INTELLIGENT MULTIMEDIA PROCESSING LABORATORY

DEPAUL CTI

# Texture Analysis for Computed Tomography Studies

**Stelu Aioanei, Arati Kurani, Dong-Hui Xu**
**CTI Undergraduates**

**Advisors: Jacob Furst, PhD, Assistant Professor**
**Daniela Raicu, PhD, Assistant Professor**

**CTI, DePaul University**

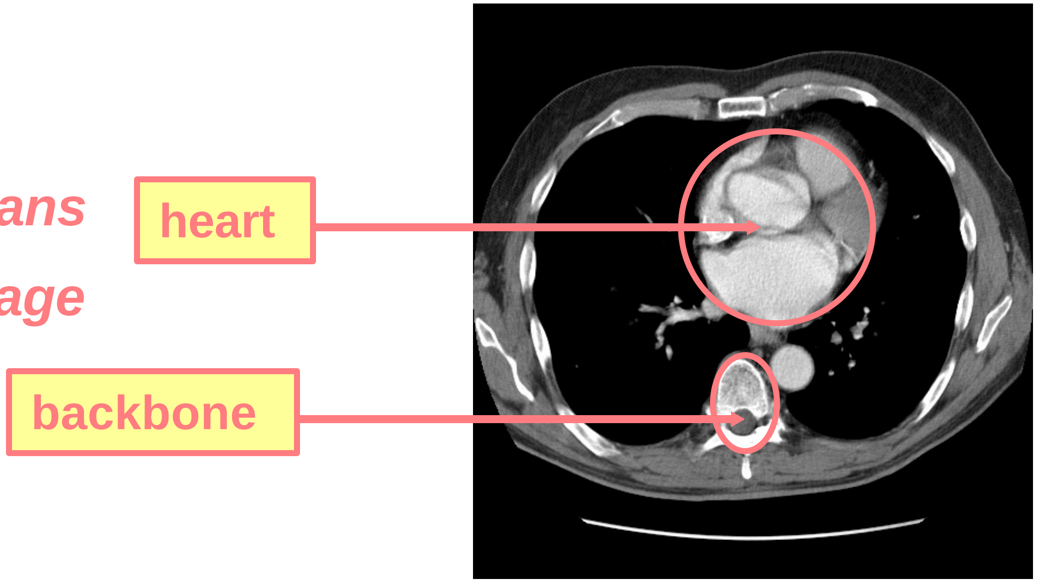
# Undergraduate Research

- **Research Interests**
  - Data Mining for United States Patent and Trademark Office
  - Data Mining for Customer Behavior Modeling
  - **Medical Imaging**
    - Collaborators: Department of Radiology, Northwestern Memorial Hospital, Chicago, IL
    - Dr. David Channin, Director of the Department of Radiology
    - Dr. Alice Averbukh, Research Associate
- **Project**
  - Classification of Tissues Using Texture Information in CT studies

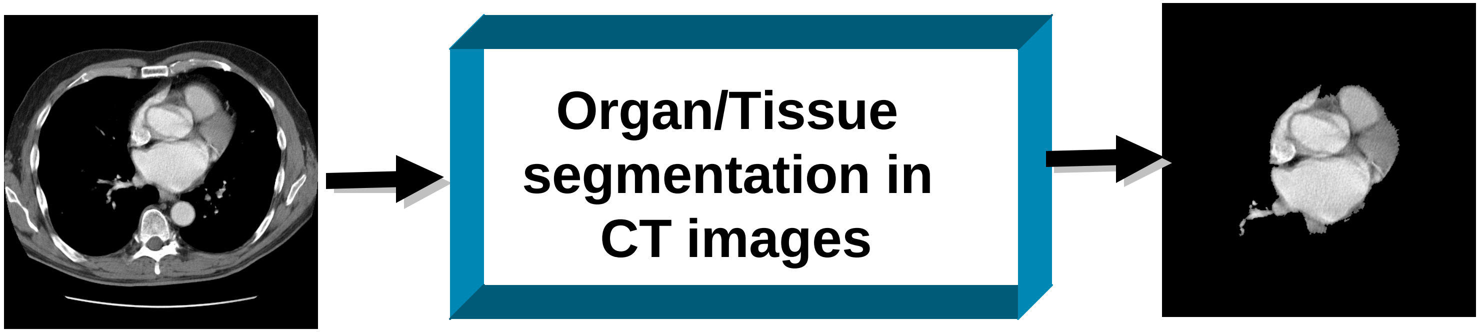
# Classification of Tissues

- **Problem Statement**
  - The classification of human organs using raw data (pixels) from abdominal and chest CT images
  - While traditional texture metrics have concentrated on 2D texture, 3D imaging modalities are becoming more and more prevalent, providing the possibility of examining texture as a volumetric phenomenon
  - We expect that texture derived from volumetric data will have better discriminating power than 2D texture derived from slice data, thus increasing the organ classification accuracy

*labels for the organs present in the image*



# Segmentation

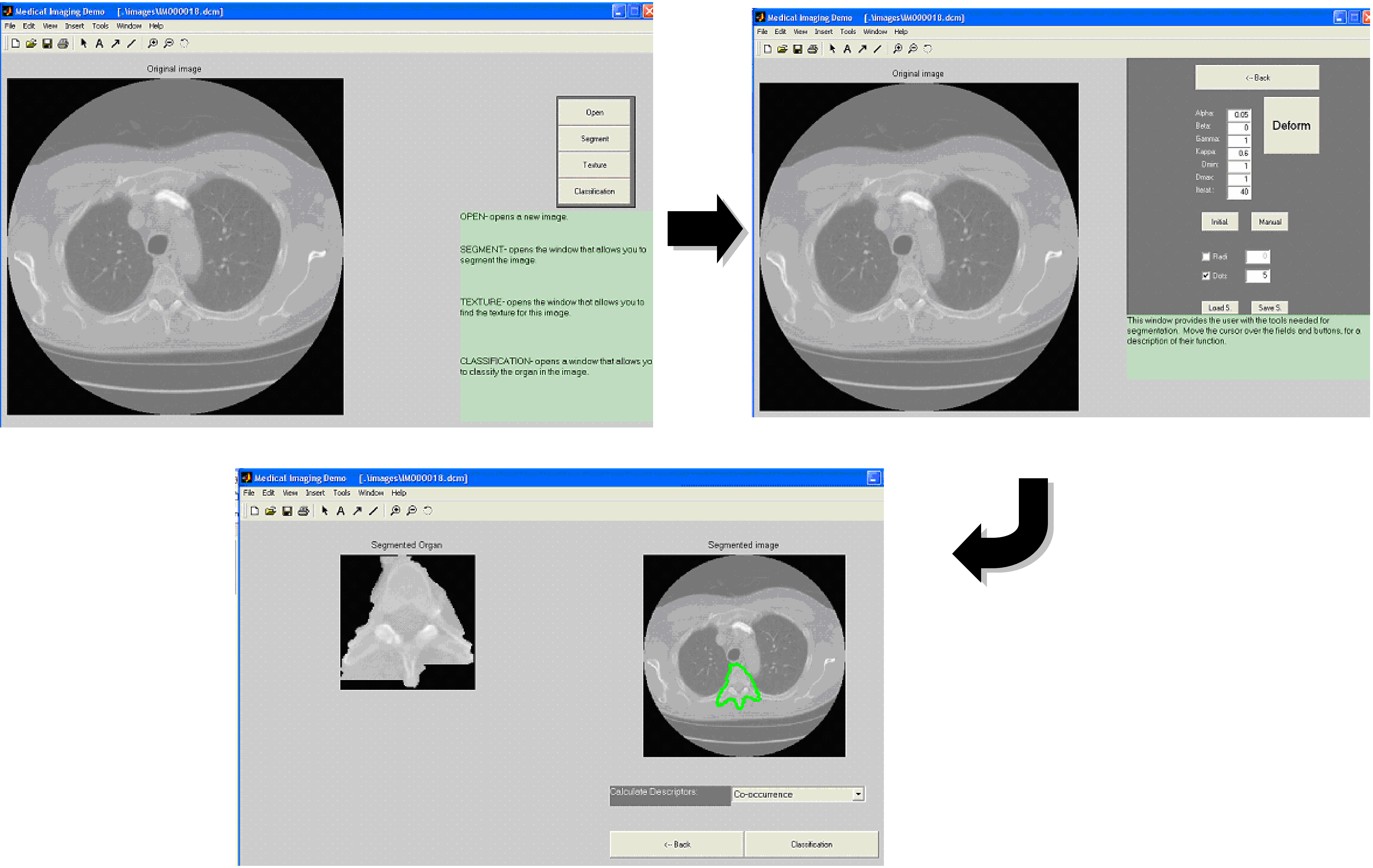



- **Data:** 344 DICOM images
- **Segmented organs**: liver, kidneys, spleen, backbone, & heart
- **Segmentation algorithm:** ***Active Contour Mappings (Snakes)***
  - -A boundary-based segmentation algorithm
  - -Input for the algorithm: a number of initial points & five main parameters that influence the way the boundary is formed
  - -The values of the five parameters simulate the action of two forces:
    - -Internal: designed to keep the snake smooth during the deformation
    - -External: designed to move the snake towards the boundary

INTELLIGENT MULTIMEDIA PROCESSING LABORATORY

DEPAULCTI

# Segmentation



# Texture Models

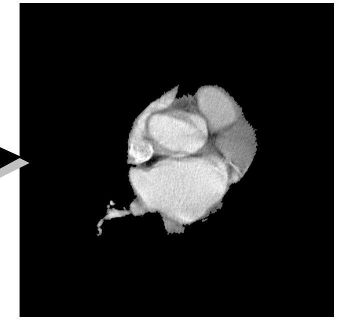


- **What is texture?**
  - *Texture* is a measure of the variation of the intensity of a surface, quantifying properties such as smoothness, coarseness, and regularity
  - *Texture* is a connected set of pixels satisfying a given gray level property which occurs repeatedly in an image region constitutes a textured region

- **Texture Models:**
  - Co-occurrence Matrices
  - Run-Length Matrices

# Co-occurrence Matrices

**Organ/Tissue segmentation in CT images**



**Calculate 2D Co-occurrence Matrices**

# Co-occurrence Matrices



- **2D Co-occurrence Matrix (Haralick 1973)**
  - Capture the spatial dependence of gray-level values within an image
  - A 2D co-occurrence matrix, P, is an n x n matrix, where n is the number of gray-levels within an image
  - The matrix acts as an accumulator so that P[*i* , *j*] counts the number of pixel pairs having the intensities *i* and *j*

Original Matrix

$$\begin{pmatrix} 7 & 5 & 3 & 5 & 7 & 9 \\ 3 & 4 & 2 & 1 & 6 & 3 \\ 8 & 6 & 3 & 4 & 5 & 6 \\ 2 & 4 & 6 & 8 & 2 & 3 \\ 1 & 6 & 5 & 7 & 9 & 3 \\ 2 & 5 & 4 & 8 & 7 & 4 \end{pmatrix}$$

| | | | |
|---|---|---|---|
| 2 | 5 | 1 | **0** |
| 4 | 5 | 2 | **1** |
| 1 | 3 | 2 | **2** |
| **0** | **1** | **2** | |

bin values

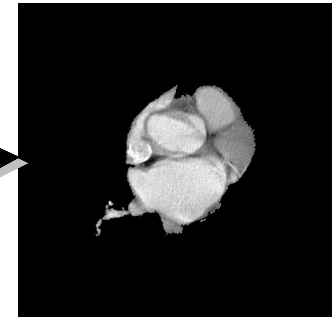
bin values

| Bin | Pixel Values |
|---|---|
| 0 | 1,2,3 |
| 1 | 4,5,6 |
| 2 | 7,8,9 |

# Feature Extraction



Organ/Tissue segmentation in CT images



Calculate 2D Co-occurrence Matrices

Calculate numerical texture descriptors for each region
$[D_1, D_2, \ldots D_{10}]$

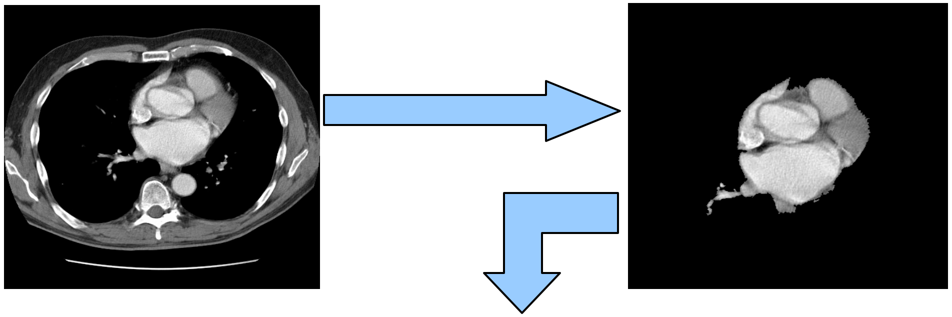
# Feature Extraction

- 2D Co-occurrence Matrix (cont.)
- In order to quantify this spatial dependence of gray-level values, we calculate various textural features proposed by Haralick:
  - Entropy
  - Energy (Angular Second Moment)
  - Contrast
  - Homogeneity
  - SumMean (Mean)
  - Variance
  - Correlation
  - Maximum Probability
  - Inverse Difference Moment
  - Cluster Tendency

# 2D Results



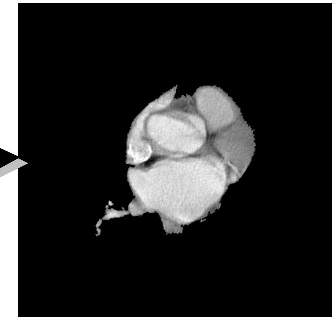
- For each slice per organ we average the feature values derived from the 20 matrices, leaving us with a single value per feature per slice:



| Entropy | Energy | Contrast | Homogeneity | SumMean | Variance | Correlation | Maximum Probability | Inverse Difference Moment | Cluster Tendency |
|---|---|---|---|---|---|---|---|---|---|
| 3.892828 | .034692 | 2.764427 | .6345745 | 11.662886 | 7.308909 | .110921 | .112929 | .44697 | 26.471211 |

# Feature Extraction

**Organ/Tissue segmentation in CT images**

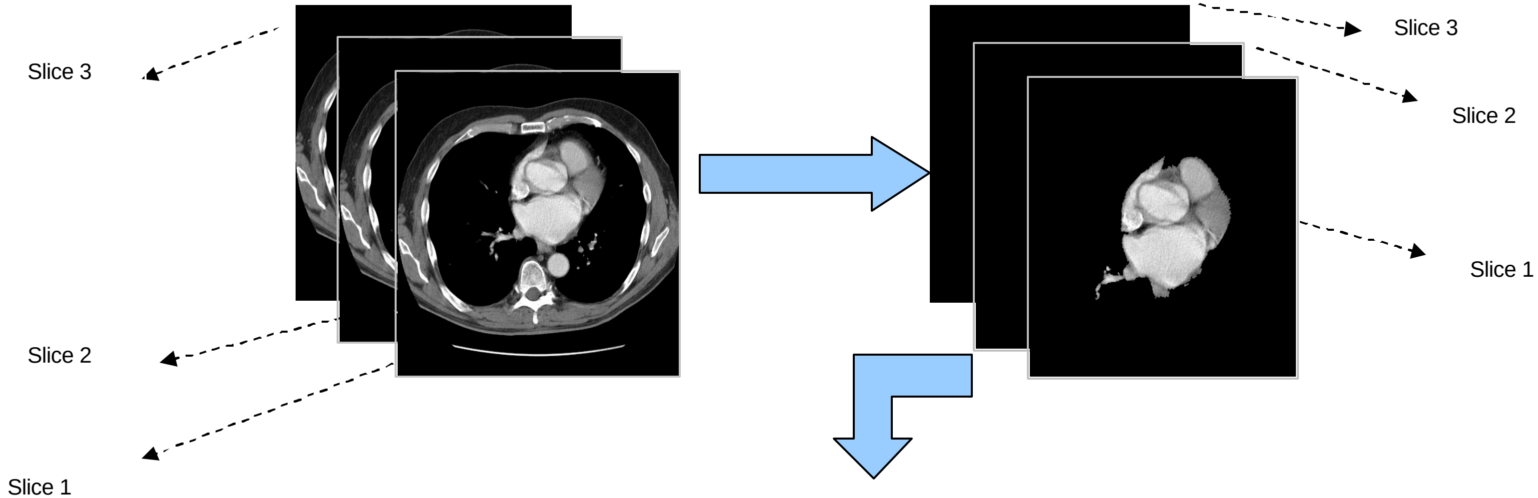


**Calculate Co-occurrence Matrices for Volumetric Data**

**Calculate numerical texture descriptors for each organ**
**$[D_1, D_2, \ldots D_{10}]$**

# Volumetric Results

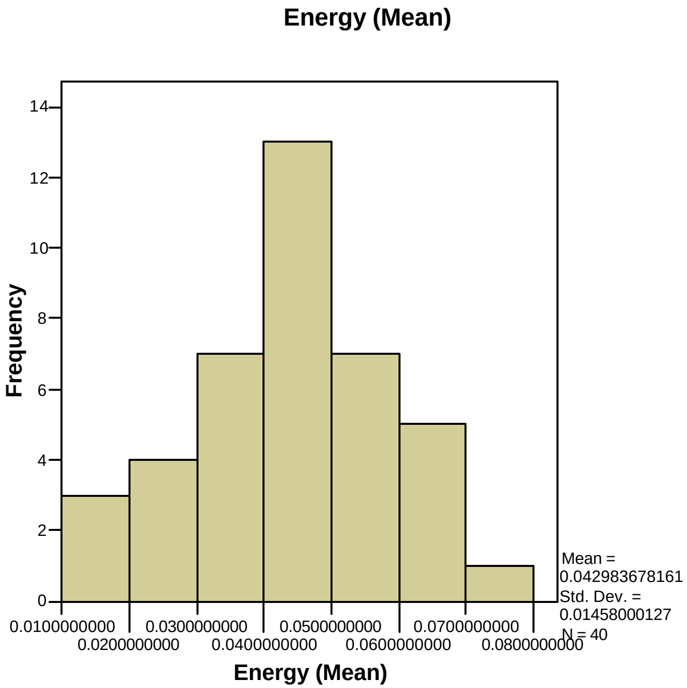
2D dependence matrices that are able to capture the spatial dependence of gray-level values in a set of three-dimensional data (i.e. a set of CT scans for a given patient = single 3D input)




| Entropy | Energy | Contrast | Homogeneity | SumMean | Variance | Correlation | Maximum Probability | Inverse Difference Moment | Cluster Tendency |
|---|---|---|---|---|---|---|---|---|---|
| 3.80848 | 0.03709 | 4.93375 | 0.54080079 | 11.07523 | 3.98559 | 0.0986871 | 0.096027 | 0.4021295 | 11.008588 |

# 3D Results - Backbone

- For each organ, we took the corresponding average feature value and plotted it on the 2D graph (does it fall within min and max?):




- Using the five number summary for 2D data, we defined another range (does the volumetric value fall within Q1 and Q3?):

| | | Energy (Mean) |
|---|---|---|
| Minimum | | 0.01230 |
| Maximum | | 0.07021 |
| Percentiles | **25.00000** | 0.03211 |
| | 50.00000 | 0.04226 |
| | **75.00000** | 0.05375 |

Volumetric Data

Patient 1 - 0.03295

Patient 2 - 0.03746

# 2D and 3D Results



**The distribution of the volumetric co-occurrence descriptors with respect to the 2D data**

| | # of Organs | $Q_1 - Q_3$ | Min - Max |
|---|---|---|---|
| **Entropy** | 5 | 2 | 5 |
| **Energy** | 5 | 2 | 5 |
| **Contrast** | 5 | 1 | 3 |
| **Homogeneity** | 5 | 1 | 3 |
| **SumMean** | 5 | 4 | 5 |
| **Variance** | 5 | 5 | 5 |
| **Correlation** | 5 | 2 | 4 |
| **Maximum Probability** | 5 | 2 | 5 |
| **Inverse Difference Moment** | 5 | 1 | 3 |
| **Cluster Tendency** | 5 | 5 | 5 |
| **TOTAL** | 50 | 25 | 43 |

# 2D and 3D Results



## The distribution of the volumetric co-occurrence descriptors per organ

| | # of Desc. | $Q_1 - Q_3$ | Min - Max |
|---|---|---|---|
| **Backbone** | 10 | 4 | 10 |
| **Heart** | 10 | 4 | 7 |
| **Kidney** | 10 | 3 | 6 |
| **Liver** | 10 | 8 | 10 |
| **Spleen** | 10 | 6 | 10 |
| **TOTAL** | 50 | 25 | 43 |

# Run-Length Encoding for Texture Characterization



## Run-length encoding

- Run-length encoding is used to represent strings of symbols in an image matrix
- For a given image *a gray level run* is defined as a set of consecutive, collinear pixels having the same gray level
- *Length of the run* is the number of pixels in the run

000011100111

- Galloway proposed the use of a run-length matrix for texture feature extraction
- Run-length statistics capture the coarseness of a texture in a specific direction

# Definition of Run-Length Matrices

The run-length matrix p (i, j) is defined by specifying direction and then count the occurrence of runs for each gray levels and length in this direction

(i) Dimension corresponds to the gray level (bin values) and has a length equal to the maximum gray level (bin values) n
(j) dimension corresponds to the run length and has length equal to the maximum run length (bin values)

0 °

$$\begin{pmatrix} 1 & 1 & 2 & 2 & 1 & 1 \\ 3 & 3 & 1 & 1 & 2 & 2 \\ 1 & 1 & 2 & 3 & 1 & 1 \\ 3 & 1 & 2 & 2 & 1 & 1 \\ 1 & 1 & 3 & 2 & 2 & 2 \\ 2 & 3 & 1 & 1 & 2 & 2 \end{pmatrix}$$

| i \ j | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | 1 | 8 | 0 | 0 | 0 | 0 |
| 2 | 2 | 4 | 1 | 0 | 0 | 0 |
| 3 | 4 | 1 | 0 | 0 | 0 | 0 |

# Feature Extraction



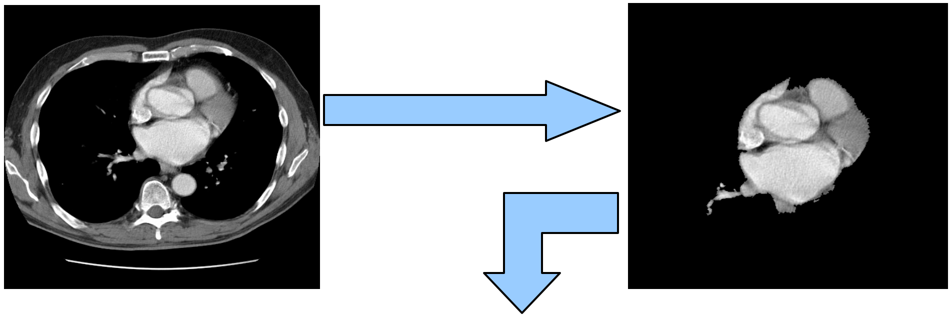
- **Run-length matrix (cont.)**
- Eleven texture descriptors are calculated to capture the texture properties and differentiate among different textures

  **- short run emphasis (SRE)**
  **- long run emphasis (LRE)**
  **- high gray-level run emphasis (HGRE)**
  **- low gray-level run emphasis (LGRE)**
  **- Short Run Low Gray-Level Emphasis (SRLGE)**
  **- Short Run High Gray-Level Emphasis (SRHGE)**
  **- Long Run Low Gray-Level Emphasis (LRLGE)**
  **- Long Run High Gray-Level Emphasis (HRHGE)**
  **- Gray-Level Non-uniformity (GLNU)**
  **- Run Length Non-uniformity (RLNU)**
  **- Run Percentage (RPC)**

# 2D Results



- For each slice per organ we calculated the feature values derived from the four matrices, leaving us with a single value per feature per slice:

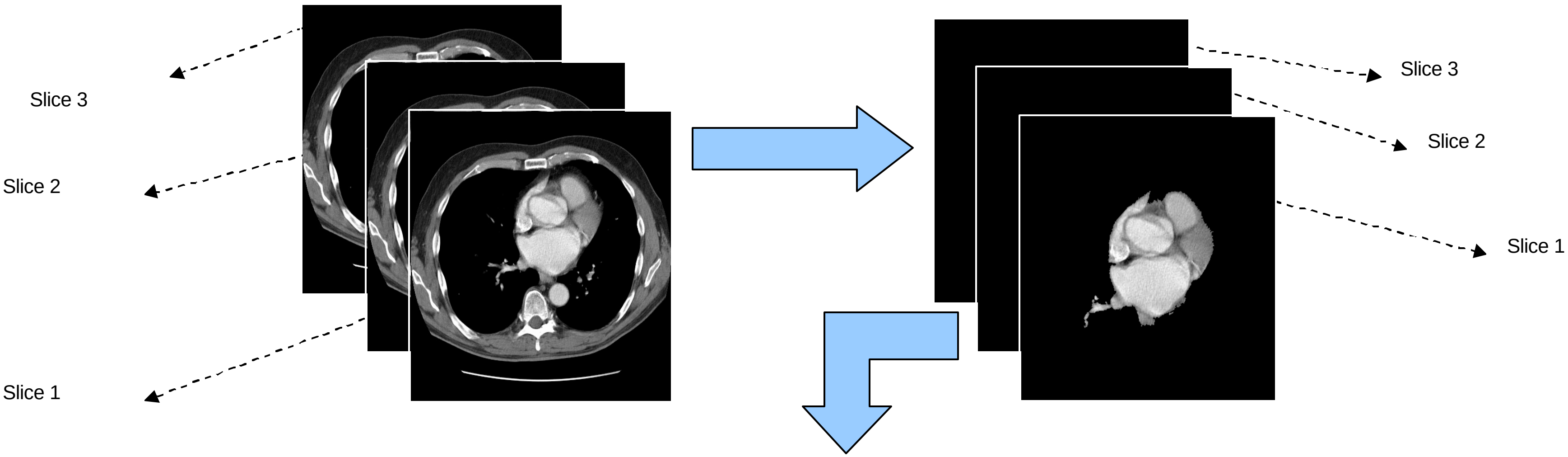


| SRE | LRE | LGRE | HGRE | SRLGE |
|---|---|---|---|---|
| 0.7946625 | 1.9866925 | 0.0049275 | 209.80237 | 0.00399 |

| SRHGE | LRHGE | LRLGE | GLNU | RLNU | RPC |
|---|---|---|---|---|---|
| 163.619125 | 434.95482 | 0.0093575 | 218.6210575 | 584.8621275 | 0.769425 |

# Volumetric Results

2D dependence matrices that are able to capture the spatial dependence of gray-level values in a set of three-dimensional data (i.e. a set of CT scans for a given patient = single 3D input)



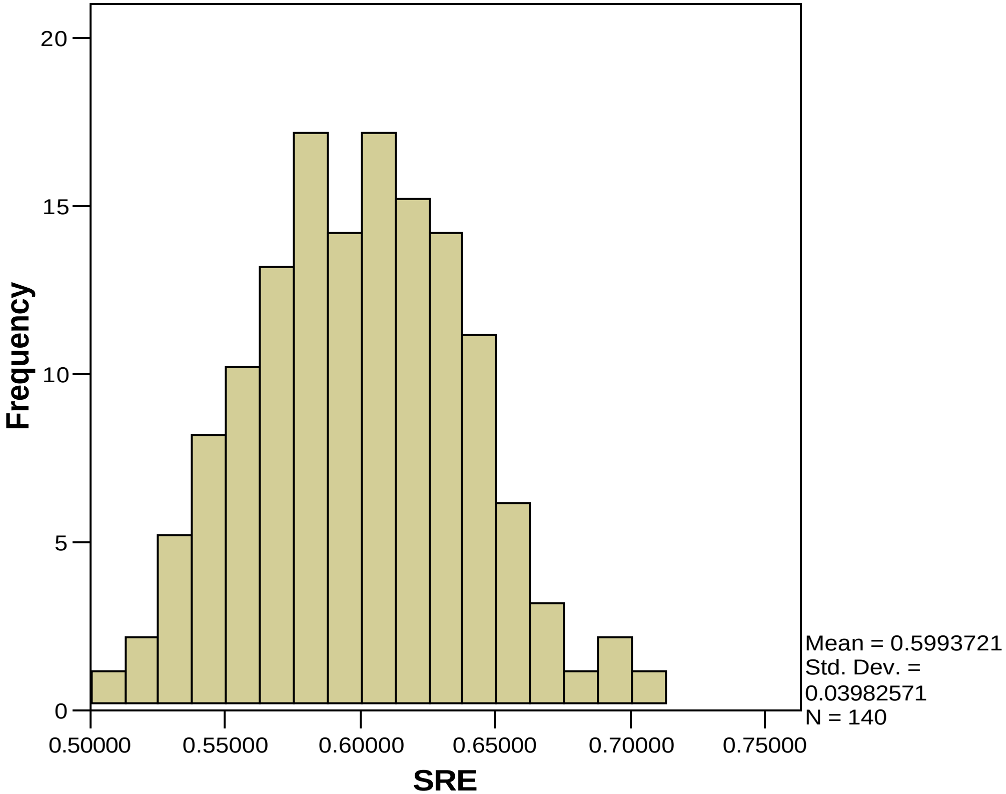

| SRE | LRE | LGRE | HGRE | SRLGE |
|---|---|---|---|---|
| 0.7946625 | 1.9866925 | 0.0049275 | 209.80237 | 0.00399 |

| SRHGE | LRHGE | LRLGE | GLNU | RLNU | RPC |
|---|---|---|---|---|---|
| 163.619125 | 434.95482 | 0.0093575 | 218.6210575 | 584.8621275 | 0.769425 |

# Results

- For each organ, we took the corresponding average feature value and plotted it on the 2D graph
- Using the five number summary for 2D data, we defined another range

**Histogram**




| | | |
|---|---|---|
| Minimum | | 0.50515 |
| Maximum | | 0.71057 |
| percentile | 25 | 0.57065 |
| | 50 | 0.60013 |
| | 75 | 0.62919 |

**Backbone SRE**

# Results



- The distribution of the volumetric run-length descriptors with respect to the 2D data

| | # of Organs | Q1-Q3 | Min-Max |
|---|---|---|---|
| sre | 5 | 0 | 0 |
| lre | 5 | 1 | 3 |
| lgre | 5 | 5 | 5 |
| hgre | 5 | 5 | 5 |
| srlge | 5 | 0 | 4 |
| srhge | 5 | 2 | 3 |
| lrhge | 5 | 1 | 2 |
| lrlge | 5 | 1 | 2 |
| rpc | 5 | 0 | 2 |
| total | 45 | 15 | 26 |

# Results



- The distribution of the volumetric run-length descriptors per organ

| | # of Desc. | Q1-Q3 | Min-Max |
|---|---|---|---|
| backbone | 9 | 3 | 5 |
| heart | 9 | 2 | 6 |
| kidney | 9 | 2 | 2 |
| liver | 9 | 6 | 8 |
| spleen | 9 | 2 | 5 |
| total | 45 | 15 | 26 |

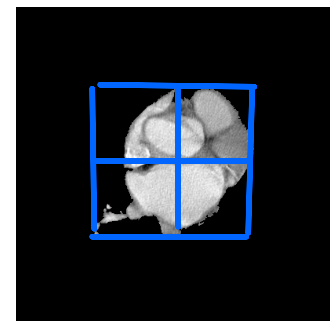
# Classification

- Motivation:

  Investigate the feasibility of automated classification of human organ tissues in a chest and abdominal CT image

- Stage 1: Texture feature extraction

**Organ/Tissue segmentation in CT images**



**Calculate numerical texture descriptors for each region $[D_1, D_2, \ldots D_{21}]$**

**Data:** 340*4 = 1360 regions of interests

**Segmented organs**: liver, kidneys, spleen, backbone, & heart

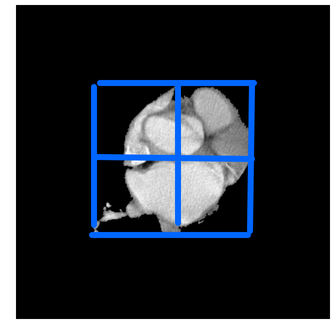
# Classification



## Stage 2: Classifier creation

Data: 911 training set (66.6%), 449 testing set (33.3%)

**Organ/Tissue segmentation in CT images**



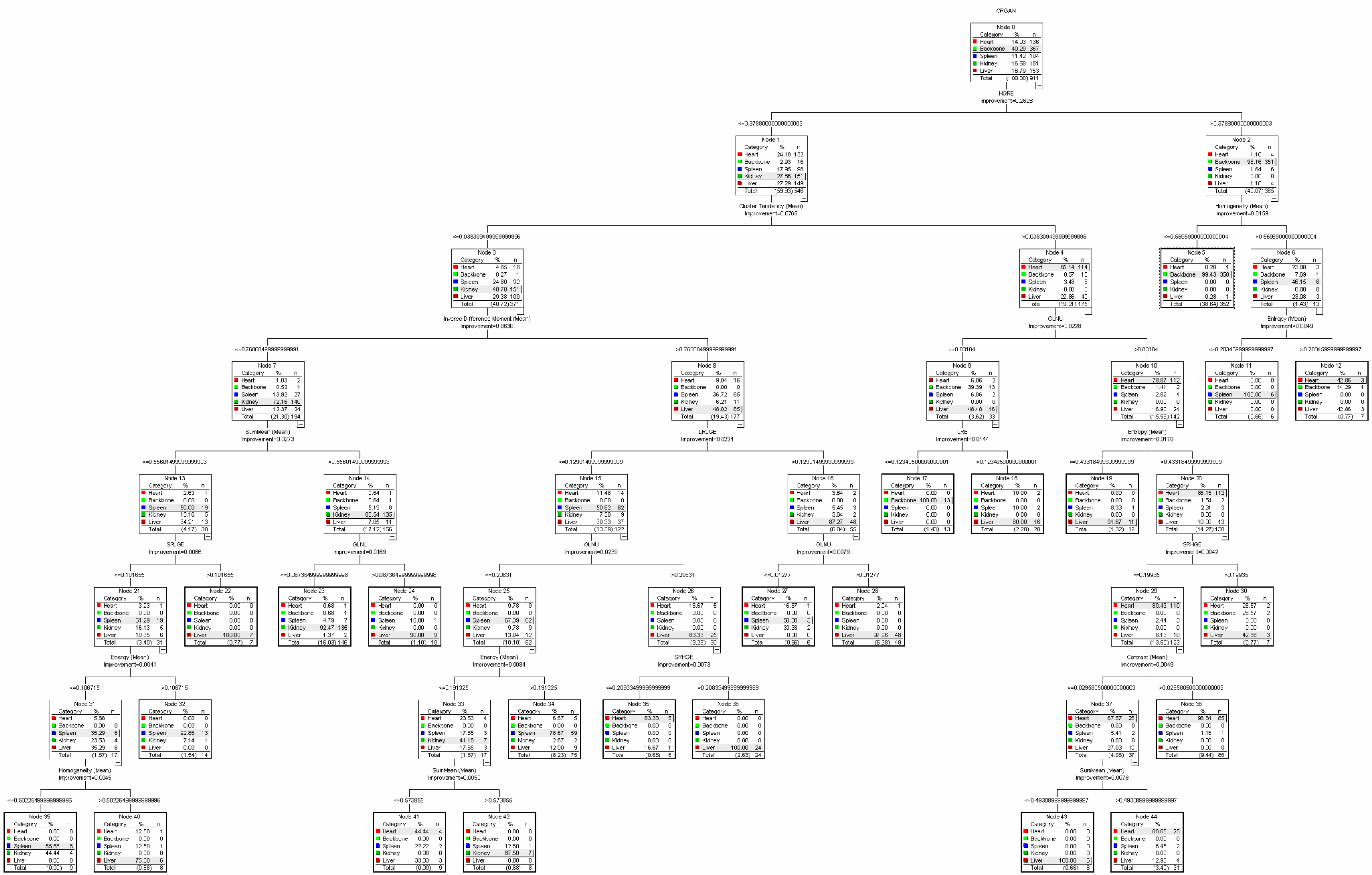
**Calculate numerical texture descriptors for each region $[D_1, D_2, \ldots D_{21}]$**

**Classification rules for tissue/organs in CT images**

**IF HGRE <= 0.38 AND**
**ENTROPY > 0.43**
**AND SRHGE <= 0.20 AND**
**CONTRAST > 0.029**
**THEN Prediction = Heart**
**Probability = 0.99**

# Classifier Creation

**Decision Tree:** A decision tree is a flow-chat-like structure that resembles a tree



# Classifier Creation

- Decision rules highlight the most relevant descriptors for the classification of chest and abdominal image

IF ((HGRE <= 0.3788) AND
(CLUSTER > 0.0383095 ) AND
(GLNU <= 0.03184) AND
(LRE <= 0.123405))
THEN
Node = 17
Prediction = Backbone
Probability = 1.000000

- Classification accuracy parameters
  - sensitivity
  - specificity
  - precision
  - accuracy

# Results



## Organ/Tissue Classification

Decision Tree Accuracy on Testing Data

(Co-occurrence, Run-length, and Combined):

| ORGAN | Sensitivity | Specificity | Precision | Accuracy |
|---|---|---|---|---|
| Backbone | 96% / 98% / 98% | 99% / 100% / 99% | 99% / 99% / 99% | 98% / 99% / 99% |
| Liver | 64% / 57% / 78% | 96% / 98% / 95% | 75% / 84% / 71% | 92% / 92% / 92% |
| Heart | 79% / 82% / 75% | 96% / 95% / 98% | 80% / 77% / 90% | 94% / 93% / 95% |
| Kidney | 89% / 89% / 89% | 96% / 93% / 96% | 80% / 67% / 77% | 94% / 92% / 95% |
| Spleen | 60% / 44% / 60% | 93% / 93% / 95% | 53% / 45% / 63% | 89% / 87% / 91% |

# Conclusion and Future Work



- **Conclusion**
  - Using only 21 descriptors it is possible to classify different organ tissues in the CT images
  - Combining some other texture models will increase the performance of the classifier
- **Future Work**
  - Improve snake segmentation
  - Compare the classification result with result obtained using neural network technique
  - Use texture descriptors derived from volumetric data to aid the classification in CT image