Texture Based MRI Segmentation with a Two-Stage Hybrid Neural Classifier

Article · June 2003

DOI: 10.1109/IJCNN.2002.1007457 · Source: CiteSeer

4 authors, including:

Alain Pitiot
University of Nottingham
63 PUBLICATIONS **1,822** CITATIONS

SEE PROFILE

Nicholas Ayache
National Institute for Research in Computer Science and Control
897 PUBLICATIONS **30,297** CITATIONS

SEE PROFILE

Paul Thompson
University of Southern California
1,654 PUBLICATIONS **71,119** CITATIONS

SEE PROFILE

Texture Based MRI Segmentation with a Two-Stage Hybrid Neural Classifier

Alain Pitiot[1,2], Arthur W. Toga[2], Nicholas Ayache[1], Paul Thompson[2]

[1] INRIA-EPIDAURE
Sophia Antipolis, France

[2] Laboratory of Neuro Imaging
UCLA School of Medicine, USA

Abstract - We propose an automated method for extracting anatomical structures in magnetic resonance images (MRI) based on texture classification. It consists of two consecutive stages. The textures of an input MRI are first classified by a network of adaptive spline neurons, organized within a hybrid master classifier/mixtures-of-experts architecture (stage I). The output map is then fed into a second neural network, which aims to better contrast the target structure and eliminate the mistakes of the first phase via local shape/texture analysis and a carefully designed learning process (stage II). Results are demonstrated on medical imagery with the segmentation of various brain structures.

I. INTRODUCTION

Image segmentation, that is, the division of an image pixel set into a number of disjoint regions that are homogeneous with respect to a set of characteristics, is a major goal of computer vision and image processing.

Effective identification and labeling of anatomical structures in possibly complex magnetic resonance images (MRI) proves to be especially challenging, given the wide variety of shapes and intensities each anatomical structure can present. Given the complex intensity distribution, general lack of contrast and rather poor resolution of MR images, most segmentation approaches rely primarily on shape information to extract the boundary of the target structure (see [11] for a review of deformable template-oriented techniques). Sometimes, tissue classification is used as a pre-processing step to facilitate the segmentation task (see [1] for a survey of segmentation techniques applied to medical images). However, a tissue classifier usually partitions the input voxels into 3 or 4 tissue classes (gray matter, white matter, cerebro-spinal fluid and "other") *independently* of the target structure.

Yet, when expert neuro-scientists manually delineate these structures, they also use texture to guide the segmentation. For instance, even though both the *fornix* and *corpus callosum* consists of white matter (Figure 4), a different orientation of the fibers yields two different textures. Similarly, even though the mean intensity of *hippocampus* (Figure 5, *middle*) is very close to that of its neighborhood, it exhibits a characteristic set of textures.

Intensity and textural information gathered from the vicinity of the target structure also helps the extraction process. It is clear from Figure 5 (left) that the typical darker and homogeneous appearance of the ventricle is a useful clue in finding the *caudate nucleus.*

This paper addresses the problem of extracting an anatomical structure from a series of MR images using texture information. Our segmentation system has 2 stages:

- *stage I : texture classification*

A hybrid master classifier/mixtures of experts neural architecture is used in conjunction with adaptive spline neurons (ASN) to classify the textures of the input MRI. The initial weights and shapes of the ASNs' activation functions are determined by an *a priori* distribution describing the selectivity and accuracy of a statistically-derived bank of texture filters.

- *stage II : local shape/texture analysis*

We then use a multi-scale neural network, subject to node pruning, to further contrast the target structure and correct the possible mistakes of the first stage. This separate stage allows us to design a learning set concentrated along the classification boundary between the desired textures and the others. It also takes advantage of a larger neighborhood to enhance the quality of the segmentation by using intensity information from nearby structures.

We detail our texture classifier in the following Section 2, before discussing in Section 3 some segmentation results.

II. ARCHITECTURE AND LEARNING PROCESS

A. Texture classification via hybrid master classifier / mixtures of experts architecture

A number of texture descriptors have been discussed in the literature (see [13] for a review). These make use of statistical measures, time-space-frequency decompositions, or fractal parameters for instance. As every descriptor has weaknesses and strengths, one usually combines them to improve the classification performance.

We chose to use linear discriminant analysis (LDA) [10] to evaluate a large number of texture descriptors on a training database and help in selecting the most pertinent ones. For each anatomical structure to be segmented, we compiled a series of MR images, along with their associated manually segmented images to form the training database. We then used Wilks' stepwise method to find the coefficients of the discriminant function: only the descriptors with the highest coefficients were selected for classification.

We give below a short description of these descriptors before detailing how we combined them within a neural architecture.

*Texture descriptors***.** Following our LDA results, we selected:

- co-occurrence matrices [3]: these embed in a single array the relative frequencies of gray level pairs of pixels for a series of translations. We use 8 translations.

- discrete cosine transform [12]: we used a 3x3 DCT without critical sampling.

- dyadic Gabor filters [5]: we chose 3 different frequency bands and 4 orientations, and derived the energy and entropy for each of the 3x4=12 filtered windows.

- local Hölder exponent [16]: this corresponds to an intuitive perception of regularity: it is computed by comparing the total variation of the gray levels in the sub-window with functions of the form: $e \mathbin{®} e^{a}$.

*Hybrid master classifier / mixtures of experts architecture***.** For each pixel in the input MRI, we consider a surrounding sub-window (5x5 pixels) and "normalize" the intensities by subtracting that of the central pixel from the others, before applying the various texture filters. Once computed, texture features can be combined in a number of ways:

- they can be concatenated into a single feature vector to be classified all together;

- each feature can be classified separately and the results subsequently combined with a voter;

- a master classifier [4] can be used to classify the results of several slave classifiers;

- or, an independent classifier (associative switch [18]) decides which feature filter is considered as providing the final result.

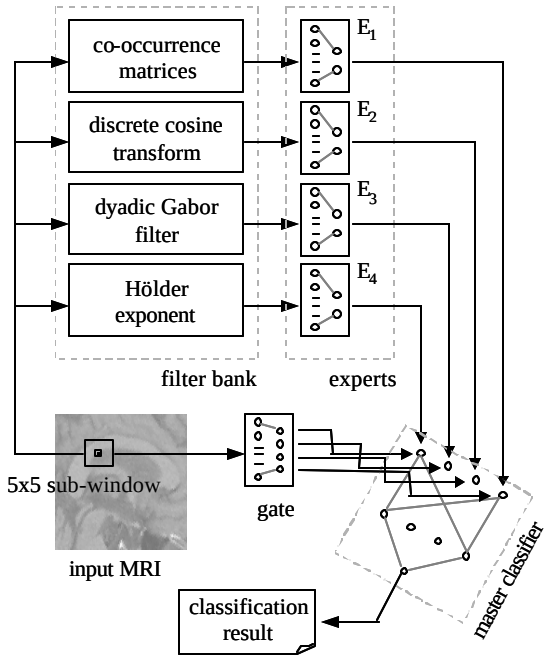

Figure 1. Stage I: texture segmentation.

Note that the LDA feature selection step we implemented yields only a coarse approximation of the optimal set of filters that can be classified. In particular, some classes may not be separable in certain feature space. In [6], Jain *et al.* suggest using a neural network, subject to node pruning, to select the appropriate feature filters and eliminate the less relevant ones. A mixtures of experts algorithm [7] can also be used to overcome the feature selection issue. It consists of a number of experts, which receive the same feature vector as input, and a decision module (the gate), with one output per expert, which controls the extent to which each expert contributes to the final classification result: the i^{th} output of the gate represents the confidence in the classification result of the i^{th} expert.

We propose a neural network architecture that blends together a master classifier and a specialized mixtures of experts [17] architecture (Figure 1). The network consists of N=4 "expert" networks $\boldsymbol{E} = \{E_i\}$, a gate G that receives input directly from the MRI, and a master classifier in the form of a 2-layered feed-forward perceptron (MLP; [14]). As the number of outputs of each texture filter varies, we use the same number of neurons in the hidden layer of each expert to "normalize" the overall feature vector.

To take full advantage of both types of architecture, we use adaptive spline neurons (see [2] for a detailed presentation) instead of standard sigmoid activated ones for the input layer of the master classifier network. The activation function of an ASN consists of a

parameterized spline, whose coefficients are adjusted during the training process. By adapting the activation function of each ASN together with its weights, we in fact implement a feature selection scheme: when a neuron's activation function comes close to the horizontal axis, for instance, the corresponding feature expert plays only a minor role in the classification process.

*Learning process***.** For each anatomical structure to be segmented, the learning set consists of a series of MR images, together with the associated manually hand-drawn structures. We select, at random, points across the images and ensure that the number of samples belonging to the target structure is roughly the same as those which do not.

The learning process then consists of 3 consecutive phases:

a. We replace the master classifier by a two-level hierarchical mixture of experts, as described by Jordan and Jacobs in [7]. We then train the network using Walter *et al.* 's modified version [17] of their EM algorithm.
b. The master classifier with ASNs is trained to mimic the behavior of the two-level mixture of experts when receiving the inputs from the expert networks and the gate. We use the gradient-based learning rules detailed in [2].
c. The entire network is re-trained by a gradient descent algorithm based on that of Guarnieri *et al.*: we use an adaptive weight learning rate (m_w in [2]) where m_w is small for the experts and the gate and high for the master classifier.

Phase (c) enables us to specialize the master classifier with the ASN (and thus increase the classification performance) while still improving on the mixture of experts. Experimental results showed that this learning method yielded better results than those obtained when using a modified version of Jordan *et al.* 's gradient descent algorithm adapted to handle the ASN.

The initial activation functions and weights of the ASN in phase (b) are set to reflect the *a priori* confidence in the corresponding feature selector. We use our LDA results to determine the confidence distribution. The initial activation functions take the form of sigmoid functions: $\mathbf{F} = \{F_i : x \mapsto F_i(x) = a_i.\tanh(b_i.x)\}$ and consist of 11 control points. $\{\alpha_i\}_i$ and $\{\beta_i\}_i$ are determined from the coefficients of the discriminant function found by the LDA.

Note that modifying the activation function while training the network not only triggers faster convergence, but also enables the learning process to better escape local minima.

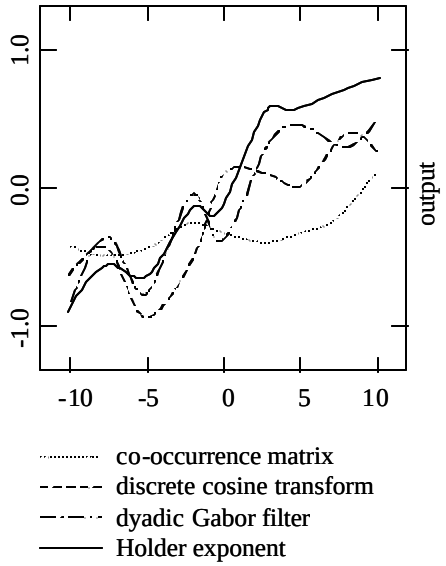

Figure 2. Typical activation functions after the learning process.

Figure 2 shows characteristic shapes of a few ASNs upon completion of the learning process, for the segmentation of *corpus callosum* (learning set: 5 MRI in mid-sagittal section). Note the flat shape of the activation function of the neuron linked to the co-occurrence filter, which reflects its poor classification capabilities (this is especially clear since the weight of the link to the gate is also very small: 3.10^{-5}).

B. Local shape/texture analysis via multi-scale neural net

The first classification stage uses textural information in the input MRI in a very local fashion: a somewhat limited 5x5 sub-window is used to compute the texture features and the points are considered independently from each other. Hence the need for a second classification stage whose task is:

a. to take into account a larger neighborhood for each point and make use of the nearby structures to help the segmentation;

b. to correct possible mistakes of the first stage;

c. to regularize the shape and texture of the segmented outcome.

*Multi-scale architecture***.** We used a multi-scale neural network with standard sigmoid neurons (see Figure 3). A larger sub-window (9x9 pixels) and two downsampled versions of it (scale 1.5 and 3) are fed into a layer of hidden units. We adapted an input connection pattern

commonly used in the speech and character recognition literature. It consists of 3 distinct blocks with 9 neurons each. The first block looks at vertical stripes (each of the 9 neurons in this hidden layer block is fully connected to the 9 neurons of the associated vertical line in the input layer); the second looks at similar horizontal stripes; and the third block looks at 3x3 input sub-regions. These specialized groups of neurons are responsible for detecting local features and relative positions of nearby structures. The blocks are fully connected to a second hidden layer (20 neurons) connected to the output.

As the overall number of connections is quite large (~5000 weights), we use node pruning [9] to further adapt the architecture of the neural network to the segmentation problem at hand.

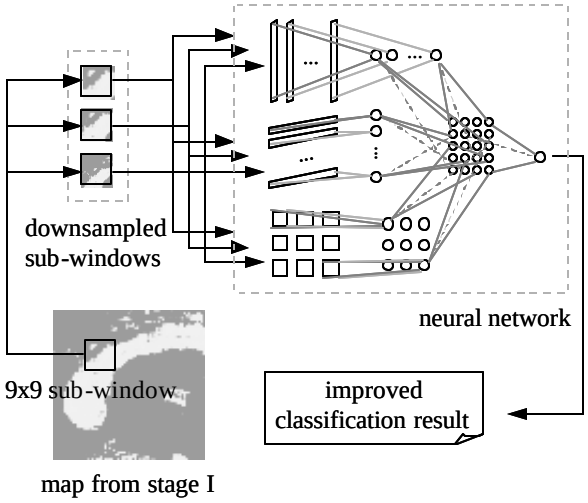

Figure 3. Stage II: local shape/texture analysis.

Learning process. A major problem of neural networks when used as classifiers lies in their lack of good rejection capabilities: a neural network has to assign every input feature vector to a class even though some vectors may not belong to any of the learnt classes. During the learning phase, classification boundaries are constructed that may acutely depend on the distribution of samples in the learning set. In our case, it is easy to select learning patterns belonging to the set of textures of the target anatomical structure and much more difficult to give a "representative" set of those which do not belong to it.

By using a separate second texture classification stage, we can design a learning set that lies along the classification boundary. This technique is similar to those described in [15] and [8].

We insert in the learning set 90% of the misclassified points from the outcome of stage I and an equal number of correctly classified points drawn at random across the entire set of images.

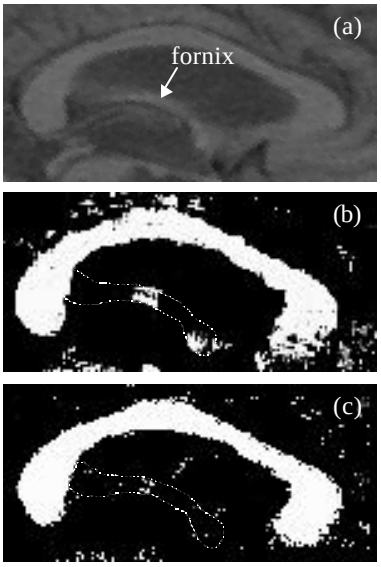

Figure 4. Texture segmentation of *corpus callosum*. The second stage (c) successfully manages to remove the *fornix* that was not completely erased in the first one (b).

Figure 4 illustrates one feature of interest in our approach. In 4(b), the *fornix*, which is not part of the *corpus callosum*, has not been properly erased by stage I as its texture closely resembles that of the target *corpus callosum*. The second stage readily removes it (4.c) and additionally smoothes the contours of the *callosum*.

III. RESULTS

We first applied our algorithms to the segmentation of the *corpus callosum*, *caudate nucleus* and *hippocampus* in T1-weighted MRI data (1mm isotropic resolution) of the brain. For each structure, the learning set consists of 5 volumetric brain MRI datasets taken from 5 different individual, at the same relative position.

Figure 5 shows the outcome of each stage for 3 additional test images (which do not belong to the learning set). As the texture of *corpus callosum* is quite different from that of the surrounding structures (with the exception of the *fornix*), stage I has no difficulty segmenting it. Stage II further regularizes the boundary. Notice how the second stage also successfully manages to remove most of the structures and tissues surrounding the *caudate nucleus* in spite of the poor contrast in the input MRI. The salt and pepper noise left around *hippocampus* can easily be removed by morphological cleaning.

Table I gives the classification rates (ratio of the number of correctly classified pixels over the total number of pixels) of the two consecutive stages for these 3 structures. We report the average rates computed over 10 test images (which do not belong to the learning set), within the windows of Figure 5 (results are given in the form mean ± standard deviation). Stage II not only substantially increases the performance of the classifier but also makes it more consistent as the standard deviations are significantly smaller. Note that the combined use of our 2 stages enables a better classification mostly in the immediate neighborhood of the structures, as depicted in Figure 5. Our approach could be most effectively used as a pre-processing step in a boundary segmentation application.

TABLE I. CLASSIFICATION RATES FOR BOTH STAGES

	caudate nucleus	hippocampus	corpus callosum
Stage I	0.81 ± 0.05	0.86 ± 0.08	0.94 ± 0.04
Stage II	0.90 ± 0.02	0.91 ± 0.03	0.98 ± 0.01

We then compared the hybrid architecture of stage I to a specialized mixture of experts (SME) without master classifier. Table II reports the classification rate for both stages. Comparison between the first rows of tables I and II demonstrates the superiority of the hybrid approach, especially on difficult structures such as *hippocampus*. Incidentally, it appears that the second stage somewhat reduces the performance differences between the methods: by using a separate learning set, it manages to compensate for the increased number of mistakes.

TABLE II. CLASSIFICATION RATES FOR SME

	caudate nucleus	hippocampus	corpus callosum
Stage I	0.77 ± 0.08	0.75 ± 0.12	0.92 ± 0.02
Stage II	0.88 ± 0.03	0.89 ± 0.04	0.95 ± 0.02

IV. CONCLUSION

We have presented an approach to segment anatomical structures from MR images via texture classification.

The hybrid master classifier / mixtures of experts architecture we developed shows good promises in that it achieves both better classification results and a faster convergence during the learning phase.

By decomposing the classification process into two consecutive sub-processes, we also give the system a greater tolerance to mistakes as the second texture classification stage can correct the errors of the first one.

Finally, it is of additional interest to analyze the neural architecture that a given anatomical structure yields (both in the first and second stages). By understanding the relationships between the complexity of the target structure (texture composition, shape variability, etc.) it may be possible to create even better systems.

ACKNOWLEDGMENTS

This work was supported by research grants from the National Center for Research Resources (P41 RR13642), National Library of Medicine (LM/MH05639), NINDS (NS38753) and by a Human Brain Project grant known as the International Consortium for Brain Mapping, which is funded jointly by NIMH and NIDA (P20 MH/DA52176).

REFERENCES

[1] J.C. Bezdek, L.O. Hall, M. Clark, D. Goldgof and L.P. Clarke, "Segmenting Medical Images with Fuzzy Models: an Update," *In: R. Yager, D. Dubois and H. Prade, Fuzzy Set Methods in Information Engineering: A Guided Tour of Applications*, John Wiley and Sons, 1996.

[2] S. Guarnieri, F. Piazza and A. Uncini, "Multilayer Feedforward Networks with Adaptive Spline Activation Function," *IEEE Trans. on Neural Networks*, vol. 10, n° 3, 1999, pp. 672-683.

[3] R.M. Haralick, "Statistical and Structural Approaches to Texture," *Proc. IEEE*, vol. 67, n° 5, 1979, pp. 786-804.

[4] Y.S. Huang, K. Liu and C.Y. Suen, "The Combination of Multiple Classifiers by a Neural Network Approach," *Int'l Journal of Pattern Recognition and Artificial Intelligence*, vol. 9, 1995, pp. 579-597.

[5] A.K. Jain and F. Farrokhnia, "Unsupervised Texture Segmentation Using Gabor Filters," *Pattern Recognition*, vol. 24, n° 12, 1991, pp. 167-186.

[6] A.K. Jain and K. Karu, "Learning Texture Discrimination Masks," *Trans. on PAMI*, vol. 18, n° 2, 1996, pp. 195-205.

[7] M.I. Jordan and R.A. Jacobs, "Hierarchical Mixtures of Experts and the EM Algorithm," *Neural Computation*, vol. 6, 1994, pp. 181-214.

[8] B. Kamgar-Parsi and B. Kamgar-Parsi, "Rejection with Multiplayer Neural Networks: Automatic Generation of the Training Set," *Proc. World Congress on Neural Networks*, vol. 2, 1995, pp. 174-177.

[9] J. Mao, K. Mohiuddin and A.K. Jain, "Minimal Network Design and Feature Selection through Node Pruning," *Proc. 12th Int'l Conf. Pattern Recognition*, vol. 2, 1994, pp. 622-624.

[10] A.M. Martinez and A.C. Kak, "PCA versus LDA," *IEEE Trans. on PAMI*, vol. 23, n° 2, 2001, pp. 228-233.

[11] T. McInerney and D. Terzopoulos. "Deformable Models in Medical Image Analysis: a Survey," *Medical Image Analysis*, vol. 1, n° 2, 1996, pp. 91-108.

[12] I. Ng, T. Tan and J. Kittler, "On Local Linear Transform and Gabor Filter Representation of Texture," *Proc. Int'l Conf. Pattern Recognition*, 1992, pp. 627-631.

[13] T. Randen and J.H. Husøy, "Filtering for Texture Classification: a Comparative Study," *IEEE Trans. on PAMI*, vol. 21, n° 4, 1999, pp 291-310.
[14] F. Rosenblatt, "The Perceptron: a Perceiving and Recognizing Automaton," *Report 85-460-1*, Project PARA, Cornell Aeronautical Laboratory, Ithaca, New York, 1957.
[15] H.A. Rowley, S. Baluja and T. Kanade, "Neural Network-Based Face Detection," *IEEE Trans. on PAMI*, vol. 20, 1998, pp. 23-38.
[16] J. Lévy-Véhel and B. Guiheneuf, "Texture Based Video Indexing," *Proc. IASTED*, 2000.
[17] P. Walter, I. Elsen, H. Muller and K.-F. Kraiss, "3D Object Recognition with a Specialized Mixtures of Experts Architecture," *Proceedings of IJCNN '99*, pp. 3563-3568, 1999.
[18] L. Xu, A. Krzyzak and C.Y. Suen, "Associative Switch for Combining Multiple Classifiers," *Journal of Artificial Neural Networks*, vol. 1, pp. 77-100, 1994.

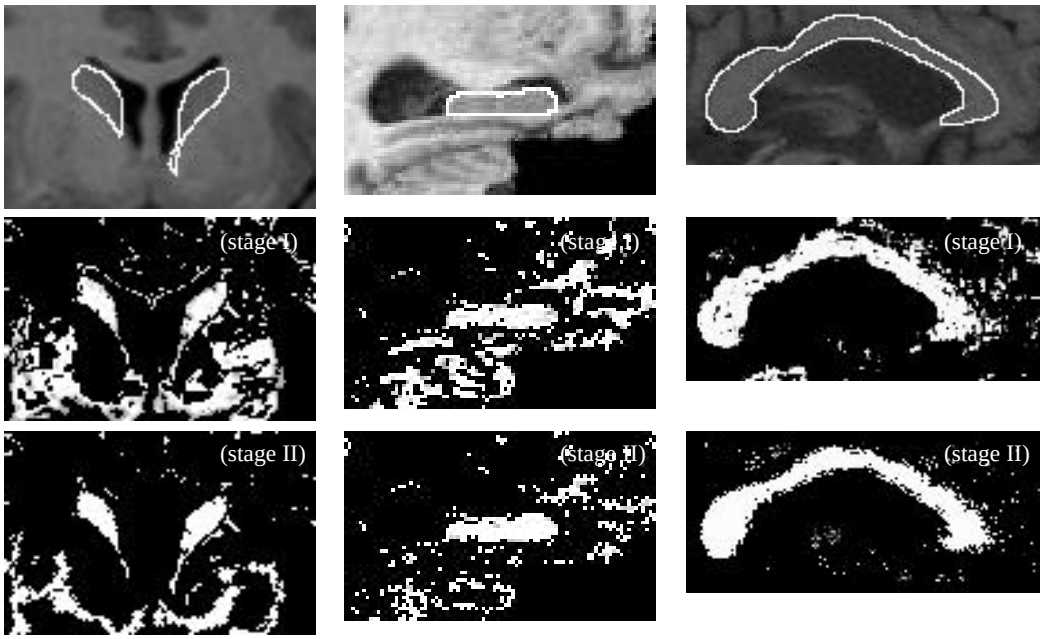

Figure 5. Outputs of the two sequential stages of our approach for 3 anatomical structures in the brain: the *caudate nucleus* (left), *hippocampus* (middle) and *corpus callosum* (right).