# Texture Classification of Proteins Using Support Vector Machines and Bio-inspired Metaheuristics

Carlos Fernandez-Lozano, Jose A. Seoane, Pablo Mesejo, Youssef S.G. Nashed, Stefano Cagnoni, Julian Dorado

# Texture Classification of Proteins Using Support Vector Machines and Bio-Inspired Metaheuristics

Carlos Fernandez-Lozano[1,*], Jose A. Seoane[2], Pablo Mesejo[3], Youssef S.G. Nashed[3], Stefano Cagnoni[3] and Julian Dorado[1]

[1] Information and Communication Technologies Department, Faculty of Computer Science. University of A Coruña. Campus de Elviña s/n, 15071, A Coruña, Spain
[2] MRC Centre for Causal Analyses in Translational Epidemiology. School of Social and Community Medicine. University of Bristol. Oakfield House, Oakfield Grove, Bristol BS82BN, UK
[3] Dipartimento di Ingegneria dell'Informazione, Universitá Degli Studi Di Parma, Viale G. Usberti 181/a, I-43100, Parma, Italy
carlos.fernandez@udc.es, j.seoane@bristol.ac.uk, {pmesejo, nashed, cagnoni}@ce.unipr.it, julian@udc.es

**Abstract.** In this paper, a novel classification method of two-dimensional polyacrylamide gel electrophoresis images is presented. Such a method uses textural features obtained by means of a feature selection process for whose implementation we compare Genetic Algorithms and Particle Swarm Optimization. Then, the selected features, among which the most decisive and representative ones appear to be those related to the second order co-occurrence matrix, are used as inputs for a Support Vector Machine. The accuracy of the proposed method is around 94%, a statistically better performance than the classification based on the entire feature set. This classification step can be very useful for discarding over-segmented areas after a protein segmentation or identification process.

**Keywords:** Texture analysis, Feature Selection, Electrophoresis, Support Vector Machines, Genetic Algorithm, Proteomic Imaging

## 1 Introduction

Proteomics is the study of protein properties in a cell, tissue or serum aimed at obtaining a global integrated view of disease, physiological and biochemical processes of cells, and regulatory networks. One of the most powerful techniques, widely used to analyze complex protein mixtures extracted from cells, tissues, or other biological samples, is two-dimensional polyacrylamide gel electrophoresis (2D-PAGE). In this method, proteins are classified by molecular weight (MWt) and iso-electric point (pI) using a controlled laboratory process and digital imaging equipment.

Since the beginning of proteomic research, 2D-PAGE has been the main protein separation technique, even before proteomics became a reality itself. The main advantages of this approach are its robustness and its unique ability to analyze complete proteins at high resolution, keeping them intact and being able to isolate them entirely [1]. However, this method has also several drawbacks like its very low effectiveness

in the analysis of hydrophobic proteins, as well as its high sensitivity to dynamic range (i.e. the quantitative ratio between the rarest protein expressed in a sample and the most abundant one) and quantitative distribution issues [2]. The outcome of the process is an image like the ones showed in [3, 4].

Dealing with this kind of images is a difficult task because there is not a commonly accepted ground truth [3, 4]. Another aspect that makes the work difficult from a computer vision point of view is that both protein images and background noise seem to follow a Gaussian distribution [5]. The inter- and intra-operator variability of the outcome of manual analysis of these images is also a big drawback [6].

The aim of this paper is to demonstrate that there is enough texture information in 2D-electrophoresis images to discriminate proteins from noise or background. In this work the most representative group of textural features are selected using two metaheuristics, namely Genetic Algorithms [7] and Particle Swarm Optimization [8].

## 2 Theoretical Background and Related Work

The method proposed in this work intends to assist 2D-PAGE image analysis by studying the textural information they carry. To do so, a novel combination of metaheuristics and Support Vector Machines [9] is presented. In this section, the main techniques used in our approach are briefly introduced and explained.

### 2.1 Texture

One of the most important characteristics used for identifying objects or regions of interest in an image is texture, which is related with the spatial (statistical) distribution of the grey levels within an image [10]. Texture is a surface's property and can be regarded as an almost regular spatial organization of complex patterns, always present even if they could exist as a non-dominant feature. Different approaches to texture analysis have been applied which are extensively reviewed in [11, 12].

There are four major issues one can relate with texture: synthesis, classification, segmentation and shape from texture [12, 13].

- Texture synthesis: this is a subjective process which creates textures in synthetic images. It is important where the goal is to obtain object surfaces as realistic as possible [14].
- Texture classification: the goal is to classify regions of interest in the image according to the kind of texture they embed [15]. This particular issue is widely studied in medical imaging, quality control and remote sensing among others. Classification algorithms can rely on a quantitative measure of success. It is necessary to have a-priori knowledge of the possible texture types in order to perform such a task.
- Texture segmentation: the principal task here is to find the texture boundaries in an image with a large number of textural types [16]. This is a blind process in the sense that there is no a priori information available about them or about how many different textures or types of textures are there.

- Shape from texture: the task is to reconstruct a three-dimensional object from a two-dimensional image based on textural information. Firstly proposed by Gibson [17].

This paper is focused on texture classification, where texture can be computed from the variations in the intensities within the image. As said before, there is no textural information in one pixel, so texture is a contextual property related with gray levels in a neighborhood. First-order statistics depend only on individual pixel values and can be computed from the histogram of pixel intensities in the image, but second-order statistics depend on pairs of grey values, concerning their relative position and spatial resolution. Commonly used second-order statistics can be derived from the so-called Grey Level Co-Occurrence Matrix (GLCM), first proposed by Haralick [9]. Common properties playing an important role with texture definition, as identified by Laws [18], are: density, coarseness, uniformity, roughness, regularity, linearity, directionality, direction, frequency, and phase. All these properties are often related [19].

### 2.2 Metaheuristics

Genetic Algorithms (GAs) are search techniques inspired by Darwinian Evolution and developed by Holland in the 1970s [7]. In a GA, an initial population of individuals, i.e. possible solutions defined within the domain of a fitness function to be optimized, is evolved by means of genetic operators: selection, crossover and mutation. The selection operator ensures the survival of the fittest individuals, crossover represents the mating between individuals, and mutation introduces random modifications into the population. GAs possess effective capabilities to explore the search space in parallel, exploiting the information about the quality of the individuals evaluated so far [20]. Using the crossover operator, GAs combine the features of parents to produce new and better solutions, which preserve the parents' best characteristics. Using the mutation operator, new information is introduced in the population in order to explore new areas of the search space. The strategy known as elitism, which is a variant of the general process of constructing a new population, allows the best organisms from the current generation to survive to the next, remaining unaltered. At the end of the process, the population of solutions is expected to converge to the global optimum of the fitness function.

Particle Swarm Optimization [8] (PSO) is a bio-inspired optimization algorithm based on the simulation of the social behavior of bird flocks. In the last fifteen years PSO has been applied to a very large variety of problems [21] and numerous variants of the algorithm have been presented [22].

During the execution of PSO a set of particles moves within the function domain searching for the optimum of the function (best fitness value). The motion of each particle is driven by the best positions visited so far by the particle itself and by the entire swarm (gbest PSO) or by some pre-defined neighborhood of the particle (lbest PSO). Consequently, each particle relies both on "individual" and on "swarm" intelligence, and its motion can be described by two simple discrete-time equations which regulate the particle's position and velocity.

### 2.3 Support Vector Machines

Vapnik introduced Support Vector Machines (SVMs) in the late 1970s on the foundation of statistical learning theory [9]. The basic implementation deals with two-class problems in which data are separated by a hyperplane defined by a number of support vectors. This hyperplane separates the positive from the negative examples, maximizing the distance between the boundary and the nearest data point in each class; the nearest data points are used to define the margins, known as support vectors [23]. These classifiers have also proven to be exceptionally efficient in classification problems of high dimensionality [24, 25], because of their ability to generalize in high-dimensional spaces, such as the ones spanned by texture patterns. SVMs use different non-linear kernel functions, like polynomial, sigmoid and radial basis functions, to map the training samples from the input spaces into a higher-dimensional feature space through a mapping function [23].

### 2.4 Related work

With respect to related work, the authors were not able to find any other work in the literature dealing with evolutionary computation in combination with texture analysis in 2D-electrophoresis images, while we did find one article describing a discriminant partial least squares regression (PLSR) method for spot filtering in 2D-electrophoresis [26]. The authors use a set of parameters to build a model based on texture, shape and intensity measurements using image segments from gel segmentation. As regards texture information, they focus on the descriptors related to the noisy surface texture of unwanted artifacts and conclude that their textural features allow them to distinguish noisy features from protein spots. In their work, five out of the eleven second-order textural features are used, along with five new textural features accounting for intensity relationships among sets of three pixels. They distinguish proteins in the image by using shape information, since cracks and artifacts in gel surface deviate from a circular shape. Besides that, a degree of Gaussian fit is calculated as an indicator of whether the image segment corresponds to a protein or to an artifact. Textural features are used for noise and crack detection and as a complement for spot segmentation. Finally, the 17 initial variables are reduced to five PLSR components to account for 85% of the total variation with respect to the response factor, and 82% of the total variation in the data matrix.

## 3 Materials

In order to generate the dataset, ten 2D-PAGE images, representative enough of different types of tissues and different experimental conditions, were used. These images are similar to the ones used by G.-Z. Yang (Imperial College of Science, Technology and Medicine, London), and are available for download at the webpage http://personalpages.manchester.ac.uk/staff/andrew.dowsey/rain/. It is important to notice that Hunt et al. [27] determined that 7-8 is the minimum acceptable number of samples for a proteomic study.

For each image, 50 regions of interest (ROIs) representing proteins and 50 representing non-proteins (noise, black non-protein regions, and background) were selected to build a training set with 1000 samples in a double-blind process in which two clinicians selected the fifty ROIs they considered and after that, within which they selected proteins which were representative of the different possible scenarios (isolated, overlapped, big, small, darker, etc.). For each ROI, as will be seen later, 296 texture features are computed.

The ROIs were selected taking into consideration that, for each manually selected protein, there is an area of influence surrounding it. It means that, once the clinician has selected a protein, the ROI is slightly larger than the visible dark surface of such a protein. This assumption is made because texture characterizes not only the darkest regions but also the lightest ones.

As said before, proteins seem to fit a Gaussian peak, and ideally the center of the protein is in the darkest zone of that peak. This approach prevents the loss of information caused by neglecting the lowest values of the inverted protein (grey levels closest to white) that also fit the Gaussian peak. This information could be useful to classify a ROI as including a protein or to discard it.

## 4 Proposed Method

This paper goes further than related work in texture analysis of 2D-electrophoresis images, studying the ability of textural features to discriminate not only cracks from proteins but background and non-protein dark spots as well.

The first step in texture analysis is texture feature extraction from the ROIs. With a specialized software called Mazda [28], 296 texture features are computed for each element in the training set. Various approaches have demonstrated the effectiveness of this software in extracting textural features from different types of medical images [29-33].

These features [34], reported in Table 1, are based on:

- Image histogram.
- Co-occurrence matrix: information about the grey level value distribution of pairs of pixels with a preset angle and distance between each other.
- Run-length matrix: information about sequences of pixels with the same grey level values in a given direction.
- Image gradients: spatial variation of grey level values.
- Autoregressive models: description of texture based on the statistical correlation between neighboring pixels.
- Wavelet analysis: information about the image frequency content at different scales.

Thus, within each ROI, texture information was analyzed by extracting first and second-order statistics, spatial frequencies, co-occurrence matrices and two other statistical methods as autoregressive model and wavelet based analysis, preserving the original gray-level and spatial resolution in all runs. Histogram-related measures conform the first-order statistics proposed by Haralick [10] but second-order statistics are those derived from the Grey Level Co-occurrence Matrices (GLCM). Additionally, a

group of features derived from the textural ones is also calculated, such as the area of the ROI, but cannot be used for texture characterization.

**Table 1.** Textural features extracted and used in this work

| Group | Features |
|---|---|
| Histogram | Mean, variance, skewness, kurtosis, percentiles 1%, 10%, 50%, 90% and 99% |
| Absolute Gradient | Mean, variance, skewness, kurtosis and percentage of pixels with nonzero gradient |
| Run-length Matrix | Run-length non-uniformity, grey-level non-uniformity, long-run emphasis, short-run emphasis and fraction of image in runs |
| Co-ocurrence Matrix | Angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance and difference entropy |
| Autoregressive Model | Theta: model parameter vector, four parameters; Sigma: standard deviation of the driving noise |
| Wavelet | Energy of wavelet coefficients in sub-bands at successive scales; max four scales, each with four parameters |

All these feature sets were included in the dataset. The normalization method applied was the one set by default in Mazda: image intensities were normalized in the range from 1 to $Ng=2^k$, where k is the number of bits per pixel used to encode the image under analysis.

Two solutions are available for decreasing the dimensionality: extraction of new features derived from the existing ones and selection of relevant features to build robust models. In order to extract a feature set from the problem data, principal component analysis (PCA) has been commonly used. In this work, we use GAs to find the smallest feature subset able to yield a fitness value above a threshold. Besides optimizing the complexity of the classifier, feature selection may also improve the classifiers' quality. In fact, classification accuracy could even improve if noisy or dependent features are removed.

The use of GAs for feature selection were first proposed by Siedlecki and Skalansky [35]. Many studies have been done on GAs for feature selection since then [36], concluding that they are suitable for finding optimal solutions to large problems with more than 40 features to select from. GAs for feature selection could be used in combination with a classifier such as SVM, k-nearest neighbor (KNN) or artificial neural networks (ANN), optimizing its performance. In terms of classification accuracy with imaging problems, SVMs have shown to yield good performance with textural features [37-39], but also KNN [40]; hybrid approaches which use a combination of both classifiers [41] have obtained good results. Other techniques use GAs to optimize both feature selection and classifier parameters [42, 43].

In our method, based on both GAs and SVMs, there is not a fixed number of variables. As GAs continuously reduce the number of variables that characterize the samples, a pruned search is implemented. Each individual in the genetic population is described by p genes (using binary encoding). The fitness function (1) considers not

only the classification results but also the number of variables used for such a classification, so it is defined as the sum of two terms, one related to the classification results and another to the number of variables selected. In (1) the number of genes with a true binary value (feature selected) is represented by *numberActiveFeatures*. Regarding classification results, taking into account the F-measure apparently gives better results than only using the accuracy obtained with image features [44, 45]. F-measure (2) is a function made up of the recall (true positives rate or sensitivity: proportion of actual positives which are correctly identified) and precision (or positive predictive value: proportion of positive test results that are true positives) measurements.

$$Fitness = (1 - F) + \frac{numberActiveFeatures}{numberTotalFeatures} \quad (1)$$

$$F = 2.\frac{precision\ .recall}{precision + recall} \quad (2)$$

Therefore individuals with fewer active features (genes) are favored. Once the reduced feature dataset is generated, a statistical parametric test is made to evaluate the adequacy of the feature selection process.

## 5 Experimental results

The test set is composed of ten representative images for the different types of proteomic available images, and for each one of them, 50 protein and 50 non-protein ROIs have been extracted to generate a dataset with 1000 elements, that was divided randomly in 800 elements, of which 600 elements are used for training and 200 elements are used for validation (inside the GA feature selection process) and finally, 200 elements for test. Once the GA finishes, the best individual found (the one with lowest fitness value) is tested, using a 10-fold cross validation (10-fold CV), to calculate the error of the proposed model using the full and reduced datasets. Then, a test set is used in order to evaluate the adequacy of the reduction process.

After a preliminary experimental study of the values suggested in the literature, the parameters used in the feature selection process were empirically set the population size to 250 individuals, with no elitism, a 95% crossover probability, a 2% mutation rate, with scattered crossover, tournament selection and uniform mutation.

To evaluate the performance of this method, there are several number of well-known accuracy measures for a two-class classifier in the literature such as: classification rate (accuracy), precision, sensitivity, specificity, F-score, Area Under an ROC Curve (AUC), Youden's index, Cohen's kappa, likelihoods, discriminant power, etc. The ROC curve is a graphical plot of the sensitivity against 1-specificity as the detector threshold, or a parameter which modifies the balance between sensitivity and specificity. An experimental comparison of performance measures for classification could be found in [46]. In [47], the authors proposed that AUC is a better measure in general than accuracy when comparing classifiers and in general. The most common measures used for their simplicity and successful application are the classification rate

and Cohen's kappa measures. Table 3 in the appendix shows the results for classification rate (accuracy), AUC, F-measure, Youden's and unweighted Cohen's Kappa for each kernel. For this problem, all the measures consider the same ranking, and the best kernel function is the linear one. For each kernel, Table 4 in the Appendix section shows the selected features in their textural membership group.

Among others, Mazda computes the area for each ROI. This feature merely indicates the number of pixels used to compute the textural features and, since it has nothing to do with the description of textures, it cannot be used for texture characterization. With linear, polynomial (order 3), and RBF (C=100 and sigma=10) kernels, non-textural features are selected for classification. The results obtained seem to indicate that the textural group with more representatives in 2D-PAGE images is the Co-occurrence matrix Group (second-order statistics).

As the proposed work intends to evaluate the textural information present in a 2D-PAGE image, the RBF(2) kernel function is selected as the most appropriate for solving this problem, since only textural features where selected for classification with this kernel, and it yields the best accuracy.

In order to compare the GA-based feature selection results, a binary Particle Swarm Optimization implementation for feature selection is used [48]. PSO is an optimization algorithm inspired by the organized behavior of large groups of simple animals and, like GAs and other Evolutionary Computation techniques; it is a derivative-free global optimum solver. Firstly proposed by Kennedy and Eberhart [8] and used optimization of non-linear functions [49].

The experiments were performed with the same final combination of common parameters (population size, stall conditions, etc.), and the same elements for training and validation separated with the same seed in order to reproduce experimental conditions with the RBF(2) kernel function. Results are shown in Table 2.

**Table 2.** Results for the GA and PSO approaches with RBF(2) kernel.

| | Accuracy | AUC | F-Measure | Number of variables | Generations |
|---|---|---|---|---|---|
| GA | 0.88 | **0.88** | 0.89 | 6 | 45 |
| PSO | 0.83 | 0.86 | 0.85 | 58 | 44 |

The results show that GA-based have better results, improving the AUC-ROC score of the PSO and is able to reduce to 6 features whilst the PSO is only able to reduce to 58 features. Both techniques reach the stall condition in a very close number of generations.

We evaluated the reduced textural feature set on the 200 patterns of the validation dataset using the RBF (2) kernel, by calculating the F-measure and the areas under the receiver operating characteristic curves and a 10-fold CV, using the Libsvm classifier implementation [50] in Weka [51] and comparing the results with the same classifier using the full dataset. Thus, we have obtained samples composed by 10 AUC measures. AUC area can be seen as the capacity to be sensitive and specific at the same time, in the sense that the larger is the AUC, the more accurate is the model. We use the RBF kernel with different gamma values (the parameter controls the width of the kernel) to check if there is a significant improvement when the reduced dataset is used.

In order to use a parametric test, it is necessary to check the independence, normality and heteroscedasdicity [52]. In statistics, two events are independent when the occurrence of one does not modify the probability of the other one. An observation is normal when its behavior follows a normal or Gaussian distribution with a certain value of mean and variance. The heteroscedasticity indicates the existence of a violation of the hypothesis of equality of variances [53].

With respect to the independence condition, we separate the data using 10-fold CV. We perform a normality analysis using the Shapiro-Wilk test [54] with a level of confidence alpha=0.05, for the Null Hypothesis that the data come from a normally distributed population, and such null hypothesis was rejected. The observed data fulfill the normality condition, a Bartlett test [55] is performed in order to evaluate the heteroscedasticity with a level of confidence alpha=0.05.

A corrected paired Student's t-test could be performed in Weka [51], with a level of confidence alpha=0.05, for the Null Hypothesis that there are no differences between the average values obtained by both methods. Results in average, with standard deviation in brackets for AUC-ROC are 0.94 (0.07) for the reduced dataset, and 0.55 (0.34) for the full dataset and the corrected paired Student's t-test determines that there is a significant improvement in using the reduced dataset. The reduced dataset has better accuracy result than the full dataset. Even more, the corrected paired Student's t-test evaluates this improvement as significant with an alpha=0.05.

## 6 Summary and conclusions

To the best of our knowledge, this is the first work in which protein classification in two-dimensional gel electrophoresis images is tackled using Evolutionary Computation, Support Vector Machines and Textural Analysis. In fact, this paper demonstrates the existence of enough textural information to discriminate proteins from noise and background, as well as to show the potential of SVMs in proteomic classification problems.

A new dataset with six features, starting from the 296 original ones, is created without loss of accuracy, and the most representative textural group has shown to be the one related to the Co-occurrence matrix Group (second-order statistics). A proper statistical test has determined that there is a significant improvement in using this reduced feature set with respect to the full feature set.

## Acknowledgements

This work is by "Development of new image analysis techniques in 2D Gel for biomedical research" (Ref. 10SIN105004PR), CN2102/217, CN2011/034 and CN2012/130 by Xunta de Galicia. Jose A. Seoane acknowledges Medical Research Council Project Grant G1000427. Pablo Mesejo and Youssef S.G. Nashed are funded by the European Commission (MIBISOC Marie Curie Initial Training Network, FP7 PEOPLE-ITN-2008, GA n. 238819).

**References**

1. Rabilloud, T., Chevallet, M., Luche, S., Lelong, C.: Two-dimensional gel electrophoresis in proteomics: Past, present and future. Journal of Proteomics 73, 2064-2077 (2010)

2. Zhang, J., Tan, T.: Brief review of invariant texture analysis methods. Pattern Recognition 35, 735-747 (2002)

3. Marten Lab Proteomics Page. http://www.umbc.edu/proteome/image_analysis.html

4. Center for Cancer Research Nanobiology Program (CCRNP), http://www.ccrnp.ncifcrf.gov/users/lemkin

5. Tsakanikas, P., Manolakos, E.S.: Improving 2-DE gel image denoising using contourlets. Proteomics 9, 3877-3888 (2009)

6. Millioni, R., Sbrignadello, S., Tura, A., Iori, E., Murphy, E., Tessari, P.: The inter- and intra-operator variability in manual spot segmentation and its effect on spot quantitation in two-dimensional electrophoresis analysis. Electrophoresis 31, 1739-1742 (2010)

7. Holland, J.H.: Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. University of Michigan Press (1975)

8. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Neural Networks, 1995. Proceedings., IEEE International Conference on, pp. 1942-1948 vol.1944. (Year)

9. Vapnik, V.N.: Estimation of dependences based on empirical data [in Russian]. English translation Springer Verlang, 1982, Nauka (1979)

10. Haralick, R.M., Shanmugam, K., Dinstein, I.: Textural features for image classification. IEEE Transactions on Systems, Man and Cybernetics smc 3, 610-621 (1973)

11. Materka, A., Strzelecki, M.: Texture analysis methods-A review. Technical University of Lodz, Institute of Electronics. COST B11 report (1998)

12. Tuceryan, M., Jain, A.: Texture analysis. Handbook of pattern recognition and computer vision, vol. 2. World Scientific Publishing Company, Incorporated (1999)

13. Levina, E.: Statistical Issues in Texture Analysis. University of California, Berkeley (2002)

14. Peitgen, H.O., Saupe, D., Barnsley, M.F.: The Science of fractal images. Springer-Verlag (1988)

15. Pietikainen, K.: Texture analysis in machine vision. World Scientific Publishing Company, Incorporated (2000)

16. Mirmedhdi, M., Xie, X., Suri, J.S.: Handbook of texture analysis. Imperial College Press (2008)

17. Gibson, J.J.: The perception of the visual world. Houghton Mifflin (1950)

18. Laws, K.I.: Textured Image Segmentation. University of Southern California (1980)

19. Tomita, F., Tsuji, S.: Computer Analysis of Visual Textures. Kluwer Academic Publishers (1990)

20. Goldberg, D.: Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley Professional (1989)

21. Poli, R.: Analysis of the publications on the applications of particle swarm optimisation. J. Artif. Evol. App. 2008, 1-10 (2008)

22. Banks, A., Vincent, J., Anyakoha, C.: A review of particle swarm optimization. Part I: background and development. 6, 467-484 (2007)

23. Burges, C.J.C.: A tutorial on support vector machines for pattern recognition. Data Mining and Knowledge Discovery 2, 121-167 (1998)

24. Moulin, L.S., Alves Da Silva, A.P., El-Sharkawi, M.A., Marks Ii, R.J.: Support vector machines for transient stability analysis of large-scale power systems. IEEE Transactions on Power Systems 19, 818-825 (2004)

25. Chapelle, O., Haffner, P., Vapnik, V.N.: Support vector machines for histogram-based image classification. IEEE Transactions on Neural Networks 10, 1055-1064 (1999)

26. Rye, M.B., Alsberg, B.K.: A multivariate spot filtering model for two-dimensional gel electrophoresis. Electrophoresis 29, 1369-1381 (2008)

27. Hunt, S.M.N., Thomas, M.R., Sebastian, L.T., Pedersen, S.K., Harcourt, R.L., Sloane, A.J., Wilkins, M.R.: Optimal Replication and the Importance of Experimental Design for Gel-Based Quantitative Proteomics. Journal of Proteome Research 4, 809-819 (2005)

28. Szczypinski, P.M., Strzelecki, M., Materka, A.: MaZda - A software for texture analysis. pp. 245-249.

29. Szymanski, J.J., Jamison, J.T., DeGracia, D.J.: Texture analysis of poly-adenylated mRNA staining following global brain ischemia and reperfusion. Computer Methods and Programs in Biomedicine 105, 81-94 (2012)

30. Harrison, L., Dastidar, P., Eskola, H., Järvenpää, R., Pertovaara, H., Luukkaala, T., Kellokumpu-Lehtinen, P.L., Soimakallio, S.: Texture analysis on MRI images of non-Hodgkin lymphoma. Computers in Biology and Medicine 38, 519-524 (2008)

31. Mayerhoefer, M.E., Breitenseher, M.J., Kramer, J., Aigner, N., Hofmann, S., Materka, A.: Texture analysis for tissue discrimination on T1-weighted MR images of the knee joint in a multicenter study: Transferability of texture features and comparison of feature selection methods and classifiers. Journal of Magnetic Resonance Imaging 22, 674-680 (2005)

32. Bonilha, L., Kobayashi, E., Castellano, G., Coelho, G., Tinois, E., Cendes, F., Li, L.M.: Texture Analysis of Hippocampal Sclerosis. Epilepsia 44, 1546-1550 (2003)

33. Létal, J., Jirák, D., Suderlová, L., Hájek, M.: MRI 'texture' analysis of MR images of apples during ripening and storage. LWT - Food Science and Technology 36, 719-727 (2003)

34. Szczypiski, P.M., Strzelecki, M., Materka, A., Klepaczko, A.: MaZda-A software package for image texture analysis. Computer Methods and Programs in Biomedicine 94, 66-76 (2009)

35. Siedlecki, W., Sklansky, J.: A note on genetic algorithms for large-scale feature selection. Pattern Recognition Letters 10, 335-347 (1989)

36. Kudo, M., Sklansky, J.: A comparative evaluation of medium- and large-scale feature selectors for pattern classifiers. Kybernetika 34, 429-434 (1998)

37. Li, S., Kwok, J.T., Zhu, H., Wang, Y.: Texture classification using the support vector machines. Pattern Recognition 36, 2883-2893 (2003)

38. Kim, K.I., Jung, K., Park, S.H., Kim, H.J.: Support vector machines for texture classification. IEEE Transactions on Pattern Analysis and Machine Intelligence 24, 1542-1550 (2002)

39. Buciu, I., Kotropoulos, C., Pitas, I.: Demonstrating the stability of support vector machines for classification. Signal Processing 86, 2364-2380 (2006)

40. Jain, A.: Feature selection: evaluation, application, and small sample performance. IEEE Transactions on Pattern Analysis and Machine Intelligence 19, 153-158 (1997)

41. Zhang, H., Berg, A.C., Maire, M., Malik, J.: SVM-KNN: Discriminative nearest neighbor classification for visual category recognition. pp. 2126-2136. (Year)

42. Huang, C.L., Wang, C.J.: A GA-based feature selection and parameters optimizationfor support vector machines. Expert Systems with Applications 31, 231-240 (2006)

43. Manimala, K., Selvi, K., Ahila, R.: Hybrid soft computing techniques for feature selection and parameter optimization in power quality data mining. Applied Soft Computing Journal 11, 5485-5497 (2011)

44. Müller, M., Demuth, B., Rosenhahn, B.: An evolutionary approach for learning motion class patterns. vol. 5096 LNCS, pp. 365-374 (2008)

45. Tamboli, A.S., Shah, M.A.: A Generic Structure of Object Classification Using Genetic Programming. In: Communication Systems and Network Technologies (CSNT), 2011 International Conference on, pp. 723-728. (Year)

46. Ferri, C., Hernádez-Orallo, J., Modroiu, R.: An experimental comparison of performance measures for classification. Pattern Recognition Letters 30, 27-38 (2009)

47. Huang, J., Ling, C.X.: Using AUC and accuracy in evaluating learning algorithms. IEEE Transactions on Knowledge and Data Engineering 17, 299-310 (2005)

48. Chen, S.: Another Particle Swarm Optimization Toolbox. Ontario (2003)

49. Perez, R.E., Behdinan, K.: Particle swarm approach for structural design optimization. Computers & Structures 85, 1579-1588 (2007)

50. Chang, C.C., Lin, C.J.: LIBSVM: A Library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, (2011)
51. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. SIGKDD Explor. Newsl. 11, 10-18 (2009)

52. Sheskin, D.J.: Handbook of Parametric and Nonparametric Statistical Procedures. Taylor and Francis (2011)

53. García, S., Fernández, A., Luengo, J., Herrera, F.: A study of statistical techniques and performance measures for genetics-based machine learning: Accuracy and interpretability. Soft Computing 13, 959-977 (2009)

54. Shapiro, S.S., Wilk, M.B.: An analysis of variance test for normality (complete samples). Biometrika 52, 591-611 (1965)

55. Bartlett, M.S.: Properties of Sufficiency and Statistical Tests. Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 160, 268-282 (1937)

# Appendix

**Table 3a.** Results with different SVM Gaussian kernels

| Kernel Type | Measure | Value |
|---|---|---|
| RBF(1) | TP | 90 |
| | FN | 10 |
| | FP | 18 |
| | TN | 82 |
| | Accuracy | 0.86 |
| | AUC | 0.86 |
| | F-Measure | 0.86 |
| | Youden's | 0.72 |
| | Kappa | 0.72 |
| | Nvar | 8 |
| RBF(2) | TP | 94 |
| | FN | 6 |
| | FP | 17 |
| | TN | 83 |
| | Accuracy | 0.88 |
| | AUC | 0.88 |
| | F-Measure | 0.89 |
| | Youden's | 0.77 |
| | Kappa | 0.77 |
| | Nvar | 6 |
| RBF (100;10) | TP | 94 |
| | FN | 6 |
| | FP | 18 |
| | TN | 82 |
| | Accuracy | 0.88 |
| | AUC | 0.88 |
| | F-Measure | 0.88 |
| | Youden's | 0.76 |
| | Kappa | 0.76 |
| | Nvar | 8 |

**Table 3b.** Results with different SVM polynomial kernels

| Kernel Type | Measure | Value |
|---|---|---|
| Linear | TP | 95 |
| | FN | 5 |
| | FP | 11 |
| | TN | 89 |
| | Accuracy | 0.92 |
| | AUC | 0.92 |
| | F-Measure | 0.92 |
| | Youden's | 0.85 |
| | Kappa | 0.85 |
| | Nvar | 6 |
| Poli (3) | TP | 87 |
| | FN | 13 |
| | FP | 19 |
| | TN | 81 |
| | Accuracy | 0.84 |
| | AUC | 0.84 |
| | F-Measure | 0.84 |
| | Youden's | 0.68 |
| | Kappa | 0.68 |
| | Nvar | 16 |

**Table 4.** Study of texture parameters between best SVM kernels in accuracy

| | Histogram | Absolute gradient | Run-length matrix | Co-occurence matrix | Wavelet | Non-textural features |
|---|---|---|---|---|---|---|
| RBF(1) | | | | S(2,0)InvDfMom<br>S(0,3)SumAverg<br>S(0,3)DifVarnc<br>S(4,-4)Contrast<br>S(0,5)SumEntrp<br>S(0,5)DifEntrp<br>S(5,5)SumEntrp<br>S(5,-5)Entropy<br>S(2,-2)DifEntrp<br>S(5,0)Correlat | | |
| RBF(2) | Perc.01% | | | S(5,0)InvDfMom<br>S(0,5)DifVarnc<br>S(5,5)SumEntrp | | |
| Linear | Skewness | | | S(2,2)Correlat<br>S(4,0)InvDfMom<br>S(5,0)Contrast | | Area_S(0,4)<br>Area_S(5,-5) |
| Poli(3) | Kurtosis | GrKurtosis | 45dgr_<br>GLev<br>NonU | S(1,-1)Contrast<br>S(1,-1)DifEntrp<br>S(0,2)DifEntrp<br>S(0,4)SumAverg<br>S(4,-4)Correlat<br>S(4,-4)SumVarnc<br>S(5,0)InvDfMom<br>S(0,5)SumOfSqs | WavEnLH_s-2<br>WavEnLH_s-4 | AreaGr |
| RBF(100;10) | | | Horzl_<br>GLev<br>NonU | S(0,5)InvDfMom<br>S(0,5)SumEntrp<br>S(2,0)InvDfMom<br>S(5,0)InvDfMom<br>S(0,5)InvDfMom<br>S(5,-5)DifEntrp | WavEnLH_s-4 | Area_S(0,1)<br>Area_S(5,0) |