

# Classifiability Criteria for Refining of Random Walks Segmentation

Steven Rysavy<sup>1</sup>    Arturo Flores<sup>1</sup>    Reyes Enciso<sup>2</sup>    Kazunori Okada<sup>1</sup>

*1 Computer Science Department, San Francisco State University*

*1600 Holloway Avenue, San Francisco, CA 94132*

*{kazokada, rysavy, aflores}@sfsu.edu*

*2 Craniofacial Sciences and Therapeutics, University of Southern California*

*925 West 34th Street, Los Angeles, CA 90089*

*renciso@usc.edu*

## Abstract

*This paper proposes a novel approach to improve the segmentation quality of a 3D random walks algorithm using classifiability criteria. We produce a range of potential threshold values by extending the decision function of a random walks algorithm using a likelihood ratio test. Optimal threshold values are quantitatively isolated using two data-driven methods: maximum total accuracy and Bayesian cross validation criteria. The proposed methods are evaluated using a dataset of 28 dental lesions in 3D cone-beam CT scans. Both methods produce viable thresholds, the first corresponding to a conservative segmentation and the second a relaxed segmentation. We qualitatively compare the results to determine the best method.*

## 1. Introduction

Segmentation is a standard pre-process for data analysis including classification and recognition. Determining the best segmentation is difficult because it is highly dependent upon problem-specific criteria. Different models can produce various distinct segmentations. Logically the segmentation algorithm should be biased to achieve the best post-processing result. This article proposes a novel approach to improve the segmentation pre-processing by incorporating classification criteria as a part of the objective function for segmentation.

We demonstrate this method using a graph-based random walks algorithm, which is an effective approach for segmentation in medical imaging [2, 3]. Specifically we extend its decision function by using the likelihood ratio test (LRT) [1] and refine the algorithm by opti-

mizing the threshold of LRT via the results of a trained linear discriminant analysis (LDA) classifier. We deduce the threshold with two data-driven methods: the maximization of total accuracy of multiple LDA classifiers (maximum total accuracy) and the maximization of Bayesian posterior distribution using the classifier's cross validation results (Bayesian cross validation).

In this paper we provide a brief overview of the random walks segmentation algorithm, the initialization method in 3D space, and the two approaches for deriving a threshold. Experimentally we evaluate 3D segmentations of dental lesions in cone-beam computed tomography (CBCT). Our study shows the benefits of each approach in this medical imaging application. Finally we compare the qualitative results and state our conclusion.

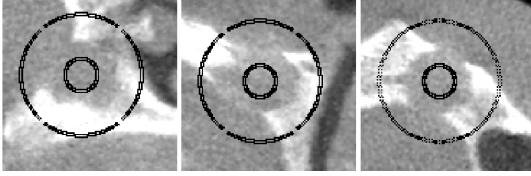
## 2. Methods

### 2.1. Random Walks Algorithm

The random walks algorithm proposed by Grady [2] is adapted to segment a 3D target lesion for medical imaging applications [3]. Random walks is a graph-theoretic, semi-automatic segmentation algorithm. This method utilizes inter-voxel weights according to

$$w_{ij} = e^{-(\beta(g_i - g_j)^2)}, \quad (1)$$

where  $g$  represents the voxel intensity and  $\beta$  is a free parameter in the algorithm. Given initial binary seed labels manually specified by users, these weights are iteratively calculated from each voxel to each seed, assigning each voxel a probability that it belongs to foreground label *in* or background label *out*. Treating this as a combinatorial Dirichlet problem results in



**Figure 1. Example seed points for random walks segmentation. The inner sphere of seed points designates a region inside the lesion while the outer sphere of seed points designates a region outside the lesion.**

a sparse, symmetric, positive-definite system of equations. See [2] for more details. Among other methods, we chose this algorithm due to its advantage in segmenting weak boundaries that are common in our target medical imaging applications.

## 2.2. Initialization Method in 3D

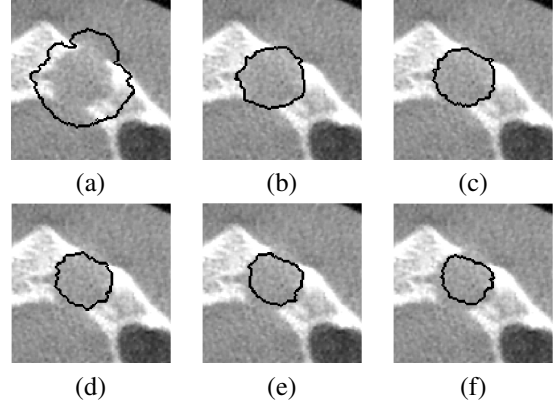
We introduce a method for the random walks algorithm to place initial seeds in 3D with minimal user-interaction. The method specifies two concentric spheres. It then places foreground and background seed labels along the surface of the inner and outer sphere, respectively. Three parameters are required to generate these seed points: the center of the spheres, an inner radius and an outer radius. The inner radius specifies a sphere small enough to be completely surrounded by the region of interest (ROI). Likewise, the outer radius specifies a sphere large enough to entirely enclose the ROI. An example is provided in Figure 1.

## 2.3. Segmentation with Likelihood Ratio Test

The original random walks segmentation assigns a label to a voxel by comparing the foreground and background probability. The label with larger probability will be assigned at each voxel. We extend this decision rule by using the LRT formalism with a threshold parameter introduced to control trade-off between under- and over-segmentations. This LRT extension assigns voxel label  $l_v$  using

$$l_v = \begin{cases} 1 \text{ (in)} & \text{if } p(x|in)/p(x|out) > k \\ 0 \text{ (out)} & \text{otherwise} \end{cases} \quad (2)$$

where  $p(x|in)$  and  $p(x|out)$  denote the spatial likelihood function, indicating the probability that the voxel



**Figure 2. Slices of 3D segmentation examples with increasing threshold values,  $k$ . (a)  $k = 0.05$ , (b)  $k = 0.50$ , (c)  $k = 1.00$ , (d)  $k = 1.50$ , (e)  $k = 2.00$ , (f)  $k = 2.50$ .**

is inside or outside the lesion, respectively. These probabilities are computed directly by the random walks algorithm.  $k$  is a threshold parameter. The parameter flexibly controls segmentation results in the standard likelihood ratio test sense. The 3D segmentation becomes more conservative as  $k$  increases, as exemplified in figure 2. The original random walks segmentation uses  $k = 1.0$ .

## 2.4. Data-Driven Threshold Selection

**2.4.1. Maximum Total Accuracy Criterion.** Our first approach derives the optimal performing threshold by maximizing the total accuracy of multiple LDA classification results. Given a set of  $N$  features, we train a set of  $K$  LDA classifiers for arbitrary feature combinations. We then perform leave-one-out cross-validation (LOOCV) for  $M$  test cases chosen randomly from a training dataset. *Total accuracy* is defined as the total number of accurately classified test cases among  $K \times M$  cases. For each value of  $k$ , the total accuracy is computed with  $N$  features extracted from the corresponding segmentation. The  $k$  value that results in the largest total accuracy provides an optimal threshold according to this maximum total accuracy criterion.

**2.4.2. Bayesian Cross Validation Criterion.** We also deduce an optimal threshold derived directly from the cross validation performance by using Bayesian inference. Using the best performing feature combination, we train a single LDA classifier. LOOCV performance statistics, such as hit rate, sensitivity, and 1-specificity

are computed at each  $k$  value. After an appropriate normalization, we treat such a function  $f(k)$  of the cross validation results as a data likelihood distribution  $P(f|k)$ . Suppose that we now have an independent prediction, as prior knowledge, of what  $k$  value we should use. For simplicity, we model the probability distribution  $P(k)$  of this prior knowledge by a Gaussian distribution with mean  $\mu$  and width  $\sigma$ .

$$P(k) = \frac{1}{\sigma(\sqrt{2\pi})} e^{-\frac{(k-\mu)^2}{2\sigma^2}}, \quad (3)$$

Using the Bayes theorem, the posterior probability distribution is define by

$$P(k|f) = \frac{P(f|k)P(k)}{P(f)}. \quad (4)$$

Bayesian cross validation criterion is defined by finding the  $k$  value that maximizes  $P(k|f)$  in the maximum a posteriori (MAP) estimation sense.

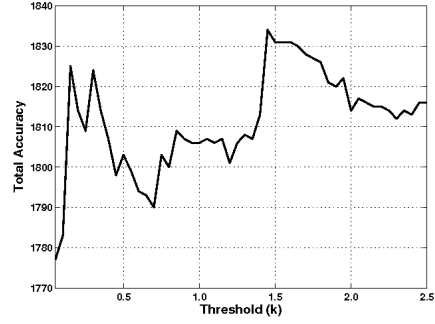
$$k^* = \operatorname{argmax}_k P(k|f) = \operatorname{argmax}_k P(f|k)P(k). \quad (5)$$

### 3. Experiments

#### 3.1. Application and Data

We consider a computer-aided diagnosis application using 3D medical scans. In such an application, a target lesion is first segmented within an input 3D scan and a feature set is next extracted from the segmented lesion. Clinical diagnosis is ascertained by training a machine learning-based classifier on expert-validated cases, assigning the unclassified lesions to clinical categories. In this scenario, a slight error in segmentation can largely influence the final classification results. For example, inclusion of bone with high intensity values near the target lesion can introduce a strong bias to the features extracted from the segment, causing mis-classification. To counteract this, we refine the segmentation results in order to minimize the inclusion of high intensity regions while ensuring the maximum number of statistical samples. This is achieved by assigning voxel label  $l_v$  using the LRT extension of the random walks with the proposed optimal threshold selection methods.

Our data is comprised of 28, 3D cone-beam computed tomography (CBCT) target lesion images. Random walks has previously been shown to outperform other graph-based segmentation algorithms for this dataset [3]. Ground-truth binary classification between cyst and granuloma cases, as defined by [4], is available for all 28 lesions and is referenced in all classification results.



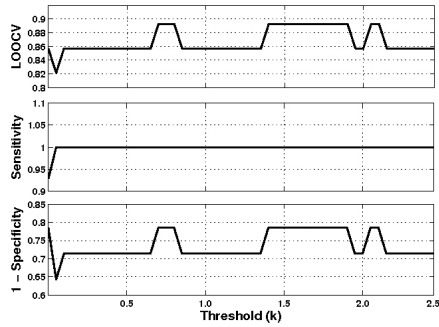
**Figure 3. Total accuracy computed for varying threshold  $k$ . See text for more information.**

#### 3.2. Methods and Results

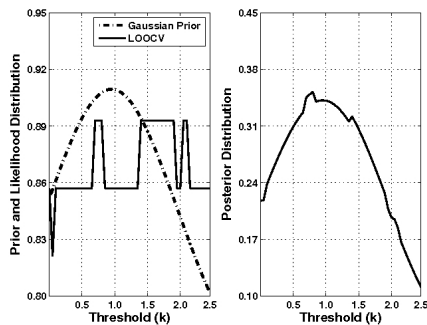
For each segmented lesion with varying  $k$ , we extract eight features ( $N = 8$ ) based on intensity statistics: mean, median, maximum, minimum, standard deviation, skewness, kurtosis, entropy. For evaluating the maximum total accuracy (MTA) criterion, a set of LDA classifiers are trained for every possible two, three and eight feature combinations of  $K = 85$ . Each classifier is tested by LOOCV with the total of  $M = 28$  test cases. Figure 3 displays the total accuracy computed for each  $k$  value. The maximum is observed at  $k = 1.45$ .

For evaluating the Bayesian cross validation (BCV) criterion, we choose the best performing two feature combination of median and minimum intensity. Figure 4 shows the LOOCV results of the hit rate, sensitivity, and 1-specificity. We treat cyst cases as positive cases. In this example, no prominent peak was observed in the sensitivity statistics. Thus we choose the LOOCV hit rate as our data likelihood distribution. We centered the Gaussian prior at  $\mu = 1.0$  as this is the baseline random walks segmentation. The width is manually set broadly to  $\sigma = 1$  so that we focus more on the data term. Figure 5 shows the data likelihood, prior and posterior distributions computed for our data. The maximum is observed at  $k = 0.75$ .

Next we qualitatively compare the MTA and BCV approaches in terms of segmentation quality. For most of the 28 cases, both approaches result in an appropriate segmentation, confirmed by visual inspection. We observe that the threshold identified by MTA corresponds to a conservative segmentation while the threshold derived from BCV results in a relaxed segmentation. However, in a few cases, both conservative and relaxed segmentation result in classification failures. Figure 6



**Figure 4. LOOCV, Sensitivity and 1-Specificity computed for varying threshold  $k$ , respectively.**

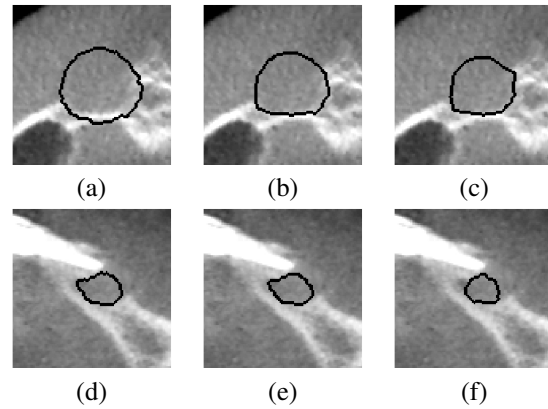


**Figure 5. Likelihood, prior and posterior distributions. Gaussian prior centered at  $k=1$  is scaled and translated for visual enhancement.**

illustrates 2D slices of such failure cases.

The first row in figure 6 shows a case whose segmentation quality improves with an increasing threshold. Figure 6a illustrates a relaxed segmentation corresponding to  $k = 0.75$ , which over-segments the lesion and incorporates a large amount of dense tissue. This may disrupt the LDA classification. A more conservative segmentation resulting from a threshold of  $k = 1.45$  by MTA is presented in figure 6c and contains only desired voxels, ensuring a better classification.

Conversely the second row in figure 6 illustrates a case that benefits from a lower threshold. The segmentations of this lesion decrease in quality as  $k$  increases. The relaxed segmentation found by BCV in figure 6d closely follows the delineation between the lesion and the dense tissue, identifying the majority of the lesion



**Figure 6. Example slices of 3D segmentations showing benefits of MTA (a-c) and BCV (d-f). (a,d)  $k = 0.75$ , (b,e)  $k = 1.00$ , (c,f)  $k = 1.45$ .**

while simultaneously obtaining from including dense-tissue. The conservative segmentation corresponding to a threshold of  $k = 1.45$  includes fewer lesion voxels, decreasing the number of statistical values and therefore hindering the performance of the LDA classifier.

## 4. Conclusion

We propose an extended random walks segmentation using a likelihood ratio test. Two selection methods are presented which uniquely derive thresholds from data-driven classifiability criteria. MTA avoids inclusion of dense tissue with high intensity values while BCV includes more voxels. Although both produce valid segmentations, the avoidance of dense tissue has a greater impact on the LDA classifier than a statistically insubstantial increase in sampled voxels. Therefore MTA is the better choice for this application.

## References

- [1] W. Feller. *An Introduction to Probability Theory and Its Applications*, volume 2. Wiley, New York, 1971.
- [2] L. Grady. Random walks for image segmentation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006.
- [3] S. Rysavy, A. Flores, R. Enciso, and K. Okada. Segmentation of large periapical lesions toward dental computer-aided diagnosis in cone-beam CT scans. In *SPIE Medical Imaging*, pages 6914–44, 2008.
- [4] J. Simon, R. Enciso, J.-M. Malfaz, M. Bailey-Perry, and A. Patel. Differential diagnosis of large periapical lesions using cone-beam computed tomography measurements and biopsy. *Journal of endodontics*, 32(9):833–837, 2006.