Computationally-efficient classification of HEp-2 cell patterns in IIF images

Luis Fernando Planella Gonzalez
Pontifical Catholic University of Rio Grande do Sul, Brazil
luisfpg@gmail.com

Duncan Dubugras Alcoba Ruiz
Pontifical Catholic University of Rio Grande do Sul, Brazil
duncan.ruiz@pucrs.br

Márcio Sarroglia Pinho
Pontifical Catholic University of Rio Grande do Sul, Brazil
marcio.pinho@pucrs.br

ABSTRACT

Classification of HEp-2 cell images is seeing an increasing interest in the last years. Research on this subject is being further stimulated by contests in important pattern recognition conferences (ICPR-2012, ICIP-2013 and ICPR-2014). Several feature extraction methods for those images were proposed by researchers. In some cases, authors mine data with thousands of dimensions, originated by morphology, texture, pixel statistics and other features. Most of the published works focus solely on classification accuracy, suppressing the discussion about computational resources needed to extract features, train the classifier and test the instances. In this work, our research question is whether a classification accuracy comparable to other works can be achieved when using only features which are simple and fast to compute. We have demonstrated that, at least in particular scenarios, using only pixel value statistics and histogram can result in a classification accuracy which is comparable to other works.

Categories and Subject Descriptors

H.2.8 [**Database Applications**]: Image databases; I.5.1 [**Pattern Recognition**]: Models, Statistical

General Terms

Algorithms, Performance, Experimentation

Keywords

HEp-2 cells, Immunofluorescence pattern, Classification, Feature extraction, Image processing

1. INTRODUCTION

The human body produces antibodies to fight infections and external agents which cause diseases. However, in some cases, antibodies attack healthy tissue, causing autoimmune diseases, such as systemic lupus erythematosus, systemic sclerosis, mixed connective tissue disease, polymyositis, dermatomyositis, Sjögren's syndrome, Felty's syndrome and others [19, 26]. Antinuclear autoantibodies (ANA) are useful markers a physician can use to detect such diseases [26], and the preferred method to detect ANA is the indirect immunofluorescence technique (IIF) on cell line HEp-2 (from the laryngeal carcinoma substrate) [19]. The intensity level for IIF can be either negative, intermediate or positive. Observing HEp-2 cells reveals that patients with such diseases shows frequencies of ANA between 50 and 100% [26].

The standard way for identifying ANA needs at least two experts opinions, and can be affected by subjective factors, such as misinterpretation and fatigue. The task is labor-intensive and time consuming [19, 26], being a good candidate for Computer-Aided Diagnosis (CAD) systems.

Classification is one of the most commonly used data mining tasks, and is suitable for classifying IIF cell images. As classical data mining algorithms requires tabular data, comprised with records (also called rows, instances or objects) and attributes (also called columns, features or dimensions), mining images normally involves preprocessing the input images, transforming them in a dataset in such format [11].

In the context of image processing, attributes are generally called features. There are several approaches for feature extraction, but there is no standard categorization. Some examples include:

- Morphology: number of objects, number of holes, area, perimeter, roundness and convex hull;
- Texture: Grey Level Coocurrence Matrix (GLCM) [12], Local Binary Patterns (LBP) [20], Grey Level Run-Length Matrices (GLRM) [23], Gray Level Size Zone Matrix (GLSZM) [25];
- Statistics: mean, standard deviation, skewness, kurtosis and histograms;
- Some others: Discrete Cosine Transform (DCT) [2], Discrete Wavelet Transform (DWT) [18], 2D Gabor Wavelets [15], Discrete Fourier Transform (DFT), Scale Invariant Feature Transform (SIFT) [17], Speeded-Up Robust Features (SURF) [4] and Histograms of Oriented Gradients (HoG) [6]

When too many features are employed it is also expected that the demand of computational resources is increased. This applies for processing each image, training the classifier and classifying the instances. Also, the larger is the data dimensionality, the more sparse is the training data in

SAC'15 April 13-17, 2015, Salamanca, Spain.

http://dx.doi.org/10.1145/2695664.2695730

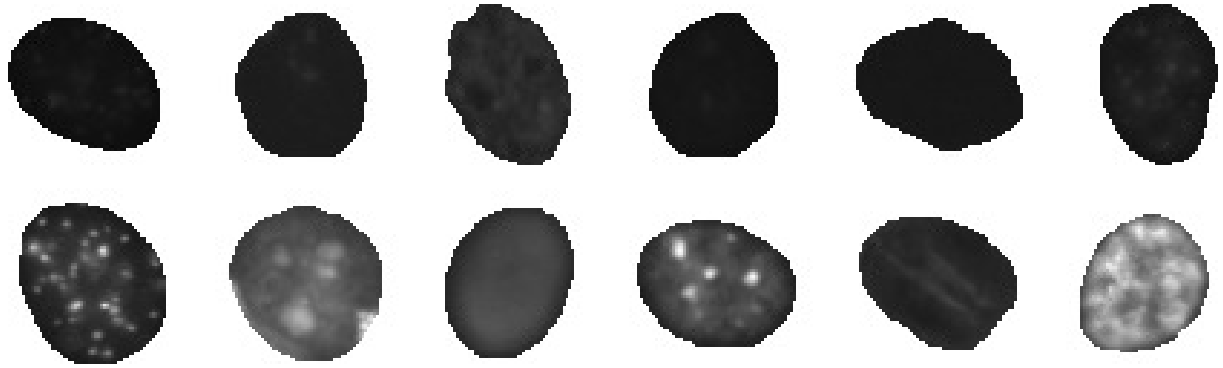

Figure 1: Zoomed samples of images in the dataset, with both intermediate (top) and positive (bottom) intensities for each class: Ctr, Glg, Hgn, Ncl, Nmn and Spk, respectively.

the dimensional space, which can lead to the classifier performance degradation due to overfitting. This is known as "curse-of-dimensionality" [14]. As such, this paper focuses on experimenting classification of HEp-2 IIF cell images using features which are simple and fast to compute. The results are contrasted to other published works, in order to determine whether a comparable accuracy can be achieved, at least in particular scenarios.

2. IIF IMAGES DATASET

This work uses the IIF images dataset from the ICIP-2013 contest[1]. This dataset is comprised of 13,596 gray scale images, being 7,448 and 6,148 of intermediate and positive intensities, respectively. No negative intensity images are provided. For each image there is also a corresponding mask image, describing the cell boundary. The dataset contains images with the following six classes: Centromere (Ctr), Golgi (Glg), Homogeneous (Hgn), Nucleolar (Ncl), Nuclear Membrane (Nmn) and Speckled (Spk). Figure 1 shows zoomed samples of each class in each intensity, and Table 1 shows the number of instances per intensity and class in the dataset.

Table 1: Number of instances per intensity and class.

Intensity	Ctr	Glg	Hng	Ncl	Nmn	Spk
Intermed.	1,363	375	1,407	1,664	1,265	1,374
Positive	1,378	349	1,087	934	943	1,457
Total	2,741	724	2,494	2,598	2,208	2,831

3. RELATED WORK

Despite being a relatively unexplored area [1], classification of HEp-2 IIF cell images has been performed by some published researches. Each work extracts a set of image features, and apply one or more data mining algorithm. The main focus of most of those papers is the classification accuracy, without mentioning computational resources needed to run the tool. Table 2 presents some of those papers, sorted by publication year in descending order. The table columns are the publication reference and year, the number of dimensions (Dim) and approximated claimed accuracy (Acc). In this table, the following abbreviations are used for classifiers: Random Forest (RF), Support Vector Machine (SVM), Regression Tree (RT), multiclass SVM (mSVM), k-Nearest Neighbors (kNN), Neural Networks (NN), Naïve Bayes (NB), Nearest Convex Hull (NCH), Logistic Regression (LR), AdaBoost (AB), Learning Vector Quantization (LVQ), Mahalanobis Distance Classifier (MDC) and Parzen Window Classifier (PWC).

The Table 2 shows that only one of the reviewed works [8] employed less than a dozen attributes. Others employed dozens [21, 24, 13], hundreds [1, 9, 7, 16, 22] or even thousands [10] of attributes. Four of the reviewed works [27, 25, 3, 5] didn't make a clear statement about the actual number of dimensions used, only describing the employed features. In terms of classification accuracy, most of the reviewed works claim to achieve 90% or more [1, 27, 21, 9, 25, 10, 16, 22, 24, 3, 8]. The highest claimed accuracy was achieved by a work which didn't specify the dimensionality employed [25]. Another data from this table is that two works which achieved almost the same accuracy, 90 and 91% employed a 9 and 2,342 attributes, respectively, indicating that accuracy and number of features are not proportional.

4. METHOD

As stated in Section 3, most published works for IIF cell classification focus on extracting a feature set from images and mining the resulting data with classification algorithms. Most of them have used more than one feature category (for example, morphological and texture) together, in order to increase accuracy, getting even to the point of thousands of features [10].

However, a question that remains unanswered is whether an acceptable classification accuracy can be achieved when using simple and computationally-efficient feature sets. To test this hypothesis, this work compares the time needed to compute distinct feature sets, in order to select those which can be computed faster. Afterwards, some well-known classifiers are applied, and the classification results are analyzed.

4.1 Feature selection

Intuitively, the features which are faster to extract from images are those which perform simple operations over raw pixel values. Some of those operations are pixel-value his-

[1] `http://nerone.diiie.unisa.it/contest-icip-2013/index.shtml`

Table 2: Some recent publications on HEp-2 IIF image classification.

Publication	Features	Classifier(s)	Dim.	Acc.
Agrawal *et al*, 2013 [1]	Morphological, statistical, LBP, HoG, GLCM, GLRM, Laws	k-NN, NB, RF, SVM	489	93%
Wiliem *et al*, 2013 [27]	Dual-region codebook, using both SIFT and DCT	NCH	?	96%
Snell *et al*, 2012 [21]	Roundness, DCT and statistical	mSVM	40	98%
Ersoy *et al* , 2012 [9]	Histogram, HoG, Hessian matrix and LBP	RT	130	93%
Di Cataldo *et al*, 2012 [7]	GLCM and DCT	SVM	372	87%
Thibault & Angulo, 2012 [25]	GLSZM and Pattern spectrum (PS)	LR, RF, NN	?	99%
Ghosh & Chaudhary, 2012 [10]	SURF, morphological features, GLCM, HoG	mSVM	2,342	91%
Li *et al*, 2012 [16]	LBP, DCT, 2D Gabor Wavelets, morphological and statistical	mSVM	180	98%
Strandmark *et al*, 2012 [22]	Morphological features and GLCM	RF	966	97%
Theodorakopoulos *et al*, 2012 [24]	Morphological and LBP	SVM	86	96%
Ali *et al*, 2012 [3]	Contrast differences, based on difference of Gaussians	K-NN	?	96%
Cordelli & Soda, 2011 [5]	Statistical, LBP and Wavelets	NN, kNN, SVM, AB	?	79%
Hsieh *et al*, 2009 [13]	Statistical, BVLC, SGLDM, GLDM, NGTDM, fractals	LVQ	51	80%
Elbischger *et al*, 2009 [8]	Statistical and morphological	MDC, PWC	9	90%

tograms and statistical central moments. In order to select which features we would use for classification, we have measured the time taken to extract the following features from the ICIP-2013 IIF image dataset: 2 central moments statistics (mean and standard deviation), 4 central moments statistics (mean, standard deviation, skewness and kurtosis) and a normalized 16-bin intensity histogram. For comparison, we have included a feature set for texture and a feature set for analysis in frequency domain: GLCM and wavelet coefficients, respectively.

Table 3 presents the time taken to calculate each feature for all 13,596 images in the training dataset (as described in Section 2). The presented times were calculated by averaging 10 executions for each feature. The time spent on loading images and reading pixel values is not included, as they are constant for any feature extraction method. All tests were executed as a Java application using OpenJDK [2] version 7u65, running on a 64 bit Linux system on an Intel core i7-4500U (4x1,8Ghz) laptop with 8G RAM.

Table 3: Time taken to extract distinct features from ICIP-2013 IIF contest cell images.

Features	Time taken
Histogram (16 bins)	0.9s
Statistics (2 central moments)	1.1s
Statistics (4 central moments)	3.4s
GLCM (size 32, 1px diff., 4 angles)	10.2s
Wavelet (size 64, 4 level DWT)	20.7s
GLCM (size 32, 1-3px diff., 4 angles)	64.2s

Confirming our initial expectations, histogram and statistics are significantly faster to calculate, when comparing to other methods, like GLCM and wavelet coefficients. As such, those were the features we have used in our tests, as reported in the next section.

4.2 Classifier selection

In order to test our hypothesis, that classification results can still be comparable to other related works when using simple and fast features, we need to test those features with at least one classifier. As most of the works cited in Table 2 have used either Random Forest and/or Support Vector Machine algorithms in their works [1, 21, 7, 25, 10, 16, 22, 24, 5], those are the algorithms we have also employed in our test. The classifiers were executed using the fastest features according to Section 4.1, extracted from images in the ICIP-2013 IIF contest dataset (described in Section 2).

As stated by Agrawal *et al* [1], classification of IIF images with intermediate and positive intensity values should be measured separately, as classifying intermediate intensity cells is harder. Figure 1 makes it visually clear that intermediate intensity images have less contrast and details, when compared to positive intensity images. With less visual discrimination between classes, it is expected to have worst classification results.

The results of our tests are presented in Table 4. It shows, for each combination of feature, intensity level and algorithm, the classification accuracy (first row), the time taken to build the model with all images in the dataset (second row) and the time taken to test the same images in the training set (third row). The features are shown as 2S (2 moments central statistics), 4S (4 moments central statistics), H (16-bin normalized histogram), 2S+H (2 moments central statistics plus 16-bin normalized histogram) and 4S+H (4 moments central statistics plus 16-bin normalized histogram). The intensity levels are presented as I (intermediate), P (positive) and C (combined). The reported accuracies were calculated using 10-fold cross validation method. For this task, we have used the implementations provided by the Weka 3.6.11[3] library, keeping their default settings.

The time taken to train the model and test the training set were measured because our objective is to perform resource-efficient classification, with classification accuracy comparable to other works. It should be noted that, as the dataset contains more intermediate than positive intensity images, both times (to train and test the dataset) are expected to be lower in positive, which has 6,148 samples, than in intermediate images, which has 7,448 samples. Consequently, the times for combined features should be larger, as the dataset has 13,596 images in total.

As shown in Table 4 the classifier which performed better in all aspects was Random Forest. It presented a higher classification accuracy and reduced times, both to build the classifier and to test the dataset, when compared to SVM. Hence, we have concluded that the Random Forest classifier is well suited for our work.

[2] `http://openjdk.java.net/`

[3] `http://www.cs.waikato.ac.nz/ml/weka/`

Table 4: Classification using the combinations of 2 and 4 central moments statistics and histogram, for Random Forest and Support Vector Machine classifiers.

	Random forest			SVM		
Feat.	I	P	C	I	P	C
2S	58.19%	75.93%	59.80%	63.76%	79.16%	64.98%
	0.90s	0.70s	1.32s	5.54s	7.86s	26.32s
	0.11s	0.09s	0.13s	4.03s	4.07s	18.97s
4S	79.82%	**92.24%**	82.74%	79.70%	88.71%	80.77%
	1.01s	0.78s	1.87s	4.30s	4.76s	18.30s
	0.11s	0.09s	0.15s	3.12s	2.64s	13.90s
H	58.67%	91.74%	68.51%	34.69%	55.81%	41.92%
	1.04s	0.78s	1.74s	7.18s	4.32s	23.44s
	0.13s	0.11s	0.21s	5.01s	3.56s	16.72s
2S+H	74.61%	**92.24%**	78.71%	62.49%	80.17%	64.84%
	0.93s	0.74s	1.86s	4.54s	2.99s	14.65s
	0.12s	0.12s	0.22s	4.14s	2.49s	15.79s
4S+H	80.03%	**94.11%**	84.10%	77.79%	89.44%	79.47%
	1.02s	0.86s	1.83s	3.68s	2.75s	12.48s
	0.13s	0.11s	0.20s	3.57s	2.20s	14.83s

4.3 Classification results

Table 4 also shows that classification accuracy is significantly higher for positive than intermediate or combined intensity images. This is also consistent with Agrawal *et al* [1]. For this reason, we have focused on positive intensity only. In this scenario, we have highlighted in the table the 3 feature combinations with highest accuracy: 4S+H (94.11%), 4S (92.24%) and 2S+H (92.24%). When all images are considered, those feature combinations were also the ones with highest accuracy: 84.10%, 82.74% and 78.71%, respectively.

The classification accuracy per class is depicted in Figure 2. All feature combinations are included, so we can analyze the impact of each feature on each class. It can be noted that 2S only have a good (+90%) accuracy for the Ctr class. Specially for the Glg class, using 2S resulted in only 22.6% of accuracy, which is not shown in the graph, whose y-axis scale ranges from 60 to 100%. Histogram was the most discriminating feature for Ctr and Ncl, as combining it with statistical features didn't improve the classification results significantly for those classes. For Glg, Hgn, Nmn and Spk classes, the most discriminating feature was 4S. In all those classes, however, the combination with histogram showed the best results. Another interesting data from Figure 2 is that Glg was the only class which didn't achieve 90% accuracy in any feature combination.

4.4 Classification errors

For a more detailed view on classification results on the best accuracy scenario, Table 5 presents the confusion matrix. In this table, rows are counters for the real class, while columns are counters for the predicted class. The table shows that for Nmn class, classification errors are relatively well distributed among other classes. For other classes, most errors were concentrated on a single class: for Glg, Spk; for Ctr, Ncl; for Spk, Hgn; for Ncl, Ctr; and for Hgn, Spk. It can also be noted that Ctr and Hgn had mutually zero false positives.

Still on the best performing combination, Table 6 shows the false positives per class. This table presents the absolute

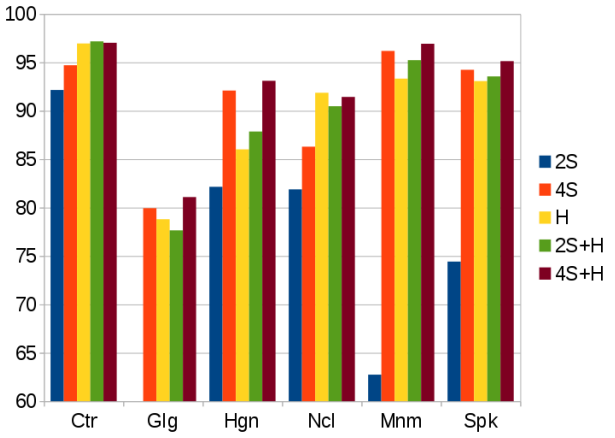

Figure 2: Accuracy per class for features with higher accuracy for positive intensity images, using Random Forest.

Table 5: Confusion matrix for the best performing combination: 4 moments central statistics plus histogram, using the Random Forest algorithm over positive intensity images.

	Nmn	**Glg**	**Ctr**	**Spk**	**Ncl**	**Hgn**
Nmn	914	3	4	9	4	9
Glg	5	283	18	26	16	1
Ctr	2	5	1,337	10	24	0
Spk	12	17	6	1,386	4	32
Ncl	6	4	60	7	854	3
Hgn	22	2	0	48	3	1,012

number of false positives (FP), the total number of instances per class (N) and the normalized false positive rate (FP / N). The highest normalized false positive rate was Glg. This may be related to the fact it was the class with less instances in the dataset. The class with less relative false positives was Hgn.

Table 6: False positives for the best performing combination: 4 moments central statistics plus histogram, using the Random Forest algorithm over positive intensity images.

	False positives	**Instances**	**Rate**
Nmn	47	943	4.98%
Glg	31	349	8.88%
Ctr	88	1378	6.39%
Spk	100	1457	6.86%
Ncl	51	934	5.46%
Hgn	45	1,087	4.14%

5. CONCLUSION

In our experiment for HEp-2 IIF cell classification, we have taken a different approach than most of the other works in the same subject (some of them reviewed in Table 2). Most of those publications extract a wide range of features from images, like morphology and texture, in the hope that it can increase classification accuracy. However, they usually

don't discuss the computational resources needed both to train the classifier and to extract features from images.

Our work, when considering positive-intensity images only, has achieved about 94% of classification accuracy. This result was obtaining using only statistics and histogram. These features are significantly faster to compute than others, as shown in Table 3. Most of the reviewed works achieved 90% or more of classification accuracy on HEp-2 IIF images [1, 27, 21, 9, 25, 10, 16, 22, 24, 3, 8]. All of them, however, used more numerous and complex features than we did. When considering images in both intensities (positive and intermediate), our classification accuracy drops to about 84%, but is still comparable to some of the reviewed works [7, 5, 13], which also applied features which are significantly more complex than the one we used.

We have demonstrated that, at least for positive intensity HEp-2 IIF images, using simple statistical and histogram features is enough to obtain a classification accuracy comparable to other, more complex approaches. Together with the Random Forest classifier, a computationally-efficient solution for automated image classification can be achieved. Computer-aided diagnosis systems could benefit from the proposed approach when classifying positive intensity images. For intermediate intensity images, such systems could rely on more complex methods, if needed.

6. REFERENCES

[1] P. Agrawal, M. Vatsa, and R. Singh. Hep-2 cell image classification: A comparative analysis. In *Machine Learning in Medical Imaging*, pages 195–202. Springer, 2013.

[2] N. Ahmed, T. Natarajan, and K. R. Rao. Discrete cosine transform. *Computers, IEEE Transactions on*, 100(1):90–93, 1974.

[3] W. ali, P. Piro, D. Giampaglia, T. Pourcher, and M. Barlaud. Biological cells classification using bio-inspired descriptor in a boosting k-nn framework. In *Computer-Based Medical Systems (CBMS), 2012 25th International Symposium on*, pages 1–6, June 2012.

[4] H. Bay, T. Tuytelaars, and L. Van Gool. Surf: Speeded up robust features. In *Computer Vision–ECCV 2006*, pages 404–417. Springer, 2006.

[5] E. Cordelli and P. Soda. Color to grayscale staining pattern representation in iif. In *Computer-Based Medical Systems (CBMS), 2011 24th International Symposium on*, pages 1–6. IEEE, 2011.

[6] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, volume 1, pages 886–893. IEEE, 2005.

[7] S. Di Cataldo, A. Bottino, E. Ficarra, and E. Macii. Applying textural features to the classification of hep-2 cell patterns in iif images. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 3349–3352. IEEE, 2012.

[8] P. Elbischger, S. Geerts, K. Sander, G. Ziervogel-Lukas, and P. Sinah. Algorithmic framework for hep-2 fluorescence pattern classification to aid auto-immune diseases diagnosis. In *Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on*, pages 562–565. IEEE, 2009.

[9] I. Ersoy, F. Bunyak, J. Peng, and K. Palaniappan. Hep-2 cell classification in iif images using shareboost. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 3362–3365. IEEE, 2012.

[10] S. Ghosh and V. Chaudhary. Feature analysis for automatic classification of hep-2 florescence patterns: Computer-aided diagnosis of auto-immune diseases. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 174–177. IEEE, 2012.

[11] L. F. P. Gonzalez, M. A. G. Pivel, and D. D. A. Ruiz. Improving bathymetric images exploration: A data mining approach. *Computers & Geosciences*, 54:142–147, 2013.

[12] R. M. Haralick, K. Shanmugam, and I. H. Dinstein. Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*, (6):610–621, 1973.

[13] T.-Y. Hsieh, Y.-C. Huang, C.-W. Chung, and Y.-L. Huang. Hep-2 cell classification in indirect immunofluorescence images. In *Information, Communications and Signal Processing, 2009. ICICS 2009. 7th International Conference on*, pages 1–4. IEEE, 2009.

[14] E. Keogh and A. Mueen. Curse of dimensionality. In C. Sammut and G. Webb, editors, *Encyclopedia of Machine Learning*, pages 257–258. Springer US, 2010.

[15] T. S. Lee. Image representation using 2d gabor wavelets. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 18(10):959–971, 1996.

[16] K. Li, J. Yin, Z. Lu, X. Kong, R. Zhang, and W. Liu. Multiclass boosting svm using different texture features in hep-2 cell staining pattern classification. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 170–173. IEEE, 2012.

[17] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004.

[18] S. G. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 11(7):674–693, 1989.

[19] P. L. Meroni and P. H. Schur. Ana screening: an old test with new recommendations. *Annals of the Rheumatic Diseases*, 2010.

[20] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 24(7):971–987, 2002.

[21] V. Snell, W. Christmas, and J. Kittler. Texture and shape in fluorescence pattern identification for auto-immune disease diagnosis. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 3750–3753. IEEE, 2012.

[22] P. Strandmark, J. Ulén, and F. Kahl. Hep-2 staining pattern classification. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, pages 33–36. IEEE, 2012.

[23] X. Tang. Texture information in run-length matrices.

Image Processing, IEEE Transactions on, 7(11):1602–1609, 1998.

[24] I. Theodorakopoulos, D. Kastaniotis, G. Economou, and S. Fotopoulos. Hep-2 cells classification via fusion of morphological and textural features. In *Bioinformatics & Bioengineering (BIBE), 2012 IEEE 12th International Conference on*, pages 689–694. IEEE, 2012.

[25] G. Thibault, B. Fertil, C. Navarro, S. Pereira, P. Cau, N. Levy, J. Sequeira, and J. Mari. Texture indexes and gray level size zone matrix application to cell nuclei classification. 2009.

[26] A. S. Wiik, M. Høier-Madsen, J. Forslid, P. Charles, and J. Meyrowitsch. Antinuclear antibodies: a contemporary nomenclature using hep-2 cells. *Journal of autoimmunity*, 35(3):276–290, 2010.

[27] A. Wiliem, Y. Wong, C. Sanderson, P. Hobson, S. Chen, and B. C. Lovell. Classification of human epithelial type 2 cell indirect immunofluoresence images via codebook based descriptors. In *Applications of Computer Vision (WACV), 2013 IEEE Workshop on*, pages 95–102. IEEE, 2013.