A Preliminary Analysis on HEp-2 Pattern Classification: Evaluating Strategies Based on Support Vector Machines and Subclass Discriminant Analysis

Ihtesham Ul Islam, Santa Di Cataldo, Andrea Bottino[(✉)], Enrico Macii, and Elisa Ficarra

Department of Control and Computer Engineering, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
{ihtesham.ulislam,santa.dicataldo,andrea.bottino, enrico.macii,elisa.ficarra}@polito.it

Abstract. The categorization of different staining patterns in HEp-2 cell slides by means of indirect immunofluorescence (IIF) is important for the differential diagnosis of autoimmune diseases. The clinical practice usually relies on the visual evaluation of the slides, which is time-consuming and subject to the specialist's experience. Thus, there is a growing demand for computer-aided solutions capable of automatically classifying HEp-2 staining patterns. In the attempt to identify the most suited strategy for this task, in this work we compare two approaches based on Support Vector Machines and Subclass Discriminant Analysis. These techniques classify the available samples, characterized through a limited set of optimal textural attributes that are identified with a feature selection scheme. Our experimental results show that both strategies have a good concordance with the diagnosis of the human specialist and show the better performance of the Subclass Discriminant Analysis (91 % accuracy) compared to Support Vector Machines (87 % accuracy).

Keywords: Indirect immunofluorescence imaging · HEp-2 staining pattern classification · Support vector machines · Subclass discriminant analysis · Pattern recognition

1 Introduction

Indirect immunofluorescence (IIF) is an imaging modality detecting abundance of those molecules that induce an immune response in the sample tissue. This technique uses the specificity of antibodies to their antigen in order to bind fluorescent dyes to specific biomolecule targets within a cell. The screening for anti-nuclear antibodies by IIF is a standard method in the current diagnostic approach to a number of important autoimmune pathologies such as systemic rheumatic diseases as well as multiple sclerosis and diabetes [1]. This screening,

which makes use of a fluorescence microscope, is typically done by visual inspection on cultured cells of the HEp-2 cell line: the specialist observes the IIF slide at the microscope (see Fig. 1 for an example), and makes a diagnosis based on the perceived intensity of the fluorescence signal and on the type of the staining pattern. Fluorescence intensity evaluation is needed for classifying between positive, intermediate and negative (*i.e.* absence of fluorescence) samples. Then, specific staining patterns on positive and intermediate samples reveal the presence of different antibodies and, thus, different types of autoimmune diseases. Therefore, a correct description of staining patterns is fundamental for the differential diagnosis of the pathologies. Examples of the six main staining patterns described by literature (homogeneous, fine speckled, coarse speckled, nucleolar, cytoplasmic or centromere) are reported in Fig. 2. They are distinguished as follows:

- **Homogeneous:** diffuse staining of the entire nucleus, with or without apparent masking of the nucleoli.
- **Nucleolar:** fluorescent staining of the nucleoli within the nucleus, sharply separated from the unstained nucleoplasm.
- **Coarse/Fine Speckled:** fluorescent aggregates throughout the nucleus which can be very fine to very coarse depending on the type of antibody present.
- **Centromere:** discrete uniform speckles throughout the nucleus, the number corresponds to a multiple of the normal chromosome number.
- **Cytoplasmic Fluorescence:** granular or fibrous fluorescence in the cytoplasm.

The manual classification of HEp-2 staining pattern suffers from usual problems in medical imaging, that is (i) the reliability of the results is subject to the specialist's experience and expertise, and (ii) the analysis of large volume of images is a tedious and time-consuming operation, translating into higher costs for the health system. Studies report very high inter- and intra-laboratory variability for this type of screening (up-to 10 %), that can be even higher in case of non-specialized structures [1]. This variability impacts on the reliability of the obtained results and, most of all, on their reproducibility.

Thus, in the last few years, reliable automatic systems for automating the whole IIF process have been in great demand and several tools have been proposed [2–6]. Nevertheless, the accurate classification of the staining patterns still remains a challenge.

Several classification schemes have been applied: among the others, learning vector quantization [3], decision tree induction algorithms [4,5], and multi-expert systems [6]. Unfortunately, direct comparison of the results presented by different works is not possible, since they were obtained on different datasets and on different classes. However, it is worth noting that textural features are generally acknowledged for being the most appropriate for staining pattern classification.

In this work, we compare two techniques we implemented that classify the cells into one of the six staining patterns addressed by literature. The first is based on Support Vector Machines (SVM). This approach was already introduced in our previous work [7], and it is described again here for the sake of

completeness. The second technique is a novel procedure based on Subclass Discriminant Analysis (SDA), a recent dimensionality reduction method that has been proven successful in different problems. SDA aims at improving the classification of a large number of different data distributions, whether they are composed by compact sets or not, by describing the underlying distribution of each class using a mixture of Gaussians. Since some of the staining patterns are characterized by a relevant within-class variance, SDA appears to be a promising method to improve their classification accuracy.

In our approach, each cell is initially characterized with a set of features based on statistical measurements of the grey-level distributions and on frequency-domain transformations. The dimension of this feature vector is then reduced applying different procedures, aiming at selecting the subsets of feature variables that are best suited to the classification with both SVM and SDA.

After a description of the dataset employed for training and testing our methods, Sect. 2, and a description of the two classification techniques, Sect. 3, this work presents in Sect. 4 the results of our experiments, aimed at identifying the best IIF classification technique. Discussion and conclusions are presented in Sects. 4.3 and 5, respectively.

2   Materials

The dataset used in our experiments contains IIF images that are publicly available at [8]. It is composed of 14 annotated IIF images acquired using slides of HEp-2 substrate at the fixed dilution of 1:80, as recommended by the guidelines in [9]. The images were acquired with a resolution of 1388×1038 pixels and a color depth of 24 bits. The acquisition unit consisted of a fluorescence microscope (40-fold magnification) coupled with a 50 W mercury vapour lamp and with a
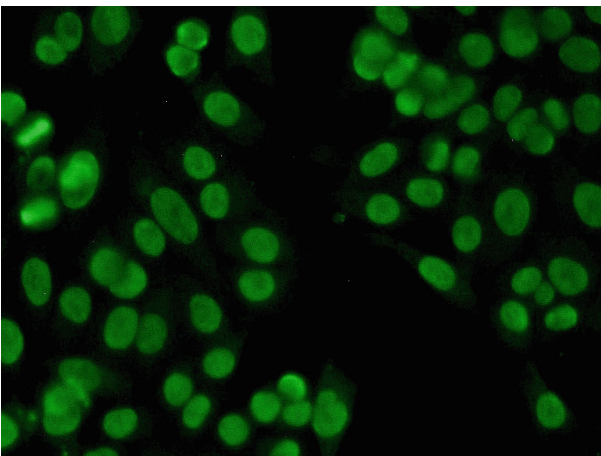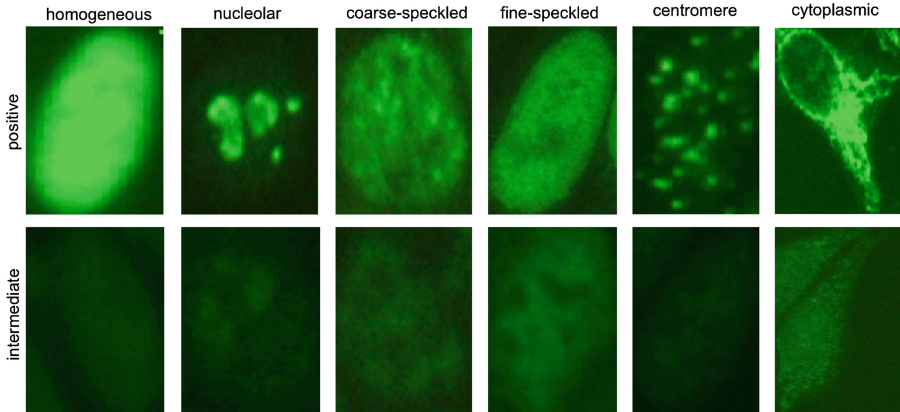

Fig. 1. HEp-2 IIF image.

Fig. 2. Examples of staining patterns that are considered relevant to diagnostic purposes, either with intermediate and positive fluorescence intensity.

Table 1. HEp-2 cell dataset characterization.

Pattern	N. of samples	Int.	Pos.
Homogeneous	150	47	103
Nucleolar	102	46	56
Coarse speckled	109	41	68
Fine speckled	94	48	46
Centromere	208	119	89
Cytoplasmic	58	24	34
Tot.	721	325	396

digital camera (SLIM system by Das srl) having a CCD with square pixel of 6.45 μm side. An example of the available images can be seen in Fig. 1.

From these images, a set of samples of HeP-2 cells have been extracted. Specialists manually segmented each cell at a workstation monitor, labelling it with the corresponding fluorescence intensity level (either intermediate or positive) and staining pattern. The latter can be distinguished in the six classes described in the introduction.

The dataset contains a total of 721 cells, 325 of which with intermediate and 396 with positive fluorescence intensity. A full characterization of dataset is reported in Table 1.

3   Methods

Our approach combines texture analysis and feature selection techniques in order to obtain a limited set of image features that is optimal for the classification task.

As already mentioned, for classification we implemented and compared two different methods, based on SVM and on SDA.

In the following subsections we provide details about all the steps of the proposed techniques.

3.1   Size and Contrast Normalization

Size and intensity normalisation of the samples is a necessary preprocessing step. In fact, small differences in the dimensions of the cell images are normal, and these differences are completely independent from their staining pattern. On the other hand, there are considerable variations of fluorescence intensity between intermediate and positive samples. Reducing such variability helps to decrease the noise in the classification process and avoids as well the necessity of training two separate classifiers for intermediate and positive samples.

Size normalization was obtained by re-sampling all the cell images to 64×64 pixels dimension. Contrast normalization consisted in linearly remapping the intensity values so that 1% of data is saturated at low and high intensities.

3.2   Feature Extraction

Textural analysis techniques have already been proven successful in HEp-2 cell staining characterization [7]. In fact, they are able to describe the most relevant image variations occurring in the cell allowing to differentiate between the staining patterns.

The two major approaches for textural analysis are either based on statistical methods describing the distribution of grey-levels in the image or on frequency-domain measurements of image variations. In our work we propose a combination of both of them in order to extract a comprehensive set of features able to fully characterize the staining pattern of the cell.

A first set of features was computed based on Gray-Level Co-occurrence Matrices (GLCM), a well established technique that extracts textural information from the spatial relationship between intensity values at specified offsets in the image. More specifically, textural features are computed from a set of grey-tone spacial dependence matrices reporting the distribution of co-occurring values between neighbouring pixels according to different angles and distances [10]. In practice, the $(i, j)_{d,\theta}$ element of a GLCM contains the probability for a pair of pixels located at a neighbourhood distance d and direction θ to have gray levels i and j, respectively.

In our work, we extracted 44 GLCM textural features reported in Table 2, based on four 16×16 GLCMs computed for a fixed unitarian neighbourhood distance and a varying angle $\theta = 0^o$, 45°, 90°, 135° (see [7] for details). The features are based on well-established statistical measurements whose characterization can be found in [10–12]. The use of 4 different directions is aimed at making the method less sensitive to rotations in the images.

Table 2. GLCM features.

N.	Feature	Ref.
1	Uniformity	[10]
2	Entropy	[11]
3	Dissimilarity	[11]
4	Contrast	[10]
5	Inverse difference	[12]
6	Correlation	[10]
7	Homogeneity	[11]
8	Autocorrelation	[11]
9	Cluster Shade	[11]
10	Cluster Prominence	[11]
11	Maximum probability	[11]
12	Sum of Squares	[10]
13	Sum Average	[10]
14	Sum Variance	[10]
15	Sum Entropy	[10]
16	Difference variance	[10]
17	Difference entropy	[10]
18	Information measures of correlation (1)	[10]
19	Information measures of correlation (2)	[10]
20	Maximal correlation coefficient	[10]
21	Inverse difference normalized (INN)	[12]
22	Inverse difference moment normalized (IDN)	[12]

Besides statistical methods, a largely used approach to extract relevant textural information for image compression and classification is based on frequency-domain transformations [13]. The underlying concept is the transformation of the image spatial information into a different space whose coordinate system has an interpretation that is closely related to the description of image texture.

In our work, we computed the two-dimensional Discrete Cosine Transform (DCT) [14] of the normalized images and then extracted 328 DCT coefficients (described in details in our previous work [7]) representing different patterns of image variation and directional information of the texture. The same approach was already successfully applied for texture classification and pattern recognition [13].

Combining GLCM and DCT sets, we obtained a total number of $44 + 328 = 372$ features to characterize each sample image.

3.3   Classification Based on Support Vector Machines

The first classification method we implemented was already introduced in our previous work [7]. It is based on Support Vector Machines (SVM), a well-established machine learning technique that has been proven successful for classification and regression purposes in many applications [15].

The classification is based on the implicit mapping of data to a high-dimensional space via a kernel function, and on the identification of the maximum-margin hyperplane that separates the given training instances in this space.

In our work we used SVM with radial basis kernel, optimizing the kernel parameters by means of ten-fold cross-validation technique and a grid search, as suggested in [15].

Feature Selection. Feature selection (FS) strategies were applied in order to select a limited set of optimal features able to improve the accuracy of the staining pattern classifier.

SVM are widely acknowledged for their built-in feature selection capability, as they implicitly map data in a transformed domain where the features that are crucial to the classification purpose are emphasized [16]. Nevertheless, the combination of SVM with feature selection strategies, besides improving training efficiency, can further enhance the accuracy of classification. In fact, although the presence of irrelevant features does not change the hyperplane margin of SVM, it may increase the radius of the training data points, impacting on SVM's generalization capability and also increasing the probability of over-fitting [17].

In our work we applied feature selection in two sequential steps. The first is based on minimum-Redundancy-Maximum-Relevance (mRMR) algorithm, whose better performance over the conventional top-ranking methods has been widely demonstrated in the literature [18]. The mRMR algorithm sorts the features that are most relevant for characterizing the classification variable, pointing at the contemporaneous minimization of their mutual similarity and maximization of their correlation with the classification variable. The number of the candidate features selected by mRMR was arbitrarily set to 50.

As for mRMR to work at its best the classification variables have to be categorical and not continuous, we preventively performed features discretization on the input data. For this purpose, we applied CAIM (class-attribute interdependence maximization) algorithm [19], which is best suited to work with supervised data, as it generates a minimal number of discrete intervals by maximizing the class-attribute interdependence.

The output of mRMR is a generic candidate feature set, which is independent from the classification algorithm [18] and not necessarily optimal for SVM. Therefore, we applied as second FS step a Sequential Forward Selection (SFS) scheme in order to iteratively construct the subset of optimal features that is best suited for SVM classification.

Classical SFS [20] works towards the minimization of the misclassification error: starting from an initial empty set, at each iteration the feature providing the greatest classification accuracy improvement is added, until no more
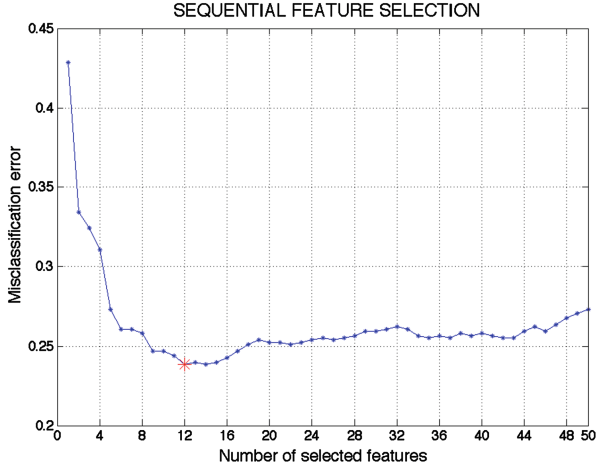
Fig. 3. Sequential Feature Selection strategy: misclassification error vs. number of selected features at each iteration. The optimal feature set size is 12.

improvement is obtained. As this implementation tends to be trapped in local minima, in our approach we proceeded with the iterations until all the available features were added, and then we selected the feature set with the best classification accuracy. The final dimension of this optimal set was found to be 12 (see Fig. 3).

3.4   Classification Based on Subclass Discriminant Analysis

SDA belongs to family of Discriminant analysis (DA) algorithms, which have been used for dimensionality reduction and feature extraction in many applications of computer vision and pattern recognition [21–24]. These algorithms map a set of samples $\boldsymbol{X} = (x_1, x_2, \ldots, x_n)$, associated to a class label $\in [1, C]$ and belonging to a high-dimensional feature space $\in \Re^D$, to a low-dimensional subspace $\in \Re^d$, with $d \ll D$, where the data can be more easily separated according to their class-labels. Therefore, DA problem can be generally stated as finding the transformation matrix $\boldsymbol{V} = (v_1, v_2, \ldots, v_d)$, with $v_i \in \Re^D$, for mapping a sample x into the final d-dimensional subspace.

In most DA algorithms, the transformation matrix V is found by maximizing the so-called Fisher-Rao's criterion:

$$J(V) = \frac{\left|V^T A V\right|}{\left|V^T B V\right|} \tag{1}$$

where A and B are symmetric and positive-defined matrices. The solution to this problem is given by the generalized eigenvalue decomposition:

$$AV = BV\Lambda \tag{2}$$

Where V is (as above) the desired transformation matrix, and Λ is a diagonal matrix of its corresponding eigenvalues.

Linear Discriminant Analysis (LDA) is probably the most well-known DA technique. This method assumes that the C classes the data belong to are homoscedastic, that is their underlying distributions are Gaussian with common variance and different means. In (1), LDA uses $A = S_B$, the between-class matrix, and $B = S_W$, the within-class scatter matrix, defined as:

$$S_B = \sum_{i=1}^{C} (\mu_i - \mu)(\mu_i - \mu)^T \tag{3}$$

$$S_W = \frac{1}{n} \sum_{i=1}^{C} \sum_{j=1}^{n_i} (x_{ij} - \mu_i)(x_{ij} - \mu_i)^T \tag{4}$$

where C is the number of classes, μ_i is the sample mean for class i, μ is the global mean, x_{ij} is the jth sample of class i and n_i the number of samples in class i.

LDA provides the (C-1)-dimensional subspace that maximizes the between-class variance and minimizes the within-class variance in any particular data set. In other words, it guarantees maximal class separability and, possibly, optimizes the accuracy in later classification.

However, the assumption of having C homoscedastic classes is the very limitation of this method. LDA works well for linear problems and fails to provide optimal subspaces for inherently non-linear structures in data. Several extensions of LDA have been introduced in literature to effectively classify data with non-linearities [25].

To this end, one of the most effective approaches is the Subclass Discriminant Analysis (SDA), proposed in [26]. The main idea of SDA it is to find a way to describe a large number of different data distributions, whether they are composed by compact sets or not, by describing the underlying distribution of each class using a mixture of Gaussians. This is achieved by dividing the classes into subclasses. Therefore, the problem to be solved is to find the optimal number of subclasses maximizing the classification accuracy in the reduced space. In SDA, the transformation matrix V is found by defining the between-subclass scatter matrix S_B in Eq. (1) as:

$$S_B = \sum_{i=1}^{C-1} \sum_{j=1}^{H_i} \sum_{k=i+1}^{C} \sum_{l=1}^{H_k} p_{ij} p_{kl} (\mu_{ij} - \mu_{kl})(\mu_{ij} - \mu_{kl})^T \tag{5}$$

where H_i is the number of subclasses of class i, μ_{ij} and p_{ij} are the mean and prior probability of the jth subclass of class i, respectively. The priors are estimated as $p_{ij} = n_{ij}/n$, where n_{ij} is the number of samples in the jth subclass of class i. In the simplest case of SDA with no class subdivisions, this equation reduces to that of LDA.

In order to select the optimal number of subclasses, in [26], the authors propose two different methods. The first is based on a stability criterion described

in [27]. However, as pointed out in [28], when data have a Gaussian homoscedastic subclass structure, the minimization of the metric used in this criterion is not guaranteed. Authors in [28] hypothesize that this is likely to happen also for heteroscedastic classes.

The second selection criterion is based on a Leave-one-object test. In practice, for each subdivision, a leave-one-out cross validation (LOOCV) is applied, and the optimal subdivision is the one giving the maximal recognition rate. The problem with this strategy is that it has very high computational costs, especially when the dataset to classify is large and the number of classes is high. This is exactly what is happening in our case, where the initial classes are 6 and the samples are 721.

Therefore, to overcome these problems, we used a different formulation of the optimality criterion, similar to the leave-one-object test, but based on a stratified 5-fold cross validation, which optimizes the accuracies obtained with a k-Nearest Neighbour (kNN) classifier. A value of 8 for k has been heuristically found to provide good classification results.

Our implementation differs from the original SDA formulation for two other details. The first concerns the clustering methods used to divide classes into subclasses. In [26] data are assigned to subclasses by first sorting the class samples with a Nearest-Neighbour based algorithm and then dividing the obtained list into a set of clusters of the same size. However, this method does not allow to model efficiently the non-linearity present in the data, as in the case of staining patterns under analysis. Therefore, we used the K-means algorithm, which partitions the samples into k clusters by minimizing iteratively the sum, over all clusters, of the within-cluster sums of sample-to-cluster-centroid distances. Since, in this method, the centroids are initially set at random, different initialization results in different divisions. Hence, we repeated the clustering 20 times and kept the solution providing the minimal sum of all within-cluster distances.

The second difference is that, instead of increasing the number of subclasses for each class of the same amount at each iteration, all the possible permutation of class subdivisions are created by iteratively incrementing by one the number of subclasses of a single class in a set of nested loops. For a specific class r, the subdivision process is stopped when the minimal number of samples in the H_r clusters obtained with K-means drops below a predefined threshold. In order to reduce the computational times, the clusters created in inner loops are computed only once and cached for further use.

The classification accuracy of our method is computed as the average accuracy of the different CV rounds. It should be underlined that, given the differences between training and test sets, different optimal subclasses subdivision are likely to be obtained at each CV iteration.

Feature Selection. As well as for the SVM classifier, we applied FS strategies to SDA too. In this case, we used only the reduced feature set obtained with mRMR. This has been done for two reasons.

First, while mRMR is independent from the classification method, SFS relies on the classifier output, which makes it unfeasible with the computational cost of SDA.

Second, it can be easily shown that the rank of matrix S_B, and therefore of the dimensionality d of the reduced subspace obtained from Eq. (2), is given by $min(H-1, rank(S_X))$, where H is the total number of subclasses and $rank(S_X)$ is equal (or minor) to the number of features characterizing each sample. While the number of features selected with mRMR (50) is a reasonable upper bound for d, reducing it further might hamper the possibility to obtain a good classification accuracy in problems, like the one tackled in this paper, in which the data present high non-linearities.

4   Results

The two classification methods presented in Sect. 3 were tested on the same annotated IIF images, using the staining pattern information provided by the specialists as ground truth for cross-validation.

4.1   SVM Classification

We recall here the experimental results on SVM classification, already reported in our previous paper [7], for comparison with SDA approach.

As for SVM classification, experiments were run on the following datasets:

dataset I, the initial 372 elements feature set;
dataset II, the 50 elements candidate set selected by mRMR feature selection;
dataset III, the final 12 elements feature vector obtained with combination of mRMR + SFS.

The accuracy results of the 10-fold cross-validation, grouped by staining pattern, are reported in Table 3. The last row of the table shows the overall accuracy obtained by SVM in each dataset.

The following initial remarks can be drawn from the analysis of the results of Table 3:

– SVM classifier obtained an average accuracy of 86.96 % in the six staining patterns. The maximum and minimum per-class accuracy were 98.17 % (coarse speckled pattern) and 71.28 % (fine speckled pattern).
– FS significantly improved the classification performances (+9.01 % on the overall average accuracy). This confirms the considerations drawn in Sect. 3.3 about the weakness of the implicit feature selection ability claimed by SVM. As it can be seen, mRMR improved the per-class accuracy of all the staining patterns (see results on dataset II compared to those on dataset I). The combination of mRMR+SFS (dataset III) further improved the average accuracy of SVM. While per-class accuracies of centromere and cytoplasmic patterns were slightly decreased (of, respectively, -1.44 % and -3.45 % w.r.t. dataset II), the fine speckle pattern, that had lowest per-class accuracy, is the class that obtained the best improvement (+9.58 % w.r.t. dataset II and +25.53 % w.r.t. dataset I). This non-uniform behaviour is not surprising, since SFS optimized the average classification accuracy in the overall dataset and not the accuracies of the single classes.

Table 3. SVM Classification results: accuracy rate (%).

Pattern	Dataset		
	I	II	III
Homogeneous	78.66	84.00	86.00
Nucleolar	89.22	93.14	93.14
Coarse speckled	92.66	95.41	98.17
Fine speckled	45.75	61.70	71.28
Centromere	84.13	88.46	87.02
Cytoplasmic	58.62	86.21	82.76
Overall	77.95	85.58	86.96

4.2   SDA Classification

Table 4, which is, again, organized by staining class, summarizes the classification results obtained with SDA. As already explained in Sect. 3.4, SFS strategy was not applied in combination with SDA classifier. Therefore, the table contains only results on dataset I (the initial 372 feature set) and dataset II (50 feature set obtained after mRMR).

LDA results (which are those obtained with SDA with no class subdivisions) are also provided for comparison, in order to demonstrate the effective capabilities of SDA to better classify datasets with high non-linearities. Finally, in the last row we show the overall accuracies obtained in the four cases.

Analysing the results, some considerations can be drawn:

- as expected, the overall accuracy of SDA outperforms that of LDA (+7.29 % on dataset I and +7.13 % on dataset II). Concerning the per-class results, better results are obtained in most of the cases (except for centromere class for dataset I, and coarse speckled class for dataset II), with, as best improvements, a +17.34 % in dataset I for homogeneous class and + 23.51 % in dataset II for fine speckled class;
- as in the SVM experiments, FS effectively improves the SDA accuracies of all classes (the best improvement being the +9.71 % of fine speckled). The overall improvement is +4.75 %
- the best average accuracy obtained is 90.79 %, with dataset II, which outperforms the best accuracy obtained by SVM with mRMR+SFS feature selection (86.96 %, dataset III). The best per-class improvements have been obtained for fine speckled (+18.14 %) and cytoplasmic class (+17.24 %), while coarse speckled and centromere class obtained slightly lower accuracies (respectively, −6.48 % and −4.85 %).

4.3   Discussion

The results presented in Tables 3 and 4 suggest that the proposed algorithm is a good solution for the automated classification of immunofluorescence cell patterns.

Table 4. SDA Classification results: accuracy rate (%).

Pattern	Dataset			
	I		II	
	LDA	**SDA**	LDA	**SDA**
Homogeneous	63.33	80.67	80.00	85.33
Nucleolar	90.29	94.29	91.19	96.14
Coarse speckled	87.19	88.05	92.55	91.69
Fine speckled	62.87	79.71	65.91	89.42
Centromere	75.96	73.53	80.78	82.17
Cytoplasmic	92.88	100.00	91.52	100.00
Overall	78.75	86.04	83.66	90.79

As a matter of facts, the accuracy rate is comparable to the one obtained by the specialists, whose inter-laboratory variability is generally assessed around 10 % or even higher [1]. Besides that, differently from human operators, our technique provides fully-repeatable results that are based on objective and quantitative features of the images.

As for the classification techniques, the same results show that SDA technique, in combination with a proper selection of the most relevant features, outperforms the best accuracy achievable with SVM on the same dataset (II) and even with those obtained by SVM on dataset III, specifically optimized for that technique with a two-step FS process. Therefore, our experiments show the capabilities of SDA to describe in a more suitable way the underlying distributions of each of the staining pattern class, improving their classification accuracies.

5   Conclusions

In this paper we proposed the comparison of two approaches, based on SVM and SDA, for the automatic classification of staining patterns in HEp-2 cell IIF images. Texture descriptors based on GLCM and DCT coefficients are first exploited to extract a 372-size characteristic vector for each cell. Then, a feature selection algorithm is applied to obtain a reduced candidate feature set that improves the classification accuracies of the two methods.

Feature selection is based on the mRMR algorithm, which sorts the features that are most relevant for characterizing the classification variable. The 50 top-ranked features were selected. In the case of SVM-based method, a two-steps feature selection procedure, coupling mRMR with SFS algorithm, is implemented in order to further improve classification accuracies.

The two approaches provide average classification accuracies of about 87 % and 91 %, respectively. These results are comparable with those of human specialists. Conversely, they are completely repeatable since our automated technique

does not depend on the subjectivity of the operator. Moreover, our experiments show the effectiveness of SDA into describing more precisely, compared to SVM, the underlying distributions of each of the staining pattern class.

As future steps, we plan to work on:

(1) a better characterization of cell patterns, which can be insensitive to changes in size, rotation and intensity;
(2) an improvement of the SDA classifier in terms of computational efficiency. For this purpose, methods selecting a priori the classes that effectively needs to be partitioned, like the one described in [29], will be investigated;

Moreover, we plan to develop a pipeline for automatic cells segmentation in IIF images and to combine it with our pattern classification algorithm in order to obtain a complete automated approach for the computer-aided diagnosis (CAD) of autoimmune diseases.

References

1. Egerer, K., Roggenbuck, D., Hiemann, R., Weyer, M.G., Buttner, T., Radau, B., Krause, R., Lehmann, B., Feist, E., Burmester, G.R.: Automated evaluation of autoantibodies on human epithelial-2 cells as an approach to standardize cell-based immunofluorescence tests. Arthritis Res. Ther. **12**(2), 1–9 (2010)
2. Creemers, C., Guerti, K., Geerts, S., Van Cotthem, K., Ledda, A., Spruyt, V.: HEp-2 cell pattern segmentation for the support of autoimmune disease diagnosis. In: Proceedings of ISABEL 2011, vol. 28, pp. 1–5 (2011)
3. Hsieh, R.Y., Huang, Y.C., Chung, C.W., Huang, Y.L.: HEp-2 cell classification in indirect immunofuorescence images. In: Proceedings of ICICS 2009, vol. 26, pp. 211–214 (2009)
4. Perner, P., Perner, H., Muller, B.: Mining knowledge for HEp-2 cell image classification. Artif. Intell. Med. **26**, 161–173 (2002)
5. Sack, U., Knoechner, S., Warschkau, H., Pigla, U., Emmrich, F., Kamprad, M.: Computer-assisted classification of HEp-2 immunofluorescence patterns in autoimmune diagnostics. Autoimmun. Rev. **2**(5), 298–304 (2003)
6. Soda, P., Iannello, G.: A hybrid multi-expert systems for HEp-2 staining pattern classification. In: Proceedings of ICIAP 2007, pp. 685–690 (2007)
7. Di Cataldo, S., Bottino, A., Ficarra, E., Macii, E.: Applying textural features to the classification of HEp-2 cell patterns in IIF images. In: 21st International Conference on Pattern Recognition (ICPR 2012), Tsukuba, Japan, 11–15 November (2012)
8. MIVIA Lab. http://nerone.diiie.unisa.it/zope/mivia/databases/db_database/biomedical/. Accessed: September 2012
9. Tozzoli, R., Bizzaro, N., Tonutti, E., Villalta, D., Bassetti, D., Manoni, F., Piazza, A., Pradella, M., Rizzotti, P.: Guidelines for the laboratory use of autoantibody tests in the diagnosis and monitoring of autoimmune rheumatic diseases. Am. J. Clin. Pathol. **117**(2), 316–324 (2002)
10. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Trans. Syst. Man Cybern. **3**(6), 610–621 (1973)
11. Soh, L., Tsatsoulis, C.: Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Trans. Geosci. Remote Sens. **37**(2), 780–795 (1999)

12. Clausi, D.A.: An analysis of co-occurrence texture statistics as a function of grey level quantization. Can. J. Remote Sens. **28**(1), 45–62 (2002)
13. Sorwar, G., Abraham, A., Dooley, L.S.: Texture classification based on DCT and soft computing. In: Proceedings of FUZZ-IEEE 01, 2–5 December 2001
14. Ahmed, N., Natarajan, T., Rao, K.R.: Discrete cosine transform. IEEE Trans. Comput. **23**, 90–93 (1974)
15. Chang, C.-C., Lin, C.-J.: Libsvm: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**(3), 27:1–27:27 (2011)
16. Temko, A., Camprubi, C.N.: Classification of acoustic events using SVM-based clustering schemes. Pattern Recognit. **39**(4), 682–694 (2006)
17. Weston, J., Mukherjee, S., Chapelle, O., Pontil, M., Poggio, T., Vapnik, V.: Feature selection for SVMs. Adv. Neural Inf. Process. Syst. **13**, 668–674 (2000)
18. Peng, H., Long, F., Ding, C.: Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy. IEEE Trans. PAMI **27**, 1226–1238 (2005)
19. Kurgan, L.A., Cios, K.J.: Caim discretization algorithm. IEEE Trans. Knowl. Data Eng. **16**(2), 145–153 (2004)
20. Ververidis, D., Kotropoulos, C.: Fast and accurate feature subset selection applied into speech emotion recognition. Els. Sig. Process. **88**(12), 2956–2970 (2008)
21. Fukunaga, K.: Introduction to Statistical Pattern Recognition, 2nd edn. Academic Press, San Diego (1990)
22. Swets, D.L., Weng, J.J.: Using discriminant eigenfeatures for image retrieval. IEEE Trans. Pattern Anal. Mach. Intell. **18**(8), 831–836 (1996)
23. Etemad, K., Chellapa, R.: Discriminant analysis for recognition of human face images. J. Optical Soc. Am. A **14**(8), 1724–1733 (1997)
24. Belhumeur, P.N., Hespanha, J.P., Kriegman, D.J.: Eigenfaces vs. Fisherfaces: recognition using class specific linear projection. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 711–720 (1997)
25. Boulgouris, N.V., Plataniotis, K.N., Micheli-Tzanakou, E.: Discriminant analysis for dimensionality reduction: an overview of recent developments. In: Biometrics: Theory, Methods, and Applications. Wiley, New York (2009)
26. Zhu, M., Martinez, A.M.: Subclass discriminant analysis. IEEE Trans. PAMI **28**(8), 1274–1286 (2006)
27. Martinez, A.M., Zhu, M.: Where are linear feature extraction methods applicable? IEEE Trans. Pattern Anal. Mach. Intell. **27**(12), 1934–1944 (2005)
28. Gkalelis, N., Mezaris, V., Kompatsiaris, I.: Mixture subclass discriminant analysis. IEEE Sig. Process. Lett. **18**(5), 319–322 (2011)
29. Kim, S.-W.: A pre-clustering technique for optimizing subclass discriminant analysis. Pattern Recogn. Lett. **31**(6), 462–468 (2010)