

Support Vector Machine Applied to Compress Medical Image

Chahinez Mérièm Bentaouza^{1,2*}, Mohamed Benyettou¹

¹ Department of Computer Science, University of Science and Technology of Oran, Oran, Algeria.

² Department of Computer Science, University of Mostaganem, Mostaganem, Algeria.

* Corresponding author. Email: chahinez5@yahoo.fr

Manuscript submitted May 25, 2017; accepted July 5, 2017.

doi: 10.17706/jcp.13.5.580-587

Abstract: This article presents the constitution of compressed image after learning by support vector machines applied to medical images. Medical images are a valuable means of decision support. However, they provide a large number of images per examination. Compression is used to reduce the size of medical images for storage and transmission. This work uses Support Vector Machines for compression to reduce the pixels of microscopic images. The compression rates are satisfactory. However, we have a loss of visual information because the compressed image has only support vectors especially when using radial function.

Key words: Compression, microscopic image, support vector.

1. Introduction

The medical images represent a large enough depending on the tool acquisition. For example, the scanner is 200 MB to 1 GB per examination; an MRI (Magnetic Resonance Imaging) is an average of 50 MB per examination [1] and the resulting images of a microscope of up to 1GB for examination [2]. In all, the annual production of a medical imaging center is estimated at several hundred or even thousands of GB.

Consequently, image compression is the solution for reducing image size to archive it for a number of years according to country law.

A lot of work we're interested by compression uses statistical methods or machine learning methods [3].

The compression or information theory dates back to Claude Shannon's work in 1948 allowing returning messages to be as short as possible, which is essential for comprehension and without redundancy [4].

This paper is the second part of the medical image compression by the Support Vector Machine (SVM) recognized by their performance in machine learning [5].

2. Image Compression

The image compression is used to reduce the size of images without losing important information while keeping an acceptable visual aspect.

There are two different types of compression: lossless compression and lossy compression.

Lossy compression allows removing redundant information. It is possible to have the same image, but some data is removed and cannot be recovered. The losses are undetectable to the eye, however, the initial image is not identical to the decompressed image [6].

Lossless compression does not eliminate any information. It is possible to find all pixels of the original image. After decompression, the reconstituted image is identical to initial image [6].

In this work, we will test a method to see if it can be used as lossless or lossy, though no loss is preferable because we are dealing with medical images where loss of information can change a diagnosis.

3. Support Vector Machine

A Support Vector Machine is a method from statistical theory [7] for supervised learning suggested by Vladimir Vapnik in 1992. The goal of SVM is to find a linear classifier to separate data and maximize distance between classes. This classifier is called a hyper plane. The closest points, which are used to determinate the hyper plane, are called support vectors (see Fig. 1).

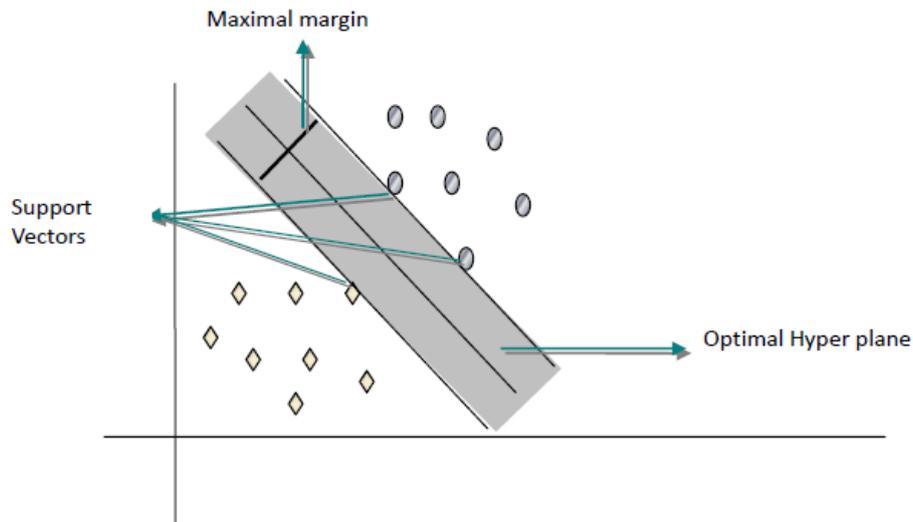


Fig. 1. Support vector machine.

The principal idea of a binary SVM is to trace data on the space of higher dimension by using kernel functions to identify the hyper plane in order to separate training examples. A kernel function used in these experiments is a polynomial function (see equation 1) and radial function (see equation 2):

$$k(x, y) = (x \cdot y + 1)^d \quad (1)$$

$$k(x, y) = \exp(-(x - y)^2 / \sigma^2) \quad (2)$$

Only the examples corresponding to the support vectors are really useful in learning [8]. Because the support vectors contain important aspects of the data set, the data set can be compressed while retaining only the support vectors.

For our project, we used multi-class method One-Vs-Rest (OVR). OVR is the simplest method of multi-class SVM or K binary SVM classifiers: one positive class against all negative classes. The decision function chooses the maximum value of K binary decision.

4. Experiments

4.1. Configuration

The images used are a type « astrocytoma » of “II” grade; it is from central nervous system tumors, of which growth is slow on different categories according to principal localization [9]. We used C++ Builder 6 for the preprocessing of medical images to have colors (Red, Green and Blue) as entry data of SVM-Light Multi-class [10] for supervised learning.

4.2. Pre-treatment

We take different images of an appropriate type and size. Each image is divided into blocks of 16 x 16 pixels. In each block, we regroup the same pixels which have the same value RGB in the same class (as shown in Table 1) for each image. Afterwards, we start to learn each bloc with SVM-Light Multi-class [10] because it is easier to learn by block than to learn an image.

Table 1. Example of Pixel Value in a Block with SVM-Light Multi-class

RGB Value
1 1:0 2:1 3:7
2 1:6 2:11 3:17
3 1:14 2:19 3:25
4 1:9 2:14 3:20
5 1:3 2:8 3:14
5 1:3 2:8 3:14
6 1:5 2:10 3:16
7 1:4 2:9 3:15
8 1:17 2:8 3:27
9 1:4 2:0 3:14
9 1:4 2:0 3:14
10 1:16 2:7 3:26

Table 2. Example of Obtained Values after Block Learning

Value	Support Vector
4.6256767998371746 e - 011	Yes
4.7474963825494089 e - 011	Yes
5.2789323068850174 e - 011	Yes
0	No
9.0041365783835314 e - 011	Yes
0	No
9.8891746822150612 e - 011	Yes
0	No
0	No
0	No
9.9783341772223327 e - 011	Yes
9.369367491988727 e - 011	Yes

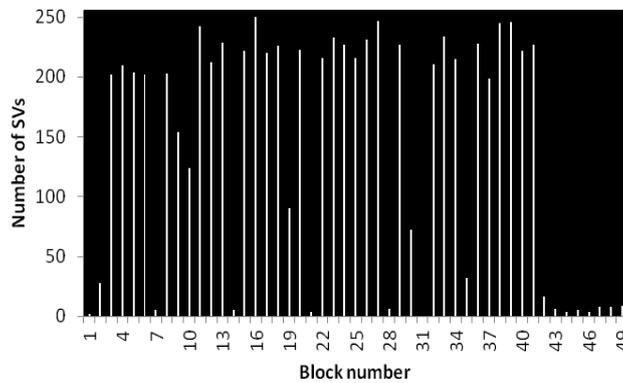


Fig. 2. Example of number of SVs for each block of a small image.

The result of the learning gives us 7052 support vectors of 9604 pixels in total (see Fig. 2) [5]. These points are not null (as shown in Table 2) and are used for image compression because they are the representative points in the image.

4.3. Results

For image compression, we must go through these steps (see Fig. 3):

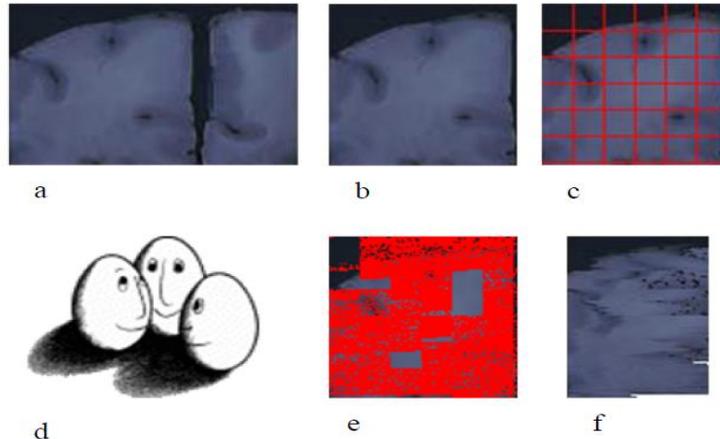


Fig. 3. Experimentation steps: a) Original image; b) Sample image; c) Blocks image; d) SVM learning; e) SVs image; f) Compressed image.

The parameters chosen for the polynomial function (1) are the parameters which have a minimized learning error and which have a maximized classification ratio as presented in support vector machines for classification (3). So, the polynomial degree (d) value is 3 and the radial gamma (σ) value is also 3 (3).

We choose for our experiment two cases. In all cases, the image is reconstructed by line using support vectors (5) generated by training. The compression ratio (see equation 3) is calculated using image size saved on disk with this equation:

$$\text{Compression ratio} = \frac{\text{compressed image size}}{\text{sample image size}} * 100 \tag{3}$$

4.3.1. Case 1

The sample used is Astrocytoma image, fibrillary subtype with magnification of 40 of 221x221 pixels for image. The image (see Fig. 4.) size is 143 KB.

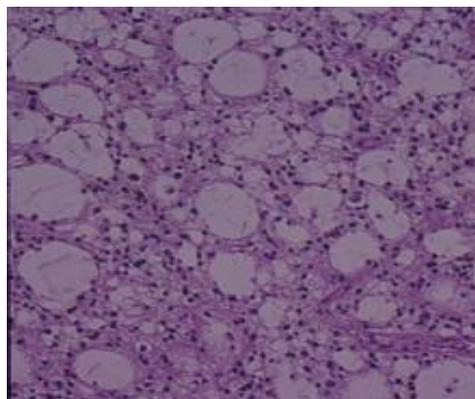


Fig. 4. Sample image in case 1.

The size of the compressed image (as presented in Fig. 5.) is 54.4 KB. The compression ratio calculated for the image with the polynomial function is equal to 38.04 %.



Fig. 5. Compressed image in case 1 with polynomial function.

The size of the compressed image (as presented in Fig. 6.) is 0.91 KB. The compression ratio calculated for the image with the radial function is equal to 0.63%.



Fig. 6. Compressed image in case 1 with radial function.

4.3.2. Case 2

The sample used is Astrocytoma image, fibrillary subtype with a magnification of 100 of 229 x 229 pixels in image. The image (see Fig. 7.) size is 153 KB.

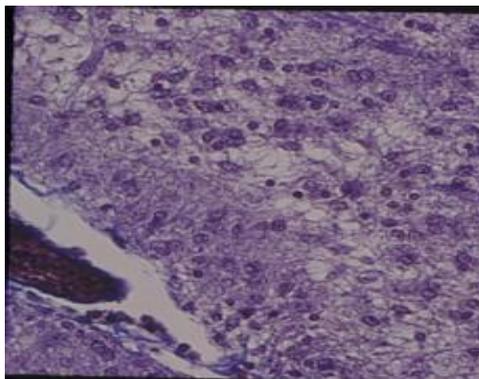


Fig. 7. Sample image in case 2.

The size of the compressed image (as presented in Fig. 8.) is 50.1 KB. The compression ratio calculated for average image with the polynomial function is equal to 32.74%.

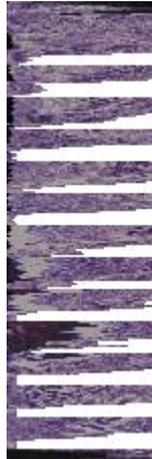


Fig. 8. Compressed image in case 2 with polynomial function.

The size of the compressed image (as presented in Fig. 9.) is 0.94 KB. The compression ratio calculated for the image with the radial function is equal to 0.61%.



Fig. 9. Compressed image in case 2 with radial function.

4.4. Discussion

Table 3. Comparison between Compression Ratios in the Two Cases

Case	Sample size	Kernel	Compressed size	Compression rate
Case 1	143 KB	Polynomial	54.4 KB	38.04 %
		Radial	0.91 KB	0.63%
Case 2	153 KB	Polynomial	50.1 KB	32.75%
		Radial	0.94 KB	0.61%

From these experiments (as shown in Table 3), we notice that:

- Two images having the same texture have the same compression ratio such as case 1 and case 2.
- The better compression ratio is in the two cases.

- The compressed image is different from the sample image because many pixels which are not support vectors were lost.
- When we compare between the polynomial kernel and radial kernel, we can say that radial function gives a better compression ratio compared to polynomial function but point of view visualization, we lose a lot of information with radial function.

5. Conclusion

In this article, we proposed the application of support vector machines applied to compression. All images are divided into several blocks to facilitate compression of medical images.

The proposed approach achieves good results of compression ratio and reduces the amount of redundant data in medical images but the compression is often reserved for authors of the applications.

However, our fast decompression step will be re-served for the end user.

Acknowledgment

I would like to offer my special thanks to Professor Mohamed Benyettou for having me accept in his modeling laboratory of industrial systems (LAMOSI) of University of Science and Technology of Oran (USTO) to carry out this research work which I am very grateful. I would like to express my very great appreciation to Professor Abderraouf Benali of Multidisciplinary Research Laboratory in Systems Engineering, Mechanical, Energy (PRISME) of National Institute of Applied Sciences Val de Loire (INSA) that pointed me in the writing of this article. Advice given by Doctor Eric Nichols has been a great help in the English language.

References

- [1] Seret, A., & Hoebeke, M. (2006). *Imagerie Médicale, Bases Physiques*. Belgique: ULG.
- [2] Brisset, F. (2008). Microscopie électrique à balayage et microanalyses. *EDP Sciences*.
- [3] Bentaouza, C. M., & Benyettou, M. (2010). Support vector machines for brain tumours cells classification. *Journal of Applied Sciences, 10*, 1755-1761.
- [4] Dumas, J., Roch, J., Tannier, E., & Varrette, S. (2013). *Théorie des Codes - 2e éd.: Compression, cryptage, correction*. Dunod.
- [5] Bentaouza, C. M., & Benyettou, M. (2014). Support vector machines for microscopic medical images compression. *Pakistan Journal of Biological Sciences, 17*, 335-345.
- [6] Sayood, K. (2000). *Introduction to Data Compression* (2nd ed.). USA: Morgan Kaufmann.
- [7] Vapnik, V. (1998). *Statistical Learning Theory*. New York: John Wiley and Sons.
- [8] Preux, P. (2011). *Fouille de Données : Notes de Cours*. Université de Lille 3.
- [9] Brucher, J. M., Chatel, M., Darcel, F., & Theaux, R. (1999). *Association of French Speaking Neuro-Oncologists (ANOCEF's)*. Atlas of Neuro-Oncologie.
- [10] Joachims, T. (1999). Making large-scale svm learning practical. In B. Scholkopf, C. J. C. Burges, & A. J. Smola (Eds.), *Advances in Kernel Methods-Support Vector Learning*. (pp. 169-184). USA: Cambridge.



Chahinez Mérièm Bentaouza was born in Oran, Algeria. She graduated from university in Oran.

She is interested in computer science and has received the Ph.D at the University of Science and Technology of Oran. She is teaching at the University of Mostaganem.

Chahinez Mérièm Bentaouza has made several short-stay placement for improved scientific research in the field of machine learning and computer security at the INT, INRIA

and INSA.



Mohamed Benyettou was born in Algeria. He studied at the University of Science and Technology of Oran.

He is a research teacher rank of professor in mathematics.

Mohamed Benyettou has held positions as dean of the Faculty of Science and he is still the head of LAMOSI laboratory.