

# Texture Classifier Robustness for Sub-Organ Sized Windows

William H. Horsthemke  
Jacob D. Furst  
Daniela Raicu  
DePaul University  
School of Computer Science  
Chicago, IL  
horsthemke@acm.org  
jfurst@cs.depaul.edu  
draicu@cs.depaul.edu

## Abstract

Automated identification of organs in medical imaging is becoming possible using texture analysis on whole organ cross-sections. However, there exists a need to identify organs using smaller sub-sections, or sub-regions, of organs. Two motivations drives this research: first, segmenting organs from the backgrounds of images requires significant manual intervention; second, this technique will be applied when a clinicial selects a region of interest and requests an identification of the organ currently selected. This research addresses the question of how well does the classifier perform as the window size becomes smaller.

This paper extends recent work by [Raicu 2004] who studied the effectiveness of grey level co-occurrence matrices (GLCM) and run-length encoding (RLE) statistics on classifying various abdominal and thoracic organs. Their methodology measured the GLCM and RLE on a set of labeled, segmented organs and trained a decision tree classifier to label a new unknown organ given its GLCM and RLE measurements. The methodology of this paper follows the same basic approach except the organs are subdivided into successively smaller regions. The texture measurement results of these smaller windows into the organ are used to train machine-learning classifiers. The goal of this paper is to quantify the classification performance over the range of window sizes.

The results suggest that classification performance will remain effective as the window size shrinks from 75 pixels per side to 35 pixels (only square windows were studied). Below size 30, the performance significantly degrades.

This study considered the effect of smaller views into an organ on the ability to measure sufficient textural information to properly identify the organ. As the window size decreases, the classification performance also reduces but remains effective above a

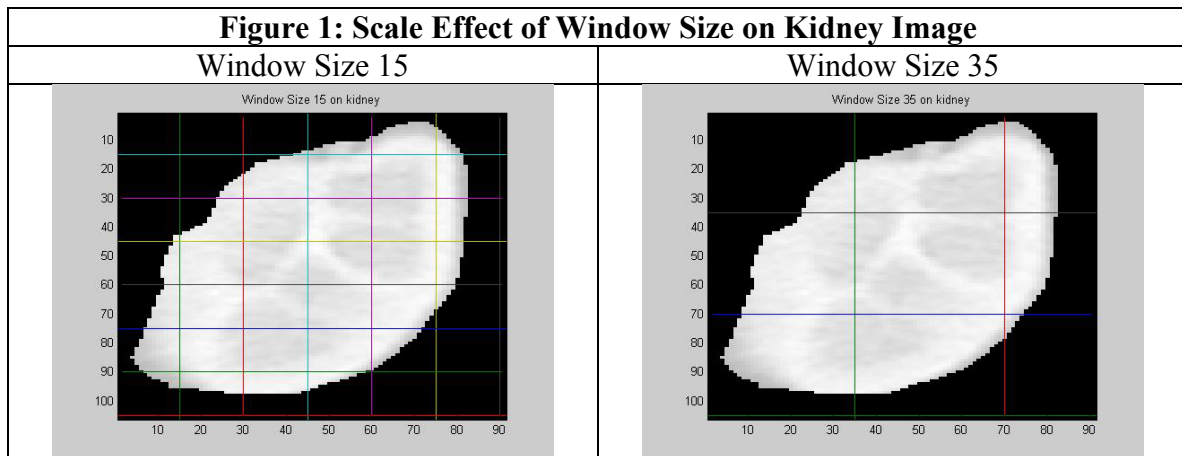
window size of 35 pixels. Smaller windows markedly degrade the texture classification performance.

## 1 Introduction

Texture analysis of medical images shows promise towards the automated identification of organs. Early work using whole organs, segmented from the background tissue, successfully identified several important thoracic and abdominal organ using grey level co-occurrence (GLCM) matrices and run-length encoding (RLE). Segmenting organs from computed-tomography images requires manual intervention by a skilled medical image analyst. Future systems will need to identify organs using only subsections of organs obtained using techniques such as windowed cursors placed over an unknown organ. This study examines the effectiveness of current texture-based organ classifications when the spatial sample (window) size of the organ is less than a whole cross-section of an organ.

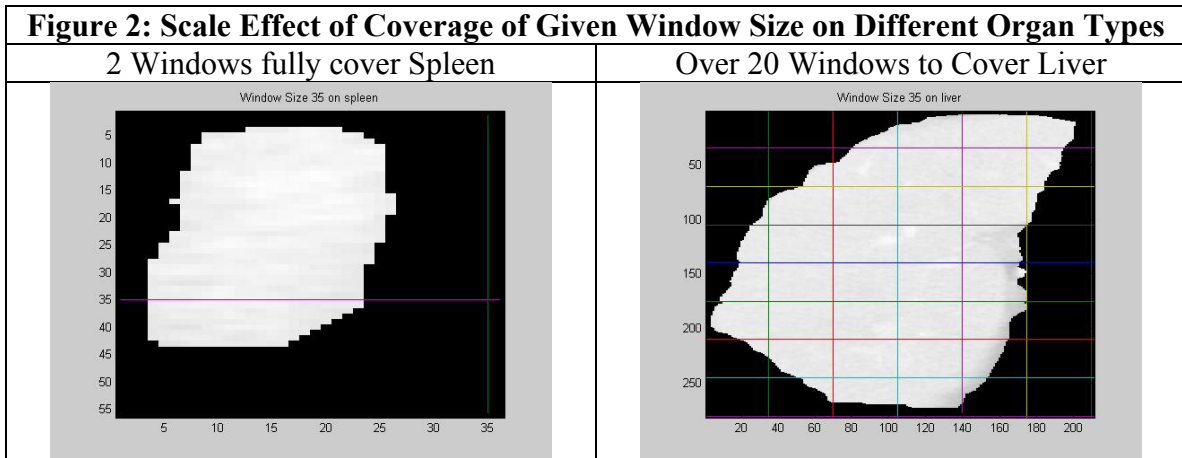
The textural features of whole organs shows notable differences between different organ types. However, sub-sections of organs may have more textural similarity to sub-sections of other organs than to other regions of the same organ, say the kidney calyces and heart valves; or the vasculature of different organs may appear more similar to each other than to remainder of the surrounding organ.

Scale presents another challenge as introduced by the use of different sized windows. The texture response of a smaller window will differ from larger windows when the underlying organ structure varies across the regions those windows sample. Consider the difference between the structure observed by windows of size 15 versus 35 in the kidney example.

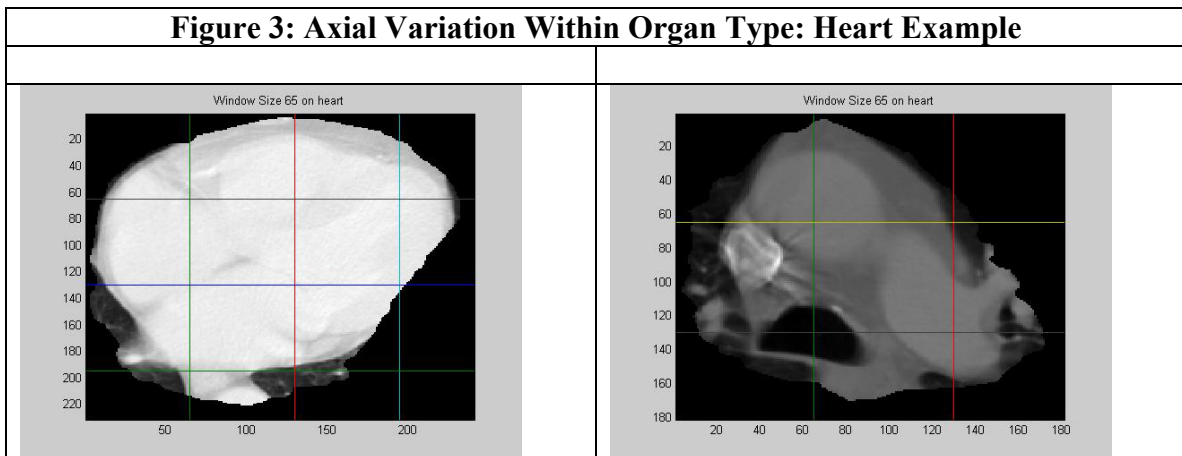


Between organ classes, the coverage proportion of the window over the sampled organ can vary markedly due to size differences between organs as well as the axial size variation of each organ. As an example of this scale effect, consider the coverage

proportion of a 35 pixel window over both a liver and spleen. This sized window fully covers the spleen with two samples, while over 20 samples are required to sample the liver.

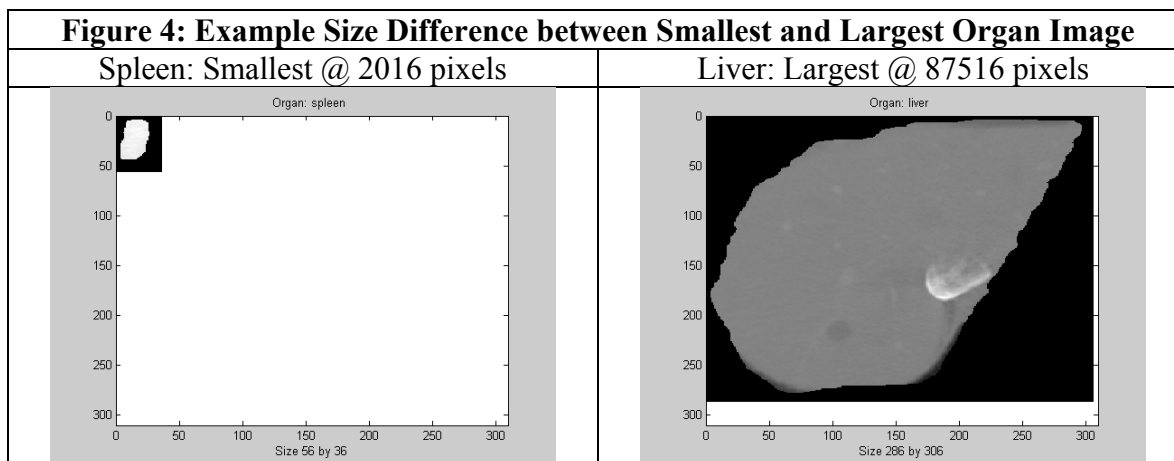


Axial variation (along the pedal-caudal axis) presents an enormous challenge; though not only for sub-organ sampling, the differences within an organ class also challenge whole organ texture classification. However, the use of sub-organ windows may confound the between-organ class differences. This question is not addressed in this paper and will require further research.

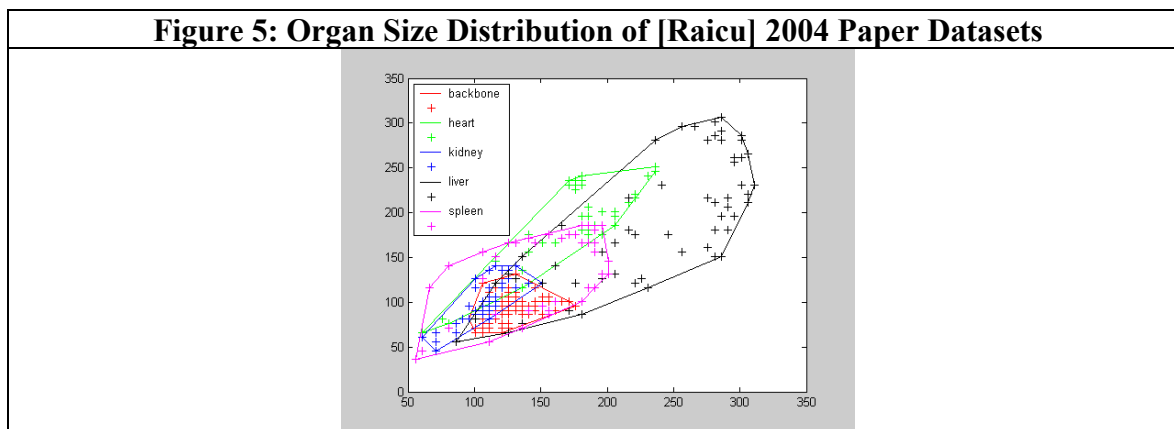


## 2 Organ Size Distribution

Medical images of organ vary in size not only between different organs, but also along the axial dimension. Within the sample set of organs for the original study, the range of image sizes varied from 2K to 87K pixels, a factor of 43 orders of magnitude. An illustration of the range compares the smallest image, a spleen sample to the largest, liver.



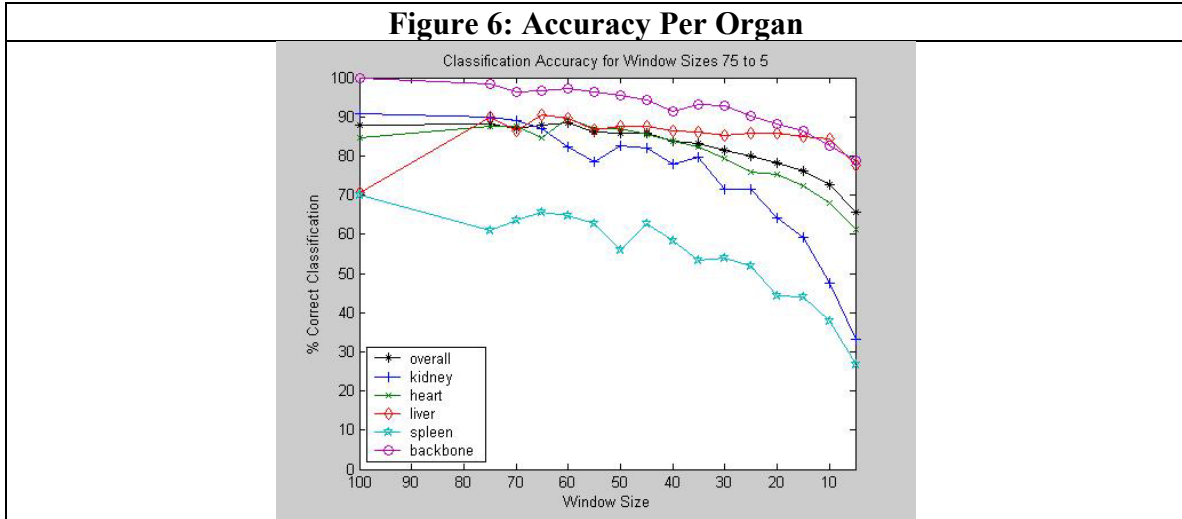
Organ size distribution is important to consider because its effect on window size sampling. Larger organs will provide many more samples for any given window size, and represent even greater proportions of the samples as the window size increases. Within this dataset, only livers and hearts will be represented in window sizes above 200 pixels. For this reason, this study placed an arbitrary upper limit on window size of 75 to ensure adequate representation by all organ types.



### 3 Results

Overall classification accuracy remains robust for (square) window sizes above 30 pixels and degrades rapidly down to small windows of size 5. Overall accuracy identifies the likelihood that any organ will be properly classified with its type. Individual accuracy rates per organ were also studied and represent the likelihood that any particular organ, say liver, is properly classified as a liver. As shown in the Accuracy Per Organ figure, the classification performance for backbone and liver remain consistently high over the range of window sizes studied, though the classification accuracy of kidneys and spleens markedly degraded.

**Figure 6: Accuracy Per Organ**



These results support the use of further studies in the use of textural measurements for classifying and identifying organs when the data extraction techniques uses windows over sub-organ sections rather than fully segmented whole organs, as long as the window sizes remain above approximately 30 pixels. Further studies will extend this analysis of classifier robustness to the use of other types of texture measurement techniques, such as Gabor filtering.

#### 4 References

- [1] D. S. Raicu, J. D. Furst, D. Channin, D. H. Xu, & A. Kurani, "A Texture Dictionary for Human Organs Tissues' Classification", Proceedings of the 8th World Multiconference on Systemics, Cybernetics and Informatics (SCI 2004)