Statistical/Morphological Cell Texture Characterization and Classification

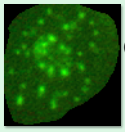
Guillaume THIBAULT and Jesús ANGULO
Guillaume.Thibault@Mines-ParisTech.fr
Jesus.Angulo@Mines-ParisTech.fr
Center for Mathematical Morphology (CMM), MINES-ParisTech
35 rue Saint Honoré, 77300 Fontainebleau, France

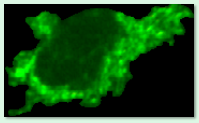
Methodology

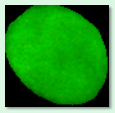
This paper presents an automatic and efficient fluorescence-labelled cell classification method, based on using two different families of descriptors and three different classifiers. The method provides a high classification rate for the dataset under study.

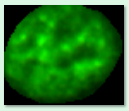
Working Data

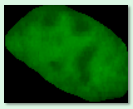
The dataset contains 1457 cells is acquired by indirect immunofluorescence (IIF), provided by the ICPR 2012 HEp-2 Cells Classification contest, and divided into 6 classes:

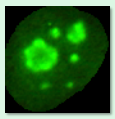

Centromeres (388)


Cytoplasmic (128)


Homogeneous (345)

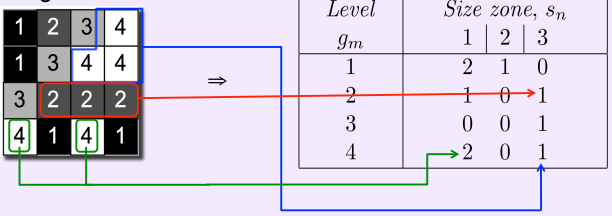

Coarse Speckled (239)


Fine Speckled (225)


Nucleolar (257)

Two Characterization Techniques and Three Classifiers

Gray Level Size Zone Matrix: [1] provides a statistical representation by the estimation of a bivariate (size / intensity) conditional probability density function of the image distribution values.



$Level$ g_m	$Size\ zone,\ s_n$ 1	2	3
1	2	1	0
2	1	0	1
3	0	0	1
4	2	0	1

Pattern Spectrum: [2] studies the size distribution of the objects of an image, with families of morphological openings φ and closings γ.

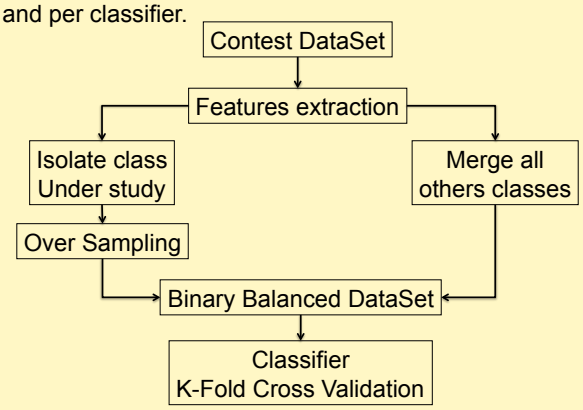
$$PS(f,n) = \frac{1}{\int f} \begin{cases} \int \gamma_n(f) - \int \gamma_{n+1}(f), for\ n \geq 0 \\ \int \varphi_{|n|}(f) - \int \varphi_{|n|-1}(f), for\ n \leq -1 \end{cases}$$

Classification:

- Logistic Regression [3] (LR)
- Random Forest [4] (RF)
- Neural Network [5] (NN)

Sub-model construction

One sub-model per class, per characterization technique and per classifier.



Sub-model Performances

Classes	LR	RF	NN
Centromere	81.97	97.86	93.28
Coarse speckles	98.2	99.59	98.72
Cytoplasmic	99.1	100	99.43
Fine speckles	97.56	98.48	94.47
Homogeneous	97.81	98.42	96.78
Nucleolar	93.46	99.53	96.92

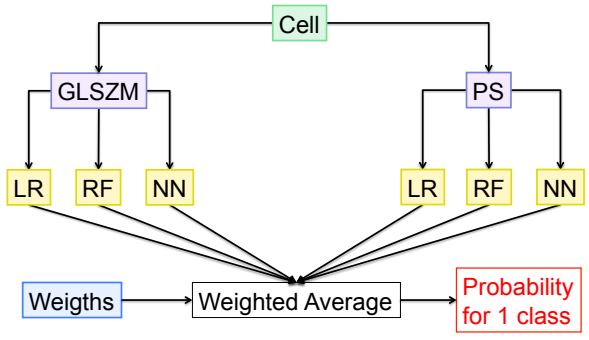
Size Zone Matrix

Classes	LR	RF	NN
Centromere	92.69	99.51	97.62
Coarse speckles	91.91	99.51	98.55
Cytoplasmic	97.06	100	99.51
Fine speckles	90.75	99.36	97.41
Homogeneous	93.61	99.04	94.61
Nucleolar	92.08	98.95	97.18

Pattern Spectrum

Weighted Integrated Model

One model per class.



Conclusions

This paper presents an application of statistical size zone matrix and morphological pattern spectrum to the automatic classification of cells. These methods provides powerful features who are then combined to three classifiers in order to provide an efficient classification of cells.

References

[1] G. Thibault, J. Angulo, and F. Meyer. "Advanced statistical matrices for texture characterization: Application to dna chromatin and microtubule network classification". In *IEEE International Conference on Image Processing (ICIP)*, pages 53–56, September 2011.
[2] J. Serra. "Image Analysis and Mathematical Morphology", volume 1. Academic Press, 1982.
[3] J. Berkson. "Application of the logistic function to bioassay". *Journal of the American Statistical Association*, 39:357–365, 1944.
[4] L. Breiman. "Random forests". In *Machine Learning*, 45(1):5–32, 2001.
[5] W. S. McCulloch and W. Pitts. "A logical calculus of the ideas immanent in nervous activity". In *Bulletin of Mathematical Biophysics*, 5:115–133, 1943.