Pakistan
Journal of Biological Sciences

Support Vector Machines for Microscopic Medical Images Compression

[1,2]Chahinez Mérièm Bentaouza and [1]Mohamed Benyettou
[1]Department of Computer Science, Modeling and Optimization of Industrial Systems Laboratory,
Faculty of Science, University of Science and Technology of Oran, P.O. Box 1505, El Menaour,
Oran, Algeria
[2]Department of Computer Science, Faculty of Science and Computer Science,
University of Mostaganem, B.P. 188, Mostaganem, Algeria

Abstract: This study presents the compression of microscopic medical images by Support Vector Machines using machine learning. The visual cortex is the largest system in the human brain and is responsible for image processing such as compression, because the eye does not necessarily perceive all the details of an image. Medical images are a valuable means of decision support. However, they provide a large number of images per examination that can be transmitted over a network or stored for several years under the law imposed by the country. To apply the reasoning of biological intelligence, this study uses Support Vector Machines for compression to reduce the pixels of medical images in order to transmit data in less time and store information in less space. The results found by using this method are satisfactory for compression though the time must be improved.

Key words: Archiving, compression, machine learning, microscopic medical image, support vector machines

INTRODUCTION

The human visual system is more very sensitive to changes but the removal of a significant portion of the information is only slightly visible and therefore not noticeable to the eye (Herellier, 1993). Today, digital images have about a million pixels, with each pixel requiring three bytes for RGB (Red, Green, Blue) images. Without compression, this is 3MB for a single image (Chirollet, 2005).

In addition, 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 (Seret and Hoebeke, 2006) and the resulting images of a microscope of up to 1GB for examination (Brisset, 2008). In all, the annual production of a medical imaging center is estimated at several hundred or even thousands of GB. So, storage for these types of files on disks and the transfer will also take considerable time (Dillenseger and Moerschel, 2009).

Storage should be done for short-term use in order to respect the laws forcing hospitals to archive medical records of patients for several years. Consequently, image compression is the only solution for archiving despite the increase in storage capacity (Hyvarinen et al., 2008).

Many researches have focused on compression by statistical learning methods or machine learning methods (Hu and Hwang, 2002). The Compression of digital images is consequently one of the most active areas research.

Saad et al. (1998) Mahmoud presented an Artificial Neural Network (ANN) and Vector Quantization for image compression. Better reconstructed images can be obtained if a variety of different activity images are used in the design phase of the proposed neural system (1993). Similarly, the Principal Component Analysis (PCA) technique and (ANN) Multi Layer Perceptron (MLP) architecture for compression of Lena image was proposed. The result shows the good quality of reconstructed image (Gaidhane et al., 2010). Also, compression was made by the Fourier Transform (FT). It can be stated that the use of FT is an important tool which offers an improvement in image compression (Singh et al., 2009). Forchhammer et al. (2011) proposed Moving Picture Experts Group (MPEG) coded images and Discrete Cosine Transform (DCT) (2011).

The growth of medical imaging involves storage and transmission of large amounts of digital information. It is possible to reduce the data amount with caution so that the result is very similar to the original, if not identical, to the human eye without losing information so as to not distort the medical diagnosis.

Corresponding Author: Chahinez Mérièm Bentaouza, Department of Computer Science, Modeling and Optimization of Industrial Systems Laboratory, Faculty of Science, University of Science and Technology of Oran, P.O. Box 1505, El Menaour, Oran, Algeria

For medical imaging, an ANN MLP architecture was applied to digital X-ray image compression with a high reconstruction quality (Steudel *et al.* 1995). The nonlinear Principal Component Analysis (PCA) technique shows the best compression results in combination with the ANN MLP of the mammographic image (Meyer-Baese 2000). According to Ziguang *et al.*, the Binary Decision Diagrams compression technique performs better than WINZIP in terms of compression ratio though the execution time is rather high (2009). The authors present Coded Excitation systems Ultrasound Medical Imaging which improves the Signal-to-Noise Ratio (SNR) and resolution (Toosi and Behnam, 2009). Alternatively, the Discrete Cosine Transform (DCT) provides medical images of the brain reconstructed in high-fidelity compression (Al-Fayadh *et al.*, 2009). Robinson and Kecman, 2003 deduced that Combining Support Vector Machine (SVM) with the Discrete Cosine Transform (DCT) produces better image quality than the JPEG compression algorithm (2003).

Finally, although all the proposed approaches are efficient , they are not sufficient to achieve very good compression. In this study, we propose Separator with Vast Margin (SVM) for performance in machine learning and classification (Sahbi and Geman, 2006; Chiou *et al.*, 2009; Chen, 2009; Bentaouza and Benyettou, 2010) applied to microscopic image of brain tumor because these images contain a lot of redundant information.

The objective of the study is to find out that SVM is the most frequently used method and can be applied to compression of microscopic images.

MATERIALS AND METHODS

Image compression: Compression is the action used to reduce the physical size of information block. By compressing the data, we can place more information in the same storage space, or use less time to transfer through a network. Since the images require high capacity, compression becomes an essential part that can reduce the number of pixels by exploiting the informational redundancy in the image (Incerti, 2003).

Physical and logical compression: The physical compression acts directly on the data and it is to watch a redundant data bit stream to another. But the logic compression is performed by substituting information in a different representation (Incerti, 2003). In this article, we treat the logical compression.

Symmetric and asymmetric compression: A symmetric compression method uses the same algorithm and requires the same computing capacity, both for compression than for decompression. Asymmetric compression methods require more work in one direction than another. Normally, the compression step requires much more time and system resources than the decompression stage. Incerti (2003) We can't discuss of this point because we are interested by the compression and the decompression is our perspectives.

Lossy and lossless compression: This is the most important criterion for comparing compression algorithms.

Lossy compression methods make it possible to find an approximation of the image. The data are removed and cannot be to recover. The losses are generally undetectable with the eye but the initial image is not identical to the decompressed version. These algorithms are applicable to data with high levels of redundancy, such as images, or sounds. Joint Photographic Expert Group (JPEG) and Moving Picture Experts Group (MPEG) are lossy compression algorithms.

Lossless compression methods make it possible to find exactly the pixels of the original image. Thus, after decompression, the image is rebuilt perfectly and identical to the initial image. Much of the compression schemes used are lossless compression algorithms. Shannon-Fano, Huffman and Run-Length Encoding (RLE) are lossless compression algorithm.

In this research, we can't tolerate the compression with information loss because we have sensitive images which are the important medical information for their archiving.

CRITERION OF COMPRESSION

Compression quality: Peak Signal to Noise Ratio (PSNR) measures the relationship between information and noise in an image. It is expressed with a logarithmic scale in decibels (dB) because it can take a very large range of values:

$$PSNR = 10 \times \log \frac{255^2}{EQM} dB \qquad (1)$$

EQM is the mean square error calculated between the original image and the reconstructed image:

$$EQM = \frac{1}{n \times \sum_i \sum_j (X_{ij} - Y_{ij})^2} \qquad (2)$$

where, n is the number of pixels in the image, X_{ij} is the value of the pixel (i, j) in the original image and Y_{ij} is the value of the pixel (i, j) in the reconstructed image. The PSNR gives only a rough idea of the quality of the

reconstructed image after decompression, does not replace the visual assessment of the image. This part is not used in our article because we are interested by the compression.

Compression ratio: The compression ratio is used to refer to the ratio between the uncompressed data size and the compressed data size. There are several ways to calculate the compression ratio:

- Bit rate:

$$D = \frac{\text{Byte No. of compressed imgae}}{\text{Pixle No. of original imgae}} \text{byte pixle}^{-1} \qquad (3)$$

- Compression ratio:

$$Q = \frac{\text{Initial size}}{\text{Final size}} \qquad (4)$$

- Compression rate:

$$T = \frac{\text{Final size}}{\text{Initial size}} = 1/Q\% \qquad (5)$$

- Gain compression:

$$G = 1\text{-}T \qquad (6)$$

We used to evaluate compression of medical images by SVM using rate compression as seen in Eq. 5.

Support vector machines: The support vector machine is a new method for classification of data. This technique of supervised classification suggested in 1992 by Vladimir VAPNIK (Vapnik, 1998) results from the statistical theory of the training. Now, it is an active zone of search for machine learning (Bentaouza and Benyettou, 2010). The fundamental idea is to build an optimal learning machine for better generalization, having:

- A definite surface in an optimal way
- In general non linear in the space of entry
- Linear in a dimensional space more raised
- Implicitly defined by a kernel function

Binary SVM: The principal idea of the binary SVM is to implicitly trace data with a dimensional space more raised by the intermediary of a kernel function and to solve a problem of optimization to identify the hyperplane whose margin is maximum which separates the examples from training (Vapnik, 1998). The hyperplane is based on the support vector i.e., a whole of examples of training being on the margin. New examples are classified according to the side of the hyperplane. The problem of optimization generally is formulating in a manner and takes account of the inseparable data while penalizing distort them classifications.

Separable linear case: The solution of the problem of optimization (Burges, 1998) is as follows:

- Result of the stage of training: The optimal hyperplane is:

$$w.x+b = 0 \qquad (7)$$

$$w = \sum_{i=1}^{n} \alpha_i x_i y_i \qquad (8)$$

$$b = \frac{1}{2n} \sum_{i=1}^{n} (w.x_i^+ + w.x_j^-) \qquad (9)$$

α_i Correspond to the closest points, i.e., the support vectors that are no null. They are found by solving the following equation:

$$\begin{cases} \max_{\alpha_i} W(\alpha_i) = -\sum_{i=1}^{n} \alpha_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i.x_j) \\ \forall i \, \alpha_i \geq 0 \\ \sum_{i=1}^{n} \alpha_i y_i = 0 \end{cases} \qquad (10)$$

- Result of the stage of test: The function of decision is obtained by:

$$g(x) = \text{sgn}(\sum_{i=1}^{n} \alpha_i y_i (x_i.x) + b) \qquad (11)$$

Non separable linear case: The algorithm established for the separable linear case cannot be used in many real problems. In general, of the noisy data will make separation linear impossible. For this there, the approach above can be extended to find a hyperplane that minimizes the error on the training example (Training error). This approach also indicated under the name of "soft margin hyperplane" (Vapnik, 1995). We introduce the "slack variable" ξ_i the error of the example x_i with $\xi_i>0$, $i = 1,...,$ n. The value of ξ_i indicates, where are x_i before a hyperplane of separation:

- If $\xi_i \geq 1$ then x_i is badly classified
- If $0<\xi_i<1$ then x_i is correctly classified but is inside margin
- If $\xi_i = 0$ then x_i is correctly classified and is outside the margin or with the band of the margin

Alternatively, the expression of the function changes from:

$$\frac{1}{2}\|w\|^2 \text{ to } \frac{1}{2}\|w\|^2 + C\sum_{i=0}^{n} \xi_i^k$$

where, C is the weight of error, it is a parameter to be chosen by the user, larger C corresponding to assign a higher penalty with the errors, is generally taken 1/n.

Non linear case: How the methods above can be generalized if the function of decision is not a linear function of the data? In the linear case space used is Euclidean space. Separation by hyperplanes is a relatively poor framework of classification, bus for much of case; the examples which one wants to separate are not separable by hyperplanes. To continue to profit from the theory which justifies the search for hyperplanes of maximum margin while using more complex surfaces of separation (nonlinear), VAPNIK (Boser *et al.*, 1992) suggests separating not items x_i but of representatives this points $X_i = \varphi(x_i)$ in a space F of dimension much larger, than VAPNIK calls: feature space. One gives oneself a nonlinear function noted φ, such as $\varphi \colon R^m \to F$. The separation nonlinear of maximum margin consists in separating the points $X_i = \varphi(x_i)$, as the function of decision is expressed exclusively using scalar products is thus, one in practice never uses the function directly but only the kernel K is defined by: $K(x, x') = \varphi(x).\varphi(x')$ it is a nonlinear scalar product. Therefore, the problem is solved; it is enough to make a transformation: $x_i.x_j \to \varphi(x_i).\varphi(x_j)$ in all the formulas of hard margin by using a type of following kernel (Cornuejols, 2002) function:

- Polynomial function:

$$k(x, y) = (x.y+1)^d \tag{12}$$

- Radial function:

$$k(x, y) = \exp\frac{(x-y)^2}{\sigma^2} \tag{13}$$

- Sigmoid function:

$$k(x, y) = \tanh(s\, x.y + c) \tag{14}$$

The solution of optimization problem in the non-linear case is:

- Result of the stage of training: The optimal hyperplane is:

$$w.\varphi(x)+b = 0 \tag{15}$$

$$w = \sum_{i=1}^{n} \alpha_i \varphi(x_i) y_i \tag{16}$$

$$b = \frac{1}{2n}\sum_{i=1}^{n}(w.\varphi(x_i^{+}) + w.\varphi(x_j^{-})) \tag{17}$$

α_i is found by solving the following equation:

$$\begin{cases} \max_{\alpha_i} W(\alpha_i) = -\sum_{i=1}^{n}\alpha_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j(\varphi(x_i).\varphi(x_j)) \\ \forall i\, 0 \le \alpha_i \le C \\ \sum_{i=1}^{n}\alpha_i y_i = 0 \end{cases} \tag{18}$$

- Result of the stage of test: The function of decision is obtained by:

$$g(x) = \operatorname{sgn}\left(\sum_{i=1}^{n}\alpha_i y_i(\varphi(x_i).\varphi(x)) + b\right) \tag{19}$$

Multiclass SVM: There are several methods for the multiclass (Platt, 1999) but in this article, we tried out One Versus Rest (OVR) which is implemented in multiclass SVM. OVR is conceptually the method of simplest SVM multiclass. Here we build of the binary classifiers of K SVM: classify 1 positive against all other negative classes, classifies 2 against all other classes..., class K against all other classes. The combined function of decision of OVR chooses the class of a sample which corresponds to the maximum value of the binary functions of decision of K indicated by the other positive hyperplane (Statnikov *et al.*, 2005). While thus making, hyperplanes of decision calculated by K SVM which call into question the optimality of multiclass classification.

RESULTS

Configuration: The data base is extracted from classification of OMS (Brucher *et al.*, 1999); the ranks are like degrees of malignity, indicating the biological behavior. In the atlas used, the various tumor entities are also grouped according to their rank which corresponds to the images taken to make our compression by SVM.

Images used are Astrocytoma of grade 2, tumors of the central nervous system, whose growth is slow on different categories according to their main localization.

The machine used to carry out these experimentation is a PC (Pentium 4, 2.80 GHz with 2 GO of RAM).

And concerning the programming two essential tools were used:

- C++ Builder 6 under Windows Virtual PC to create a graphic interface summer developed on the one hand for the pretreatment of the classification of the medical images which consists in extracting significant information in RGB (Red, Green and Blue) which will make the object of the supervised training, so, the input of the Multi-classes SVM
- SVM Multiclass (OVR) under Windows Virtual PC is a free tool (Joachims, 1999) which was developed for the scientific research whose problem of optimization was included

Pretreatment: We have several pictures of different types and sizes where each image is divided into several blocks of 16×16 pixels, this block will be grouping identical pixels in a class and then we'll start learning about the block using SVMs because it is easier to make learning as a block to learning a whole image.

The result of learning gives us the support vectors and it is these points that there will be compression.

The compression rate for each block is calculated in this form:

$$\text{Compression rate} = \frac{\text{No. of SVs in block}}{\text{No. of pixels in block}} \qquad (20)$$

The total compression rate for sample image is calculated in this form:

$$\text{Compression rate} = \frac{\text{No. of SVs}}{\text{No. of pixels}} \times 100 \qquad (21)$$

The number of pixels in sample image is calculated in this form:

$$\text{No. of pixels in sample image} = \text{Height} \times \text{Width} \qquad (22)$$

The parameters chosen for the polynomial function and the radial function are parameters which minimize the learning error and maximize the classification rate as presented in the article (Bentaouza and Benyettou, 2010).

EXPERIMENTATION

Case 1: The sample used is Astrocytoma image with frontal section of 98×98 pixels for small image and 230×230 pixels for average image (Table 1).

The result obtained after the learning of all the blocks with the SVM on a small image is presented on figures (Fig. 1-4).

The result obtained after the learning of all the blocks with the SVM on an average image is presented on figures (Fig. 5-8).

Table 1: Features of a small and an average image in case 1

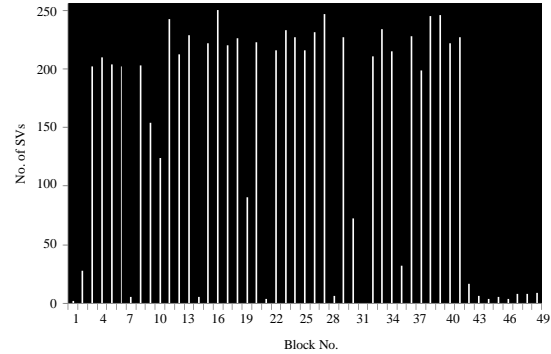|	Small image	Average image
No. of pixels	9604	52900
No. of classes	2258	9761

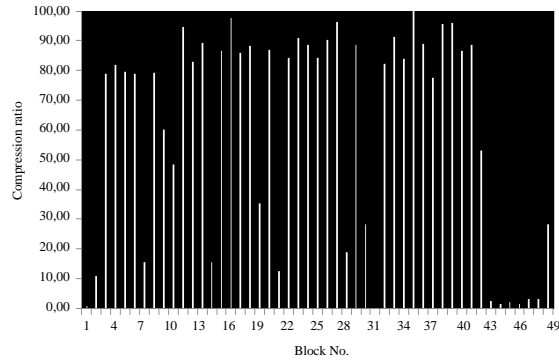Fig. 1: Number of SVs for each block of a small image in case 1

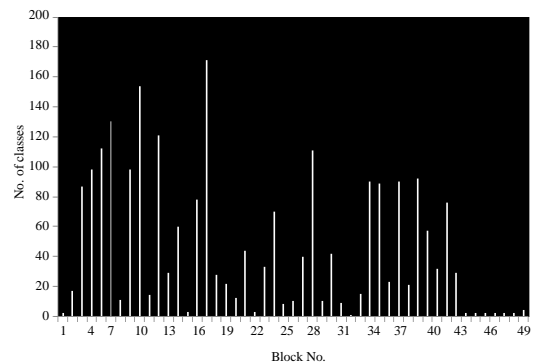Fig. 2: Compression ratio for each block of a small image in case 1

Fig. 3: Number of classes for each block of a small image in case 1

Note that in this case, the compression rate of the same image with different size is almost the same as it is
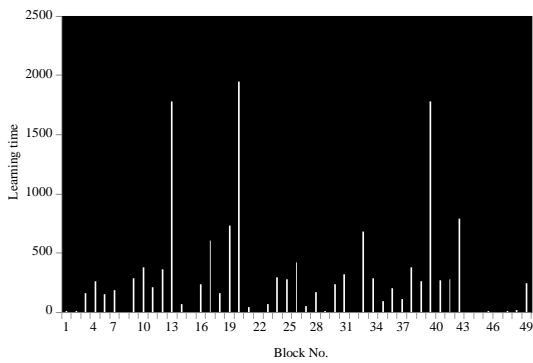
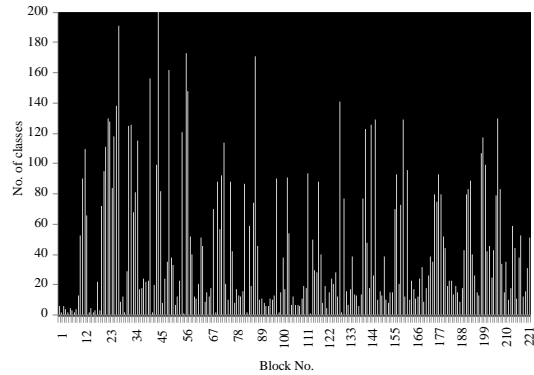Fig. 4: Learning time for each block of a small image in case 1

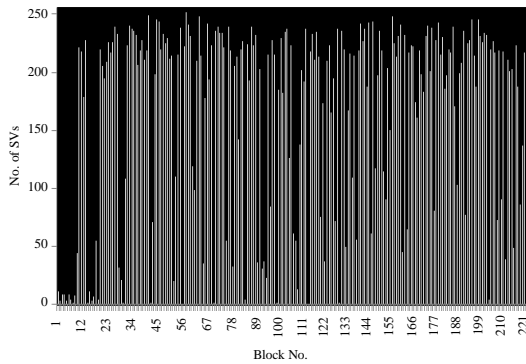Fig. 5: Number of SVs for each block of a average image in case 1

Fig. 6: Compression ratio for each block of an average image in case 1

Fig. 7: Number of classes for each block of an average image in case 1

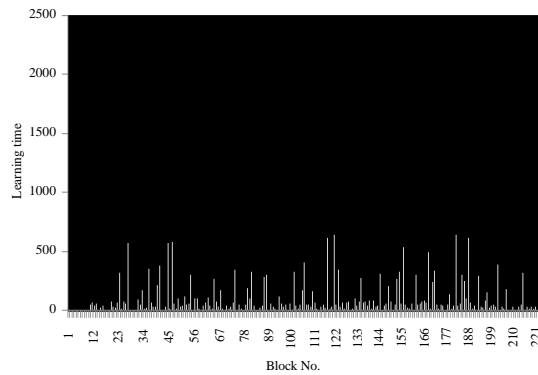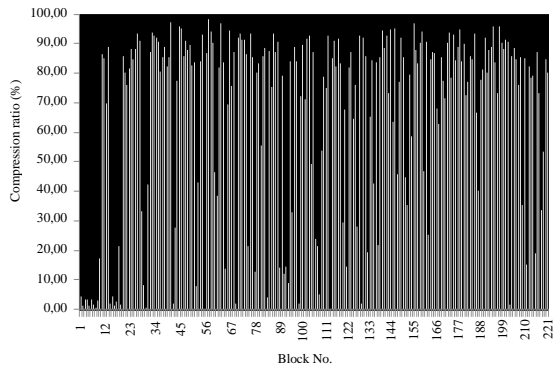Fig. 8: Learning time for each block of an average image in case 1

Table 2: Compression results for a small image in case 1

	Polynomial kernel	Radial kernel
No. of SVs	7052	37
Compression rate	73,42%	0,003%
Learning time	14864	22

Table 3: Compression results for an average image in case 1

	Polynomial kernel	Radial kernel
No. of SVs	37443	197
Compression rate	70,78%	0,003%
Learning time	21159	142

Table 4: Features of a small and an average image in case 2

	Small image	Average image
No. of pixels	9604	52441
No. of classes	486	2465

for the polynomial function and for the radial function. But it took a long time for the polynomial function but a less time for the radial function (Table 2 and 3).

Also, more than the number of pixels increases, more than the learning time is high because we have more blocks including more classes to learn. And if we make a forecast for a third image pixels larger than the second, we can expect to have the same compression rate with a time of higher learning.

Case 2: The sample used is Astrocytoma image, fibrillary subtype with magnification of 100 of 98×98 pixels for small image and 229×229 pixels for average image (Table 4).
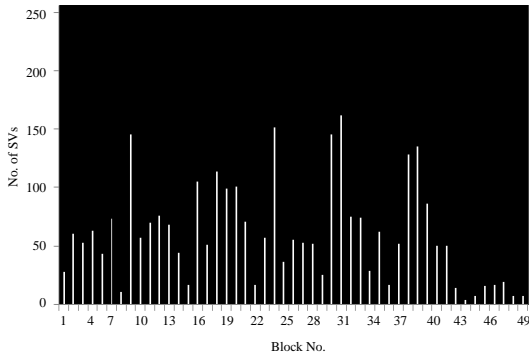
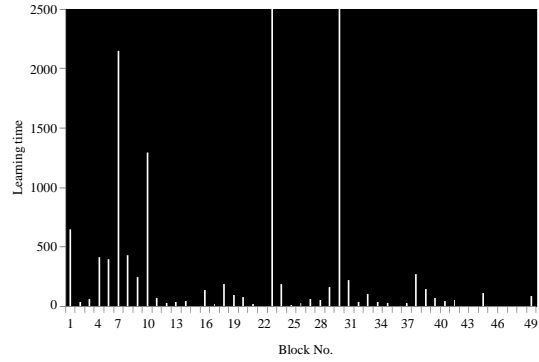Fig. 9: Number of SVs for each block of a small image in case 2

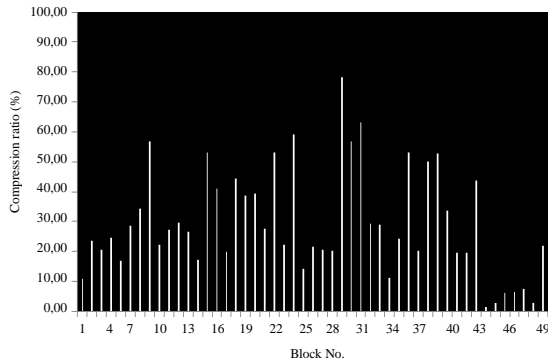Fig. 12: Learning time for each block of a small image in case 2

Fig. 10: Compression ratio for each block of a small image in case 2
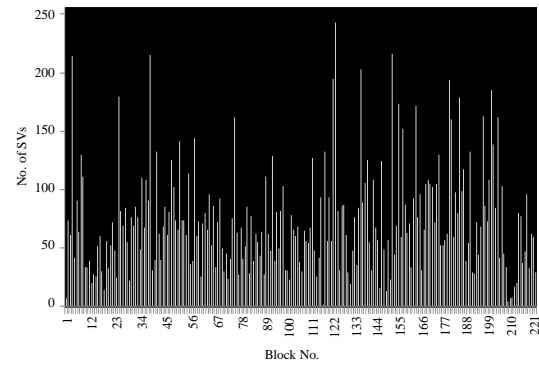

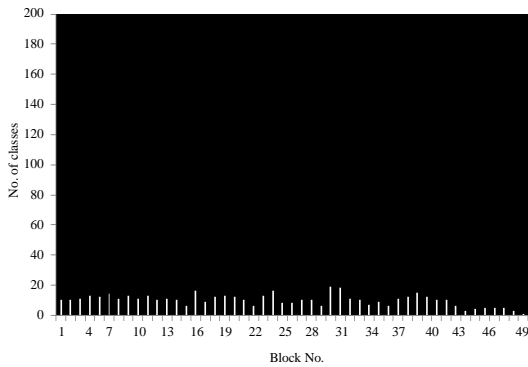Fig. 13: Number of SVs for each block of an average image in case 2

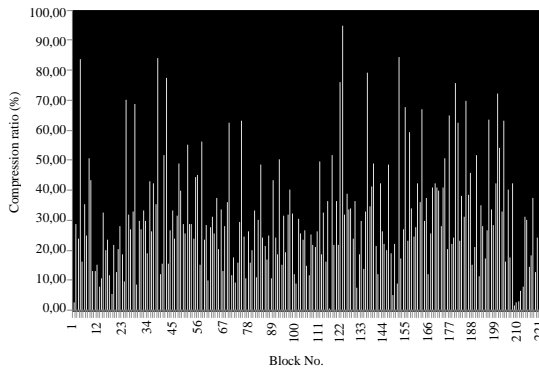Fig. 11: Number of classes for each block of a small image in case 2

Fig. 14: Compression ratio for each block of a small image in case 2

The result obtained after the learning of all the blocks with the SVM on a small image is presented on figures (Fig. 9-12).

The result obtained after the learning of all the blocks with the SVM on an average image is presented on figures (Fig. 13-16).

This case is identical to the first case point of view comparison between the polynomial function and the radial function except the difference is that the compression rate is smaller compared to the first case. Also learning time increased if the pixels is larger as the first case (Table 5 and 6).
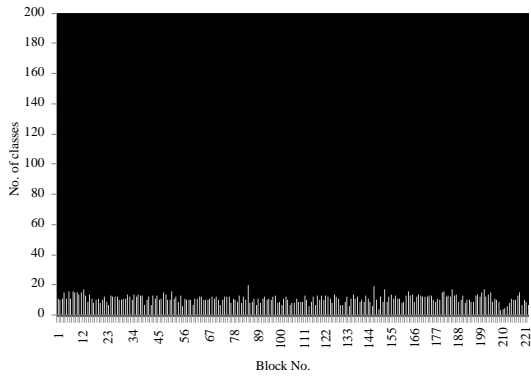
Fig. 15: Number of classes for each block of an average image in case 2
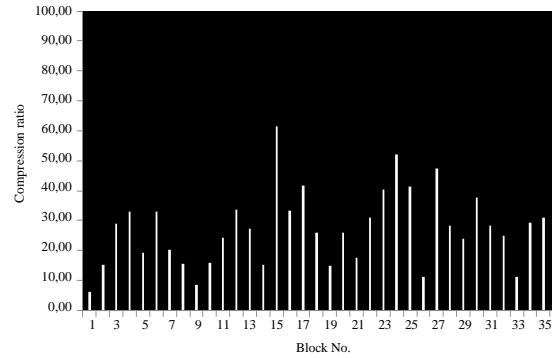

Fig. 18: Compression ratio for each block of a small image in case 3
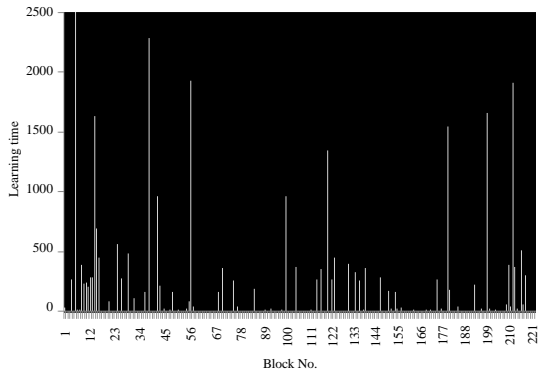

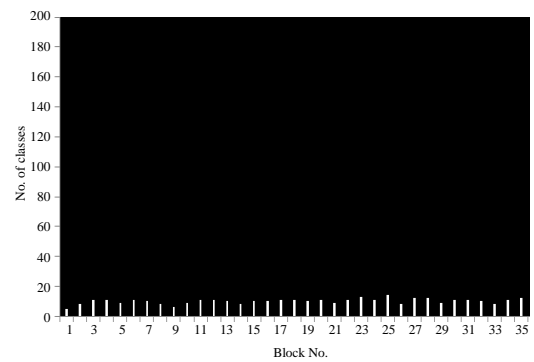Fig. 16: Learning time for each block of an average image in case 2

Fig. 19: Number of classes for each block of a small image in case 3
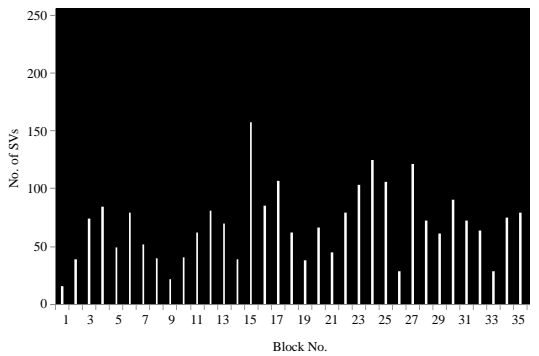

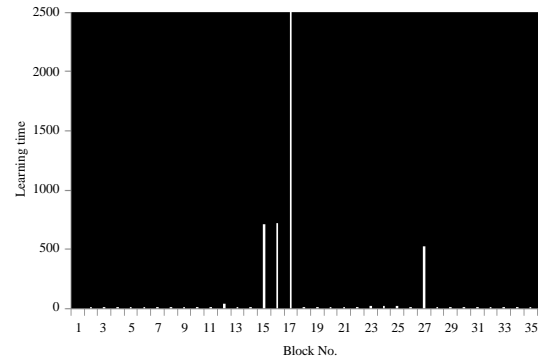Fig. 17: Number of SVs for each block of a small image in case 3

Fig. 20: Learning time for each block of a small image in case 3

We notice despite the number of pixels of the first image is identical to the number of pixels of the second image, the number of Svs found is higher in the first image that in the second image because the number of classes learned is higher in the first case.

Case 3: The sample used is Astrocytoma image, fibrillary subtype with magnification of 40 of 95×95 pixels for small image and 221×221 pixels for average image (Table 7).

The result obtained after the learning of all the blocks with the SVM on an small image is presented on figures (Fig. 17-20).
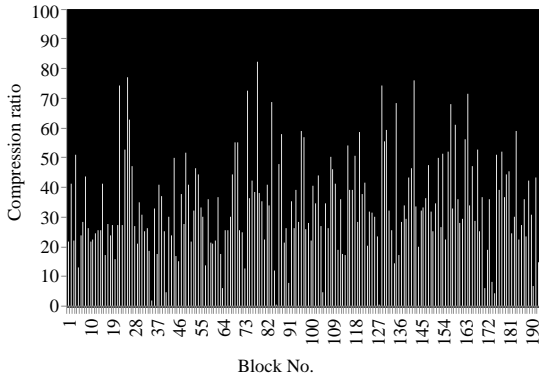
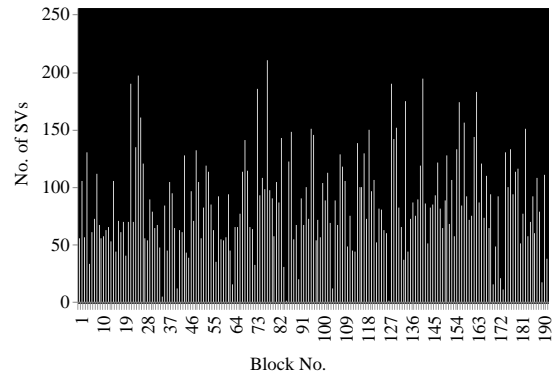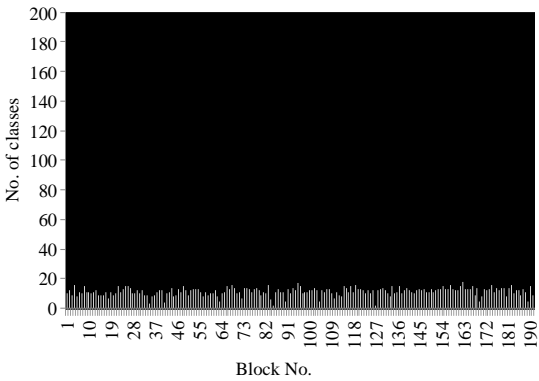Fig. 21: Number of SVs for each block of an average image in case 3

Fig. 22: Compression time for each block of an average image in case 3
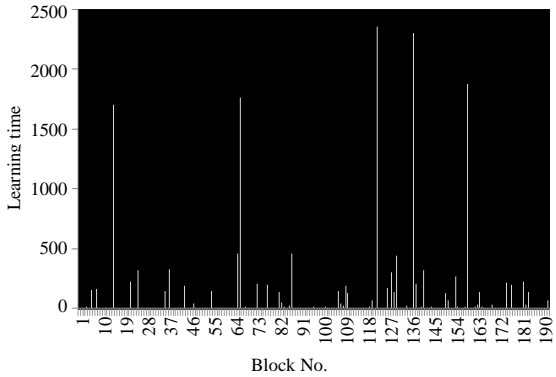

Fig. 23: Number of classes for each block of an average image in case 3

The result obtained after the learning of all the blocks with the SVM on a average image is presented on figures (Fig. 21-24).

The results found using the image resembles the results generated by the second image point of view number of SVs, compression rate and learning time

Fig. 24: Learning time for each block of an average image in case 3

Table 5: Compression results for a small image in case 2

	Polynomial kernel	Radial kernel
No. of SVs	2951	37
Compression rate	30.72%	0.003%
Learning time	13488	22

Table 6: Compression results for an average image in case 2

	Polynomial kernel	Radial kernel
No. of SVs	16425	197
Compression rate	31.32%	0.003%
Learning time	30260	142

Table 7: Features of a small and an average image in case 3

	Small image	Average image
No. of pixels	9025	48841
No. of classes	364	2236

Table 8: Compression results for a small image in case 3

	Polynomial kernel	Radial kernel
No. of SVs	2522	37
Compression rate	27.94%	0.003%
Learning time	10839	22

Table 9: Compression results for an average image in case 3

	Polynomial kernel	Radial kernel
No. of SVs	16825	197
Compression rate	34.44%	0.003%
Learning time	17732	142

because the number of classes is similar and the pixels in the same block are homogeneous, so we get fewer classes than in the first case (Table 8 and 9).

DISCUSSION

We note that as the number of SVs decreases, we obtain a satisfactory compression rate except when using the radial function as the rate is too small and can cause great loss of information especially for medical images such loss is not permitted.

This number of SVs depends on the number of classes learned by block, so that we have more classes, more we have a high number of SVs and therefore the compression rate increases.

And if we make a forecast for a third image number of pixels larger than the second image, we can expect to have the same compression rate with a time of higher learning because we have more classes.

We compare the rates found by the different images, we can say that the optimal compression with large margin separator up to 3:1 with the use of the polynomial function with a period that exceeds three hours for the compression of a single sample image.

Therefore, we can tolerate this compression time as long as the compression will be done only once when the consultation will occur several times so we can't allow this time to decompress.

NB: Much research has been done on image compression by different methods with lossy and lossless even using methods from artificial intelligence but this article presented efficient method.

In addition, the introduction discusses the state of the art of all that has been done before. Especially the article (Robinson and Kecman, 2003) but we can't compare the results found with respect to that section because the scope is different and our research treats sensitive images.

CONCLUSION

SVM is one of classification models to score the pattern recognition by providing a theoretical framework for statistical learning theory.

The purpose of this article stems from the need to test the SVM compression. In particular, the need to say whether SVMs are made for compression as does the human being.

Point of view compression rate, we can say that SVM have proven again their success against the compression of brain tumor cells.

We suggest some research avenues such as the reconstruction of the original image through the SVs generated by learning, optimization of time learning and testing the SVMs on other medical images.

ACKNOWLEDGMENT

REFERENCES

Al-Fayadh, A., A.J. Hussain, P.J.G. Lisboa, D. Al-Jumeily and M. Al-Jumaily, 2009. A hybrid image compression method and its application to medical images. Proceedings of the 2nd IEEE International Conference on Developments in Esystems Engineering, December 14-16, 2009, Abu Dhabi, pp: 237-240.

Bentaouza, C.M. and M. Benyettou, 2010. Support vector machines for brain tumours cells classification. J. Applied Sci., 10: 1755-1761.

Boser, B.E., I.M. Guyon and V.N. Vapnik, 1992. A training algorithm for optimal margin classifiers. Proceedings of the 5th Annual Workshop on Computational Learning Theory, July 27-29, 1992, Pittsburgh, Pennsylvania, USA., pp: 144-152.

Brisset, F., 2008. Microscopie Electrique a Balayage et Microanalyses. [Microscopy and Scanning Electron Microanalysis]. EDP Sciences, France, ISBN: 978-2-7598-0082-7.

Brucher, J.M., M. Chatel, F. Darcel and R. Theaux, 1999. Association of french speaking neuro-oncologists (ANOCEF's). Atlas of Neuro-Oncologie. http://www.anocef.org/atlas/index.html.

Burges, C.J.C., 1998. A tutorial on support vector machines for pattern recognition. Data Min. Knowledge Discovery, 2: 121-167.

Chen, C.J., 2009. 3D breast tumor classification using image registration framework. Proceedings of the IEEE International Conference on Future Computer and Communication, April 3-5, 2009, Kuala Lumpar, pp: 153-157.

Chiou, Y.J., H.M. Chen, J.W.W. Chai, C.C.C. Chen and Y.C. Ouyang *et al.*, 2009. Brain tissue classification using independent vector analysis (IVA) for magnetic resonance image. Proceedings of the 9th IEEE International Conference on Bioinformatics and Bioengineering, June 22-24, 2009, Taichung, pp: 324-329.

Chirollet, J.C., 2005. Numeriser, Reproduire, Archiver Les Images D'art.[Scan, Copy, Archive Images of Art]. L'Harmattan, Paris, ISBN: 2-7475-8724-X.

Cornuejols, A., 2002. [A new method of learning: SVM, separators wide margin]. Bulletin de L'afia, 51: 14-23.

Dillenseger, J.P. and E. Moerschel, 2009. Guide des Technologies de L'imagerie Medicale et de la Radiotherapie: Quand la Theorie Eclaire la Pratique. [Guide Technologies of Medical Imaging and Radiotherapy: When Theory Informs Practice]. Elsevier Masson, Issy-Les-Moulineaux, ISBN: 978-2-294-70431-4.

Forchhammer, S., H. Li and J.D. Andersen, 2011. No-reference analysis of decoded MPEG images for PSNR estimation and postprocessing. J. Visual Commun. Image Represent., 22: 313-324.

Gaidhane, V., V. Singh and M. Kumar, 2010. Image compression using PCA and improved technique with MLP neural network. Proceedings of the IEEE International Conference on Advances in Recent Technologies in Communication and Computing, October 16-17, 2010, Kottayam, pp: 106-110.

Herellier, J.M., 1993. Le Multimedia. [Multimedia]. Sybex, Paris, ISBN: 9782736112127.

Hu, Y.H. and J.N. Hwang, 2001. Handbook of Neural Network Signal Processing. 1st Edn., CRC Press, ISBN 0-8493-2359-2,2001, USA.

Hyvarinen, A., J. Hurri and P.O. Hoyer, 2008. Natural Image Statistics: A Probabilistic Approach to Early Computational Vision. Springer, London.

Incerti, E., 2003. Compression D'image: Algorithme et Standards. [Image Compression: Algorithms and Standards]. Vuibert, Paris, ISBN-10: 2711748154 .

Joachims, T., 1999. Making Large-Scale SVM Learning Practical. In: Advances in Kernel Methods-Support Vector Learning, Scholkopf, B., C.J.C. Burges and A.J. Smola (Eds.). MIT Press, Cambridge, MA., ISBN-10: 0-262-19416-3, pp: 169-184.

Meyer-Baese, A., 2000. Medical image compression by Neural-gas network and principal component analysis. Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks, Volume 5, July 24-27, 2000, Como, pp: 489-493.

Platt, J., 1999. Fast Training of Support Vector Machines Using Sequential Minimal Optimization Advances in Kernel Methods, Support Vector Learning. 1st Edn., MIT Press, Cambridge, ISBN: 0-262-19416-3.

Robinson, J. and V. Kecman, 2003. Combining support vector machine learning with the discrete cosine transform in image compression. IEEE Trans. Neural Networks, 14: 950-958.

Saad, E.S.M., A.A. Abdelwahab and M.A. Deyab, 1998. Using feed forward multilayer neural network and vector quantization as an image data compression technique. Proceedings of the 3rd IEEE Symposium on Computers and Communications, June 30-July-2, 1998, Athens, pp: 554-558.

Sahbi, H. and D. Geman, 2006. A hierarchy of support vector machines for pattern detection. J. Mach. Learn. Res., 7: 2087-2123.

Seret, A. and M. Hoebeke, 2006. Imagerie Medicale, Bases Physiques. [Medical Imaging, Physical Fundamentals]. ULG, Belgique, ISBN-10: 2-87456-022-7.

Singh, K., N. Singh, P. Kaur and R. Saxena, 2009. Image compression by using fractional transforms. Proceedings of the IEEE International Conference on Advances in Recent Technologies in Communication and Computing, October 27-28, 2009, Kottayam, Kerala, pp: 411-413.

Statnikov, A., F. Aliferis, I. Tsamardinos, D. Hardin and S. Levy, 2005. A comprehensive evaluation of multicategory classification methods for microarray gene expression cancer diagnosis. Bioinformatics, 21: 631-643.

Steudel, A., S. Ortmann and M. Glesner, 1995. Medical image compression with neural nets. Proceedings of the 3rd IEEE International Symposium on Uncertainty Modelling and Analysis, September 17-19, 1995, College Park, MD., pp: 571-576.

Toosi, T.K. and H. Behnam, 2009. Combined pulse compression and adaptive beamforming in coded excitation ultrasound medical imaging. Proceedings of the IEEE International Conference on Signal Processing Systems, May 15-17, 2009, Singapore, pp: 210-214.

Vapnik, V.N., 1995. The Nature of Statistical Learning Theory. 1st Edn., Springer-Verlag, New York, USA.

Vapnik, V.N., 1998. Statistical Learning Theory. 1st Edn., John Wiley and Sons, New York.