Statistical Validation of Automated Probabilistic Segmentation against Composite Latent Expert Ground Truth in MR Imaging of Brain Tumors

Kelly H. Zou, William M. Wells III, Michael R. Kaus, Ron Kikinis,
Ferenc A. Jolesz, and Simon K. Warfield

Surgical Planning Laboratory,
Harvard Medical School and Brigham and Women's Hospital,
75 Francis St., Boston, MA 02115 USA
{zou,sw,kaus,kikinis,jolesz,warfield}@bwh.harvard.edu

Abstract. The validity of segmentation is an important issue in image processing because it has a direct impact on surgical planning. Binary manual segmentation is not only time-consuming but also lacks the ability of differentiating subtle intensity variations among voxels, particularly for those on the border of a tumor and for different tumor types. Previously we have developed an automated segmentation method that yields voxel-wise continuous probabilistic measures, indicating a level of tumor presence. The goal of this work is to examine three accuracy metrics based on two-sample statistical methods, against the estimated composite latent ground truth derived from several experts' manual segmentation by a maximum likelihood algorithm. We estimated the distribution functions of the tumor and control voxel data parametrically by assuming a mixture of two beta distributions with different shape parameters. We derived the resulting receiver operating characteristic curves, Dice similarity coefficients, and mutual information, over all possible decision thresholds. Based on each validation metric, an optimal threshold was then computed via maximization. We illustrated these methods on MR imaging data from nine brain tumor cases, three with meningiomas, three astrocytomas, and three other low-grade gliomas. The automated segmentation yielded satisfactory accuracy, with varied optimal thresholds.

1 Introduction

Surgical planning and image-guided intervention procedures increasingly employ semi-automated segmentation algorithms. MR imaging of the brain provides useful information about its anatomical structure, enabling quantitative pathological or clinical studies. Brain segmentation frequently assigns unique labels to several classes, e.g., skin, brain tissue, ventricles and tumor, representing an anatomic structure to each voxel in an input gray-level image.

Binary (two-class) manual segmentation is a simple and yet time-consuming procedure. It also has the difficulty of differentiating subtle intensity variations among voxels, particularly for those on the border of a tumor. However, the results of such manual segmentations may ultimately influence the amount and degree of tumor removal.

Recently, Warfield et al. have proposed an automated segmenter that yields voxel-wise continuous probabilistic measures indicative of malignancy (see [1] for an example). Thus, methods for validating continuous segmentation data are required. The aim

T. Dohi and R. Kikinis (Eds.): MICCAI 2002, LNCS 2488, pp. 315–322, 2002.

of this study was to evaluate the performance of this segmenter by examining three validation metrics, compared against combined experts' manual segmentations as the ground truth.

The most important element in validating the accuracy of a segmentation algorithm is the ground truth, which is the classification truth of each voxel. For simplicity, we assume a two-class truth by labeling the non-tumor class as C_0 and tumor class as C_1.

For the purpose of comparing two sets of binary segmentation results, several accuracy and reliability metrics may be found in the literature [2]. For example, Jaccard (JSC) [3] and Dice (DSC) [4] similarity coefficients are typically used as a measure of overlap; DSC ranges from 0, indicating no similarity between these two sets of binary segmentation results, to 1, indicating complete agreement.

In order to evaluate the performance of a "continuous classifier", the distributions in the two distinct classes, C_0 and C_1, may be directly compared using two-sample statistics such as a Student's t-test or a nonparametric Mann-Whitney U-statistic. Alternatively, a Komogorov-Smirnov test may be used to directly compare the two underlying distributions. Other distance measures between the two sets may also be considered.

Several statistical methods may be adopted for assessing the performance of a continuous classifier. A popular method for assessing the overall classification accuracy is a receiver operating characteristic (ROC) curve, a function of sensitivity vs. (1-specificity). Zou et al. developed several methods, including nonparametric, semiparametric, and parametric, for estimating and comparing ROC curves derived from continuous data [5,6]. The goal of this work is to examine and illustrate three accuracy metrics, ROC curve, mutual information, and Dice similarity coefficient, to validate automated probabilistic brain tumor segmentations.

2 Notations and Assumptions

For simplicity, we assume that individual voxels belong to one of two distinct and independent populations (i.e., non-tumor control class, C_0 vs. tumor class, C_1), determined by the ground truth, T. Consider two random samples, X_i $(i = 1, ..., m)$ and Y_j $(j = 1, ..., n)$, drawn from C_0 and C_1, respectively. The observed continuous random variable is labeled Z of prosbabilistic segmentation measures. The continuous random variable Z generates our probabilistic segmentation data, while the ground truth T determines the true voxel-wise classes. Stratified by the truth, for each member of class C_0, there is a measurement $X \sim (Z|T = 0)$ assumed to have cumulative distribution function (c.d.f.) F, with probability density function (p.d.f.) f and survival function $\overline{F} = 1 - F$. Similarly, for each member of class C_1, there is a measurement $Y \sim (Z|T = 1)$ assumed to have c.d.f. G with p.d.f. g and survival function $\overline{G} = 1 - G$.

We assume that the ground truth, T, has a Bernoulli distribution, with a probability of $\Pr(T = 0) = \pi = m/(m + n)$ for class C_0 and the tumor probability of $\Pr(T = 1) = \overline{\pi} = 1 - \pi = n/(m + n)$ for class C_1. By Bayes' Theorem, the marginal distribution of Z is a mixture of F and G, with mixing proportions π and $\overline{\pi}$. That is, CDF $K = \pi \cdot F + \overline{\pi} \cdot G$ with pdf k, where the p.d.f. of Z is

$$k(z) = \pi f(z) + \overline{\pi} g(z), \quad \text{with} \quad \pi + \overline{\pi} = 1 \ (\forall z \in [0, 1]). \tag{1}$$

Specifying any arbitrary threshold, $\gamma \in (0,1)$ for Z, yields a discretized version of a decision random variable, D_γ. This implies the equivalence of the following events: $\{D_\gamma = 0\} \equiv \{Z \leq \gamma\}$ and $\{D_\gamma = 1\} \equiv \{Z > \gamma\}$. Thus, we may construct Table 1:

Table 1. A two-by-two table of the joint probabilities of the truth (T) vs. the corresponding segmentation decision (D_γ) at each possible threshold γ.

Decision vs. Truth	$T = 0$ (non-tumor)	$T = 1$ (tumor)
$D_\gamma = 0$ (non-tumor)	p_{11}	p_{12}
$D_\gamma = 1$ (tumor)	p_{21}	p_{22}
Marginal Total	π	$\overline{\pi}$

where

$$p_{11} = P(D_\gamma = 0, T = 0) = P(Z \leq \gamma | T = 0)P(T = 0) = \pi F(\gamma),$$
$$p_{21} = P(D_\gamma = 1, T = 0) = P(Z > \gamma | T = 0)P(T = 0) = \pi \overline{F}(\gamma),$$
$$p_{12} = P(D_\gamma = 0, T = 1) = P(Z \leq \gamma | T = 1)P(T = 1) = \overline{\pi} G(\gamma),$$
$$p_{22} = P(D_\gamma = 1, T = 1) = P(Z > \gamma | T = 1)P(T = 1) = \overline{\pi}\overline{G}(\gamma).$$

Note that the marginal totals, $p_{11} + p_{21} = \pi$ and $p_{12} + p_{22} = \overline{\pi}$, are related to the Bernoulli parameter of T. Let $p_\gamma = \overline{F}(\gamma)$ and $q_\gamma = \overline{G}(\gamma)$ (see Section 4.1) [5,6].

3 Estimation of the Composite Latent Binary Ground Truth and Modeling Assumptions for the Probabilistic Segmentations

3.1 Composite Latent Ground Truth Based on Experts' Manual Segmentations

Instead of directly observing the ground truth, T, we conduct manual segmentations by having R expert readers, each perform binary manual segmentation B_{lr} ($l = 1, ..., N = m+n;\ r = 1, ..., R$). Each expert gave a_{0r} and a_{1r} correct counts (agreements with the truth) for classes C_0 and C_1, respectively. Let Q_{0r} and Q_{1r} represent the true accuracy rates under these two classes. The experts' decisions are assumed to be conditionally independent, given the latent truth. We only observe binary classification decision B_{lr}, i.e., $(B_{lr}|T_l, Q_{0r}, q_{1r}) \perp (B_{lr'}|T_l, Q_{0r'}, Q_{1r'})$, for any two different experts, $r \neq r'$.

We wish to estimate the latent vector $\mathbf{T}$, of length N, by $\hat{\mathbf{T}} = \arg\ \max_T\ p(\mathbf{B}|\mathbf{T}, \mathbf{Q_0}, \mathbf{Q_1})$, for all $N = m+n$ voxels. However, these classification probabilities $\mathbf{Q_0}$ and $\mathbf{Q_1}$, each a vector of length R, are unknown quantities. An iterative maximum likelihood algorithm [7,8] has been developed by realizing that the quality fractions $(Q_{0r}|\mathbf{T})$ and $(Q_{1r}|\mathbf{T})$ have independent beta distributions with modes (a_{0r}/m) and (a_{1r}/n) as their estimates.

3.2 A Beta Mixture Model of Probabilistic Segmentation Data

Recall that the continuous random variables, X and Y, are the probabilistic segmentation results for classes C_0 and C_1, stratified by the ground truth T. Because both X and Y

take values between $[0, 1]$, it is conventional and flexible to assume independent beta distributions, i.e., $F(x) \sim \text{Beta}(\alpha_x, \beta_x)$ and $G(y) \sim \text{Beta}(\alpha_y, \beta_y)$. The expectation and variance of a Beta(α, β) distribution are known to be $\alpha/(\alpha+\beta)$ and $\alpha\beta/\{(\alpha+\beta)^2(\alpha+\beta+1)\}$, respectively. Thus, the estimates $(\hat{\alpha}_x, \hat{\beta}_x)$ of the shape parameters based on the $\mathbf{x}$-sample of C_0 may be obtained by matching the first two moments (mean and variance). To match the sample mean $\overline{x}$ and standard deviation s_x of the $\mathbf{x}$-sample, it can be shown that $\hat{\alpha}_x = \overline{x}\{\overline{x}(1-\overline{x})/s_x^2 - 1\}$, $\hat{\beta}_x = (1-\overline{x})\{\overline{x}(1-\overline{x})/s_x^2 - 1\}$. Similarly for $\hat{\alpha}_y$ and $\hat{\beta}_y$, computed based on the two moments, $\overline{y}$ and s_y, of the $\mathbf{y}$-sample of C_1. Three validation metrics are presented with a higher value in [0,1] indicating higher accuracy.

4 Three Validation Accuracy Metrics

4.1 Sensitivity, Specificity, and ROC Analysis

The accuracy of a diagnostic test can be summarized in terms of an ROC curve [5,6]. It is a plot of sensitivity (true tumor fraction) vs. (1-specificity) (true non-tumor fraction) based on Z and T, at all possible thresholds.

Conventionally, $p_\gamma = \overline{F}(\gamma)$ is labeled as false positive rate (FPR or $1 -$ specificity), on the x-axis of an ROC curve. True positive rate (TPR or sensitivity) is $q_\gamma = \overline{G}(\gamma)$ at the specified γ, or $q_p = \overline{G} \circ \overline{F}^{-1}(p)$ at any specified p, on the y-axis of an ROC curve. The ROC curve is $(\overline{F}(\gamma), \overline{G}(\gamma))$ for $\gamma \in [0, 1]$, or $(p, \overline{G} \circ \overline{F}^{-1}(p))$ for $p \in [0, 1]$. There is always a tradeoff between these two error rates, false positive and false negative rates, both taken values in $[0, 1]$.

An overall summary accuracy measure is the area under the ROC curve, AUC:

$$AUC = P(X < Y) = \int_{\gamma=0}^{1} \overline{G}(\gamma)\, d\overline{F}(\gamma) = \int_{p=1}^{1} q(p)\, dp. \tag{2}$$

4.2 Dice Similarity Coefficient

At any arbitrary threshold γ, Dice similarity coefficient (DSC), D_γ may be computed as a function of the sensitivity and specificity. Following the convention of an ROC plot, label the false positive rate $p_\gamma = P(Z > \gamma | T = 0) = P(D_\gamma = 1 | T = 0)$ and the true positive rate $q_\gamma = P(Z > \gamma | T = 1) = P(D_\gamma = 1 | T = 1)$. According to the definition of DSC_γ [4] for the tumor class and Bayes' Theorem, the Jaccard similarity coefficient at γ, JSC_γ,is defined as the voxel ratio of union and intersection between the two tumor classes determined separately by D_γ and by T [3]:

$$\begin{aligned} JSC_\gamma &\equiv \frac{\#\{(D_\gamma = 1) \cap (T = 1)\}}{\#\{(D_\gamma = 1) \cup (T = 1)\}} \\ &= \frac{P(D_\gamma = 1 | T = 1) P(T = 1)}{P(D_\gamma = 1) + P(T = 1) - P(D_\gamma = 1 | T = 1) P(T = 1)} \\ &= \frac{P(D_\gamma = 1 | T = 1) P(T = 1)}{P(D_\gamma = 1 | T = 0) P(T = 0) + P(T = 1)} = \frac{\overline{\pi} q_\gamma}{\pi p_\gamma + \overline{\pi}} = \frac{\overline{\pi}\overline{G}(\gamma)}{\pi \overline{F}(\gamma) + \overline{\pi}}. \end{aligned}$$

(Note that the DSC for the non-tumor may be computed similarly but may not be of interest.) An overall DSC, DSC, based on JSC_γ is defined by integrating over γ:

$$DSC = \int_{\gamma=0}^{1} 2JSC_\gamma/(JSC_\gamma + 1)\, d\gamma. \tag{3}$$

4.3 Entropy and Mutual Information

The mutual information between the binary decision D_γ at any threshold γ and the ground truth T can be computed as follows [9]:

$$MI_\gamma = MI(D_\gamma, T) = H(D_\gamma) + H(T) - H(D_\gamma, T), \tag{4}$$

where

$$\begin{aligned}
H(D_\gamma) &= -(p_{11} + p_{12}) \log_2(p_{11} + p_{12}) - (p_{21} + p_{22}) \log_2(p_{21} + p_{22}) \\
&= -(\pi\overline{p}_\gamma + \overline{\pi}\overline{q}_\gamma) \log_2(\pi\overline{p}_\gamma + \overline{\pi}\overline{q}_\gamma) - (\pi p_\gamma + \overline{\pi} q_\gamma) \log_2(\pi p_\gamma + \overline{\pi} q_\gamma), \\
H(T) &= -(p_{11} + p_{21}) \log_2(p_{11} + p_{21}) - (p_{12} + p_{22}) \log_2(p_{12} + p_{22}) \\
&= -\pi \log_2(\pi) - \overline{\pi} \log_2(\overline{\pi}), \\
H(D_\gamma, T) &= -p_{11} \log_2(p_{11}) - p_{12} \log_2(p_{12}) - p_{21} \log_2(p_{21}) - p_{22} \log_2(p_{22}) \\
&= -\pi\overline{p}_\gamma \log_2(\pi\overline{p}_\gamma) - \overline{\pi}\overline{p}_\gamma \log_2(\overline{\pi}\overline{p}_\gamma) - \pi p_\gamma \log_2(\pi p_\gamma) - \overline{\pi} q_\gamma \log_2(\overline{\pi} q_\gamma),
\end{aligned}$$

with the joint probabilities, $(p_{11}, p_{12}, p_{21}, p_{22})$, given in Table 1.

The mutual information between the continuous random variable Z and T may also be computed using a conditioning entropy approach (with proof omitted):

$$\begin{aligned}
MI(Z, T) &= H(Z) - H(Z|T) \\
&= -E_Z\left[\log_2\{k(Z)\}\right] - \pi E_Z\left[\log_2\{f(Z)\}\right] - \overline{\pi} E_Z\left[\log_2\{g(Z)\}\right] \\
&= -\int_{z=0}^{1} \left[k(z) \log_2\{k(z)\} - \pi f(z) \log_2\{f(z)\} - \overline{\pi} g(z) \log_2\{g(z)\}\right] dz,
\end{aligned}$$

where $k(z) = \pi f(z) + \overline{\pi} g(z)$ as in (1).

4.4 Determination of an Optimal Threshold

Each of the above criteria, e.g., the square-root of the sum of squared sensitivity and specificity, $\sqrt{q_\gamma^2 + (1 - p_\gamma)^2}$, mutual information MI_γ, and Dice similarity coefficient (DSC_γ) may be maximized numerically over the entire range of γ in order to obtain an optimal threshold $\hat{\gamma}_{opt}$. Computations and optimizations were performed on a SunMicrosystem SunBlade 100 Workstation and in Matlab6, S-Plus6.0 and C languages.

5 A Clinical Example: MRI of Three Types of Brain Tumors

5.1 Materials and Methods

(1) The Cases: A total of nine patients were selected from a neurosurgical database of 260 brain tumor patients, of which three had meningiomas (M), three astrocytomas (A),

and three other low-grade gliomas (G)[10]. The meningiomas enhanced well but the gliomas did not.
(2) Imaging Protocol: Patient heads were imaged in the sagittal planes with a 1.5T MR imaging system (Signa, GE Medical Systems, Milwaukee, WI), with a postcontrast 3D sagittal spoiled gradient recalled (SPGR) acquisition with contiguous slices (flip angle, 45°); repetition time (TR), 35 ms; echo time (TE), 7 ms; field of view, 240 mm; slice-thickness, 1.5 mm; $256 \times 256 \times 124$ matrix). The acquired MR images were transferred onto a UNIX network via Ethernet.
(3) Automated Probabilistic Segmentation: The automated probabilistic segmentation was the relative tumor probability of lesion per voxel with signal intensity modeled as a Gaussian mixture of the two classes based on an initial semi-automated binary segmentation (left panel of Fig 1, in an example case) [10].
(4) Manual Binary Segmentation and Composite Ground Truth: An interactive segmentation tool (MRX, GE Medical Systems, Schenectady, NY) was employed and ran on an Ultra 10 Workstation (Sun Microsystems, Mountain View, CA). The structures were outlined slice-wise by expert operators using a mouse. The program connected consecutive points with lines. An anatomical object was defined by closed contour, and the program labeled every voxel of the enclosed volume. For the purpose of validation, we randomly selected one single 2D slice for each case from the subset of the MR volume with the tumor. Manual segmentation was performed independently by 3 expert operators (blinded to the machine segmentation results) to outline the brain and the tumor. An M.L.E of voxel-wise composite ground truth was determined. The remaining voxels were labeled as background. Stratified analyses are conducted by tumor type based on the estimated composite voxel-wise ground truth.

5.2 Results

We show semi-automated binary segmentations of a meningioma to derive the probabilistic results, with the empirical and approximated beta densities by truth (Fig 1).

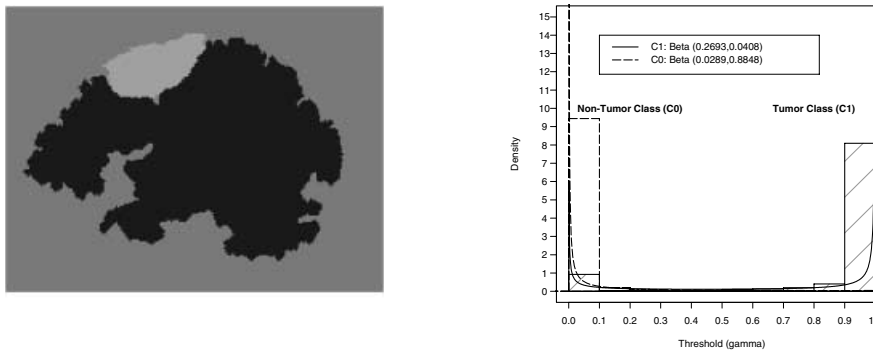

Fig. 1. Left Panel: Automated Binary segmentation of a slice of meningioma used as a basis for probablistic segmentation. Right Panel: The empirical (relative frequency histograms) and the approximated beta distributions (smooth lines) of the continuous segmentation data.

For all cases, we reported the voxel counts (m, n), stratified by the ground truth. The sample means and SD's of the non-tumor and tumor probability data were reported and were used to estimate the shape parameters of the beta distributions (Table 2).

Table 2. Sample Statistics and Estimated Beta Parameters for 9 Cases.

Tumor	Voxel Counts			Sample Means and SD's				Estimated Beta Parameters			
Type	m	n	$\overline{\pi} = n/N$	$\bar{x}$	s_x	$\bar{y}$	s_y	$\hat{\alpha}_x$	$\hat{\beta}_x$	$\hat{\alpha}_y$	$\hat{\beta}_y$
M	10534	1175	10%	0.0316	0.1264	0.8683	0.2954	0.0289	0.8848	0.2693	0.0408
	15363	1503	8.9%	0.0207	0.0890	0.8479	0.3344	0.0321	1.5227	0.1301	0.0233
	12891	1045	7.5%	0.1797	0.2746	0.7775	0.2619	0.1716	0.7832	1.1835	0.3387
A	10237	268	2.6%	0.3682	0.1548	0.6347	0.2703	3.2081	5.5044	1.3790	0.7937
	11579	1428	11.0%	0.1812	0.2496	0.7684	0.2773	0.2500	1.1303	1.0098	0.3043
	7148	1379	16.2%	0.0621	0.1229	0.9613	0.1742	0.1773	2.6790	0.2173	0.0087
G	8952	1417	13.7%	0.0112	0.0908	0.8693	0.3177	0.0038	0.3394	0.1090	0.0164
	12679	1177	8.5%	0.1564	0.2803	0.7398	0.2731	0.1063	0.5732	1.1691	0.4112
	9635	1873	16.3%	0.2275	0.2630	0.7369	0.2765	0.3505	1.1903	1.1314	0.4040

The overall validation accuracies were generally high but were variable, and generally the highest for meningiomas but lowest for astrocytomas. Furthermore, the recommended optimal thresholds varied by metric and case (Table 3).

Table 3. Estimated Accuracy metrics (ROC, MI and DSC) and Optimal Thresholds.

Tumor	Validation Metrics			Optimal Thresholds					
Type	AUC	MI	DSC	$\sqrt{(1-p)^2+q^2}$	$\hat{\gamma}_{opt}$	MI	$\hat{\gamma}_{opt}$	DSC	$\hat{\gamma}_{opt}$
M	0.9842	0.2888	0.8154	1.3255	0.4709	0.3107	0.8625	0.8730	0.8734
	0.9684	0.3012	0.8415	1.2834	0.8448	0.3065	0.8521	0.8931	0.8268
	0.9242	0.1572	0.4220	1.1844	0.2622	0.1098	0.4657	0.5185	0.8414
A	0.7860	0.0557	0.1970	1.0050	0.7713	0.0415	0.7728	0.4871	0.7808
	0.9255	0.2319	0.5146	1.1881	0.4469	0.1598	0.6843	0.6321	0.8005
	0.9858	0.4649	0.8708	1.4142	1.0000	0.5669	0.8553	0.9724	0.8385
G	0.9829	0.4018	0.8961	1.3720	0.0120	0.4032	0.4905	0.8992	0.6736
	0.9157	0.1595	0.4396	1.1735	0.0773	0.1276	0.2232	0.4897	0.6511
	0.8956	0.2505	0.5276	1.1417	0.4547	0.1693	0.6191	0.6197	0.7113

6 Summary

In this work, we have presented systematic approaches to validating the accuracy of automated segmentation results that generates voxel-wise probabilistic interpretation of the tumor class. We developed an M.L.E. algorithm for estimating the latent ground truth. In addition, we modeled the probabilistic segmentation results using a mixture of two beta distributions with different shape parameters. Summary accuracy measures, including ROC curve, mutual information, and Dice similarity coefficient, are estimated. An optimal threshold was derived under each metric.

The example data showed satisfactory accuracy of our automated segmentation algorithm. The recommended optimal threshold, however, was significantly case- and task

(metric)-dependent. The main advantage of our approaches is that the parametric modeling is simple and probabilistic. The estimation procedures are straightforward and are generalizable to similar statistical validation tasks of segmentation methods.

Acknowledgements

This work was supported by NIH grants P41 RR13218, P01 CA67165, R01 RR11747, R01 CA86879, R21 CA89449-01, a research grant from the Whitaker Foundation, and a New Concept Award from the Center for Integration of Medicine and Innovative Technology. We thank the three experts who performed manual segmentations of the brain tumor cases.

References

1. Warfield S. K., Westin C.-F., Guttmann, C. R. G., Albert, M., Jolesz, F. A., Kikinis, R.: Fractional segmentation of white matter. In Proceedings of Second International Conference on Medical Imaging Computing and Computer Assisted Interventions, Cambridge, UK (1999) 62-71.
2. Zijdenbos, A. P., Dawant, B.M., Margolin, R. A., Palmer, A. C.: Morphometric analysis of white matter lesions in MR images: method and validation. IEEE Transactions on Medical Imaging 13 (1994) 716-724.
3. Jaccard, P.: The distribution of flora in the alpine zone. New Phytologist 11 (1912) 37-50.
4. Dice, L. R.: Measures of the amount of ecologic association between species. Ecology, 26 (1945) 297-302.
5. Zou, K. H., Hall W. J., Shapiro, D. E.: Smooth nonparametric receiver operating characteristic curves for continuous diagnostic tests. Statistics in Medicine 16 (1997) 2143-2156.
6. Zou, K. H. , Hall W. J.: Two transformation models for estimating an ROC curve derived from continuous data. Journal of Applied Statistics 27 (2000) 621-631.
7. Dempster, A. P, Laird, N. M., Rubin, D. B.: Maximum-likelihood from incomplete data via the EM algorithm. J. Royal Statistical Society (Ser. B) 39 (1977) 34-37.
8. McLachlan G. J., Krishnan, T.: The EM Algorithm and Extensions. Wiley, New York (1997).
9. Cover, T. MM, Thomas J. A.: Elements of Information Theory. John Wiley & Sons, Inc., New York (1991).
10. Kaus, M., Warfield S. K., Nabavi A., Black, P. M., Jolesz, F. A., Kikinis, R.: Automated segmentation of MRI of brain tumors. Radiology 218 (2001) 586-591.