Segmentation by Regions of Interest of Color Images Using Self-Organizing Maps Neural Networks*

*Note: Sub-titles are not captured in Xplore and should not be used

1st Zemihi Lila
dept. automatic
University of Mouloud Mammeri
Tizi-Ouzou, Algeria
z_automatique@live.fr

2nd Alkama Sadia
dept. automatic
University of Mouloud Mammeri
Tizi-Ouzou, Algeria
sadia_alkama@yahoo.fr

3rd Elmoataz Abderrahim
dept.Computer Science
University of Caen Basse Normandy
Caen, France
abderrahim.elmoataz-billah@unicaen.fr

***Abstract*—In many image processing applications, regions of interest are analyzed in order to extract the relevant information contained in the image. In this paper, we propose to segment regions of interest of a color image where the regions can be segmented into different number of classes. The Kohonen Self-Organizing Maps (SOM) algorithm, which is an unsupervised neural network, is used for this purpose. After, the neurons of the organized map are regrouped into clusters by using a new procedure based on the ascending hierarchical clustering which takes into account the connectedness of the neurons of the topological Kohonen map. The number of neurons groups obtained corresponds to the number of clusters in the region of interest; it can be different for each region.**
Experiments on different kinds of color images showed the efficiency of the proposed segmentation method.

***Index Terms*—Image Segmentation, Kohonen Self-Organizing Maps, Region of Interest.**

I. Introduction

Nowadays, the use of images has become indispensable in various sectors such as medicine, biology, remote sensing, etc. Images processing and analysis have become a rapidly expanding field of research. Segmentation is an important step in image processing as it conditions the interpretation of the results. Numerous image segmentation approaches have been developed [2]–[4], [6].

In this work, we propose an approach based on SOM neural networks. The proposed approach consists of two steps. The first step consists of training the SOM neural networks (Kohonen topological organizing maps) which have the ability to detect similarities, regularities and correlations between big data, grouping them into classes. The second step is to perform a grouping of the neurons using an ascending hierarchical classification [1] that we have modified for the purpose of our work and whose details are given in section III in the paper.

We propose in this work to segment the image not in the classical way by using all the pixels to obtain a segmentation result for the whole image but to work in regions of interest. Indeed, an image has often one or more regions that represent useful information for a user and a background that generally provides no information. The regions of interest may also have different interests for a user. For example, in one region he wants to have all the details while for others a coarse segmentation may be sufficient.

In this work, we propose a tool that allows the user to choose the number of regions of interest and the fineness of the segmentation for each region. To do so, we start in Section II by detailing the principle of Kohonen's self-organizing maps. In Section III, we explain the neurons grouping method we have developed. In Section IV, we will discuss the approach adopted and interpret the obtained results. We end the paper with a conclusion and some perspectives.

Identify applicable funding agency here. If none, delete this.

II. Kohonen self-organizing maps

A. Definition

Self-Organizing Maps (SOM) or topological are artificial neural networks whose learning takes place in an unsupervised manner. They were first developed in 1981 by Tenvo Kohonen [1], [8], [9]. This model has largely demonstrated its capabilities in performing discretization, classification, clustering, quantization and visualization tasks on multidimensional data

B. Concept and principle

The neural model of self-organizing maps proposed by Tenvo Kohonen has a main role which is to project the space of multidimensional observations (data) onto a discrete space of low dimension (generally 2). This space, which is called the map, will be noted φ. The projection allows the extraction of a set of vectors called neurons or prototypes $W = [w_{1i}, ..., w_{ni}]$. They are positioned in such a way that they preserve the topological form and neighbourhood relations of the observation space, i.e. data that are close in the input space will have close representations in the output space and will be classified in the same class or in neighboring classes [5] (see Fig. 1).

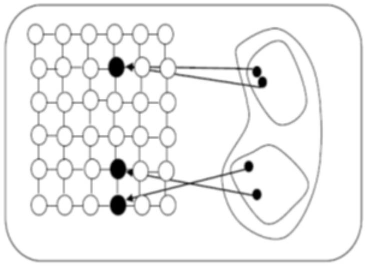

Fig. 1: Neighboring neurons on the map represent fairly "close" objects in the input data space.

The Kohonen network is therefore composed of two layers of neurons, an input layer and an output layer. The input layer is used to present the observations, which are in our case the pixels of the color image,$Y = \{y_i, i = 1, 2, ...3\}$. The output layer (also called adaptation layer) is composed of a lattice of neurons where each of them is connected to all the elements of the input layer and these neurons are organized according to a certain topological structure.

In this work we use a rectangular neighborhood topology with a 2D architecture which consists of interconnecting each neuron to its adjacent neighbors in the row and column to which it belongs (see Fig. 2).

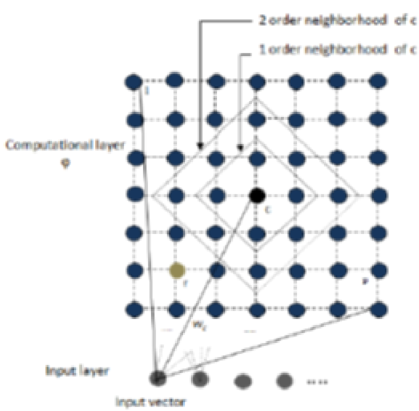

Fig. 2: Representation of a 2D topographic map of Kohonen.

The structure of this topology induces the distances $\delta(c, r)$ that link the winning neuron c to the other neurons r in the map φ. The neuron index $c \in \varphi$ is a winning neuron if its neuron vector w_c is closest to the input vector $y \in Y$.

The distances $\delta(c, r)$ allow us to vary the relative influence of different neurons. This importance is quantified by the neighborhood function $V(\delta(c, r))$ which forces neurons in the neighborhood of c to move their neuron vectors closer to the input vector y. A neuron which is closer to the winning neuron will have a smaller displacement. So the neighborhood function $V^T(\delta(c, r))$ decreases as the distance $\delta(c, r)$increases.

Several neighborhood functions can be used [1].The function used in this paper is the Gaussian one, given by equation 1 and illustrated in Fig. 3

$$V(\delta) = \exp \frac{|\delta|}{T} \tag{1}$$

This function is parameterized by a factor T to take into account the size of the neighborhood (see Fig. 3).

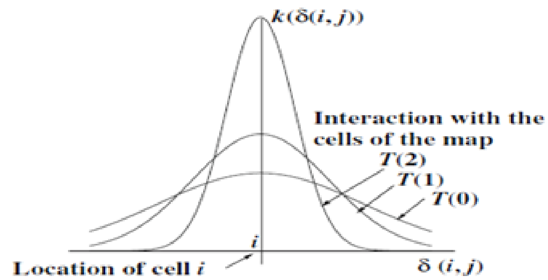

Fig. 3: Gaussian neighborhood function.

C. SOM learning

The objective of a SOM algorithm is to adapt the values of the weights of the neurons of the self-organizing map to the topology of the input space. This is done by minimizing the cost function noted J which is described by the following formula:

$$\mathcal{J}_{som}^T(\mathcal{F}, W) = \sum_{y_i \in Y} \sum_{c \in \varphi} V^T(\delta(c, \mathcal{F}(y_i))) \|y_i - w_i\|^2 \tag{2}$$

The minimization of this function is performed by successive iterations where each of these iterations realized in two phases:

- Competition (or affectation) phase: all the neurons of the grid (map) compete to determine the winning neuron corresponding to the neuron closest to the input vector. The affectation function used by Kohonen is given by :

$$\mathcal{F}(y_i) = argmin_c \|y_i - w_c\|^2 \tag{3}$$

- Adaptation phase: once the winning neuron is found, its neighboring neurons update their weights according to the function:

$$w_c^t = w_c^{t-1} - \mu^t V^T(\delta(c, \mathcal{F}(y_i)))(w_c^{t-1} - y_i) \tag{4}$$

Where μ^t represents the step size of the gradient which decreases according to the iterations t. In this work the gradient step function used is given by the following equation [2].

$$\mu^t = \frac{1}{\sqrt{t}} \tag{5}$$

The Kohonen algorithm is given as follows [1], [7] :

- **Initialization phase**
  - choose the structure, size of the map and the initial neurons (randomly).
  - set the values of T_{max} and T_{min} and the number of iterations N_{iter} , let $t = 0$.
- **t iterative step**
  - choose an observation y_i (usually randomly).
  - calculate the new value of T by applying the formula:
    $T = T_{max} \cdot (\frac{T_{min}}{T_{max}})^{\frac{t}{N_{iter}-1}}$
  - For this parameter perform the following two phases:
    - Affectation phase: affect the observation y_i to the neuron $\mathcal{F}_t(y_i)$ defined by the relation (3).
    - Minimization phase: change the values of W^t by the relation (4).
- **t=t+1**,Repeat the step 2 iteratively until $t = N_{iter}$

III. Grouping of Neurons

The segmentation of color images is the main objective of this paper. Therefore, we have proposed an unsupervised method which consists in a two-level classification. Indeed, each subset of observations is associated in the first level to a neuron of the map via the SOM algorithm. Once the self-organization of the map is completed, it must be completed by a second phase which consists in labeling each neuron of the map in order to obtain the required number of classes.

The labeling can be done by grouping these neurons as best as possible. We propose to group these neurons using a method based on the hierarchical ascending classification method [1].

Due to the topology of the map used in our work, it is not always possible to group neurons or groups of neurons that have a minimum vector distance. In order to group them together, they must also be neighbors in the topological map. We will use an example corresponding to a map containing 3x3 neurons to describe the clustering procedure (see Fig. 4). For an easier illustration, each neuron will be represented by a scalar value instead of a vector containing three values corresponding to the color components of the image.

Thus we can observe in Fig. 4.a a representation of a topological map of size 3x3 with neuron values between 0 and 1. In order to perform the first grouping, all distances between neurons i (i ranging from 1 to 9) and neurons j (j ranging from 1 to 9) of the map are computed. The measure used is a Euclidean distance. The results are displayed in Table Ia.

The minimum distance shown in this table is 0.01, which is achieved between the neurons N1 and N9. This minimum distance will not allow these two neurons to be grouping because they do not verify the neighborhood condition. In fact they are not neighbors (4 neighbors).

We then look for the smallest distance that verifies the neighborhood condition of the neurons in the map. In this example, this is the distance 0.05 measured between the referents N1and N2.

These two neurons will then be grouped together. This grouping is noted N1-N2 and its value is calculated using the following formula:

$$wg_j = \frac{1}{ng_j} \sum_{i \in g_j} w_i \tag{6}$$

where:

g_j is the set of indices of the neurons contained in grouping number j

wg_j is the weight of grouping j

ng_j is the number of neurons in grouping j

w_i is the weight of neuron i

This first grouping will then have a weight : $wg_1 = 0.52$.

A new distance table is calculated (see Table1.b) to determine the next grouping to be performed. And so on, the process is iterated until the desired number of classes is obtained (see Fig. 4 and Table I).

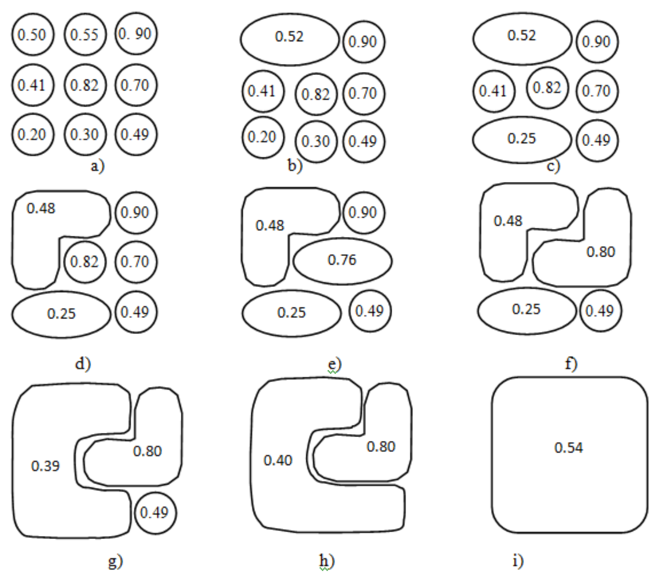

Fig. 4: Example showing the different stages of neurons grouping.

a) Example of a Kohonen map of size 3x3 after self-organization b) Map after the 1st grouping. c) Map after the 2nd grouping. d) Map after the 3rd grouping. e) Map after the 4th grouping. f) Map after 5th grouping. g) Map after 6th grouping. h) Map after the 7th grouping. i) Map after the 8th grouping

TABLE I: Tables of distances between groupings of neurons for the example of Fig. 4.

(a) Distance table for the choice of the first grouping

	N1	N2	N3	N4	N5	N6	N7	N8	N9
N1	0	0.05	0.4	0.09	0.32	0.2	0.3	0.2	0.01
N2		0	0.35	0.14	0.27	0.15	0.35	0.25	0.06
N3			0	0.49	0.08	0.20	0.7	0.6	0.41
N4				0	0.41	0.29	0.21	0.11	0.08
N5					0	0.12	0.62	0.52	0.33
N6						0	0.5	0.4	0.21
N7							0	0.10	0.29
N8								0	0.19
N9									0

(b) Distance table for the choice of the second grouping

	N1-N2	N3	N4	N5	N6	N7	N8	N9
N1-N2	0	0.38	0.11	0.3	0.18	0.32	0.22	0.03
N3		0	0.49	0.08	0.20	0.7	0.6	0.41
N4			0	0.41	0.29	0.21	0.11	0.08
N5				0	0.12	0.62	0.52	0.33
N6					0	0.5	0.4	0.21
N7						0	0.10	0.29
N8							0	0.19
N9								0

(c) Distance table for the choice of the 3rd grouping

	N1-N2	N3	N4	N5	N6	N7-N8	N9
N1-N2	0	0.38	0.11	0.3	0.18	0.27	0.03
N3		0	0.49	0.08	0.2	0.65	0.41
N4			0	0.41	0.29	0.16	0.08
N5				0	0.12	0.57	0.33
N6					0	0.45	0.21
N8						0	0.24
N9							0

(d) Distance table for the choice of the 4th grouping

	N1-N2-N4	N3	N5	N6	N7-N8	N9
N1-N2-N4	0	0.42	0.34	0.22	0.23	0.01
N3		0	0.08	0.2	0.65	0.41
N5			0	0.12	0.57	0.33
N6				0	0.45	0.21
N7-N8					0	0.24
N9						0

(e) Distance table for the selection of the 5th grouping

	N1-N2-N4	N3	N5-N6	N7-N8	N9
N1-N2-N4	0	0.42	0.28	0.23	0.01
N3		0	0.14	0.65	0.41
N5-N6			0	0.51	0.27
N7-N8				0	0.24
N9					0

(f) Distance table for the selection of the 6th grouping

	N1-N2-N4	N3-N5-N6	N7-N8	N9
N1-N2-N4	0	0.32	0.23	0.01
N3-N5-N6		0	0.55	0.31
N7-N8			0	0.24
N9				0

(g) Distance table for the selection of the 7th grouping

	N1-N2-N4-N7-N8	N3-N5-N6	N9
N1-N2-N4-N7-N8	0	0.32	0.01
N3-N5-N6		0	0.31
N9			0

(h) Distance table for the selection of the 8th grouping

	N1-N2-N4-N7-N8-N9	N3-N5-N6
N1-N2-N4-N7-N8-N9	0	0.4
N3-N5-N6		0

After the organization of the topological map and the grouping of neurons into nc classes, the image pixels are then assigned to the different classes. Each pixel is assigned to the class that is closest to it, i.e. to the grouping of neurons whose weight is closest to its value.

IV. Tests and results

Our goal is to segment the color images with a different fineness (different number of classes) for different regions of interest. Indeed, for example, in the medical field, the doctor only examines a part (region) of the image. This is the part where the disease is located. It is this part that is of interest and for which a fine segmentation must be carried out, whereas for the background it is not necessary to carry out segmentation with several classes because it does not provide any useful information (coarse segmentation). The approach adopted to carry out this segmentation is developed below.

A. Adopted approach

The choice of regions of interest and the number of classes for each region is made by the user according to the analysis to be carried out and the objective sought. To better understand the approach adopted, we illustrate its implementation using the image shown in Fig.5 which contains four regions of interest.

Fig. 5: Example of an image to be segmented into regions of interest .

For this image we want to segment the four regions corresponding to a ball and three players, and consider the rest as a background. The first task is to select the four regions of interest manually. The selection is made by delimiting these regions with the mouse. The boundary can have any shape. In the case of the image in Fig. 5, we perform the selection of the regions and the delimitation is shown in green in Fig. 6

Fig. 6: Image of Figure 5 with manually selected regions .

The next task is to set the desired number of classes for each region. In the case of this example we choose 8 classes for region 1, which corresponds to the player on the right side, 8 classes for region 2, which corresponds to the player in the center of the image, 2 classes for region 3, which corresponds to the ball, and 9 classes for the player on the left side of the image. For the background, which contains no useful information, we set the number of classes to 1. The result of the proposed segmentation is shown in the figure below.

Fig. 7: Segmented image.

B. Segmentation results

In this section we will apply the segmentation developed in the previous sections to test images from the Berkeley [11] and Kaggle [10] databases.
To better illustrate the results, we present in the following the segmentations obtained on images with one and several regions of interest.
In Fig. 8, we show the results of the segmentation on images with a single region of interest. Indeed, the image presents often only one object of interest for a user. For example, in the case of traffic cameras, the interest is to read the registration numbers of vehicles that have committed a traffic violation as in the case of Fig. 8.1.a. In the medical field, to establish a diagnosis on the skin lesion, the doctor only focuses on the lesion and ignores the rest of the image as in the case of Fig. 8.2.a.
It also happens that the image presents several regions of interest for the user (see Fig. 9).
For example, in the medical field, several regions of the skin may have lesions, as in the case of the images in Figs. 9.1 and 9.2, where the skin has 2 and 3 lesions respectively.
In the sports field, to identify athletes, it is necessary to read the numbers of their bibs as in the case of Fig. 9.4 where the 3 regions of interest correspond to the 3 bibs.

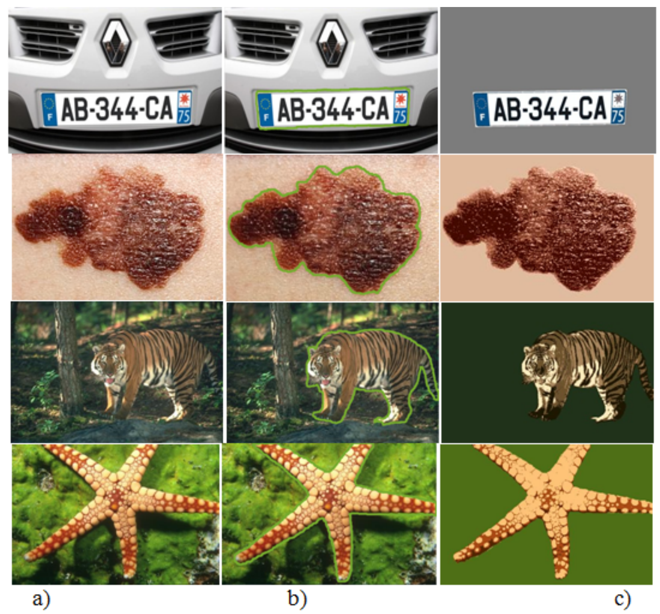

Fig. 8: Result for test images with a single region of interest. a) Original images, b) Images with a selected region, c) Segmented images.

For each manually selected region in the images of Fig. 8 and Fig. 9, we set the number of classes according to tables II and III, knowing that the regions are numbered from left to right in the images with several regions.

TABLE II: Number of classes of the manually selected regions of the images in Fig. 8

image number	nc
8.1	5
8.2	15
8.3	4
8.4	3

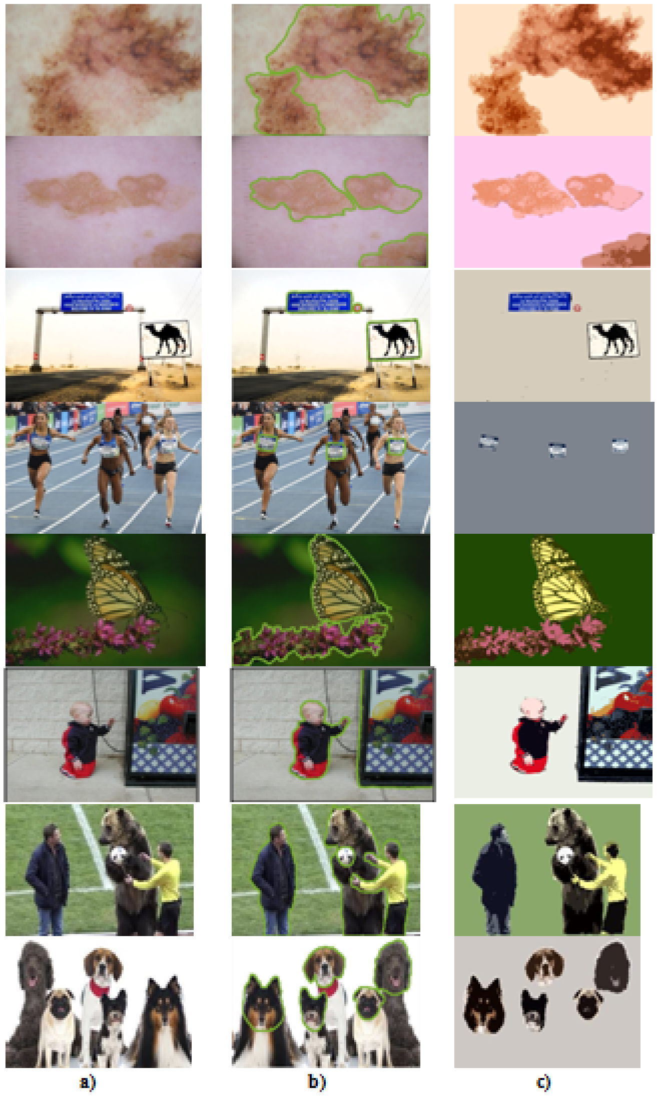

Fig. 9: Result for test images with several regions of interest. a)Original images, b) Images with several selected regions, c) Segmented images.

TABLE III: Number of classes of the manually selected regions of the images in Fig. 9

image number	nc Region1	nc Region2	nc Region3	nc Region4	nc Region5
9.1	5	10	/	/	/
9.2	3	4	2	/	/
9.3	2	3	2	/	/
9.4	5	5	5	/	/
9.5	3	5	/	/	/
9.6	4	7	/	/	/
9.7	3	7	2	3	/
9.8	6	4	5	5	3

The results of the segmentation of images 8.1.a, 8.2.a, 8.3.a and 8.4.a in Fig. 8 with the numbers of classes given in Table II are given in images 8.1c, 8.2c, 8.3c and 8.4c in the same figure. In this figure, each image shows one region of interest. The results of the segmentation of images 9.1.a, 9.2.a, 9.3.a, 9.4.a, 9.5.a, 9.6.a, 9.7.a and 9.8.a in Fig. 9 with several regions of interest are shown in images 9.1.c, 9.2.c, 9.3.c, 9.4.c, 9.5.c, 9.6.c, 9.7.c and 9.8.c of the same figure with the numbers of classes given in Table III. We can see, for example, that in the case of Fig. 8.1, the vehicle number plate is well segmented and reading it becomes easier. An automatic or semi-automatic system can more easily use the image in Fig. 8.1.c to recognize the numbers and letters on the plate. Indeed, the information to be processed in this image is greatly reduced compared to the case where the whole segmented image is used. The number of classes we have chosen in this example (which is 5) is more than sufficient since all the details of this plate are found in the segmented image. In the case of the following images in Fig. 8, the number of classes for the regions of interest is chosen according to the details we wish to have. In the case of the starfish (Fig. 8.4), three classes are sufficient to detect the epidermal growths (i.e. pedicellar) covering the central disc and the five arms of the star. In the case of image 8.3, we have performed the segmentation of the tiger with four classes corresponding to the black stripes, the white fur, the light brown fur and the dark brown fur (or shaded regions). We can see that the segmentation is well done. For image 8.2, we opted for a larger number of classes (15 classes) because in the medical field detail is important for diagnosis. The different nuances are clearly visible, although the number of color levels is greatly reduced. It also happens, in the medical field, that the skin presents more than one lesion, which is the case of images 9.1 and 9.2 of Fig. 9 which present two and three lesions. We have therefore selected respectively two and three regions of interest and the number of class assigned for each region is given in Table III. These numbers are of course chosen according to the details we wish to have of the lesions. We note that, for these two cases, the results of the segmentations are satisfactory. Indeed they allow us to visualize and read easily the different nuances in the lesions. In image 9.3 of Fig 9, the regions of interest are the three road signs. The segmentation is performed with a reduced number of classes for each sign (2, 3 and 2 (see Table. III)) and the result is satisfactory since the letters, numbers and the pictures are clear. The same is true for the other images in Fig. 9 where the number of regions of interest is between two and five and where the segmentation provides images with a reduced number of information while designing the information useful for the application domain. Indeed, for example, in the case of Fig. 9.8, the heads of the dogs are the five regions of interest and the result of the segmentation given by image 9.8.c allows identifying the family of each of them.

V. Conclusion

The main objective of this work is to allow a user to choose the regions on the image to be processed that have an interest for his application. We allow him to delimit as many regions as he wishes and give him the possibility to perform segmentations with a different number of classes from one region to another. To achieve these objectives, we implemented the SOM algorithm, which is one of the artificial neural networks. The learning is performed in an unsupervised manner. We then developed a technique for grouping the neurons of the map to obtain the desired number of classes for each region of interest. Tests carried out on images from different domains of use and presenting one or more regions of interest have shown the effectiveness of the proposed method.

REFERENCES

[1] S. Alkama, Y. Chahir and D. Berkani, "Label maps fusion for the marginal segmentation of multi-component images," Neural Network World, vol. 25, pp. 405–426, 2015.
[2] K. Qin, K. Xu, F. Liu and D. Li, "Image segmentation based on histogram analysis utilizing the cloud model," Computers & Mathematics with Applications," vol. 62, pp. 2824–2833, 2011.
[3] O. Severino and A. Gonzaga," A new approach for color image segmentation based on color mixture," Machine vision and applications,vol. 24, pp. 607–618, 2013
[4] H. Min, X. F. Wang, D. S. Huang, J.Jin, H. Z. Wang and H. Li, "Level set method for image segmentation based on moment competition," Journal of Electronic ,vol. 24, 033020 ,2015.
[5] S. khedairia and M. T. Khadir, "Impact of clustered meteorological parameters on air pollutants concentrations in the region of Annaba, Algeria," Atmospheric research, vol. 113, pp. 89-101, 2012.
[6] R. Gothwal,S. Gupta,D. Gupta and A. K. Dahiya," Color image segmentation algorithm based on RGB channels," Proceedings of 3rd International Conference on Reliability, Infocom Technologies and Optimization, pp. 1–5, IEEE, India, 2014.
[7] G . Dreyfus, J. M. Martinez,M. Samuelides,M. B. Gordon,F. Badran,S. Thiri and al. Hérault, "Réseaux de neurones ," Vol. 39, Paris: Eyrolles, 2002.
[8] T. KOHONEN, Self organization and associative memory, Springer , Heidelberg ,1984.
[9] T.KOHONEN, Self Organizing Maps, 3rd. ed. Springer, 2000.
[10] Kaggle dataset, https://www.kaggle.com/datasets, last accessed 2021/01/25.
[11] Berkeley dataset,https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/, last accessed 2021/01/25.