Hidden Markov Random Fields and Direct Search Methods for Medical Image Segmentation

El-Hachemi Guerrout[1], Samy Ait-Aoudia[1], Dominique Michelucci[2] and Ramdane Mahiou[1]
[1]*Laboratoire LMCS, Ecole Nationale Supérieure en Informatique, Oued-Smar, Algiers, Algeria*
[2]*Laboratoire LE2I, Université de Bourgogne, Dijon, France*

Keywords: Medical Image Segmentation, Hidden Markov Random Field, Nelder-Mead, Torczon, Kappa Index.

Abstract: The goal of image segmentation is to simplify the representation of an image to items meaningful and easier to analyze. Medical image segmentation is one of the fundamental problems in image processing field. It aims to provide a crucial decision support to physicians. There is no one way to perform the segmentation. There are several methods based on HMRF. Hidden Markov Random Fields (HMRF) constitute an elegant way to model the problem of segmentation. This modelling leads to the minimization of an energy function. In this paper we investigate direct search methods that are Nelder-Mead and Torczon methods to solve this optimization problem. The quality of segmentation is evaluated on grounds truths images using the Kappa index called also Dice Coefficient (DC). The results show the supremacy of the methods used compared to others methods.

1 INTRODUCTION

MRI - Magnetic Resonance Imaging, CT - Computed tomography, Radiography, digital mammography, and other imaging modalities have indispensable role in disease diagnosis. However, they produce a huge number of images for which the manual analysis and interpretation has become a difficult task. The automatic extraction of meaningful information is one among the segmentation challenges. In the literature several approaches have been proposed such as active contour models (Kass et al., 1988), edge detection (Canny, 1986), thresholding (Sahoo et al., 1988), region growing (Adams and Bischof, 1994), MRF - Markov Random Fields (Manjunath and Chellappa, 1991), etc.

HMRF - Hidden Markov Random Field, a generalization of Hidden Markov Model (Baum and Petrie, 1966) provides an elegant way to model the segmentation problem. Since the seminal paper of Geman and Geman (1984), Markov Random Fields (MRF) models for image segmentation have been investigated intensively (Manjunath and Chellappa, 1991; Panjwani and Healey, 1995; Held et al., 1997; Hochbaum, 2001; Zhang et al., 2001; Deng and Clausi, 2004; Kato and Pong, 2006; Yousefi et al., 2012). The main idea underlying the segmentation process using HMRF is: the image to segment (called also the observed image) and the segmented image (called also the hidden image) are seen like Markov Random Field. The segmented image is computed according to the MAP (Maximum A Posteriori) criterion (Wyatt and Noble, 2003). MAP estimation leads to the minimization of energy function (Szeliski et al., 2008). This problem is computationally intractable.

In this paper we examine HMRF to model the segmentation problem and the direct search methods: Nelder-Mead (Nelder and Mead, 1965) and Torczon (Kolda et al., 2003), as the optimization techniques. In order to evaluate the segmentation quality we use the Kappa index criterion (Dice, 1945) also called Dice Coefficient (DC). The Kappa Index criterion gives us how much the segmentation result is close to the ground truth. The images have been obtained from the Brainweb[1] database (Cocosco et al., 1997), widely used by the neuroimaging community, where the ground truth is known.

The achieved results are very satisfactory and show a superiority of the our proposed methods: HMRF-Nelder-Mead and HMRF-Torczon, compared to other methods (Ouadfel and Batouche, 2003; Yousefi et al., 2012) that are Classical MRF, MRF-ACO (Ant Colony Optimization) and MRF-ACO-Gossiping.

This paper is organized as follows. In section 2, we provide some concepts of Markov Random Field

[1]http://www.bic.mni.mcgill.ca/brainweb/

Guerrout, E-H., Ait-Aoudia, S., Michelucci, D. and Mahiou, R.
Hidden Markov Random Fields and Direct Search Methods for Medical Image Segmentation.
DOI: 10.5220/0005658501540161
In *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2016)*, pages 154-161
ISBN: 978-989-758-173-1

model. Section 3 is devoted to Hidden Markov Field model and direct search methods combination to perform the segmentation. We give in section 4 some experimental results. Finally, the section 5 is devoted to conclusions.

2 HIDDEN MARKOV RANDOM FIELD MODEL (HMRF)

2.1 Markov Random Field

Let $X = \{X_1, X_2, \ldots, X_M\}$ be a family of random variables on the lattice S. Each random variable takes values in the discrete space $\Lambda = \{1, 2, \ldots, K\}$. The family X is a random field with configuration set $\Omega = \Lambda^M$.

A random field X is said to be an MRF (Markov Random Field) on S, with respect to a neighborhood system $V(S)$, if and only if:

$$\begin{cases} \forall x \in \Omega, P[X = x] > 0 \\ \forall s \in S, \forall x \in \Omega, P[X_s = x_s | X_t = x_t, t \neq s] = \\ \qquad P[X_s = x_s | X_t = x_t, t \in V_s(S)] \end{cases} \quad (1)$$

The Hammersley-Clifford theorem (Hammersley & Clifford, 1971) establishes the equivalence between Gibbs fields and Markov ones. The Gibbs distribution is characterized by the following relations:

$$\begin{cases} P[X = x] = Z^{-1} e^{\frac{-U(x)}{T}} \\ Z = \sum_{\xi \in \Omega} e^{\frac{-U(\xi)}{T}} \\ U(x) = \sum_{c \in C} U_c(x) \end{cases} \quad (2)$$

where T is a global control parameter, called temperature, and Z is a normalizing constant, called the partition function. Calculating Z is prohibitive. $Card(\Omega) = 256^{512 \times 512} = 2^{2097152}$ for a 512x512 gray level image. $U(x)$ is the energy function of the Gibbs field defined as a sum of potentials over all the possible cliques C.

The local interactions between the neighboring sites properties (gray levels for example) can be expressed as a clique potential.

2.2 Standard Markov Random Field

2.2.1 Ising Model

This model was proposed by Ernst Ising for ferromagnetism studies in statistical physics. The Ising model involves discrete variables s (spins), placed on a sampling grid. Each spin can take two values, $\Lambda = \{-1, 1\}$. The spins interact in pairs. The first order clique potential is defined by $-Bx_s$ and the second order clique potential is defined by:

$$U_{c_2=\{s,t\}}(x_s, x_t) = \begin{cases} -\beta & \text{if} \quad x_s = x_t \\ +\beta & \text{if} \quad x_s \neq x_t \end{cases} \quad (3)$$

$$U_{c_2=\{s,t\}}(x_s, x_t) = -\beta x_s x_t \quad (4)$$

The total energy is defined by:

$$U(x) = -\sum_{c_2=\{s,t\}} \beta x_s x_t - \sum_{c_1=\{s\}} B x_s \quad (5)$$

The coupling constant β, between neighboring sites, regularize the model and B represents an extern magnetic field.

2.2.2 Potts Model

The Potts model is a generalization of the Ising model. Instead of $\Lambda = \{-1, 1\}$, each spin is assigned an integer value $\Lambda = \{1, 2, \ldots, K\}$. In the context of image segmentation, the integer values are gray levels or labels. The total energy is defined by:

$$U(x) = \beta \sum_{c_2=\{s,t\}} (1 - 2\delta(x_s, x_t)) \quad (6)$$

where δ is the Kronecker's delta:

$$\delta(a, b) = \begin{cases} 1 & \text{if} \quad a = b \\ 0 & \text{if} \quad a \neq b \end{cases} \quad (7)$$

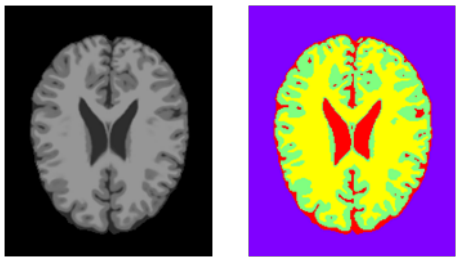
When $\beta > 0$, the probable configurations correspond to neighboring sites with same gray level or label. This induces the constitution of large homogeneous regions. The size of these regions is guided by the value of β.

2.3 Hidden Markov Random Field

HMRF is a strong model for image segmentation. The image to segment is seen as a realization $y = \{y_s\}_{s \in S}$ of a Markov Random Field $Y = \{Y_s\}_{s \in S}$ defined on the lattice S. Each realization y_s of the random variable Y_s takes its values in the space of gray levels $\Lambda_{obs} = \{0 \ldots 255\}$. The configuration set of Y is noted as Ω_{obs}. The segmented image is seen as the realization $x = \{x_s\}_{s \in S}$ of another Markov Random Field $X = \{X_s\}_{s \in S}$, defined on the same lattice S. The realization x_s of the random variable X_s takes its values in the discrete space $\Lambda = \{1, 2, \ldots, K\}$. K is the number of classes or homogeneous regions in the image. The configuration set of X is noted as Ω.

Figure 1 shows an example of the image to segment as a realization y of Y and its segmented image with $K = 4$ seen as a realization x of X.

In the context of image segmentation, we face a problem with incomplete data (Dempster et al., 1977). To every site $s \in S$ is associated different information:



y: Observed Image x: Hidden Image

Figure 1: Observed and hidden image.

observed information expressed by the random variable Y_s; missed or hidden information, expressed by the random variable X_s. The Random Field X is called Hidden Markov Random Field.

The segmentation process consists in finding a realization x of X by observing the data of the realization y, representing the image to segment.

We seek a labeling $\hat{x}$, which is an estimate of the true labeling x^*, according to the MAP (Maximum A Posteriori) criterion (maximizing the probability $P(X = x|Y = y)$).

$$x^* = \arg\max_{x\in\Omega} \{P[X = x|Y = y]\} \quad (8)$$

$$P[X = x|Y = y] = \frac{P[Y = y|X = x]\,P[X = x]}{P[Y = y]} \quad (9)$$

The first term of the numerator describes the probability to observe the image y, knowing the x label. Based on the assumption of conditional independence, we get the following formula:

$$P[Y = y|X = x] = \prod_{s\in S} P[Y_s = y_s|X_s = x_s] \quad (10)$$

The second term of the numerator describes the existence of the labeling x. The denominator is constant and independent of x. We have then:

$$P[X = x|Y = y] = AP[Y = y|X = x]\,P[X = x] \quad (11)$$

where A is a constant, from the equation 2 and 11 we will have:

$$P[X = x|Y = y] = Ae^{\ln(P[Y=y|X=x]) - \frac{U(x)}{T}} \quad (12)$$

$$P[X = x|Y = y] = Ae^{-\Psi(x,y)} \quad (13)$$

$$x^* = \arg\min_{x\in\Omega} \{\Psi(x,y)\} \quad (14)$$

Maximizing the probability $P(X = x|Y = y)$ is equivalent to minimizing the function $\Psi(x,y)$.

$$\Psi(x,y) = -\ln(P[Y = y|X = x]) + \frac{U(x)}{T} \quad (15)$$

We replace the equation 10 in the equation 15 and we obtain:

$$\Psi(x,y) = -\ln\left(\prod_{s\in S} P[Y_s = y_s|X_s = x_s]\right) + \frac{U(x)}{T} \quad (16)$$

$$\Psi(x,y) = -\sum_{s\in S} \ln(P[Y_s = y_s|X_s = x_s]) + \frac{U(x)}{T} \quad (17)$$

We assume that the variable $[Y_s = y_s|X_s = x_s]$ follows a Gaussian distribution with parameters μ_k (mean) and σ_k (standard deviation). By giving the class label $x_s = k$, we have:

$$P[Y_s = y_s|X_s = k] = \frac{1}{\sqrt{2\pi\sigma_k^2}} e^{\frac{-(y_s-\mu_k)^2}{2\sigma_k^2}} \quad (18)$$

The Potts model is often used in image segmentation to privilege large regions in the image. The energy is then:

$$U(x) = \beta \sum_{c_2=\{s,t\}} (1 - 2\delta(x_s, x_t)) \quad (19)$$

According to equations 17, 18 and 19 we get:

$$\Psi(x,y) = \sum_{s\in S} G_s + \frac{\beta}{T} \sum_{c_2=\{s,t\}} L_{s,t} \quad (20)$$

$$G_s = [\ln(\sigma_{x_s}) + \frac{(y_s - \mu_{x_s})^2}{2\sigma_{x_s}^2}]$$

$$L_{s,t} = (1 - 2\delta(x_s, x_t))$$

Optimization techniques to get the estimation $\hat{x}$ of the labeling x^* are presented in next sections.

3 HMRF AND DIRECT SEARCH METHODS

Applying the optimization techniques is not obvious. Thus this section first gives the big picture of our approach.

Let $y = (y_1, \ldots, y_s, \ldots, y_M)$ be the image to segment into K classes. Instead of looking for the segmented image $x = (x_1, \ldots, x_s, \ldots, x_M)$ we look for the vertex μ which contains the means of K classes $\mu = (\mu_1, \ldots, \mu_j, \ldots, \mu_K)$. The segmented image $x = (x_1, \ldots, x_s, \ldots, x_M)$ is computed by classifying y_s to nearest mean μ_j *i.e.,* if the nearest mean of y_s is μ_j then $x_s = j$, in the case of tie between two means then y_s is classified to the smallest mean.

Let S_j be the set all the sites $s \in S$ in which the nearest mean to y_s is μ_j.

$$S_j = \{s \in S \mid x_s = j \Leftrightarrow \text{the nearest mean of } y_s \text{ is } \mu_j\}$$

Let $|S_j|$ be the cardinal of S_j. The standard deviation σ_j is calculated as follow:

$$\sigma_j = \sqrt{\frac{1}{|S_j|}\sum_{s\in S_j}(y_s-\mu_j)^2}$$

The image to segment y is known, so if we know the means μ_j, then we can compute: its standard deviation $\sigma = (\sigma_1,\ldots,\sigma_j,\ldots,\sigma_K)$, the segmented image x and the value of the function $\Psi(x,y)$.

To segment the image, we look for the point $\mu = (\mu_j, j = 1\ldots K)$ where the function Ψ is minimal. This time, Ψ is seen as a function of μ:

$$\Psi(\mu) = \sum_{j=1}^{K}\sum_{s\in S_j} G_{j,s} + \frac{\beta}{T}\sum_{c_2=\{s,t\}} L_{s,t} \qquad (21)$$

$$G_{j,s} = [\ln(\sigma_j) + \frac{(y_s-\mu_j)^2}{2\sigma_j^2}]$$

$$L_{s,t} = (1-2\delta(x_s,x_t))$$

3.1 Direct Search Methods

In this section we explain what we need to apply the direct search methods for solving our problem defined above.

The popular Nelder-Mead method or the simplex-based direct search method was proposed by John Nelder and Roger Mead (1965). The method starts with $(n+1)$ vertices in $\mathbb{R}^n$ that are viewed as the vertices of a simplex. The process of minimizing is based on the comparison of the function values at $(n+1)$ vertices of the simplex, followed by replacing the vertex which has the highest value of the function by another vertex that can be obtained by operations of reflection, expansion or contraction relatively to the center of gravity of the n best vertices of the simplex. When the substitution conditions are not satisfied, the method will make the simplex shrink.

Torczon method (1989) (Kolda et al., 2003) comes to fix gaps of Nelder-Mead like degenerated flat simplices. In the Torczon method, operations are applied to all vertices of the current simplex, thus by construction, all simplexes in Torczon method are homothetic to the first one, and no degeneracy (*i.e.,* ., flat simplex) can occur.

The stopping criterion in direct search methods is satisfied when the simplex vertices or their function values are close.

In our case, we are looking for the best $\mu \in [0\ldots255]^K$ which minimizes the function Ψ. For that we start with a non degenerate simplex with $K+1$ vertices (coordinates are given in Table 1). When some vertex is out the bounds (*i.e.,* $\notin [0\ldots255]^K$), the function Ψ is set to $+\infty$ (in the practice a very big value).

3.2 HMRF and Nelder-Mead Combination

We summarize hereafter the Nelder-Mead method and give examples in two-dimensional space.

HMRF-Nelder-Mead Algorithm.

Repeat.

1. Evaluate
   Compute $\Psi_i := \Psi(V_i)$
   Determine the indices *h, s, l*:
   $\Psi_h := \max_i(\Psi_i)$, $\Psi_s := \max_{i\neq h}(\Psi_i)$, $\Psi_l := \min_i(\Psi_i)$
   Compute $\bar{V} := \frac{1}{K}\sum_{i\neq h} V_i$
   Example in $\mathbb{R}^2$ (See figure 2)

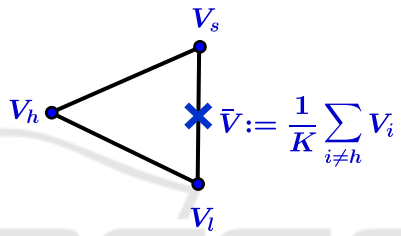


Figure 2: Center of gravity calculation.

2. Reflect
   Compute the reflection vertex V_r from
   $$V_r := 2\bar{V} - V_h$$
   Evaluate $\Psi_r := \Psi(V_r)$. If $\Psi_l \leq \Psi_r < \Psi_s$, replace V_h by V_r and terminate the iteration.
   Example in $\mathbb{R}^2$ (See figure 3)

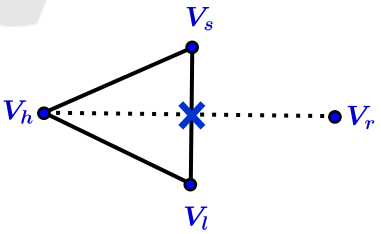


Figure 3: Reflection.

3. Expand
   If $\Psi_r < \Psi_l$, compute the expansion vertex V_e from
   $$V_e := 3\bar{V} - 2V_h$$
   Evaluate $\Psi_e := \Psi(V_e)$.
   If $\Psi_e < \Psi_r$ replace V_h by V_e and terminate the iteration; otherwise (if $\Psi_e \geq \Psi_r$), replace V_h by V_r and terminate the iteration.
   Example in $\mathbb{R}^2$ (See figure 4)

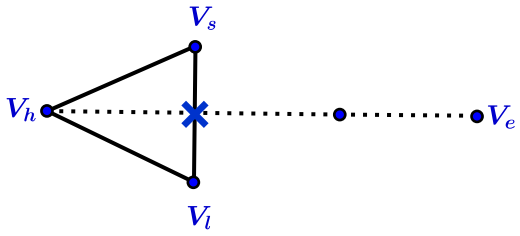


Figure 4: Expansion.

4. Contract
   If $\Psi_r \geq \Psi_s$, compute a contraction between $\bar{V}$ and the best of V_h and V_r.

   (a) Outside contraction
   If $\Psi_r < \Psi_h$, compute an outside contraction V_c from
   $$V_c := \frac{3}{2}\bar{V} - \frac{1}{2}V_h$$
   Evaluate $\Psi_c := \Psi(V_c)$.
   If $\Psi_c < \Psi_r$ replace V_h by V_c and terminate the iteration; otherwise (If $\Psi_c \geq \Psi_r$), go to the step 5 (shrink).
   Example in $\mathbb{R}^2$ (See figure 5)

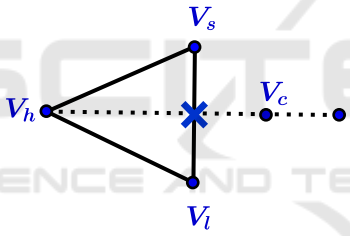


Figure 5: Outside contraction.

   (b) Inside contraction
   If $\Psi_h \leq \Psi_r$, compute an inside contraction V_c from
   $$V_c := \frac{1}{2}(V_h + \bar{V})$$
   Evaluate $\Psi_c := \Psi(V_c)$.
   If $\Psi_c < \Psi_h$ replace V_h by V_c and terminate the iteration; otherwise (If $\Psi_c \geq \Psi_h$), go to the step 5 (shrink).
   Example in $\mathbb{R}^2$ (See figure 6)

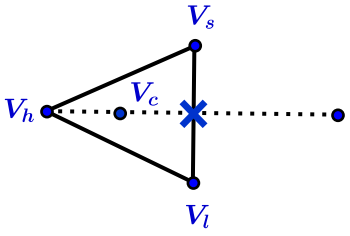


Figure 6: Inside contraction.

5. Shrink
   Replace all vertices according to the following formula $V_i := \frac{1}{2}(V_i + V_l)$
   Example in $\mathbb{R}^2$ (See figure 7)

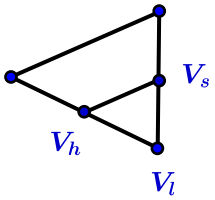


Figure 7: Shrink.

Until Satisfying a Stopping Criterion.

3.3 HMRF and Torczon Combination

We summarize hereafter the Torczon method and give examples in two-dimensional space.

HMRF-Torczon Algorithm.

Repeat.

1. Evaluate
   Compute $\Psi_i := \Psi(V_i)$
   Determine the indices l from $\Psi_l := \min_i(\Psi_i)$
2. Reflect
   Compute the reflected vertices, $V_i^r := 2V_l - V_i$
   Evaluate $\Psi_i^r := \Psi(V_i^r)$
   If $\min_i\{\Psi_i^r\} < \Psi_l$ go to step 3; otherwise, go to the step 4
   Example in $\mathbb{R}^2$ (See figure 8)

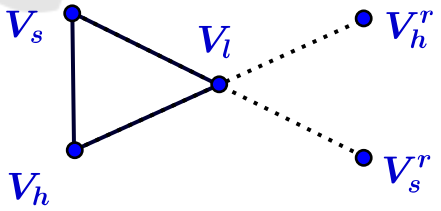


Figure 8: Reflection.

3. Expand
   Compute the expanded vertices, $V_i^e = 3V_l - 2V_i$
   Evaluate $\Psi_i^e := \Psi(V_i^e)$
   If $\min_i\{\Psi_i^e\} < \min_i\{\Psi_i^r\}$, replace all vertices V_i by the expanded vertices V_i^e; otherwise replace all vertices V_i by the reflected vertices V_i^r. In either case, terminate the iteration.
   Example in $\mathbb{R}^2$ (See figure 9)

Table 1: Parameters used in our tests.

Method	Parameters
Classical MRF	T: Temperature=4
MRF-ACO	T: Temperature=4, a: Pheromone info. Influence=1, b: Heuristic info. Influence=1, q: Evaporation rate=0.1, w: Pheromone decay coefficient=0.1
MRF-ACO-Gossiping	T: Temperature=4, a: Pheromone info. Influence=1, b: Heuristic info. Influence=1, q: Evaporation rate=0.1, w: Pheromone decay coefficient=0.1, c_1: Pheromone reinforcing coefficient=10, c_2: Pheromone reinforcing coefficient=100
HMRF-Nelder-Mead and HMRF-Torczon	T=4, $\beta = 1$, vertices of the initial simplex are $\begin{cases} V_1 = (196,127,127,127) \\ V_2 = (127,\ 63,127,127) \\ V_3 = (127,127,127,127) \\ V_4 = (127,127,188,127) \\ V_5 = (127,127,127,195) \end{cases}$

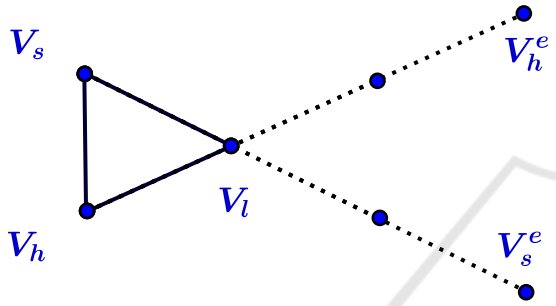


Figure 9: Expansion.

4. Contract
   Compute the contracted vertices, $V_i^c = \frac{1}{2}(V_i + V_l)$
   Replace all vertices V_i by the contracted vertices V_i^c
   Example in $\mathbb{R}^2$ (See figure 10)

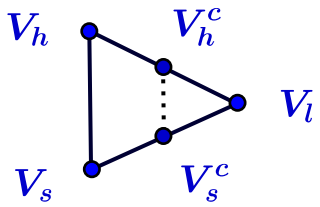


Figure 10: Contraction.

Until Satisfying a Stopping Criterion.

4 EXPERIMENTAL RESULTS

To assess the two investigated combination methods referred to as HMRF-Nelder-Mead and HMRF-Torczon, we made a comparative study with three segmentation algorithms operating on brain images that are Classical MRF, MRF-ACO (Ant Colony Optimization) and MRF-ACO-Gossiping (Yousefi et al., 2012). To perform a meaningful comparison, we use the same medical images with ground truth from Brainweb database. The comparison is based on the Kappa Index (or Dice coefficient). We give in table 1 parameters setting for each tested algorithm. The Dice Coefficient DC (Dice, 1945) or Kappa Index KI given hereafter allows the visualization of the performance of an algorithm. The Dice coefficient between two classes A and B equals 1 when they are identical and 0 in the worst case, i.e. no match between A and B.

$$KI = \frac{2|A \cap B|}{|A \cup B|} = \frac{2TP}{2TP + FP + FN} \quad (22)$$

where TP stands for true positive, FP for false positive and FN for false negative (See figure 11).

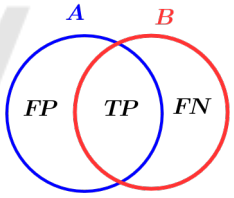


Figure 11: TP, FP and FN.

The tables 2 and 3 show respectively the mean segmentation time and the mean Kappa Index values of the three classes: GM (Grey Matter), WM (White Matter), CSF (Cerebro Spinal Fluid) . We have used the Brainweb database with the parameters: Modality= T1, Slice thickness = 1mm, Noise = 0% and Intensity non-uniformity = 0%. The slices chosen are used in (Yousefi et al., 2012) which are: 85, 88, 90, 95, 97, 100, 104, 106, 110, 121 and 130.

Figure 12 shows the segmentation result of HMRF-Nelder-Mead and HMRF-Torczon methods on a sample of BrainWeb database.

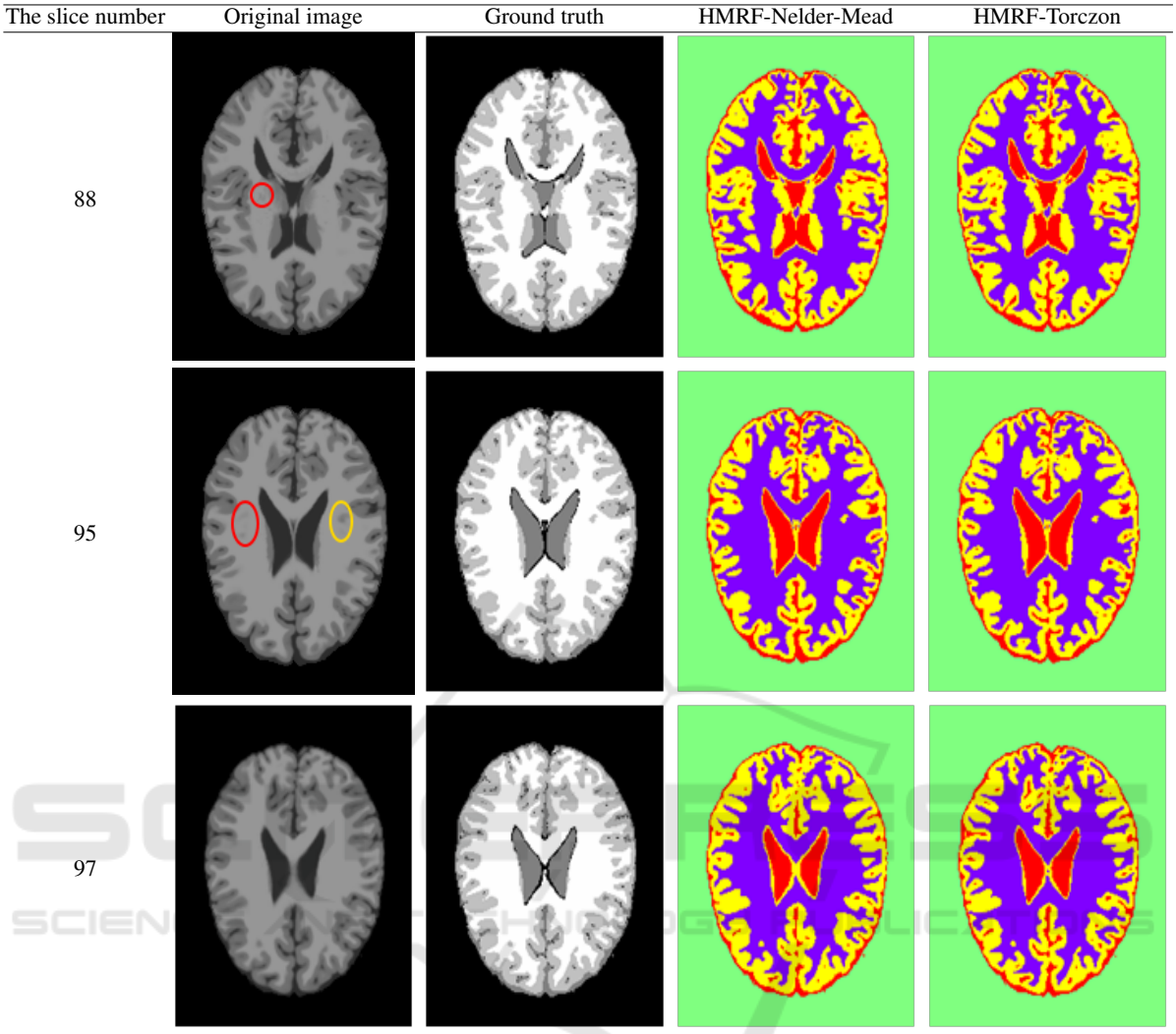


Figure 12: Segmentation result of HMRF-Nelder-Mead and HMRF-Torczon methods on a sample of BrainWeb database.

Table 2: The mean segmentation time.

Methods	Time (s)
Classical-MRF	3318
MRF-ACO	418
MRF-ACO-Gossiping	238
HMRF-Nelder-Mead	12.24
HMRF-Torczon	5.55

Table 3: The mean Kappa Index Values.

Methods	Kappa Index			
	GM	WM	CSF	Mean
Classical-MRF	0.763	0.723	0.780	0.756
MRF-ACO	0.770	0.729	0.785	0.762
MRF-ACO-Gossiping	0.770	0.729	0.786	0.762
HMRF-Nelder-Mead	0.952	0.975	0.939	0.955
HMRF-Torczon	0.975	0.985	0.956	0.973

5 CONCLUSION

In this paper, we have described methods that combine Hidden Markov Random Fields and Direct Search methods that are Nelder-Mead and Torczon optimisation algorithms to perform segmentation. Performance evaluation was carried out on sample medical images from the Brainweb database. From the tests we have conducted, the combination methods outperform classical methods. The results are very promising. Nevertheless, the opinion of specialists must be considered in the evaluation when no ground truth is available to have a more synthetic view of the whole segmentation process.

REFERENCES

Adams, R. and Bischof, L. (1994). Seeded region growing. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 16(6):641–647.

Baum, L. E. and Petrie, T. (1966). Statistical inference for probabilistic functions of finite state markov chains. *The annals of mathematical statistics*, pages 1554–1563.

Canny, J. (1986). A computational approach to edge detection. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, (6):679–698.

Cocosco, C. A., Kollokian, V., Kwan, R. K.-S., Pike, G. B., and Evans, A. C. (1997). Brainweb: Online interface to a 3d mri simulated brain database. In *NeuroImage*. Citeseer.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society. Series B (methodological)*, pages 1–38.

Deng, H. and Clausi, D. A. (2004). Unsupervised image segmentation using a simple mrf model with a new implementation scheme. *Pattern recognition*, 37(12):2323–2335.

Dice, L. R. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302.

Held, K., Kops, E. R., Krause, B. J., Wells III, W. M., Kikinis, R., and Muller-Gartner, H.-W. (1997). Markov random field segmentation of brain mr images. *Medical Imaging, IEEE Transactions on*, 16(6):878–886.

Hochbaum, D. S. (2001). An efficient algorithm for image segmentation, markov random fields and related problems. *Journal of the ACM (JACM)*, 48(4):686–701.

Kass, M., Witkin, A., and Terzopoulos, D. (1988). Snakes: Active contour models. *International journal of computer vision*, 1(4):321–331.

Kato, Z. and Pong, T.-C. (2006). A markov random field image segmentation model for color textured images. *Image and Vision Computing*, 24(10):1103–1114.

Kolda, T. G., Lewis, R. M., and Torczon, V. (2003). Optimization by direct search: New perspectives on some classical and modern methods. *SIAM review*, 45(3):385–482.

Manjunath, B. and Chellappa, R. (1991). Unsupervised texture segmentation using markov random field models. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (5):478–482.

Nelder, J. A. and Mead, R. (1965). A simplex method for function minimization. *The computer journal*, 7(4):308–313.

Ouadfel, S. and Batouche, M. (2003). Ant colony system with local search for markov random field image segmentation. In *Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference on*, volume 1, pages I–133. IEEE.

Panjwani, D. K. and Healey, G. (1995). Markov random field models for unsupervised segmentation of textured color images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 17(10):939–954.

Sahoo, P. K., Soltani, S., and Wong, A. K. (1988). A survey of thresholding techniques. *Computer vision, graphics, and image processing*, 41(2):233–260.

Szeliski, R., Zabih, R., Scharstein, D., Veksler, O., Kolmogorov, V., Agarwala, A., Tappen, M., and Rother, C. (2008). A comparative study of energy minimization methods for markov random fields with smoothness-based priors. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 30(6):1068–1080.

Wyatt, P. P. and Noble, J. A. (2003). Map mrf joint segmentation and registration of medical images. *Medical Image Analysis*, 7(4):539–552.

Yousefi, S., Azmi, R., and Zahedi, M. (2012). Brain tissue segmentation in mr images based on a hybrid of mrf and social algorithms. *Medical image analysis*, 16(4):840–848.

Zhang, Y., Brady, M., and Smith, S. (2001). Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *Medical Imaging, IEEE Transactions on*, 20(1):45–57.