Automatic Design of Window Operators for the Segmentation of the Prostate Gland in Magnetic Resonance Images

Marco E. Benalcázar[1,2,3], Marcel Brun[1], and Virginia Ballarin[1]

[1] Grupo de Procesamiento Digital de Imágenes, Universidad Nacional de Mar del Plata (UNMDP), Argentina.
[2] Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT), Ecuador.
[3] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.

Abstract— **W-operators are nonlinear image operators that are translation invariant and locally defined inside a finite spatial window. In this work, we consider the problem of automatic design of W-operators for the segmentation of magnetic resonance (MR) volumes as a problem of classifier design. We propose to segment the objects of interest in an MR volume by classifying each pixel of its slices as either part of the objects of interest or background. The classifiers used here are the artificial feed-forward neural networks. The proposed method is applied to the segmentation of the two main regions of the prostate gland: the peripheral zone and the central gland. Performance evaluation was carried out on the volumes of the Prostate-3T collection of the NCI-ISBI 2013 Challenge. The results obtained show the suitability of our approach as a marker detector of the prostate gland.**

Keywords— **W-operator, segmentation, magnetic resonance, prostate gland, feed-forward neural network.**

I. Introduction

Mathematical morphology is one of the most important techniques used to design nonlinear operators for image processing and image analysis. Based on this technique, we can design complex and powerful image analysis and processing tools, called morphological operators, by combining two fundamental operators: erosions and dilations [1]. The most common approach used to design morphological operators is to combine these operators heuristically [1, 2]. One of the main issues with this approach is finding the appropriate number and the best sequence of operations to achieve good performance. A breakthrough in this area is shifting from the heuristic-based approach to the statistical design using training examples, pattern recognition theory, and machine learning systems [3, 4].

Under the statistical approach, the family of morphological operators consists of those Window operators, or W-operators, that are locally defined by a finite spatial window and invariant to translations. In this context, the problem of statistical design of W-operators using training examples is reduced to the problem of finding the optimal discrete classifier (or one close to the optimal) on a finite set of random variables [3-5]. The main characteristic of W-operators is to label a given pixel based only on the values observed within a given window neighborhood. The pixel to process is usually the center pixel of the window [5]. The training examples used to design W-operators are composed of pairs of observed and ideal images. Observed images are samples of the problem to solve; whereas, ideal images represent the desired output after processing [3].

The image segmentation problem consists of partitioning the set of pixels of a given image into two or more disjoint subsets [1, 2]. The similarities among pixels in the same subset are stronger than the similarities among pixels from different subsets. The segmentation problem of an image can be seen as a classification problem of its pixels in the field of pattern recognition [3, 5]. In this context, we label each pixel of a given image either as part of the object of interest or as background based on a window observation for that pixel. In this context, W-operators are well suited for image segmentation as long as enough discriminative information can be captured by a finite spatial window. The size and shape of the window are usually chosen *a priori* based on the characteristics of the objects to segment (e.g., size, shape, texture) [3].

The complexity of the design of W-operators depends mainly on three factors: the number of gray levels in the observed images l, the number of gray levels in the ideal images m, and the size of the window to be used n [3]. For a given problem, the search space for the best W-operator is composed of m^{l^n} hypotheses or possible operators. Thus, increasing the size of the window causes an exponential increase of the search space. For example, for binary segmentation ($m = 2$) of grayscale images with $l = 256$ gray levels, using a small window of 3×3 pixels the search space has $2^{256^9} \approx 10^{1.5*10^{21}}$ hypotheses. For the segmentation of medical images, which is the focus of this paper, we usually need larger windows (e.g., 15×15) [5, 6].

To overcome the high computational complexity of designing W-operators, aperture filters are proposed [4]. These filters constrain both the domain of the image to process and the range of the pixels of the window observation from l to $2k+1$ gray levels, where $k<<l$. Even in these circumstances, finding the best filter is a very difficult task because

the search space is still huge [3]. In our example, if we set $k = 5$, the search space has $m^{(2k+1)^n} = 2^{11^9} \approx 10^{2.35*10^9}$ possible filters. Several approaches have been proposed to design W-operators based on aperture observations including the use of decision trees, multi-mask filters, and pyramidal multiresolution [3-5]. The main issue with these approaches is their high computational cost for searching and representing the best operators for problems with real data.

To make feasible the design of W-operators for solving real problems, our approach is to constrain the search space to those operators that can be implemented by a well known machine learning system: artificial feed-forward neural networks (NNs) [7]. In this case, the search space is given by the set of functions that a given NN architecture can implement. The total number of filters we can implement is given by m^{VC}, where VC denotes the Vapnik-Chervonenkis (VC) dimension that measures the diversity or richiness of a given set of functions [7].

Here, we propose to segment the prostate gland in T2-weighted magnetic resonance images (MRIs) [6] based on ensembles of W-operators designed with feed-forward NNs. This segmentation task consists of detecting the two major parts of the prostate: the peripheral zone and the central gland. The prostate anatomy resembles an ice cream cone with a scoop of ice cream, where the cone is the peripheral zone and the ice cream is the central gland [8]. Some of the ultrasound-guided biopsies of the prostate return negative results despite the high levels of the prostate-specific antigen (PSA) marker and the actual presence of cancer. The distinction between these two parts of the prostate is important because such failures have been attributed to lack of sampling the central gland [8], where about 30% of the adenocarcinomas arise [9].

The segmentation problem addressed here provides important information for several medical tasks: localization of the prostate boundaries both for radiotherapy treatment and to guide biopsy procedures; volume estimation for tracking the progress of prostate diseases; initialization of multimodal registration algorithms; and establishing regions of interest for automatic methods of detection and evaluation of prostate cancer [8, 9]. Moreover, the automatic segmentation of the prostate gland is important because its manual segmentation is a tedious and time-consuming task and requires medical expertise.

When dealing with the automatic segmentation of the prostate gland in T2-weighted MRIs, we must be aware that the pixels/voxels belonging to the peripheral zone have high intensity values. The pixels/voxels belonging to the central gland usually have lower intensity values. However, the former characteristic can change with the presence of prostate cancer or benign pathologies such as hemorrhages, prostatitis, hyperplastic nodules, and post-treatment sequelae of radiotherapy. The automatic segmentation of the prostate gland is a very challenging problem, not only because of the noise and inhomogeneities of the MRIs, but also because of the complex anatomical structures of the prostate and surrounding organs and glands. Following this introduction, in Section 2, we describe the source of images used in this work and the proposed methodology. In Section 3, we present and analyze the results, and in Section 4, we draw some conclusions and outline future work.

II. Materials and Methodology

In this section, we briefly describe the source of images and the methodology used in this paper.

A. Images

The MRIs used in this work come from the Prostate-3T collection of the NCI-ISBI 2013 Challenge [10]. This collection consists of 30 prostate transversal T2-weighted volumes and their manual segmentations. The number of slices for each volume varies from 15 to 24. Each slice contains 320×320 pixels, which take their values from the set {0,1,…,2965}. For this work, we randomly divided the dataset into two subsets: one for training and the other for testing, where each subset contains 15 volumes.

B. Automatic design of W-operators

Let a magnetic resonance (MR) *volume* be a function $\mathbf{V}$: $\boldsymbol{D} \rightarrow L$, where $\boldsymbol{D}$ is a subset of $\mathbb{Z}^3$ and $L = \{0,1,\ldots,l\text{-}1\}$ is a set containing l gray levels. A *voxel* $r = (x,y,z)$ is a 3D coordinate of $\mathbf{V}$. An MR *slice* of $\mathbf{V}$ can be seen as a grayscale image O: $\boldsymbol{E} \rightarrow L$, with $\boldsymbol{E} \subset \mathbb{Z}^2$, taken at a fixed value of z. A *pixel* $t = (x,y)$ is a 2D coordinate of the slice O. We denote the set of all possible grayscale images by $L^{\boldsymbol{E}}$. We also define a *spatial window* $W = \{w_1,\ldots,(0,0),\ldots,w_n\}$ as a small subset of $\mathbb{Z}^2$ and a *range window* $K = \{-k,\ldots,0,\ldots,k\}$ as a set of $2k + 1$ gray levels with $k \in \mathbb{Z}^+$ and $k << l$.

For the design of W-operators, we consider a pair of grayscale images (O,I), which is assumed to be generated by a random, stationary, joint process $(\boldsymbol{O},\boldsymbol{I})$. $\boldsymbol{O}$ generates observed images (i.e., original image), and $\boldsymbol{I}$ generates ideal images (i.e., desired output). A W-operator is a function $\boldsymbol{\Psi}$: $L^{\boldsymbol{E}} \rightarrow L^{\boldsymbol{E}}$ that maps grayscale to grayscale images. We define the W-operator $\boldsymbol{\Psi}$ by a classifier ψ: $\boldsymbol{\mathcal{X}} \rightarrow \{1,\ldots,c\}$ that maps a feature vector $\mathbf{X} = (X_1,\ldots,X_{2n})$ to a label from the set $\{1,\ldots,c\}$. Here, we consider $c = 3$, where the pixels of the background, the peripheral zone, and the central gland are labeled with 1, 2, and 3, respectively.

C. Feature extraction:

The feature vector **X** is composed of a window observation and an aperture observation. For an arbitrary pixel t of the slice O, a *window observation* **u** is obtained by translating W by t: $W_t = \{w_i + t: w_i \in W, i = 1,\ldots,n\}$, and then copying into the vector $\mathbf{u} = (u_1,\ldots,u_n)$ the pixel values of O inside W_t. An *aperture observation* $\mathbf{v} = (v_1,\ldots,v_n)$ is obtained by subtracting $r = median(u_1,\ldots,u_n)$ from each component of the vector **u**. Then we constrain each $u_i - r$, with $i = 1,\ldots,n$, to have the domain K. For this we apply $v_i = min(max(-k, u_i - r), k)$. Thus, for the vector **X**, we have $(X_1,\ldots,X_n) = (u_1,\ldots,u_n)$ and $(X_{n+1},\ldots,X_{2n}) = (v_1,\ldots,v_n)$. The label Y of **X** is $I(t)$ [3]. In this work, we used a squared window $W = A \times A$ composed of 19×19 points, where $A = \{-18,-16,\ldots,16,18\}$ and $k = 150$.

D. Classifier design

For classifier design, in this work we use an artificial feed-forward NN of three layers. The input layer is composed of 722 units, the hidden layer contains 100 neurons, and the output layer is composed of 3 neurons. For the neurons of the hidden layer, we used sigmoid transfer functions f: $\mathbb{R} \rightarrow [0,1]$ defined as $f(a) = 1/(1 + \exp(-a))$. The output layer is a softmax classifier [7], where each neuron predicts the conditional probability $\mathbb{P}(Y = i|\mathbf{X})$, with $i = 1,\ldots,c$, and $\Sigma_{i=1,\ldots,c}\mathbb{P}(Y = i|\mathbf{X}) = 1$. The network architecture was determined by using model selection with 5-fold cross-validation [4] based on the volumes of the training set. We used artificial feed-forward NNs because of their ability to implement complex decision boundaries [7].

We assign the label with maximum $\mathbb{P}(Y = i|\mathbf{X})$ to **X**, with $i = 1,\ldots,c$. For training the NN, we used the Newton gradient-descend method [7] to minimize $-ln(\mathbb{P}(\boldsymbol{\mathcal{D}}|\boldsymbol{\beta}))$. The function $\mathbb{P}(\boldsymbol{\mathcal{D}}|\boldsymbol{\beta}) = \prod_{i=1,..,N}\prod_{j=1,\ldots,c} \mathbb{P}(Y_i = j \mid \mathbf{X}_i)^{a_i^{(j)}}$ is the likelihood of the training set $\boldsymbol{\mathcal{D}}$ given the parameters $\boldsymbol{\beta}$ of the NN. The set $\boldsymbol{\mathcal{D}} = \{(\mathbf{X}_i, a_i^{(1)},\ldots, a_i^{(c)})\}$, with $i = 1,\ldots,N$, contains the feature vectors and the frequencies $a_i^{(j)}$ with which $\mathbf{X}_i$ was observed having $Y_i = j$ in the training images, where $j = 1,\ldots,c$,

E. Artificial balancing of the frequencies of the training set

Since the prostate gland occupies only a small region of each MRI volume, the training set $\boldsymbol{\mathcal{D}}$ is composed mostly of feature vectors belonging to the background. If we trained the NN with this unbalanced set, new feature vectors would be classified as background most of the time, leading to a small training error but poor generalization. Therefore, artificial balancing of the training set is necessary. We balanced the training set by replacing the values of $a_i^{(j)}$ with $x_i^{(j)}a_i^{(j)}$ such that $\sum_{i=1,\ldots,N} x_i^{(j)}a_i^{(j)} = \sum_{i=1,\ldots,N} x_i^{(k)}a_i^{(k)}$ for all $j,k = 1,\ldots,c$ and $i = 1,\ldots,N$. By doing so, we avoid subsampling the feature vectors of the majority class or oversampling the feature vectors of the minority classes. Subsampling methods discard information valuable for classifier training. Oversampling methods considerably increase the size of the training set, slowing down the training of classifiers [7].

III. Results and Discussion

In this section, we present the results of evaluating the performance of the proposed approach for segmentation of the prostate gland. The results shown below were computed using the 15 volumes in the testing subset described in Section II. Table 1 shows the confusion matrix obtained by comparing the voxels of the processed and the ground truth volumes. Each value of this table is normalized to the total number of processed voxels (29237248).

Table 1 Confusion matrix of the proposed method

		Actual labels (%)		
		Background	Peripheral zone	Central gland
Predicted labels (%)	Background	93.7229	0.2264	0.7224
	Peripheral zone	0.4112	0.9229	0.1299
	Central gland	1.6098	0.0599	2.1947

Table 2 shows the estimated values of precision and recall of the proposed method computed based on the values from Table 1.

Table 2 Precision and recall values of the proposed method

Metric	Value (%)
Precision (*Predicted labels* | *Actual labels*)	
Prostate gland | Prostate gland	77.7071
Central gland | Central gland	72.0275
Peripheral zone | Peripheral zone	76.3224
Recall (*Actual labels* | *Predicted labels*)	
Prostate gland | Prostate gland	62.0712
Central gland | Central gland	56.7926
Peripheral zone | Peripheral zone	63.0397

Fig. 1 shows 4 examples of slices processed with the proposed approach. In the processed and ground truth images, the peripheral zone appears in green and the central gland appears in red. Both the images of Figure 1 and the results of Tables 1 and 2 evidence that the proposed method is able to detect most of the central gland and the peripheral zone. On the other hand, the values of recall indicate that our approach also detects spurious objects as either central gland or peripheral zone.

The obtained results indicate that our method is not the final solution for the segmentation of the prostate gland in MRIs. However, it can be used as a part of a more sophisticated method of segmentation, which should include prior information about the shape of the prostate and its position relative to the surrounding organs and glands. In fact, our method could be used as a marker detector of the peripheral zone and the central gland. This is one of the main advantages of the proposed method compared with other state-of-the-art methods [11], most of which detect the prostate gland as a whole and cannot distinguish between the two regions considered here.

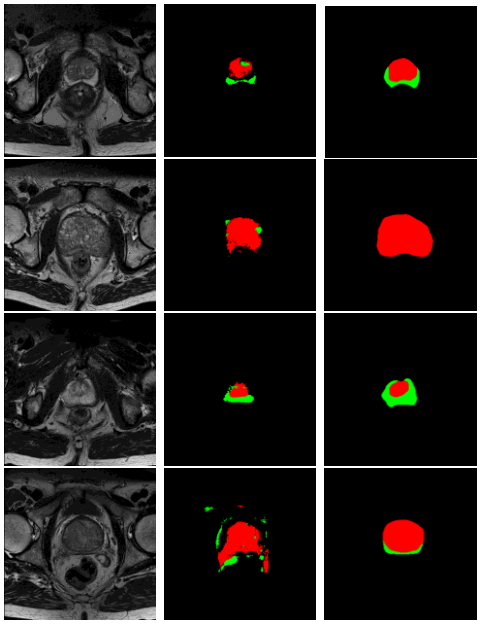

Fig. 1 Examples of original, results, and ground truth slices obtained applying the proposed method. Green = peripheral zone; Red = central gland

IV. Concluding Remarks

In this work, we have presented a new method for the segmentation of the two main parts of the prostate gland in MRI volumes, the peripheral zone and central gland. The proposed method is based on the automatic design of W-operators for the segmentation of T2-weighted MRIs. The segmentation task of MRI volumes using W-operators is addressed as a classification task of the pixels belonging to each slice that compose a given volume. Therefore, W-operator design for image segmentation consists of designing a classifier that labels each pixel of the slice to be processed as either part of the objects of interest or background.

The family of classifiers considered in this work is artificial feed-forward NNs. Each pixel is classified using the one-versus-all approach based on a feature vector composed of a window observation and an aperture observation. Results show the suitability of the proposed approach as a marker detector for a more sophisticated segmentation method.

Further work includes the development of a postprocessing method to filter spurious objects and to refine the shape of the segmented prostate gland. Additionally, performance evaluation should include metrics that take into account the 2D and 3D shapes of the prostate.

References

1. Shih F (2009) Image processing and mathematical morphology. CRC Press, New York
2. Gonzalez R, Woods R (2008) Digital image processing, Prentice Hall, New Jersey
3. Dougherty E R, Barrera J (2002) Pattern recognition theory in nonlinear signal processing. Mathematical Imaging and Vision 16:181-197
4. Hirata R, Dougherty E R, Barrera J (2000) Aperture filters. Signal Processing 80:697-721
5. Benalcázar M E, Brun M, Ballarin V L (2013) Artificial neural networks applied to statistical design of window operators. Pattern Recognition Letters 34:970-979
6. Dougherty G (2009) Digital image processing for medical applications. Cambridge Press, Cambridge
7. Bishop C M (2005) Neural Networks for Pattern Recognition. Oxford
8. Verma S, Rajesh A (2011) A clinically relevant approach to imaging prostate cancer: review. American Journal of Roentgenology 196:S1-S10
9. Claus F G, Hricak H, Hattery R R (2004) Pretreatment evaluation of prostate cancer: role of MR imaging and 1HMR spectroscopy. Radiographics 24:67-80
10. https://wiki.cancerimagingarchive.net/display/Public/Prostate-3T
11. Litjens G, Toth R, Van de Ven W et al. (2014) Evaluation of prostate segmentation algorithms for MRI: the PROMISE12 challenge. Medical Image Analysis 18:359–373

Corresponding author:
Author: Marco E. Benalcázar
Institute: Facultad de Ingeniería, UNMDP
Street: Juan B. Justo 4302
City: Mar del Plata
Country: Argentina
Email: marco_benalcazar@hotmail.com