PRO OR CONS LOCAL VS. GLOBAL IMAGERY INFORMATION FOR IDENTIFYING CELL MIGRATORY POTENTIAL

V. Vigneron[1,2]*, S. Lelandais*[1]*, C. Charriere-Bertrand*[1]*, M. Malo*[1]*, A. Ugon*[1]*, G. Barlovatz-Meimon*[1]

[1]IBISC, CNRS FRE 2873, 91020 Evry-Courcouronnes, France
email:{vincent.vigneron,sylvie.lelandais}@ibisc.univ-evry.fr,{gbm,malo}@lami.univ-evry.fr
[2]équipe SAMOS-MATISSE, CES CNRS UMR 8174, 90 rue de Tolbiac, 75634 Paris cedex 13, France
email:vigneron@univ-paris1.fr

ABSTRACT

Cell migration is a complex process involving adhesion, anchorage and de-adhesion. It is also a reliable indicator of the outcome of cancer. Can a migratory cell behavior be reliably associated to a cell morphology ? In the favourable case, it could become a visual indicator when coupled with image analyzis. In this work, microscopy images is much use to characterize the morphology of cells placed in various environments. Features were processed and cells were classified taking into account the biological expertise. Linking these results to experimental parameters and in vitro *data, we proposed to evaluate the cell migratory potential. The results give the expert new insights into the most useful features and show the feasability of an automated inspection system. A less common aspect is discussed throughout the paper regarding the relevance of the local (cell image) or global (whole image) information processing.*

1. INTRODUCTION

Nowadays, cell migration is a complex process well identified during development, present during wound healing and representing an indicator of the cancer outcome. Classically, when not circulating, cells are “embedded” within a proteic scaffold. A migrating cell destroys part of this matrix to acquire the necessary space for its translocation : the cell adopts an ellipsoidal shape and finds its way through the matrix. These events, in a “grip, stick and slip” repeated sequence, lead to the movement of the cell. This sequence is considered as essential in cell migration. The conventional mode of migration is called “*mesenchymal migration*”. Yet another mode of migration, the “*amoeboid type*” migration does not fit into the conventional cell migration model and could be utilised by metastatic cells [7]. Indeed, for the amoeboid migration, cells adopt rounded shape, migrate as individual cells, do not use the same type of adhesion/de-adhesion (via integrins molecules), develop weaker links between the cell membrane and the environment. However cell migration is still poorly understood, because it is a multi-scale phenomenon, and also because it puts at work forces, molecules, interactions for which no global mathematical model exists. Hence, there is a need for better prognosis factors. A better knowledge of this process is essential for proposing more subtle therapeutic approaches.
Recently attention has focus on the role in cell migration of the PAI-1 molecule, an inhibitor of proteolysis. Because of its presence in high amounts in the immediate vicinity of very invasive tumors, it is considered as an independent marker of bad prognosis (metastasis). PAI-1 has already been shown to influence cellmorphology [2, 4, 9] and could be implicated in the induction of the observed amoeboid movement of cancer cells.
The pro-migratory role of this molecule can be evaluated especially in terms of cell morphology. This role could lead to some (quantitative and reproducible) information of prognostic value. This is the main objective of the present work. Extraction of morphological “profiles” were applied at first to cells resulting from established and well known cell lines. The morphological approach consists in a binarization of the image followed by a morphological analysis of the cells shapes in order to distinguish between roundish and spread out cells. This approach is effective on well contrasted images. However, in a future work, features extraction based on wavelets decomposition should make it possible to obtain better results even on more noisy images.
Starting from a knowledge-acquisition process with a human operator (see section 3), we developped an image analysis and a feature extraction algorithm described in section 4. We present our results on image analysis, feature extraction and classification in section 6 : bayesiean classification on independent training and test sets has been performed. We estimated the parameters by utilizing 50% of the cells of each image in the training data and then tested the classification rule on the remaining 50% of the cells from the same image.

2. MATERIAL

The same colorectal line of cells (SW20) have been studied in two situations : a promigratory environment (PAI-1) and a nonpermissive environment. In this line, cells are fixed and colored with Crystal violet. The material included 24 256 gray-levels images of 1300×1030 pixels, in TIF format. They have been acquired though a CCD camera adapted to an optical microscope in conditions very far from optimal. Only SW20 metastatic colorectal cell line was included in this study.
The cells were studied with a `Zeiss AXIOVERT 200` inverted microscope. coupled with a Siemens CCD camera and digitized on a image processing unit at a final magnification of $5000\times$ and $0.05\mu m$ per pixel. Each image contain about 300 cells. We will try to establish their morphology in PAI-1 ($20\ \mu g/cm^2$) (Fig. 1.a) and collagen ($20\ \mu g/cm^2$) enriched environments (Fig. 1.b). On these test images, the cells are in dark colour because the colouring felt-tip is dark itself.

3. KNOWLEDGE ACQUISITION

The activation of the "moving cell machinery" leads to particular morphologies, *e.g.* “blebbing” cells [1, 7] (this cell

TAB. 1 – Human experts classification results. The percent of cells on which the experts are in accordance is given.

		non-specialists				specialists				Together[a]
Class		A (%)	B (%)	C (%)	$A \cap B \cap C^b$	D (%)	E (%)	F (%)	$D \cap E \cap F^a$	(%)
Binary	$\mathcal{C}_1$	68,9	36,52	15,61	13,7	46,66	34,44	50,27	29,45	11,74
	$\mathcal{C}_2$	28,94	59,34	84,0	27,6	46,50	61,07	49,37	38,06	24,82
	$\mathcal{C}_3$	2,12	4,14	0,39	0	6,84	4,49	0,36	0	0
Gray	$\mathcal{C}_1$	67,28	36,84	22,11	20,38	52,90	35,15	52,35	33,18	18,49
	$\mathcal{C}_2$	30,6	58,91	77,69	29,49	47,06	60,42	36,99	30,83	19,44
	$\mathcal{C}_3$	2,12	4,25	0,20	0	0,04	4,43	10,66	0	0

[a] $\approx A \cap B \cap C \cap D \cap E \cap F$
[b] $\approx$ percent of accordance.

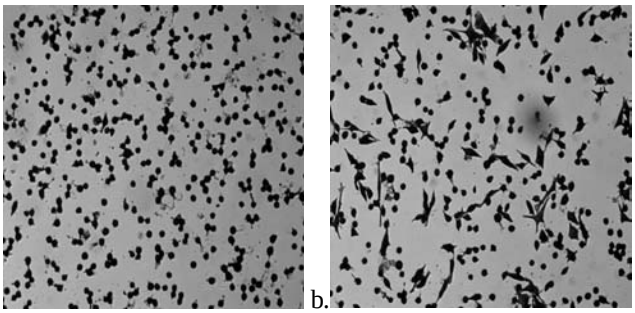

FIG. 1 – Cell samples (a) in PA1-1 environment, (b) on collagen.

morphology was recently described as "the" morphology of metastatic escaping cells[1]). In a knowledge-acquisition process with a human operator, using an interview technique, we acquired the knowledge of this operator, while classifying the different cell types. Some of this knowledge is shown in Table 2. The operator uses the cell's morphology as well as some texture information.

Class	Class name	Description
mesenchymal	$\mathcal{C}_1$	round
amoeboid	$\mathcal{C}_2$	streched out
undecided	$\mathcal{C}_3$	–

TAB. 2 – Some knowledge about the class description given by a human operator.

In addition, the appearance of the cell parts within the cells is of importance, like "blebbing edge", which also requires spatial information. We started out to develop the image analysis procedure and construct a feature set, which seems to be powerful enough to describe this symbolic knowledge. It is left to the data mining experiment to find out the relevant features for classification and to show us gaps in our description of the domain. Prior to feature extraction from imagery, round/streched cells have to be visually differentiate from their appearances and counted. The result of a human operator account is given for instance in the table 3.

[1] On this site `http ://www.lami.univ-evry.fr/~gbm/` a film on *in vitro* amoeboid migration of cancer cells on PAI-1 (`SW620 PAI-1 titre.avi`) and the control on collagen (`SW620 collagen titre.mpg`), can be seen.

Cells	round	streched
PAI-1 env.	**99,7%**	0,93%
collagen env.	43,17%	**56,9%**

TAB. 3 – Expert account of cells in PAI-1 and collagen environments.

Similar results were collected from images presenting the whole field seen by the microscope. Six experts (3 biologists and 3 non experienced biologists) inspect the images and classify the cells manually as "round, streched or undecided", or equivalently $\{\mathcal{C}_1, \mathcal{C}_2, \mathcal{C}_3\}$. The well-known problem in image interpretation "the difference between showing and naming" makes this methodology hard to follow by novices. The operator uses the gray-level intensity as well as some texture information.
Even for an experienced biologist it is often hard to decide the right class. From Fig. 1, we note that in the collagen environment, the cells are spaced, there is a not very significant number of round cells and the spread out cells have a form very characteristic. On the contrary, the number of round cells is very important in PAI-1 environment. Nevertheless, it is difficult to distinguish between a cluster of round cells and lengthened cells.
2,546 vignettes supposedly containing each a single cell steming from one of the 24 images have been analyzed. Tab. 1 illustrates the "gap" between experts and non experts judgements. Further analysis shows that the experienced biologists agree only on 36,56% of the whole set of binary vignettes, but this number raises to 37,9% in the case of gray-levels vignettes.

4. GLOBAL PROCESSING BY IMAGE ANALYSIS

4.1 Image processing

It is typical to divide the operations required into different stages (see Fig. 2). Clutter reduction refers to removing non-object data, this operation alters generally object data and thus it is not necessarily a separate stage. Automatic thresholding has been performed by the algorithm of Otsu [10]. The algorithm can localize the cells with their cytoplasmic structure, but not the nucleus itself. We then applied *morphological filters* like dilation and erosion to the image in order to get a binary mask for cutting out the cells from the image [5] (Fig. 3).

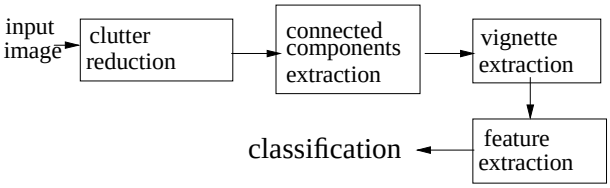

FIG. 2 – Hierarchical levels for cell images processing.

Then the number of objects in the class image is calculated. From the objects the area, some shape factors, the length of the contour, the eccentricity are calculated and these are fed to a classifier that determines the class (which type of object).

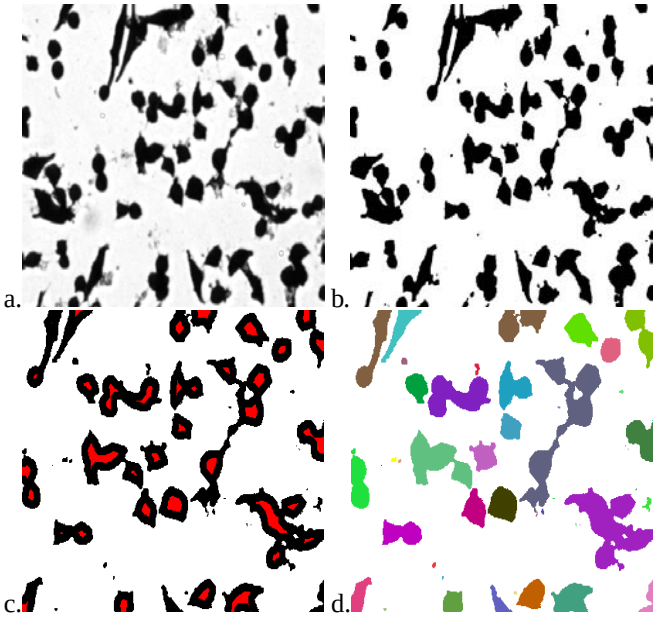

FIG. 3 – a) Initial image, b) binarisation, c) contours extraction, d) connected components extraction.

Characterization depends upon recognizing a cell as a member of some well-known class [6]. Note that the images $f(x,y)$ considered from now for further calculations are supposed to contain only one cell or possibly a cluster of cells. It is left to the data-mining experiment to find out the most relevant features for classification. We use a *sequential feature selection* algorithm (SFS) for feature selection before presenting them to the classifiers. Features were ranked in decreasing order by their discriminative power (their ability to distinguish classes). The most discriminating is the "*correlation*" character.

4.2 Radii histogram and correlation

The correlation feature can be obtained on the basis of the binarized image. We used the fact that the required cells are supposed to be (almost) round and we compute the correlation between the related components extracted from the image and a circle. The radius of this circle was obtained after determination of an average radius using a histogram of the radii (Fig. 4). From the binary image of the related components present in the image, we work starting from the germs obtained following the operation of erosion. Each germ is supposed to be at the origin of a cell. The histogram of the radii is then built by calculating the distance separating the contour points of the related component of the center to the barycentres of the germs. Indeed, the same related component can lead to several germs at the end of erosion (Fig. 4).

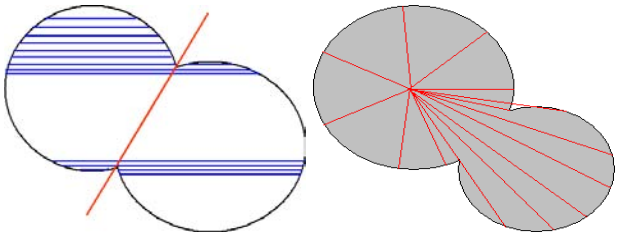

FIG. 4 – Calculus of the cell mean radius from a radius histogram.

Once all the related components are analyzed, the average radius is obtained in the class corresponding to the mode of this histogram. One thus builds a small image, containing a circle of this radius. The *correlation value* for one cell is the number of pixels belonging to the intersection cell/circle of average radius (Fig. 5).

FIG. 5 – Principle of the correlation calculus : the cell to analyze (left) is correlated with the mean radius cell (middle) providing an geometric intersection (right).

Any connected components which lead to a correlation value greater than a given threshold is supposed to satisfy the roundness criteria. The table below gives the results obtained for some values od the correlation threshold.

Corr. threshold	0,65	0,75	0,83	0,90
PAI-1 env.	**98,17%**	92,12%	71,90%	35,61%
collagen env.	84,76%	67,67%	**44,63%**	23,74%

TAB. 4 – Correlation results in terms of percent of round cells.

We test the hypothesis H_0 of no difference in environment. H_0 is rejected if $|T| \geq t_{n_x+n_y-2}(1-\frac{\alpha}{2})$. Here $\bar{y}-\bar{x}=0.1924$, $s^2 = 0.0376$ and $T = \frac{\bar{y}-\bar{x}}{s} = 5.117$. Because $t_6(0.975) = 2.7763$, we conclude at the 5% level of significance that there is a significant difference between the environment (see for instance more details on testing and confidence regions in [13, Chapters 20-22]. The level 0.95 confidence interval for the *difference in mean* round cells percent is

$$19.42\% \pm 3,8\%.$$

On the basis of the results of these experiments, we would conclude that environment induce significant change in the morphology of the cells. These results depend strongly of

	Description	**Name**	**Formula**
1	Hough parameter	H	–
2	Deviation between the length L and the width ℓ	DL	$\frac{\lvert L-\ell\rvert}{\min(\ell,L)}$
3	Ratio surface of the picture to surface of the cell	RS	$\frac{S}{L\times\ell}-\frac{\pi}{4}$
4	Sides equilibrium	LRE	$\lvert\sum_y p_{1j}-\sum_y p_{Lj}\rvert-\lvert\sum_k p_{k1}-\sum_k p_{k\ell}\rvert$
5	Distance between middle of the picture and cell (geometric) center	DOG	$\sqrt{(x_G-x_O)^2+(y_G-y_O)^2}$
6	Square of the perimeter to the area	$SP2$	$\frac{\left\{\sum_x\sum_y g(x,y)\right\}^2\lvert(x,y)\in\text{edge}}{\sum_x\sum_y g(x,y)\lvert(x,y)\in\text{ins}}$
7	Diagonal moment of the cell	SYM	$\sum_x(\sum_y g(x,y)^{\text{left}}-\sum_y g(x,y)^{\text{right}})^2$ $+\sum_y(\sum_x g(x,y)^{\text{left}}-\sum_x g(x,y)^{\text{right}})^2$
8	Measure of convexity	CNV	$\text{mean}\left(\frac{\sum_x\sum_y g(x,y)\lvert(x,y)\in\text{out}}{\sum_x\sum_y g(x,y)\lvert(x,y)\in\text{chord}}\right)$
9	Mean deviation between the observed number of pixels of the cell and the expected number	DST	$\frac{\sum_x(g(x,y)^{\text{obs}}-g(x,y)^{\text{exp}})^2}{L}+\frac{\sum_y(g(x,y)^{\text{obs}}-g(x,y)^{\text{exp}})^2}{\ell}$

TAB. 5 – List of features in a picture and their calculation.

the chosen threshold value. However, the algorithm find more round cells in collagen than with PAI-1.

Remember that Class 1 corresponds to the PAI-1 environment and the Class 2 to "collagenous" environment. The first conclusion to be drawn is that the result is very sensitive to the correlation threshold value. Nevertheless, whatever can be the tested threshold value, one finds always much less cells "labelled" as round in the collagenous environments than in the PAI-1, and this with a variation between 12 and 25%. The assumption according to which the environments induce different cell behaviors appears rather plausible. This operator of correlation could thus allow a first *differentiation* of the PAI-1 and collagenous environments. More interesting is the connection between these results and those presented in the table 3 which exposes the account of cells in the whole image (section 3).

First, there is a gap between the results obtained by this *correlation method* and those proposed by the human expert : this can be explained by the nondependence between the environmental context and the (automatic) method, which is by evidence not the case for the human expert. More precisely, the expert found a rate of round cells in PAI-1 comparable to the correlation threshold of $0,65$ (see Tab. 4). With the collagenous environment, the expert corresponds to the correlation threshold of $0,83$. This illustrates the fact that the expert analyzis is based on the whole image to decide if a cell is round or not. In the collagenous environment, cells are spaced, then it is easier to distinguish between a streched and a round cell. Do only well insulated cells with cleared roundness be "good" candidates for the "round" label[2] ? A comparison with the result of Class 2 for threshold $0,83$ should be made at this stage to answer the question. On the other hand, cells population in PAI-1 environment is much denser. One easily distinguishes spread out cells, including cells being not perfectly round. See for instance similar results for threshold $0,65$ of the Class 1.

Second, when the cells are extracted from the complete image and presented in small (square/rectangular) images to the expert, the decision to classify a cell as round/streched requires a mental process that one can probably put together with a high threshold correlation value. To decide that a circle is inside a square supposes that this circle is "perfect".

These remarks could explain the difference between the obtained results when the cells labelling is made in a local or a global context. More interesting, this correlation operator could be retained as a possible way to reach a compromise for stating on the roundness of the cell by using an intermediate correlation threshold. In our experiments, a threshold of $0,75$ guarantees a good differentiation between collagenous and PAI-1 environments.

[2]The doubt about the clusters benefitting the long cells.

5. NEURAL NETWORK CLASSIFIERS

We started out to develop the image analysis procedure and construct a feature set, which seems to be powerful enough to describe this symbolic knowledge. The features to be used for classification are classical morphological measures describing the cell geometry such as those described in [3]. The list of features included in the feature vector $\mathbf{x}=(x_1,\ldots,x_9)^T$ and their calculation is shown in Table 5. The whole data set has $2,546$ samples. Based on that data set, we acquired the knowledge for classification. The training set is subdivided into two sets : the calibration set and the validation set. We consider two classifier architectures : so-called *multi-layer perceptron* (MLP) and *belief network* (BN) as shown resp. in Fig. 6 and 7.

Both networks are designed for *supervised* learning of the diagnostic task and calculates the *posterior* probabilities $\{P(e_i|\mathbf{x}), i=1,\ldots,6\}$ for a particular $\mathbf{x}$. A set of "input" units is used to encode the feature vector. Classification problems require that each input vector $\mathbf{x}$ be assigned to one of 3 classes $\mathcal{C}_i, i=1,\ldots,3$, in which the target variables $\mathbf{t}$ represent class labels.

In Fig. 6, a layer of hidden units performs a weighted sum of teh inputs followed by a nonlinear sigmoid transformation. It is assumed there are no connection among the units within each layer. The ith hidden unit has the output $s_i=\sigma\left(\sum_{j=1}^{n_h}(w_{ij}x_j-b_i)\right)$. The sums extends over the input units. w_{ij} is the weight of the connection from input unit j to the hidden unit i, and b_i is the bias, $\sigma(\cdot)$ is the sigmoid or *logistic function* $\sigma(y)=\frac{1}{1+e^{-y}}$. The output units are receiving inputs

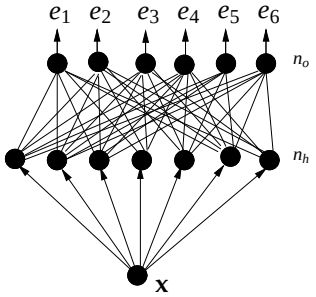

FIG. 6 – Deterministic two-layers feedforward neural network : the outputs – one output representing one expert – are probability distributions over the 3 possible classes.

from the hidden units.
The type of networks examined are feedforward, fully connected back-propagation networks as described in [12]. The back-propagation paradigm uses the *generalized delta rule* to determine the weights that will minimize the root mean square error (RMS) between desired and network outputs. The total RMS error is defined as :

$$E = \sqrt{\frac{\sum_{p=1}^{K}(t_p - e_p)^2}{K}} \qquad (1)$$

where t_p is the target pattern, K is the number of examples in the training set and $(t_p - e_p)$ is the difference between the target and the network outputs. The network is initialized with random weights and trained. The training is stopped before the error on the evaluation set starts to increase.
For classification purpose the output units are performing a sigmoid transformation which is a useful representation of the posterior probability over two classes [8].
The network in Fig. 6 is forced to produce the *posterior* probability of the *i*th expert $e_i = P(\mathcal{C}_j|x)$ by encoding the output values with 0/1-values units : 0 for a streched cell, 1 for a round cell. Hence, values in the "vicinity" of the value 0,5 indicate *undecision*. Such classifiers have numerous problems and issues. One is the training set size required with respect to the number of adaptive weights. Another concern is generalization (test performance vs training set performance). Note that this network encodes the experts know-how : such network provides better generalization (especially when the training data set is small) : the final assignment of a cell to a class will result from some combination of the outputs of the netwok, but the combination rule is not given here : it can be a majority vote, a maximum vote, etc.

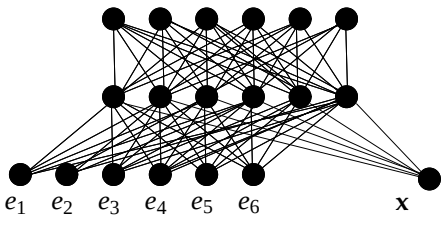

FIG. 7 – Bayesian network.

The *bayesian network* (BN) or *belief network* [11] represented Fig. 7 is designed for modeling the data without any particular task in mind. The input units are fed with *both* pattern set $\mathbf{x}$ and the desired predictions values $\{e_1,\ldots,e_6\}$ with $e_i = P(\mathcal{C}_j|\mathbf{x})$. As a result of training, these units may come to model correlations among features, among classes or between features and classes. If the network succeeds in modeling the total distribution $\{P(e_i|\mathbf{x}), i = 1,\ldots,6\}$ perfectly, it will be capable of performing any sort of pattern completion task. For example, one could clamp the attribute $\mathbf{x}$ and then observe the most likely class as the Gibbs sampling procedure is running, or conversely one could clamp a set of probabilities $\{P(e_i|\mathbf{x}), i = 1,\ldots,6\}$ (or even a part of them[3]) and observe the most likely feature $\mathbf{x}$. However, if the number of hidden units is *unsufficient* to model the total distribution, the network will end up modeling whichever correlations are strongest and these might not be the ones that are the most important for prognosis. The arrows in a belief network are a device for expressing probabilities, and need not correspond to real influences.
At the end, this involves a decision about the class to be obtained for a cell from the set probabilities $\{P(e_i|\mathbf{x}), i = 1,\ldots,6\}$. Note that in general the true class is *unknown*. One has not only to make decision about the final class, but also about the result obtained. Decision making has essentially to do with uncertainty.
Empirical data, while valuable, may be of limited extent. The need in such applications to integrate knowledge derived from experts with that derived from empirical data has been recognized by workers in the area. A group of tools are those that enable sets of data to be condensed and sumarized in ways that are clear and helpful – so-called *summary statistics*. For instance, the commonest measure appropriate for data of the form of a set of scalars $z_1,\ldots,z_n$ is the average defined by $\bar{z} = \sum_i^n z_i/n$, but it is inefficient in our case. This work being considered as non-crucial is left for future work.

6. RESULTS

To understand where our classifiers are good or failing, we give the *class-conditional error* rates, that is the error rate amongst examples of a class *i*. Further, we may want to know which classes are being confused, and so we may wish to know the set of probabilities

$$p_{ij} = P(\text{decision } j|\text{class } i), \forall 1 \leq i,j \leq 3 \qquad (2)$$

whose collection forms the *confusion matrix*.

MLP classifier (type 6)

	Class 1	Class 2	Class 3
numbers	1616	834	96
round	67,4 %	0 %	0 %
streched	32,5 %	77,1 %	100 %
undecided	0.1 %	23,9 %	0 %

Baysian network (type 7)

	Class 1	Class 2	Class 3
numbers	1616	834	96
round	78,4 %	0 %	10,4 %
streched	12,1 %	80,6 %	37,6 %
undecided	3,9 %	19,4 %	52,0 %

TAB. 6 – Confusion matrices.

[3] Missing data problem.

To assess the accuracy of the classifier, we applied the classifier to the training dataset. Experiences have been performed on the 2,546 vignettes splitted into a training and a test sets. Results are given in Tab. 6. The confusion matrices show the frequencies and conditional fractions of correct and incorrect classifications. Although the classifier has reasonnable high rates of correct classification for each population, the misclassification rates are not negligible. Fig. 8 visualizes some classification results (performed by the bayesian network).

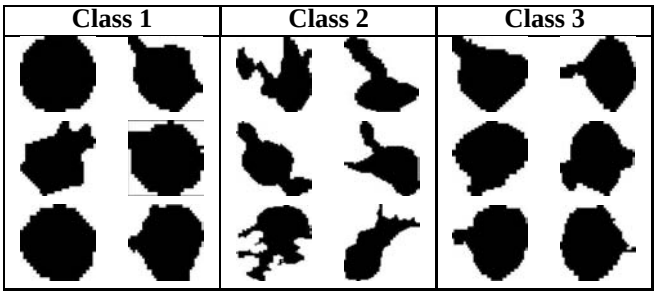

FIG. 8 – Binary images of cells picked up in predicted classes.

7. CONCLUSION

In this paper we put forward the difficulty of obtaining a truth reliable ground : experts opinions can diverged and can be influenced by the global/local view of the image. It is thus necessary to privilege methods whose results confirm the differentiation of the environments without inevitably obtaining the same results as the experts. The two methods suggested,– the method of global analysis based on the correlation and the total method based on the characterization of each related component by a set of parameters which are used as entry with a network of neurons – make it possible to obtain first promising results. The combination of these two approaches is certainly the way which will make it possible to obtain a reliable evaluation of the proportion of cells of each nature in the image.

REFERENCES

[1] S. Bourdoulous, G. Orend, D.A. MacKenna, R. Pasqualini, and E. Ruoslahti. Fibronectin matrix regulates activation of RHO and CDC42 GTPases and cell cycle progression. *J Cell Biol.*, 143(1) :267–7, Oct 1998.

[2] B. Chazaud, S. Bonavaud, A. Plonquet, M. Pouchelet, R.K. Gherardi, and G. Barlovatz Meimon. Involvement of the [μPAR :μPA :PAI-1 :LRP] complex in human myogenic cell mobility. *Exp Cell Res.*, 258(2) :237–44, Aug 2000.

[3] J.-P. Coquerez and S. Philipp. *Analyse d'images : filtrage et segmentation*. Masson, 1995.

[4] U. Dammer, O. Popescu, P. Wagner, D. Anselmetti, H.J. Guntherodt, and G.N. Misevic. Binding strength between cell adhesion proteoglycans measured by atomic force microscopy. *Science*, 24(267) :1173–5, 1995.

[5] E. Davies. *Machine Vision : Theory, Algorithms, Practicalities*. Academic Press, 1990.

[6] R.O. Duda and P.E. Hart. *Pattern classification and scene analysis*. Johen Wiley & Sons, 1973.

[7] P. Friedl and K. Wol. Tumour-cell invasion and migration : diversity and escape mechanisms. *Nat Rev Cancer*, 3(5) :362–374, 2003.

[8] M.I. Jordan and C.M. Bishop. *CRC Handbook of Computer Science*, chapter Neural Networks. CRC Press, Boca Raton, FL., 1996.

[9] M. Malo, C. Charrière-Bertrand, E. Chettaoui, C. Fabre-Guillevin, F. Maquerlot, A. Lackmy, A. Vallée, F. Delaplace, and G. Barlovatz-Meimon. The PAI-1 swing : Microenvironment and cancer cell migration. *C.R. Biol.*, 329(12) :938–944, 2006.

[10] N. Otsu. A threshold selection method from gray-level histograms. *IEEE Trans.*, SMC-9 :38–52, Jan 1979.

[11] J. Pearl. *Probabilistic reasoning in intelligent system : Networks of plausible Inference*. Morgan Kaufman, San Mateo, CA, 1988.

[12] D.E. Rumelhart, G.E. Hinton, and R.J. Williams. Nature. *Learning representations by backpropagating errors*, 323 :533–536, 1986.

[13] A. Stuart, K. Ord, and S. Arnold. *Kendall's Advanced Theory of Statistics*, volume Classical inference and the linear model. Oxford University Press, London, sixth edition, 1999.