A CLASSIFICATION APPROACH FOR ANATOMICAL REGIONS SEGMENTATION

[1]Mikhail Kalinin, [1]Daniela S. [1]Raicu, [1]Jacob D. Furst,[2]David S. Channin

[1]Intelligent Multimedia Processing Laboratory,
School of Computer Science, Telecommunications, and Information Systems,
DePaul University, Chicago, Illinois, 60604, USA
{mkalinin}@students.depaul.edu, {jfurst, draicu}@cs.depaul.edu
[2]Department of Radiology
Northwestern University Medical School
448 East Ontario Street, Ste. 300, Chicago, Illinois 60611
dsc@radiology.northwestern.edu

ABSTRACT

In this paper, a supervised pixel-based classifier approach for segmenting different anatomical regions in abdominal Computed Tomography (CT) studies is presented. The approach consists of three steps: texture extraction, classifier creation, and anatomical regions identification. First, a set of co-occurrence texture descriptors is calculated for each pixel from the image data sample; second, a decision tree classifier is built using the texture descriptors and the names of the tissues as class labels. At the conclusion of the classification process, a set of decision rules is generated to be used for classification of new pixels and identification of different anatomical regions by joining adjacent pixels with similar classifications. It is expected that the proposed approach will also help automate different semi-automatic segmentation techniques by providing initial boundary points for deformable models or seed points for split and merge segmentation algorithms. Preliminary results obtained for normal CT studies are presented.

1. INTRODUCTION

Automatic segmentation of tomograms from various medical imaging modalities is highly desirable for visualization of soft tissues or extraction of individual organs for treatment planning in radiation oncology.

Since the shape of each organ is not consistent throughout all slices of a 3D medical image and the gray level intensities overlap considerably for soft tissues, texture is especially important in medical image segmentation because of its homogeneity within the same tissue and across different slices. Once textures have been calculated and their scalar values assigned to pixels, the pixels can be clustered or classified (when the tissues' labels are available) for the purpose of segmentation.

There are a large number of texture-based segmentation algorithms in the literature. Texture segmentation usually involves the combination of texture feature extraction techniques with a suitable segmentation algorithm. Among the most popular feature extraction techniques used for texture segmentation are Gabor filters and wavelets transforms [1-3]. Among the most commonly used segmentation algorithms based on these features are clustering techniques [4, 5], region growing and split-and-merge [6, 7].

Segmentation using the traditional techniques previously mentioned requires considerable amounts of expert interactive guidance. In the medical imaging field, deformable models are commonly applied because of their capability to capture the irregular shapes and shape deformations found in anatomical structures. The deformable model that has attracted the most attention to date is popularly known as "snakes" [8] and it has been used for different applications such as the segmentation of the heart from cardiac imagery, of neural tissue textures, and of the bone and cartilage in clinical knee MRI.

However, the application of snakes and other similar deformable contour models to extract regions of interest is not without limitations. One of the limitations is that snakes were designed as interactive (semi-automatic) models. In order to increase the speed, accuracy, and consistency of the segmentation, automatic segmentation is a desirable, albeit very difficult, long-term goal.

The approach proposed in this paper does not require any selection of initial points in order to perform the organ segmentation; therefore, it can be used as a first-pass segmentation followed by the snake algorithm. Furthermore, the segmentation of the normal tissues/organs will have a significant impact on the current research efforts for providing *automatic context* (i.e. that the cursor is hovering over "liver" in CT images) and thus, allowing *automatic selection of context sensitive*

tools for various tissues (i.e. do not provide lung nodule analysis tools when the context is "liver").

2. PIXEL-LEVEL TEXTURE EXTRACTION

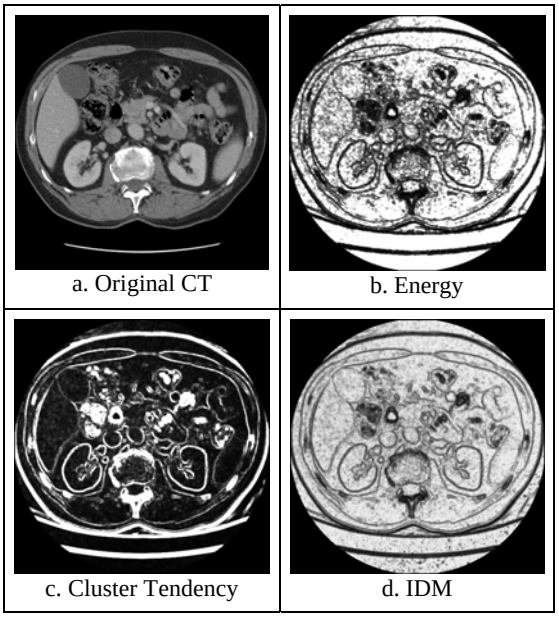
The process of calculating the texture features that capture the properties of the texture within a small neighborhood of a pixel is called pixel-level texture extraction. For each image pixel, our approach calculates a set of ten Haralick co-occurrence texture features [9], but other texture features can be incorporated in the presented approach. We started with the co-occurrence features for pixel-based texture classification because of the good results we obtained for the classification of normal tissues in CT images of the chest and abdomen using organ-level co-occurrence texture features [10].

A two dimensional co-occurrence matrix, P, is an n x n matrix, where n is the number of gray-levels within an image. For reasons of computational efficiency, the number of gray levels can be reduced if one chooses to bin them, thus reducing the size of the co-occurrence matrix. The matrix acts as an accumulator so that P[i , j] counts the number of pixel pairs having the intensities i and j; co-occurrence matrices are normally defined for a fixed distance and direction between the two pixels in a pair. In our approach, since the neighborhood of the pixel is small (our current implementation makes use of a 5 x 5 neighborhood), we do not calculate the co-occurrence along fixed directions and displacements, but instead consider all the pixel pairs within that neighborhood. Thus, our implementation produces a co-occurrence matrix for each pixel rather than for each choice of distance and direction. Then, for each co-occurrence matrix (each pixel), we calculate ten Haralick features which can be related to specific characteristics in the image: Entropy, Energy, Contrast, Homogeneity, SumMean, Variance, Correlation, Maximum Probability, Inverse Difference Moment (IDM), and Cluster Tendency. Table 1 (b-c) illustrates the image representations of different texture features for the original CT image from Table 1.a.

3. PIXEL-BASED TEXTURE CLASSIFICATION AND SEGMENTATION

Pixel-based texture classifiers are meant to recognize the texture patterns to which the pixels of an input image belong. There are many classifiers that can be used to discriminate among the organ tissue classes in the feature space. In our approach, we evaluate a Classification and Regression Tree (C&RT) classifier because: 1) it does not make any assumptions of the distribution of the data; 2) it has a relatively faster learning speed than other classification methods, while still producing classification accuracy comparable with those methods; and 3) it has a good ability to generate rules that can be easily interpreted and used to annotate various tissues in future CT scans.

Table 1: Example of texture feature images



The CR&T tree is constructed by splitting subsets of the data set using all descriptors as predictors to create two child nodes repeatedly, beginning with the entire data set. The best predictor is chosen using the Gini impurity index which works by choosing a split at each node such that each child node is more pure than its parent node. The goal is to produce subsets of the data which are as homogeneous as possible (producing pure nodes in the tree) with respect to the class label. For each split, each predictive descriptor is evaluated to find the best cut point (our descriptors being continuous predictors) based on improvement score or reduction in impurity. Then, the predictors are compared, and the predictor with the best improvement is selected for the split. The process repeats recursively until one of the stopping rules is triggered: 1) the maximum tree depth, *d*, has been reached; 2) there is no significant predictive descriptor left to split the node; 3) the number of cases in the terminal node is less than the minimum number, *np*, of cases for parent nodes; 4) if the terminal node were to split, the number of cases in one or more child nodes would be less than the minimum number, *nc*, of cases for child nodes; and 5) minimum change in impurity, *imp*, is reached. Depending on the values set for the parameters (*d*, *np*, *nc*, *imp*), a different tree will be obtained. The 'best' tree will be chosen to be the one with the highest classification accuracy.

Once the decision tree is generated, its decision rules can be used to classify each pixel within the CT images. Pixels with the same classification labels and being adjacent will form connected components and thus, the regions of interest within the corresponding CT images.

4. PRELIMINARY RESULTS

4.1. Data Set

Our preliminary results are based on data extracted from two normal CT studies from Northwestern Memorial Hospital. The data consisted of multiple, serial, axial CT images derived from helical, multi-detector CT acquisitions. The images were in DICOM (Digital Imaging and Communications in Medicine) format of size 512 by 512 and having 12-bit gray level resolution. The organs of interest with respect to segmentation were liver, renal and splenic parenchyma, and bone.

4.2. Texture Features

We computed ten co-occurrence descriptors for each pixel from the sample data using a 5 by 5 neighborhood. Several issues had to be taken into account when applying the co-occurrence to the DICOM images. While it is possible to use all intensity levels in computing co-occurrence features, this increases the computational load and reduces the validity of the co-occurrence matrix as a probability model. According to [10], a 16 level gray scale is enough to compute co-occurrence matrices and thus, the gray levels can be grouped into 16 bins. For example, for an image of 256 gray levels, bin 1 would contain gray levels 0 to 15, and bin 2 would contain gray level 16 to 31, and so on. However, DICOM images have 12 bits of resolution (4096 gray levels) and therefore, in our implementation, we use 256 bins (each bin contains 16 levels) to calculate the matrices within each neighborhood.

4.3. Decision Tree Classifier

.In order to generate the decision tree (DT), we manually selected patches of pure organ tissues from three consecutive slices. The number of patches and their sizes were chosen such that we have an equal number of pixels for each organ of interest. The pixels received the class label of the patch to which they belonged to. Since we want to use the decision tree for the segmentation of entire CT images, additional patches (not containing the organs of interest) were selected and their pixels were labeled as 'unknown'. We ended up selecting around 1500 pixels for each of the four organs and the 'unknown' class.

Since the results of a classifier can be influenced by scaling of the feature space, we normalized all descriptors independently using min-max criteria [10]. The normalized data sample was further divided into training set (66%) and testing set (34%); the training set was used to build the classifier, while the second set was used to estimate the accuracy of the classifier when used for tissue/organ annotation of previously unseen pixels.

In order to select the optimal decision tree (DT) for our data sample, we varied the number of observations (pixels) per node from 25 (number of pixels in a neighborhood) to 1000 and each time we estimated the overall accuracy (Figure 1) of the classifier (number of pixels correctly classified divided by the total number of pixels). Based on the accuracy of the testing set (Figure 2), the optimal tree was selected. The empirically found optimal parameter for the "observations per parent" was in the range from 274 to 289; any of those values would result in the combined accuracy of the testing set of over 85.%. We decided to use a tree with the "observations per parent" of 289 since it resulted in the smallest and most efficient tree.

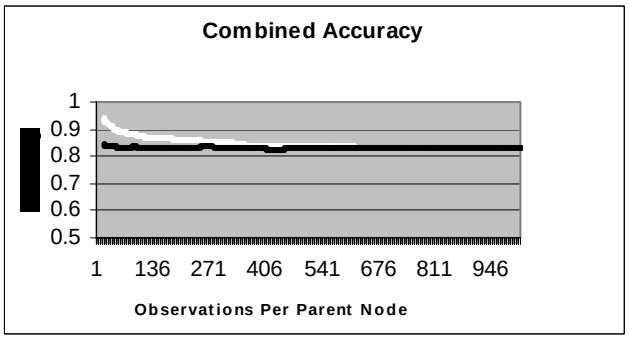


Figure 1: Accuracy of the DT classification versus number of observations per parent

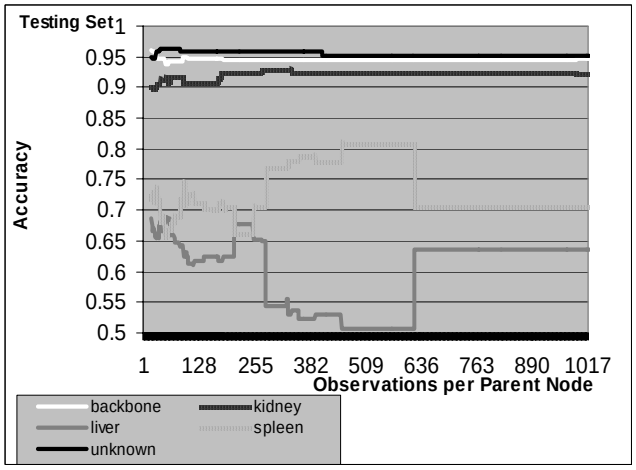


Figure 2: Accuracy of the DT classification for each class

4.4. Segmentation Results

After we generated and tested the decision tree on the sample data, we applied the tree on several consecutive

entire slices in order to segment the organs of interest. We noticed that the unknown class, the kidneys, and the bones were accurately classified while spleen and liver were very often misclassified (liver pixels were classified as spleen and vice-versa). This was not surprising given the fact that liver and spleen have homogeneous similar texture, and therefore, the current texture features could not highly differentiate them. In order to improve the segmentation, either a median filter or spatial information (liver is always in the anatomical left hand side of the abdomen and the backbone can be used as a point of reference to find the orientation of the CT scan) can be used as a post processing step. For visualization purposes, Figure 3.c-d. shows the classification image before and after post processing with a median filter; different colors represent the organs of interest **(red – liver, green – kidney, white – backbone, blue – spleen)**, gray represents the unknown class (organs which were not of interest for the current study) and black is the region outside the body which has not been included in the analysis.

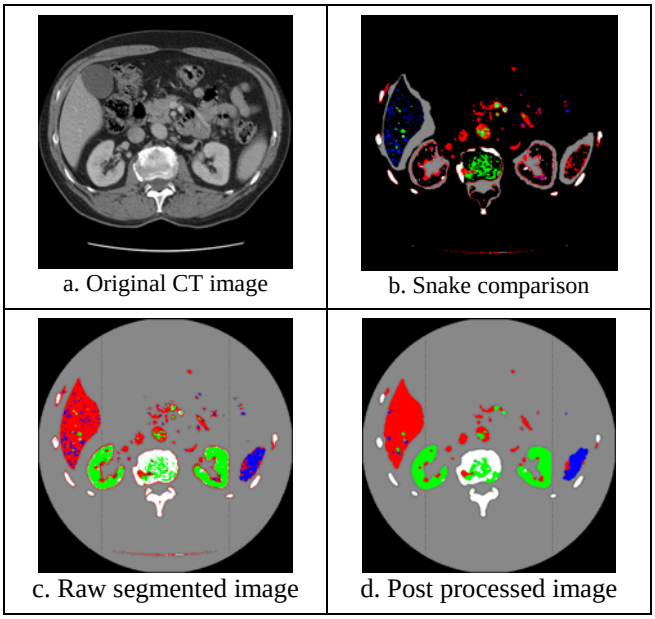


Figure 3: Visual representation of classification image, snake comparison, and median filtered image (5 by 5 filter).

5. CONCLUSIONS AND FUTURE WORK

Our preliminary results show that it is possible to segment abdominal organs. By just using 7500 pixels from pure tissues of three consecutive slices (each of size 512 by 512), we accurately classified four organs with an overall accuracy of 85%. Furthermore, the accuracy for bones and kidneys was above 92%. The results show that the approach is valid. However, there is more work that needs to be done before the approach can be used clinically. We will pursue further research with respect to:

1. *Variable window size for calculating the pixel-level texture.* For large organs (such as liver), we will use larger window sizes than for smaller organs. We expect that this will improve the features and allow increasing the current accuracy (65% for liver, 70% for spleen). 2. *Other texture features* in order to improve the segmentation and also further segmenting other organs of interest. 3. *Feature reduction step* to find all uncorrelated features and thus, speed up the entire segmentation. 4. *Automated snake algorithm*. The initial points required by the snake algorithm will be given by points automatically selected on the boundary of the objects identified by our approach. Figure 3.b. shows that the results of our approach and of the snake superimposed. Black pixels represent pixels classified by both algorithms the same way; a color pixel represents a pixel that was classified by our algorithm as a certain organ, but was classified by the snake as a different organ; a grey pixel represents a pixel that was classified as any organ by the snake but background by our algorithm. We expect that the combination of the two approaches will improve the segmentation results by removing the clearly extraneous grey areas surrounding and within organs (compare Figure 3.b with 3.b) that the snake generates and eliminating the noise our algorithm generates.

6. REFERENCES

[1] A.K. Jain and F. Farrokhia, “Unsuper. Text. Segment. using Gabor Filters”, *Patt. Recog.*, vol. 24,, pp. 1167—1186, 1991.

[2] T. Chang and C.C. Kuo, “Texture segmentation with tree-structured wavelet transform”, Proceedings *of the IEEE-SP Int’l Symp. on Tine-Freq. and Time-Scale Anal.*, pp. 543–546, 1992.

[3] M. Unser, “Texture classif. and segmentation using wavelet frames”, *IEEE Trans. on Im. Proc.*, 4(11):pp. 1549–1560, 1995.

[4] K.I. Chang, K.W. Bowyer, and M. Sivagurunath, “Evaluation of texture segmentation algorithms”, *IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 294–299, 1999.

[5] R. Porter and N. Canagarajah, “A robust automatic clustering scheme for image segmentation using wavelets”, *IEEE Trans. on Image Processing*, vol. 5(4), pp. 662–665, 1996.

[6] J. R. Beveridge, J. Gri, R.R. Kohler, A.R. Hanson, and E.M. Riseman, “Segmenting images using localized histograms and region merging”, *Int'l J. Comp. Vis.*, vol. 2, pp. 311-347, 1989.

[7] R. Adams and L. Bischof, “Seeded region growing”, *IEEE Trans. Pattern Analysis and Mac. Int.*, 16(6), pp. 641-647, 1994.

[8] M. Kass, A. Witkin and D. Terzopoulos, “Snakes: Active contour models”, *Int’l J. of Comp. Vis.* vol. 1, 321–331, 1988.

[9] R.M. Haralick, K. Shanmugam, and I. Dinstein, “Textural Features for Image Classification”, *IEEE Trans. on Systems, Man, and Cybernetics*, vol. Smc-3, no.6, pp. 610-621, 1973.

[10] D. Channin, D.S. Raicu, J.D. Furst, D.H. Xu, L. Lilly, C. Limpsangsri, "Classification of Tissues in Computed Tomography using Decision Trees", *RSNA*, December, 2004.