1000	1,92 E-05	28,40	1,85	96,3 %

Table 9. Sigmoïd kernel one-vs-one results

Param C	Param. σ	coef	Training CPU time (s)	Test CPU time (s)	Prediction rate
1000	1,92 E-05	1	22,23	1,82	58,6 %
1000	1,92 E-05	0	22,75	1,71	65,2 %
100	1,92 E-05	0	22,42	1,84	68,4 %
1000	1,00 E-06	0	22,81	1,54	97,2 %

Table 10. Linear kernel one-vs-one results

Param. C	Training CPU time (s)	Test CPU time (s)	Prediction rate
10	22,23	1,28	96,9 %
1000	22,14	1,34	96,9 %

6.3 Discussion

The experiments show the effectiveness of our system and particularly the effectiveness of our strategy for selecting hyper-parameters values based on tabu search by scanning a large area of parameters value. This choice has a great influence on the performance of the final classifier, and also on computation time.

We can notice first that the value of the regularization parameter C does not have a large influence on the results, though this parameter is quite critical for other problems where some of the data is not linearly separable even with using kernel functions.

The representation of letters we detailed in section 5.2.4. appears to be efficient. Indeed, even if the corpus contain some letters with many similarities like 'ج', 'ح' and 'خ', the prediction rates we obtained are very satisfying. The additional representation of diacritical points associated with data of distribution matrix has reduced considerably the misclassification rate.

As expected, SVM one-against-one strategy leads to better results than SVM one-against-all even if the CPU time is greater. We also found that the RBF kernel is best suited to the recognition of Arabic manuscripts. Indeed, this kernel gave better results than other kernels and has a recognition rate of 97.0% (for SVM one-against-all, σ=1.83E-05) and a recognition rate equal to 97.9% (for the SVM one-against-one, σ=1.83E-05).

We also notice that linear kernel gives significant results in one-against-one strategy. This kernel is stable, while sigmoid kernel is sensitive to any change of the parameter σ and requires an extensive search for good results. For the polynomial kernel, the parameter coef must be different from zero for best results, and has no importance for the sigmoid kernel. Laplacian kernel gives similar results between the two approaches one-against-one and one-against-all.

It is interesting to notice the improvement of prediction rates compared to those obtained with manual selection of SVM model. This improvement is due to the efficiency of automatic selection based on the local approach combined with tabu search metaheuristic to optimize the parameters of each binary SVM instead of assigning the same value for all binary SVM used in multi-class learning.

Another point is to emphasize and concerns the importance of kernel parameter initialization. Indeed, inadequate values often lead to over-fitting as outlined in [14], or may impede the convergence of SVMs due to a reduced ability of the classifier, particularly when the value of regularization parameter C is set to 1000 (when misclassifications are strongly penalized).

Finally, we note through numerous experiments performed that optimal selection of SVM models leads not only to minimize classification error, but also reduces the classifier complexity through the number of support vectors. The following table shows that the total number of support vectors obtained with automatic selection is about three times smaller than that obtained with manual selection.

Table 11. Comparison of total number of SVs

Kernel	Manual selection	Automatic selection
Polynomial	2432	1076
RBF	1621	679
Laplacian	1799	788
Sigmoïd	1832	845
Linear	2945	1311

7 Concluding Remarks

SVM is one of the machine learning techniques that has the greatest impact on pattern recognition by providing a theoretical framework. In this paper, we propose a system for recognizing handwritten Arabic letters based on SVM. The system was tested on a corpus containing 4840 examples and has given very good results in terms of recognition rate; it shows the effectiveness of the method for the extraction of primitives and the strategy used for SVM multi-class model selection based on tabu search. The system allowed us to make a comparative study of different SVM models used in the recognition of Arabic characters.

Future extensions are possible:

- Use a corpus generated from the segmentation of Arabic words datasets.

- Check other types of features to improve the recognition rate.

Finally, note the need to have a common protocol for validation of results between different approaches in recognition of handwritten Arabic script.

References:

[1] Vapnik V. N., “Statistical learning theory”. Wiley, New York, 1998.
[2] Boser B., Guyon I., and Vapnik V., “A training algorithm for optimal margin classifiers”. In *Fifth Annual Workshop on Computational Learning Theory*, Pittsburg,1992.
[3] Zaïz F., “SVM pour la reconnaissance de caractères manuscrits arabes”, LESIA Laboratory, Computer Science Department, University Mohamed Khider Biskra, Algeria, 2010.
[4] Kadri M., “Méthode de reconnaissance de l’écriture arabe manuscrite en utilisant les réseaux neuronaux”, Master Thesis, USTO MB, Oran, Algeria, 2010.
[5] Bouslimi R., “Système de reconnaissance hors-ligne des mots manuscrits arabes pour multi-scripteurs”, Master thesis, USTOMB, Oran, Algeria, 2006.
[6] Lebrun G., “Sélection de modèles pour la classification supervisée avec des SVM (Séparateurs à Vaste Marge). Application au traitement et analyse d’images”. Phd Thesis, University of Caen/Basse-Normandie, 2006.
[7] Hsu C.W., Chang C.C., and Lin C.J., “A practical guide to support vector classification”. Technical Report, National Taiwan University, 2009.
[8] Lee J. and Lin C., “Automatic Model Selection for Support Vector Machines”, Technical Report. csie.ntu.edu.tw/˜cjlin/papers/modelselect.ps.gz, 2000.
[9] Friedman J. H., “Another approach to Polychotomous classification”. Technical report, Department of Statistics, Stanford University, 1996.
[10] Knerr S., Personnaz L., and Dreyfus G.. “Single-layer learning revisited: a stepwise procedure for building and training a neural network”. In *Neuro-computing: Algorithms, Architectures and Applications*, J. Fogelman, editor, Springer-Verlag, 1990.
[11] Kreßel U. H. G., “Pairwise classification and support vector machines”. In B. Schölkopf, C. J. C. Burges, and A. J. Smola, editors, *Advances in Kernel Methods Support Vector Learning*, pages 255-268, The MIT Press, Cambridge, 1999.
[12] Platt J., Cristianini N., and J. Shawe-Taylor. “Large margin DAGs for multi-class classification”. *Advances in Neural Information Processing System*, 12, MIT Press. 2000.
[13] Hsu C. W. and Lin C. J., “A comparison of methods for multi-class support vector machines”, *IEEE Transactions on Neural Networks*, Vol. 13 415–425, 2002.
[14] Ayat N.E., Cheriet M., and Suen C.Y., “Optimization of the SVM kernels using an empirical error minimization scheme”. In S.W. Lee and A. Verri, editors, *Pattern Recognition with Support Vector Machines*, volume 2388 of *LNCS*, pages 354–369. Berlin Heidelberg, July 2002.
[15] Dréo J., Petrowski A., Siarry P., and Taillard E., “Métaheuristiques pour l’optimisation difficile”. Eyrolles Group, 2003.
[16] Glover F., “Tabu search: part I”. In *On Computing*, 1(3), pages 190–206, 1989.
[17] Glover F., “Tabu search: part II”. In *On Computing*, 2(1), pages 4–32, 1989.
[18] Cawley G. C., “Model Selection for Support Vector Machines via Adaptive Step-Size Tabu Search”. In *ICANNGA*, 2001.