the repeatability of the test across different physicians is another issue that makes the AD diagnosis an even more challenging problem.

Different HEP-2 cell classification techniques have been reported and most are based on overlapping windows with the same sizes, which are called patch based image processing. The approach first extracts various visual features such as local binary pattern (LBP) [2] and Scale-invariant feature transform (SIFT) [3–5] from each image patch, and then represents them through different transformations such as locally-constrained linear coding [6], spacial pyramid matching [7], independent component analysis [8], non-parametric Bayesian model [9] etc. The processed features are further classified by support vector machines or similar classifiers [10–12] which group HEP-2 cells into different categories as illustrated in Fig. 3.

Sparse Coding technique is widely used in different machine learning problems [13–16], where image patches are used to calculate the features to train the dictionary and the classifiers.

Although some success has been achieved for HEP-2 cell classification as reported in the recent benchmarking tests at ICPR2012 [17] and ICIP2013 [18], the best accuracy obtained is generally lower than what physicians can provide. The major constraints of these image patch techniques include the high computational cost

http://dx.doi.org/10.1016/j.patrec.2016.02.007
© 2016 Elsevier B.V. All rights reserved.
due to a huge amount of overlapped image patches to be processed and the tedious parameter tuning (for patch size, scanning step size, etc.) for optimal cell classification performance.

In this paper we propose a novel superpixel based HEp-2 cell classification technique by using sparse coding scheme, which is widely used in different image processing problems [19]. We call our proposed method the Sparse Coding of Superpixels (SCS).

In the superpixel approach which is widely used in segmentation problems [20], the input image is divided to relatively small and non-overlapped regions. Each superpixel contains many connected pixels which have similar features. In segmentation problems, the features of each superpixels are used to decide whether the superpixel should aggregate with the neighboring superpixels to make a bigger region and the process continues until the final segmentation result is obtained. However, for classifying images, not only the features of pixels within the superpixels but also superpixels boundaries which are usually aligned with the high gradient regions should be analyzed.

Our proposed superpixel based method is different from those earlier superpixel based strategies for classification problem and to the best of our knowledge, is the first time that is applied to the HEp-2 cell classification problem (see Section 2.2). These differences are highlighted below:

- superpixels are used instead of regular sampling of overlapped image patches to guide the selection of the right image patches to use in sparse coding method;
- superpixels are used to obtain the image patches which contain more ‘informative’ features rather than solely using overlapped patches;
- we introduce “extended superpixels” that are derived by dilating the boundary of each superpixel and are designed to capture the most discriminative gradient information across the HEp-2 cell boundary;
- we introduce a cell extraction method to extract a certain number of cells from each specimen image whose shapes, sizes and morphological features are similar together.

Experiments on two benchmarking datasets show that the proposed technique obtains superior HEp-2 cell classification performance.

The rest of this paper is organized as follows. The proposed SCS method is described in Section 2, including the cell extraction, superpixel and dictionary learning scheme. In Section 3 the experiments on two publicly available datasets are investigated and validation on the parameters of the proposed superpixel method is discussed. Finally, we provide our conclusions in Section 4.

2. Sparse Coding of Superpixels (SCS)

Our proposed SCS technique comprises the following four stages as illustrated in Fig. 1: cell extraction, superpixel extraction, dictionary learning and cell classification. The cell extraction stage is specially needed when the bounding boxes of the cells are not provided (as in specimen classification of ICIP2013). In the superpixel extraction stage, the superpixels are first extracted to form an informative image patch. SIFT and SURF features are then extracted from each superpixel to learn an over-complete dictionary. Finally, a linear SVM classifier is trained for HEp-2 cell classification.

2.1. Cell extraction

In the cell extraction stage, the goal of finding cells with similar shapes and sizes is achieved using morphological operations which result in connected pixels extracted from the image masks that are analyzed to select those that correspond to real cells for cell classification as shown in the ‘Cell Extraction’ stage of Fig. 1. There is a need to distinguish those large connected pixels which could have been wrongly segmented or represent overlapped cells.

The area and solidity morphological features are extracted from each connected pixel for cell classification. The histogram of the area features are quantized into bins and the maximum bin \(b_{mn}\) is taken to represent the area for most connected pixels that are likely to be proper cells with no overlaps with other cells. We then proceed to select those cells with area that are close to \(b_{mn}\) based on the variance of the histogram (Var) resulting in range \(R_{mn}^{max}; R_{mn}\) respectively) as follows:

\[
R_{mn} = \max(0, b_{mn} - \text{Var})
\]

\[
R_{mx} = \min(b_{mx}, b_{mx} + \text{Var})
\]

Only those connected pixels with areas more than a threshold \(A_t\) are considered as correctly identified cells. The value of \(R_{mn}\) is not constant and will change at each iteration by decreasing the number of quantized bin \(nb\) and accordingly calculated variance (Var).

In some cases, for instance, Centromere class, there are small shiny dot patterns in each cell or in Nucleolar class, there are bright regions inside the cell. Therefore, those small regions are very likely to be chosen wrongly as cells. To avoid such situations, we need a constant value \(A_t\) which can avoid selection of the small components as cells.

Next, cells with similar shape are identified using the solidity, \(S\), property which is related to the roundness of the cells and is defined as follows:

\[
S = \frac{A}{CA}
\]

Fig. 1. Framework of the proposed HEp-2 cell classification technique.
Algorithm 1: Cell extraction algorithm.

Input: Specimen images and masks
Output: Extracted cells

1 begin
2 \( nb = 12, S_{mn} = 0.98, A_t = 45; \)
3 for each mask image do
4 \( \) Get the connected pixels;
5 \( \) Get the morphological properties (Area, Solidity);
6 \( \) Fit a Gaussian probability distribution on the
7 \( \) histogram of the areas with nb bins;
8 \( \) Calculate the variance (Var) of fitted Gaussian
9 \( \) distribution;
10 \( R_{mn} = \min(0, b_{max} - Var); \)
11 \( R_{max} = \min(b_{max}, b_{max} + Var); \)
12 \# selected cells \( \rightarrow 0; \)
13 for each area(cell) \( \in [R_{mn}, R_{max}] \) do
14 \( \) if solidity > \( S_{mn} \) then
15 \( \) if area > \( A_t \) then
16 \( \) \( \) if area \& boundary = \( \{\} \) then
17 \( \) \( \) Get the cell;
18 \( S_{mn} \leftarrow 0.95 \times S_{mn}; \)
19 \( nb \leftarrow nb - 1; \)
20 Go to 6;

where \( A \) is the area of the connected pixel and \( CA \) (Convex Area) is the number of pixels in the convex hull of the area. The solidity \( S \) would be close to one if the cells are of circular shape.

As shown in Algorithm 1, in each iteration for one specimen image, the number of bins (\( nb \)) and the values of \( S_{mn} \) (minimum solidity) are decreased gradually to select at least 5 cells in each image. By decreasing these values, we gradually relax the constraints for selecting the cells, because in some mask images, the connected pixels have irregular shapes rather than circular shapes. The initial values are selected with a cross validation strategy: \( nb = 12; S_{mn} = 0.98; A_t = 45 \). This helps choose those cells that of average size (i.e., area) and circular in shape.

2.2. Superpixel extraction

We extract the superpixels based on the Simple Linear Iterative Clustering (SLIC) [21] due to its distinct properties of low computational cost and close adherence to the object boundaries in comparison with similar methods [22,23]. The original SLIC method, places grid points (\( P \) points) to be the initial superpixel centers. If an initial superpixel center lies along the cell boundary, another cell image pixel with the minimum gradient value lying within the \( 3 \times 3 \) neighborhood of the original center is selected as the initial superpixel center. K-means clustering is then performed to calculate the distance between the cluster centers and the neighboring pixels according to their intensity (color or gray scale) values and positions. The distance calculation is accomplished in the \( 2W \times 2W \) window size (\( W \) denotes the superpixel size which can be determined by \( \sqrt{s} \), where \( s \) is the number of desired pixels in each superpixel) to find similar neighboring pixels instead of the whole image area in order to increase the speed of the algorithm.

This method tries to minimize the distance of the color and pixel positions with the cluster centers \( [I, a, b, x, y]^T \), where the first three elements are the CIELAB color space parameters and the last two are the position of pixels. These different distances are normalized in order to be aggregated together in one formulation as in Eq. (3).

\[
d_c = \sqrt{(l_i - l_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2}
\]

\[
d_s = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}
\]

\[
D = \sqrt{\frac{d_c}{N_c}} + \left(\frac{d_s}{N_s}\right)^2
\]

where \( N_c \) and \( N_s \) are color and spatial proximities by their maximum distances within a cluster [21].

Our modifications to the SLIC superpixel algorithm are as follows.

i. besides the color and spatial proximities, we added the gradient information to also enable the algorithm to evaluate the texture information resulting in the following extended set of parameters: \( [l, a, b, x, y, g]^T \); where \( g \) is the magnitude of gradient in each pixel. The gradient distance \( d_g \) is normalized with \( N_g \) which is the maximum gradient distance between the pixels and the cluster center in one cluster as formulated in Eq. (4).

\[
D = \sqrt{\frac{d_c}{N_c}} + \left(\frac{d_s}{N_s}\right)^2 + \left(\frac{d_g}{N_g}\right)^2
\]

where \( d_g \) is

\[
d_g = |g_j - g_i|
\]

ii. while the SLIC uses the number of the desired superpixels \( P \) to control the superpixel size which does not work well as different images have different sizes and accordingly very different superpixel sizes \( s \) (the number of pixel within a superpixel), we use the superpixel size \( s \) as the input parameter to ensure that superpixels from images of different sizes will have a similar superpixel size. Note that the superpixel size \( s \) can be translated to the number of superpixels \( P \) by \( P = [N/s] \), where \( N \) denotes the number of pixels within the image.

iii. the SLIC method adheres to the boundaries (i.e., high gradient regions) which is desirable for object segmentation but for HEp-2 cell classification, it could cause serious problem as many discriminative features lie across the high gradient regions. We solve this problem by introducing an “extended superpixel” which is derived by dilating the original superpixel to include the “cross-boundary” information. We extend each superpixel to include the high gradient information to be discussed in Section 3.4.

The proposed superpixel extraction method is applied on the cells, before convolving with masks. Then those superpixels which are outside of the masks, are omitted. The extended superpixels can then be used for dictionary learning and cell classification.

2.3. Dictionary learning

SIFT and SURF are used as visual features for dictionary and classification model learning. In particular, SIFT and SURF features are first extracted from each extracted superpixel patch and then processed through the max-pooling of the extracted feature histograms. The processed features are then concatenated to form the feature description. The features of the superpixels are finally sampled from each input image and creates the \( D \)-dimensional feature matrix \( F = [f_1, f_2, ..., f_n]^T \in \mathbb{R}^{N \times D} \) as illustrated in Fig. 1.

The sparse coding with the BoW model is applied to learn the optimal dictionary and codes [24]. We aim to learn a dictionary \( D = [d_1, d_2, ..., d_K]^T \in \mathbb{R}^{K \times D} \) which has \( K \)-words, in other words,
D is a mapping matrix from feature space to sparse code space. Therefore, the reconstruction error should be minimized in order to get the dictionary and the sparse codes \( Z = [z_1, z_2, \ldots, z_N]^T \in \mathbb{R}^{(N \times K)} \).

\[
\min_{D,Z} \sum_{n=1}^{N} \| f_n - z_nD \|_p^2 + \lambda \| z_n \|_\phi
\]
\[
\text{s.t.} \quad \| d_k \| \leq 1, \quad \forall k = 1, 2, \ldots, K
\]

(6)

The \( \| \cdot \|_\phi \) in Eq. (6) is \( p \)-norm where \( \| x \|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p} \). Eq. (6) is not convex, which can be solved iteratively fill a sparsest solution is obtained [25]. In particular, the values of the dictionary is initialized by using k-means method which produces the cluster centers of the input features as the dictionary words. By keeping the dictionary values fixed, the weights are calculated by using conjugate gradient method. The weights are then fixed and the dictionary words are optimized [7].

2.4. Classification

As HEP-2 cells have multiple classes, the multi-class linear SVM is trained for the cell classification. A strategy of one-versus-all is applied which learns \( L \) (number of classes) linear binary classifiers and concatenated to form the final classifier.

\[
\min_{w_i} \left\{ J(w_i) = \| w_i \|_2^2 + C \sum_{i=1}^{L} \epsilon(w_i; y_i^f; x_i) \right\}
\]
\[
\epsilon(w_i; y_i^f; x_i) = \max(0, 1 - y_i^fx_i + 1)^2
\]

(7)

where \( y_i^f \in \{+1, -1\} \) are the class label of input images. This optimization problem uses the differentiable hinge loss function \( \epsilon \), which can be solved using conjugate gradient method. By learning the parameters of the classifier (\( w_i \)) we can assign the class label of a test image (with feature vector \( x \)) to one of the \( L \) labels by:

\[
y = \max_{c \in \{1, \ldots, L\}} \langle w_c, x \rangle
\]

(8)

Particularly, when the label prediction of a test image is required, the feature vector (\( x \)), which is the sparse codes of the image, are calculated by the pre-learned dictionary. Then, (8) maximizes the inner product of the \( w_c \) and \( x \) to calculate the predicted label for the test image.

Fig. 2. The cell level images of ICPR2012 (a) and ICIP2013 (b) for positive (top rows) and intermediate (third rows) intensity levels and their heat-maps to show the underlying pattern.

3. Experiments and results

3.1. Datasets

We evaluate our proposed technique on two publicly available datasets. The first one is the MIVIA HEP-2 Image Dataset as published in the International Conference on Pattern Recognition 2012 (i.e., the ICPR2012 dataset) [17]. The second one is published in the International Conference on Image Processing 2013 (i.e., the ICIP2013 dataset) [26]. Both these datasets have a cell classification task (Task 1) and a specimen image classification task (Task 2). It should be noted that each specimen image contains many cells and its classification is based on the majority voting of the classified cells within the specimen image. All the images are provided with the prior knowledge of intensity levels. Figs. 2 and 3 show the Cell and Image levels of both datasets.

The ICPR2012 dataset contains 28 specimen images (half for training and half for test) and 1455 cell images (721 for training and 734 for test) [17]. It has 6 cell classes including Centromere (C), Coarse-speckled (Cs), Cytoplasmatic (Cy), Fine-speckled (Fs), Homogeneous (H) and Nucleolar (N). The ICIP2013 dataset contains 252 specimen images for training and the test set is not publicly available. It has 6 cell classes for Task 1 including Centromere (C), Golgi (G), Homogeneous (H), Nucleolar (N), Nucleolar membrane (NuMem) and Speckled (S), but one more class (Mitotic Spindle (MItSp)) for task 2.

3.2. Evaluation metric and protocol

The proposed technique is evaluated by using the Mean Class Accuracy as suggested by the contest organizers. It is defined by \( \frac{1}{c} \sum_{c=1}^{c} CC_{ck} \) where \( CC_{ck} \) is the classification rate for class \( c \) and \( C \) denotes the number of classes.

For the ICPR2012 dataset, the Leave-One-Specimen-Out (LOSO) strategy is adopted for the fair comparison with the state-of-the-art results which are reported by the dataset organizers [18]. In LOSO strategy all the training and test images are used. Separately, the accuracies on 'test set' is also evaluated by training the classification model using the training image set.

In the ICIP2013 dataset, because we only have access to the training set, two protocols which are used in the literature are exploited including Leave-One-Specimen-Out strategy as used in [27,28] and the evaluation protocol used in [29], which we named HSM representing the title of the paper (High-order Statistics of Microtexton for HEP-2 Staining Pattern Classification) for cell classification problem. All the training cells which are provided by the ICIP2013 dataset are extracted from 83 specimen images. In the first method, Leave-One-Specimen-Out strategy, in each run, the cells of one specimen image is used for testing and the rest cells of 82 specimen images are used for training. In the second evaluation method (HSM) [29], 600 cells are randomly selected from each of the classes (except 300 cell images for Golgi

\[1\] http://mivia.unisa.it/datasets/biomedical-image-datasets/hep2-image-dataset/.

class) for training and the rest for testing. This task is performed 20 times and the average accuracies of all iterations are reported. Particularly, in each iteration, $5 \times 600 + 300 = 3300$ cells are selected for training and $13,596 - 3300 = 10,296$ cells for testing ($13,596$ is the total number of training cells in the ICIP2013 dataset).

To obtain the image level classification result, a modification of the masks should be applied to get the correct cell information from the images as described in Section 2.1. By performing the proposed cell extraction method on 252 specimens (1008 images), 5012 cells are extracted. The Leave-One-Specimen-Out strategy is also used where all the cells from one specimen image are used for testing and the rest cells of 251 specimen images are used for training. To obtain the class label of each specimen image, maximum voting is performed on the cell labels of that specimen image.

The calculation of the average accuracies are in the weighted fashion, where the number of positive and intermediate images are participated to obtain the average accuracy. For example, there are 829 and 636 positive and intermediate images, respectively in ICPR2012 where the weights are calculated as $829/(829 + 636)$ and $636/(829 + 636)$.

In our experiments we extract superpixels of size 100 pixels and extend them for 9 and 12 pixels from each side for ICPR2012 and ICIP2013 datasets respectively, which results in better accuracies as shown in Fig. 4. It should be noted, because the ICIP2013 dataset contains gray scale images, the intensity values of the pixels are used instead of CIELAB color space values to calculate color distance ($d_c$) in [3].

3.3. Classification results

ICPR2012 – Table 1 shows experimental results on the ICPR2012 dataset. As Table 1 shows, our method obtains the best accuracy among other methods as reported in [18]. For the Cell Level classification, an accuracy of 79% is obtained when the model is trained by using all training images including positive and intermediate images and then evaluated on the test set as shown in second column of Table 1. For the Leave-One-Specimen-Out evaluation strategy, an interesting result is obtained for the intermediate level images (forth column) where 92% accuracy is obtained which is 12% higher than other methods. On average, we got an accuracy of 94% which is 6% higher than the-state-of-art based as shown in fifth column of Table 1.

---

Table 1

<table>
<thead>
<tr>
<th>SCs</th>
<th>SCs (no gradient)</th>
<th>Kastaniotis</th>
<th>Shen</th>
<th>DiCataldo</th>
<th>Kazanov</th>
<th>Faraki</th>
<th>Nosaka</th>
<th>Wiliem</th>
<th>Xiangfei</th>
<th>Stoklasa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test set</td>
<td>Positive</td>
<td>Intermediate</td>
<td>Weighted average</td>
<td>Leave-One-Specimen-Out</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>accuracy (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICPR2012</td>
<td>SCs</td>
<td>SCs (no gradient)</td>
<td>Kastaniotis</td>
<td>Shen</td>
<td>DiCataldo</td>
<td>Kazanov</td>
<td>Faraki</td>
<td>Nosaka</td>
<td>Wiliem</td>
<td>Xiangfei</td>
</tr>
<tr>
<td></td>
<td>79</td>
<td>76</td>
<td>75</td>
<td>74</td>
<td>72</td>
<td>71</td>
<td>69</td>
<td>67</td>
<td>64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>95</td>
<td>90</td>
<td>72</td>
<td>71</td>
<td>95</td>
<td>63</td>
<td>81</td>
<td>89</td>
<td>89</td>
<td></td>
</tr>
<tr>
<td></td>
<td>92</td>
<td>86</td>
<td>55</td>
<td>65</td>
<td>80</td>
<td>49</td>
<td>59</td>
<td>71</td>
<td>74</td>
<td></td>
</tr>
<tr>
<td></td>
<td>94</td>
<td>88</td>
<td>65</td>
<td>68</td>
<td>89</td>
<td>57</td>
<td>72</td>
<td>71</td>
<td>83</td>
<td></td>
</tr>
<tr>
<td></td>
<td>93</td>
<td>85</td>
<td>78</td>
<td>78</td>
<td>93</td>
<td>75</td>
<td>78</td>
<td>85</td>
<td>89</td>
<td></td>
</tr>
</tbody>
</table>

---

Fig. 3. The image level images of ICPR2012 (a) and ICIP2013 (b) for positive (top rows) and intermediate (second rows) intensity levels.
Fig. 4. Cell classification accuracy for ICPR2012 (left graphs) on the ‘test set’ and ICIP2013 (right graphs) by using (HSM) [29] evaluation protocol. The accuracies are changed when superpixel sizes are adopted (a, b) and when different superpixel extensions are applied (c, d).

Table 2
The confusion matrices for positive (a) and intermediate (b) images and image level (c) by using Leave-One-Specimen-Out method.

<table>
<thead>
<tr>
<th></th>
<th>Ce</th>
<th>CS</th>
<th>Cy</th>
<th>FS</th>
<th>H</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ce</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>CS</td>
<td>0.0</td>
<td>66.7</td>
<td>0.0</td>
<td>33.3</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Cy</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>FS</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>H</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
<td>0.0</td>
</tr>
<tr>
<td>N</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

(a) Positive Cells  (b) Intermediate Cells  (c) Image Level

The intermediate cell classification plays an important role in HEp-2 cell classification problem. Because the intensity values of intermediate cells are much lower than positive cells, they may affect the final evaluation if they are considered together for training the classifier. In fact, those methods which have separate models for these two categories, usually perform better than that of using both categories simultaneously. Although the number of training data in the Leave-One-Specimen-Out strategy is increased, improved accuracies were not observed because the intensity levels in the test set are ignored.

We also measured the effect of our modified superpixel extraction method which incorporates the gradient information (see Section 2.2). Table 1 shows the results when our method without gradient information named ‘SCS (no Gradient)’ is applied on the dataset. Although the accuracies are high for some scenarios, the average accuracy is lower than that of using the gradient information in our method.

The confusion matrices of Image Level, Intermediate and Positive images in ICPR2012 by using Leave-One-Specimen-Out method are also shown in Table 2. As can be seen in Table 2, two classes of Fine- and Coarse-speckled are hard to classify because of the similar patterns that they have as can be seen in the heat-map show in Fig. 2a.

ICP2013 – Tables 3 and 4 show experimental results on the ICIP2013 dataset. For Cell Level classification task, the (HSM) [29] and Leave-One-Specimen-Out method (see Section 2.2) are used.

Table 3
Accuracy on ICIP2013 dataset for cell level by using (HSM) [29] evaluation method.

<table>
<thead>
<tr>
<th>Cell level</th>
<th>HSM [5]</th>
<th>[9] SCS (no gradient)</th>
<th>SCS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>95.5</td>
<td>95.8</td>
<td>96.8</td>
</tr>
<tr>
<td>Intermediate</td>
<td>80.9</td>
<td>87.9</td>
<td>88.8</td>
</tr>
<tr>
<td>Weighted average</td>
<td>87.5</td>
<td>91.5</td>
<td>92.4</td>
</tr>
</tbody>
</table>

In the first method, as shown in Table 3, where all the reported methods have used (HSM) [29] evaluation method, we obtained 97.73% and 90.54% accuracies for positive and intermediate level images which are higher than the state-of-the-art results. In the Leave-One-Specimen-Out method, we achieved higher accuracies in comparison with other state-of-the-arts as shown in Table 4 in Cell Level columns. We have also studied the effect of using gradient information to our superpixel extraction method. As can be seen in the Table 3, the accuracy is lower without gradient information (SCS(no Gradient)) than when this information is used. Although the achieved average accuracy of 90.7% is higher than by (HSM) [29] method, it is 3% lower than the SCS proposed method.

For the Image Level classification task, when we compare our proposed SCS results with two other state-of-the-art methods including [30] and [31] by using Leave-One-Specimen-Out evaluation strategy, we achieved higher accuracies for the Image Level part.
as seen in Table 4. We also evaluated our proposed cell extraction method by comparing the results of proposed SCS with and without cell extraction stage; the latter is called ‘All Cells’ strategy. In ‘All Cells’ strategy, all connected pixels of the specimen images are extracted by using the provided masks without filtering them by properties of area and solidity as explained in Section 2.1. This gives 110,000 cells in total. Then the proposed SCS method (without cell extraction part) is performed. As can be seen from the last two columns of Image Level part in Table 4, the cell extraction method achieves significantly better results. Specifically, we obtained an average accuracy of 90.39%, which is 10% higher than that of using all cells of the specimen images. Additionally, the time complexity of the proposed method on extracted cells is an order of magnitude lower than using all the cells due to the lower amount of training data.

The confusion matrices for Cell and Image Levels by using (HSM)[29] and Leave-One-Specimen-Out methods are shown in Table 5. For the Cell Level classification, the Homogenous and Speckled classes are misclassified more than the others due to their similar patterns as evident from Table 5a and b. For the Image Level classification, where one new class (MitSp) is added to the dataset, the confusion matrix in Table 5c shows that the misclassification rate between MitSp, NuMem and Speckled are high. To achieve better results, more informative features are needed which is an interesting topic for further research.

The superior cell classification accuracy can be largely explained by the use of informative patches of cell images that are obtained by proposed superpixel method. Note that accuracies on the ICPR2012 and ICI2013 datasets are very different because the quality and amount of images within the two datasets are very different.

### 3.4. Superpixel parameters

In this sub-section, we present our study on the robustness of the proposed method to changes in superpixel size and extension using the test set of the ICPR2012 dataset and the evaluation set of ICI2013 (see Section 3.2). As evident from Fig. 4, the accuracy increases from very small superpixel sizes to a value which performs the best (100 pixels in one superpixel). In contrast, the accuracy decreases when larger superpixels are extracted. In other words, when the superpixel size is very small, it contains less information for classification. Therefore, the accuracy is not acceptable for very small size superpixels. On the other hand, the large superpixels contain many informative features but categorizing them by the dictionary learning process may increase the reconstruction error. Additionally, when the large superpixels are used, the number of superpixels will decrease.

The correlation of training and test accuracies are shown in Fig. 4a and b, when the superpixel size is increased. For example in ICIP2013 dataset, where we have enough representing data, the training accuracy is also drops by increasing the superpixel size. This study shows that the proposed technique prefers a larger number of small superpixels instead of a smaller number of larger superpixels.

In addition, applying the extension to the superpixel improves the cell classification accuracy clearly which can be observed when the superpixel boundaries extend from 0 pixel to 3 pixels (see Fig. 4c and d). It shows that using the original superpixels (with no extension) results in low accuracy as the informative features of the superpixels are the edges of the images which now overlap with the boundaries of superpixels and are thus omitted in the feature extraction process. Therefore, by extending the superpixel sizes, we bring these important information into the superpixels and provide better image patches for classification purposes. At the same time, the accuracy stabilizes when the extension lies around 6-14 pixels.

### 3.5. Timing

The proposed superpixel based technique is much faster than the traditional overlapped patched based methods for both training and testing tasks. As Table 6 shows, the dictionary learning is around 12 and 6 times faster for ICIP2013 and ICPR2012 dataset,
respectively. In addition, the testing time for feature sparse coding and SVM classification is around 2 times fast. This improved speed is mainly due to extracting lower number of patches (Superpixels) than overlapping patches. For instance, it is evident from Table 6, 10,725 and 137,595 seconds are spent for the SCS dictionary learning and the overlapping patches based method, respectively in the ICIP2013 dataset. Totally, the SCS is 12.83 times faster than the overlapping based method in this dataset. The lower number of extracted patches influences the testing time as well, which is more than 2 times faster than overlapping based method. These measurements are accomplished in a machine with Intel Core i7 CPU and 16GB of ram with 64-bit operating system.

4. Conclusion

This paper presents a superpixel based HEP–2 cell classification technique. Unlike traditional image patch based approaches, the proposed technique makes use of superpixels to select image patches in a more intelligent way. In addition, several adaptations of the superpixel such as the minimizing the gradient distance and extension idea have been carefully introduced for the optimal cell classification. Extensive experiments on two public datasets show superior HEP–2 cell classification performance. The executable code of proposed method on ICIP2013 dataset for both cell and image level classification can be found in Supplementary materials and our website.3

Supplementary material

Supplementary material associated with this article can be found, in the online version, at 10.1016/j.patrec.2016.02.007.

References


3 https://sites.google.com/site/shahabensafi/publications.