



# NEW GENERAL FEATURES BASED ON SUPERPIXELS FOR IMAGE SEGMENTATION LEARNING

V Machairas, T Baldeweck, T Walter, Etienne Decencière

## ► To cite this version:

V Machairas, T Baldeweck, T Walter, Etienne Decencière. NEW GENERAL FEATURES BASED ON SUPERPIXELS FOR IMAGE SEGMENTATION LEARNING. International Symposium on Biomedical Imaging, Apr 2016, Prague, Czech Republic. hal-01276132

**HAL Id: hal-01276132**

**<https://hal.science/hal-01276132>**

Submitted on 18 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# NEW GENERAL FEATURES BASED ON SUPERPIXELS FOR IMAGE SEGMENTATION LEARNING

V. Machairas<sup>‡</sup>

T. Baldeweck<sup>\*</sup>

T. Walter<sup>†,††,†††</sup>

E. Decencière<sup>‡</sup>

<sup>‡</sup>MINES ParisTech, PSL-Research University, CMM-Centre for Mathematical Morphology, Fontainebleau, France

<sup>\*</sup>L'Oréal Research and Innovation, Aulnay-sous-Bois, France

<sup>†</sup>MINES ParisTech, PSL-Research University, CBIO-Centre for Computational Biology, Fontainebleau, France

<sup>††</sup>Institut Curie, 75248 Paris Cedex, France ; <sup>†††</sup>INSERM U900, 75248 Paris Cedex, France

## ABSTRACT

Segmenting an image is usually one of the major and most challenging steps in the pipeline of biomedical image analysis. One classical and promising approach is to consider segmentation as a classification task, where the aim is to assign to each pixel the label of the objects it belongs to. Pixels are therefore described by a vector of features, where each feature is calculated on the pixel itself or, more frequently, on a sliding window centered on the pixel. In this work, we propose to replace the sliding window by superpixels, i.e. regions which adapt to the image content. We call the resulting features SAF (Superpixel Adaptive Feature). Their contribution is highlighted on a biomedical database of melanocytes images. Qualitative and quantitative analyses show that they are better suited for segmentation purposes than the sliding window approach.

**Index Terms**— Image features, segmentation learning, superpixels, mathematical morphology

## 1. INTRODUCTION

The segmentation of an image is a partition of the latter into meaningful regions, i.e. regions which correspond to real objects in the image scene. Segmenting an image is usually one of the major and most challenging steps in the pipeline of image analysis. Numerous segmentation methods have been previously proposed. One interesting approach considers segmentation as a classification task, that aims at assigning to each pixel of the image the label of the object it belongs to. This method comprises two steps : first, each pixel is described by a set of features (such as color, position or values obtained through the application of an operator), second this vector of features is given as input to a machine learning method, which will output its corresponding label. This strategy allows the system to learn the rules according to which a pixel is assigned the object label from a set of annotated segmentation examples.

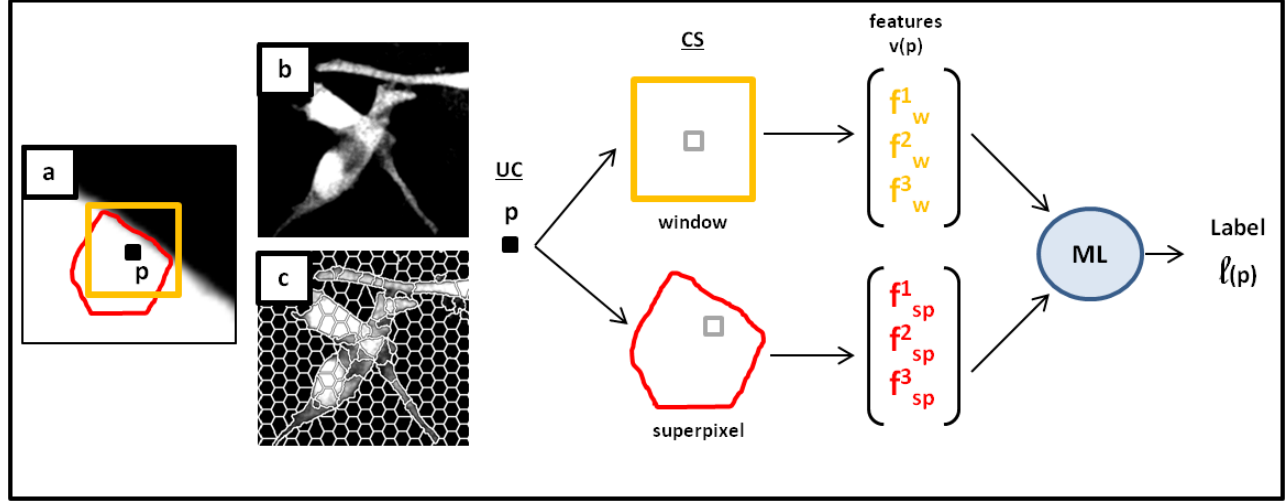
Many general or application-dependent features are proposed to classify pixels. They are often defined as the result of the application of an image operator (*e.g.* Gabor filter or local binary pattern) on a given *computational support* linked to the pixel. This computational support is usually a given neighborhood/fixed window centered on it. However, such computational supports which are independent of the image content can lead to biased information and hence poor classification. As practical example, let us picture a pixel that is close to an object contour in view of classifying it as belonging to object 1 or object 2, separated by this contour (see Fig.1.a for an illustration). A computational support such as sliding window is likely to overlap the two objects, giving an inappropriate feature. A natural strategy to overcome this problem is to use a computational support which adapts to the image contents. For instance, Morard et al. [1] and Gonzalez et al. [2] used adaptive structuring elements, defining a pixel-specific region on which to apply a given operator. They showed that indeed using an adaptive support to compute features is a promising strategy. However, computing an adaptive structuring element on each image pixels is too time consuming for the applications we are interested in.

The main contribution of this paper is to use superpixels as computational support. Note that our aim is not to classify the superpixels, as has already been done in the literature. Instead, we propose to compute pixel-level features based on the superpixel they belong to. As we will see, this allows to naturally combine different sizes or even types of superpixels to build features. Given that our final objective is to develop a fast automatic method for segmentation learning, we have chosen a superpixel approach which brings a good trade-off between quality and speed : the recently introduced waterpixels [3, 4]. Note however that this approach is general, and could be used with any sort of superpixels, such as superpixels based on simple linear iterative clustering (SLIC) [5]. We call *SAF (Superpixel Adaptive Feature)* the resulting features.

The paper is organized as follows : Section 2 explains how support adaptive features work and how they are integrated in the pixel classification pipeline. Section 3 shows the interest

---

Thomas Walter has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreements number 258068 (Systems Microscopy).



**Fig. 1. Illustration of the proposed method :** (a) regions for feature calculation (orange : sliding window, red : superpixel); (b) example image from the LORéal database (this image has been simplified for illustration purposes, see Fig.2 for a real case); (c) waterpixels (superpixels) computed on simplified image (b); computation of features on different computational supports, *i.e.* sliding window and superpixel, and their integration in the pixel's vector of features. “ML” : *Machine Learning method*.

of these new features for segmentation purposes. Section 4 concludes.

## 2. SUPERPIXEL ADAPTIVE FEATURES (SAF)

In the classification pipeline, each pixel  $p$  is described by a vector of features  $v(p) \in \mathbb{R}^P$  which will be given as input to the machine learning method. The elements that define each feature  $f^i$  are the computational support (CS), which is a connected set of pixels  $p$  belongs to, and an image operator (such as identity, linear filters or mathematical morphology operators). There are two ways to apply an operator with respect to a  $CS$  : either it is first applied on the whole image and then we evaluate the pixel values in the  $CS$  (non-geodesic way), or it is applied only on the  $CS$  (geodesic way) which enables not to take into account what happens outside this very  $CS$ . After applying the operator (either in a geodesic or a non-geodesic way), pixel values are averaged over the computational support to obtain a single value,  $v_i(p)$ . Note that other statistics than the mean could be used.

In this work, we propose to use superpixels as a new computational support for pixel classification. Superpixels are a special case of low-level segmentation : these regions form a connected partition of the image, they are homogeneous and rather regular, and they adhere well to object boundaries (*i.e.* object contours are included in superpixels contours). They have raised increasing interest in the vision community as they constitute interesting primitives for image analysis, carrying a higher level of information than the pixels themselves. Most of the time, they are used as unit of classification instead of pixels (*i.e.* each superpixel will be assigned a label,

or, in other words, all pixels belonging to the same superpixel will be assigned the same label). The idea being to alleviate the computational cost of treating each pixel of the image, especially as state-of-the-art superpixel generation methods show low complexity (linear with the number of pixels in the image), with fast implementations available. As far as classification is concerned, one drawback of using such approach though is that a misclassified unit of classification (here a group of pixels) will impact more classification performance than just a single pixel. Here, we propose to take advantage of both approaches by performing pixel classification, but using superpixels as computational support to enrich the vector of features of each pixel. We will call *SAF (Superpixel Adaptive Feature)* all features whose computational support is a superpixel.

In this paper, we will use *waterpixels*, a recently introduced superpixels generation method, for its good trade-off between computational efficiency and segmentation quality[3, 4]. Waterpixels are superpixels based on the marker-controlled watershed transformation (gradient-based approach), a powerful tool from mathematical morphology for segmentation. Two parameters  $\sigma$  and  $k$  enable the user to choose respectively the size and the spatial regularity of resulting superpixels.

The proposed strategy is illustrated in Fig. 1. Fig. 1.c shows a partition into superpixels of image Fig.1.b. For each superpixel  $s_j$ , one feature (or more) is calculated and the resulting value is stored in the vector of features of every pixel belonging to  $s_j$ . This enforces similarity between vectors of pixels belonging to the same superpixel and hence improves their chance to be classified with the same label. Actually, if we use only one partition, vectors of pixels belonging to the

same superpixel will be identical, which is equivalent to using the superpixel as classification unit. But here the process is repeated on different partitions of superpixels (computed with different parameters such as size, or even with different generation methods) in order to enrich the pixel’s vector of features with information captured at different scales. Eventually, using the pixel itself as *CS* is only one special case of SAF as it can be seen as a superpixel of size one pixel.

Once they are computed, vectors of features of all pixels coming from the database images are concatenated into a matrix  $X \in \mathbb{R}^{N \times P}$  (where  $N$  is the number of samples and  $P$  is the number of features), which will be given as input to a Random Forest classifier. Originally introduced by [6], Random Forest is an ensemble learning method which outputs the consensus between the individual classification results of a large number of decisions trees. They are widely used for their good performance for a large variety of applications, as well as for their robustness with respect to parameters choice.

### 3. EXPERIMENTS

#### 3.1. Evaluation process

We evaluate the interest of SAF features on a database of 2D grey level images (size 511x511 pixels) from the biomedical field and used in [7]. This database, provided by L’Oréal laboratories, contains eight images of reconstructed skin used in cosmetic research, acquired with multiphoton microscopy. The main visible structures in these images correspond to melanocytes, which appear as moderately bright, with elongated shapes (see examples in Fig.2.a, Fig.2.f and Fig.2.h).

The SAF approach is benchmarked against the sliding window approach on different operators, namely identity (to obtain the gray levels of the image) and a set of operators coming from mathematical morphology (erosion, dilation, opening, closing, top-hat and morphological gradient). The aim here is to compare the sliding window based features with SAF. Six evaluation criteria are used. The five first ones are usual measures for classification evaluation : precision  $P$ , recall  $R$ , f-score  $F$ , Jaccard index  $J$  and overall pixel Accuracy  $Acc$  (defined as the percentage of correctly labeled pixels in the image). Note however that these measures do not take into account the spatial structure of the result. In order to have a first quantitative measure of the spatial coherence of the result, we have also computed the number of connected components of the results, denoted by  $nb\_cc$ . As our images contain only a few cells, we can assume that the lower this value, the better the result.

A leave-one-out procedure is used to estimate the results quality, *i.e.* the training phase is performed on seven images and the prediction on the other (and last) one ; this process being repeated for each of the eight images of the database. Eventually, obtained values are averaged for each evaluation criteria.

The classification pipeline has been implemented in Python, with the help of in-house developed libraries Smil [8] and Morphm [9] (image analysis tools), as well as Scikit-learn [10]. For the Random Forest (Scikit-learn implementation), the number of trees  $n\_estimators$  has been set to 100.  $min\_sample\_leaf$  (the minimal number of data to preserve a leaf during training phase), a parameter controlling the tree depth, has been chosen equal to 100 in order to avoid the potential overfitting arising when the number of features is low as in our case. The training set has been randomly reduced to 10 000 samples, in such a way to obtain balanced classes.

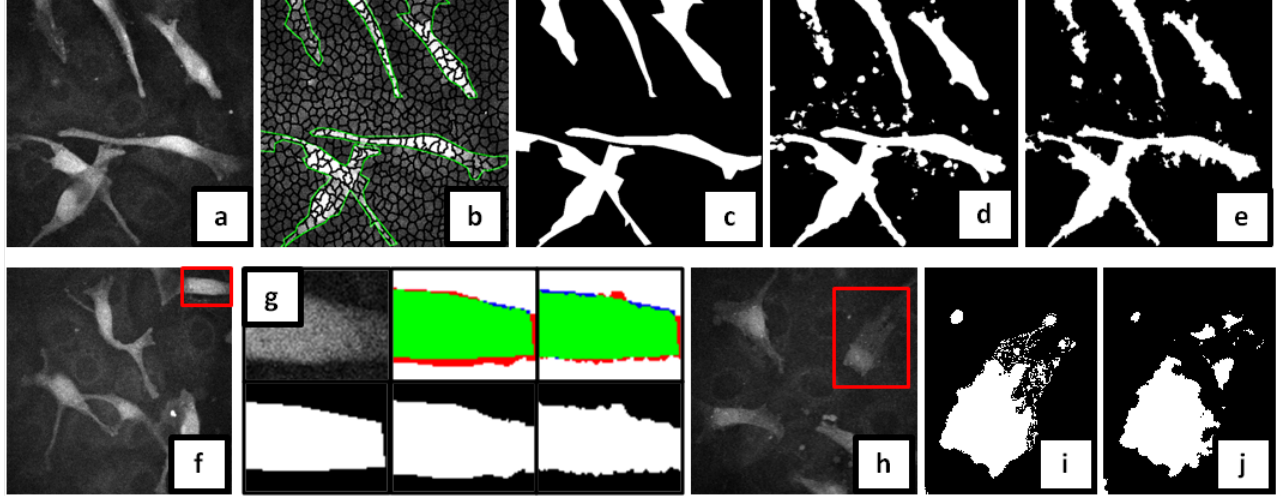
#### 3.2. Experiments and results

Results are presented in table 1. Quantitative comparisons will be performed on the basis of the three following measures : f-score  $F$  (summarizing the compromise between precision  $P$  and recall  $R$ ), Jaccard index  $J$ , and the number of connected components. Indeed, the overall pixel accuracy  $Acc$  is given as an indicative basis only as it tends to overestimate performances when classes are unbalanced.

For each operator, four sizes  $\sigma$  of waterpixels (with regularity parameter  $k$  set to 4 and hexagonal shape) are combined to fill the vector of features of each pixel ( $\sigma \in \{12, 15, 20, 25\}$ ), as well as for sliding windows scales (corresponding radius  $r_w \in [6, 7, 10, 12]$  and usual square shape). For morphological operators, the structuring element has been set to a 4-neighborhood with size 5. When specified, the “single scale” case corresponds to waterpixels of size  $\sigma = 20$  and windows of radius  $r = 10$ . This case enables us to see that SAF performances are better if we apply the morphological operators in a geodesic way rather than a non-geodesic way. The opposite phenomenon occurs for the sliding window approach (data not shown). This is coherent with the fact that a superpixel is a region which has a meaning (designed to be a piece of a unique object) whereas the window does not. In the following experiments (multi-scales), we will hence compare the superpixel approach with geodesic operators to the window approach with non-geodesic operators. We can see that both methods show similar classification performances, with a f-score of 73% and 72%, and a Jaccard index of 58% and 56% respectively for superpixels and sliding windows. However, SAFs enable to obtain a better spatial coherence between detected pixels, with an average number of connected components  $nb\_cc$  equal to 63, compared to 176 for sliding windows. The same observation can be done for the identity operator, with an even bigger difference in  $nb\_cc$  (68 and 183 respectively).

This spatial coherence advantage is illustrated in Fig.2, comparing Fig.2.d (sliding windows) and Fig.2.e (superpixels), as well as Fig.2.i (sliding windows) and Fig.2.j (superpixels).

Figure2.g, zoom of Fig.2.f, presents on the other hand the limitations of both approaches. Colors have the following



**Fig. 2. Experimental results :** (a), (f), (h) : original images from the L'Oréal database (with contrast enhanced for visualization purposes); (b) waterpixels of (a) with ground truth contours superimposed in green; (c) ground truth of (a); (d) : predicted image with the sliding window approach; (e) : predicted image with the SAF approach (superpixels); (g) : zoom of (f) with predicted image (second row) for sliding window approach (second column) and SAF approach (third column). The first row shows true positives (green), true negatives(white), false positives (red) and false negatives (blue). (i), (j) : zoom of predicted image of (h) respectively for sliding window and SAF approaches.

Feature	<i>CS</i>	Geodesic	scale	<i>P</i>	<i>R</i>	<i>F</i>	<i>J</i>	<i>Acc</i>	<i>nb_cc</i>
Morphological	superpixel	True	single scale	68±15	79±9	71±7	56±9	90±3	-
	superpixel	False	single scale	62±18	81±8	68±10	52±11	88±6	-
	superpixel	True	multi-scale	67±14	84±8	73±7	58±9	90±4	63±23
	window	False	multi-scale	63±16	87±8	72±9	56±11	89±5	176±61
Identity	superpixel	-	multi-scale	62±16	85±7	70±10	55±11	89±5	68±21
	window	-	multi-scale	62±18	87±8	70±11	55±12	88±6	183±78

**Table 1.** Experimental results on the L'Oréal database

meaning : green for true positives, white for true negatives, red for false positives, blue for false negatives. In the sliding window approach, pixels on either side of the contour will have very similar windows and hence be likely to be classified with the same label, that is why we can see both false positives and false negatives on the upper contour (Fig.2.g, second column). For SAFs features, limitations are of a different kind. Classification fails where contours are blurred as waterpixels are a gradient-based approach to compute superpixels. In this case, a superpixel may overlap a fragment of contour between two objects. We can observe that, most of the time, this overlapping superpixel is classified (or a part of it since we are in multi-scale approach) as background (thus the arrival of false negatives on borders). Figure 2.g, third column, illustrates this case, as well as when it is classified as foreground (middle of the upper contour, lower contour).

In conclusion, we have seen that the SAF approach can offer similar classification performances as the usual sliding window method, while presenting a better spatial coherence between detected pixels, which is more convenient for further

analysis of image objects.

#### 4. CONCLUSION

In this paper, we have proposed to use superpixels, instead of the usual sliding windows, as computational support for features used in pixel classification. We have shown that the resulting features perform better, mainly thanks to an improved spatial coherence, than the usual sliding windows-based features used in computer vision.

Future work will focus on combining multiple scales and operators in order to obtain a general segmentation learning method performing well on different bioimaging databases.

#### 5. REFERENCES

- [1] V. Morard, E. Decencière, and P. Dokladal, "Region growing structuring elements and new operators based on their shape," in *Signal and Image Processing (SIP 2011)*, Etats-Unis, 2011, ACTA Press.

- [2] Víctor González-Castro, Johan Debayle, and Jean-Charles Pinoli, “Color Adaptive Neighborhood Mathematical Morphology and its application to pixel-level classification,” *Pattern Recognition Letters*, vol. 47, pp. 50–62, Oct. 2014.
- [3] V. Machairas, E. Decenci re, and T. Walter, “Waterpixels : Superpixels based on the watershed transformation,” in *2014 IEEE International Conference on Image Processing (ICIP)*, Oct. 2014, pp. 4343–4347.
- [4] V. Machairas, M. Faessel, D. C rdenas-Pe a, T. Chabardes, T. Walter, and E. Decenci re, “Waterpixels,” *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3707–3716, Nov. 2015.
- [5] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Su sstrunk, “SLIC Superpixels Compared to State-of-the-Art Superpixel Methods,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, Nov. 2012.
- [6] Leo Breiman, “Random Forests,” *Machine Learning*, vol. 45, no. 1, pp. 5–32, Oct. 2001.
- [7] Andr s Serna, Beatriz Marcotegui, Etienne Decenci re, Th r se Baldeweck, Ana-Maria Pena, and S bastien Brizion, “Segmentation of elongated objects using attribute profiles and area stability : Application to melanocyte segmentation in engineered skin,” *Pattern Recognition Letters*, vol. 47, pp. 172–182, Oct. 2014.
- [8] Matthieu Faessel and Michel Bilodeau, “SMIL : Simple Morphological Image Library,” in *S minaire Performance et G n ricit , LRDE*, Villejuif, France, Mar. 2013.
- [9] “Morph-m website,” <http://cmm.enscm.fr/Morph-M/>, October, 2015.
- [10] Fabian Pedregosa, Ga l Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and  douard Duchesnay, “Scikit-learn : Machine Learning in Python,” *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, Oct. 2011.