Report - Exercise 02

Medical Image Computing

MT-M-3-ILV-IM2

Master Medical Technologies

Semester 3

Lecturer: Marco Augustin

Group: MA-MED-21-VZ

Authors: Peter Rott, Filip Slatković, Markus Wernard

21st December 2022

Contents

1 Image texture descriptions **1**

1.1 Size of the original image and blocks for further processing 1

1.2 Gray level co-occurrence matrix (GLCMs) 2

1.3 Features: Correlation, contrast, energy and homogeneity of GLCM . . . 2

1.3.1 Correlation 2

1.3.2 Contrast 3

1.3.3 Energy 4

1.3.4 Homogeneity 4

1.4 Design Matrix 5

1.5 Extra task: change distance in GLCM to 3 5

1.5.1 Correlation 5

1.5.2 Contrast 6

1.5.3 Energy 7

1.5.4 Homogeneity 7

2 Texture-based clustering **9**

2.1 Clustering using k-means clustering 9

2.2 Overlay to the original image 9

1 Image texture descriptions

1.1 Size of the original image and blocks for further processing

Figure 1.1: Raw image from which several features gets extracted

First task is to load the raw image breastXray.tif, it is shown in figure 1.1. The image has a size of 560x480 pixels, which results in 672 blocks with the size of 20x20 pixels. These are arranged in the format of 28x24 blocks. These blocks are used for further steps.

1.2 Gray level co-occurrence matrix (GLCMs)

The gray level co-occurrence matrix (GLCM) is calculated for each of the 672 blocks. That happens by running through two for-loops, where one defines the position operator D_x and one defines D_y. The distance is set to D = 1 and the GLCM is extracted in four different directions [0°, 45°, 90°, 135°] with 16 gray levels. Because of these 16 gray levels, each GLCM is a matrix of the size [16x16].

1.3 Features: Correlation, contrast, energy and homogeneity of GLCM

In the following figures processed images based on four different descriptors and four different orientations each are shown. Therefore following descriptors are extracted in four different directions [0°, 45°, 90°, 135°]:

- Figure 1.2: Correlation
- Figure 1.3: Contrast
- Figure 1.4: Energy
- Figure 1.5: Homogeneity

1.3.1 Correlation

Correlation describes the occurrence probability of pixel pairs. This can be seen in figure 1.2 where the background becomes white, because of a high occurrence probability of pixel pairs. Inside the breast tissue the correlation hardly depends on the direction what also can be seen in the figure.

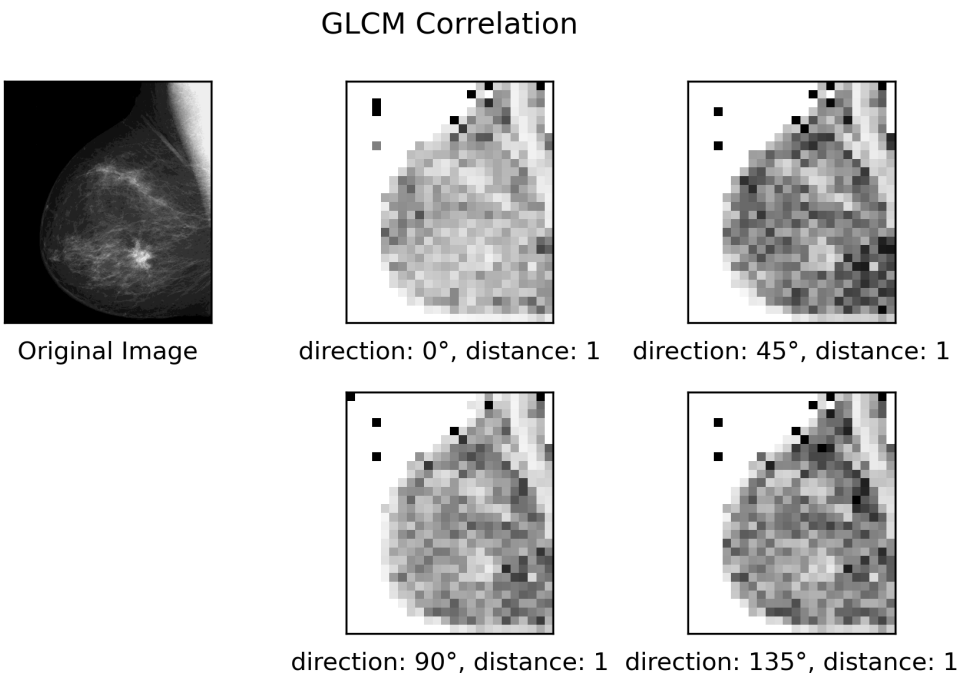

Figure 1.2: Correlation of GLCM calculated for the angels: 0°, 45°, 90° and 135°

1.3.2 Contrast

Contrast describes the local variation of pixel values and kind of amplifies high contrast areas. This can be seen in figure 1.3 where the bright spot inside the breast is highlighted through the GLCM calculation.

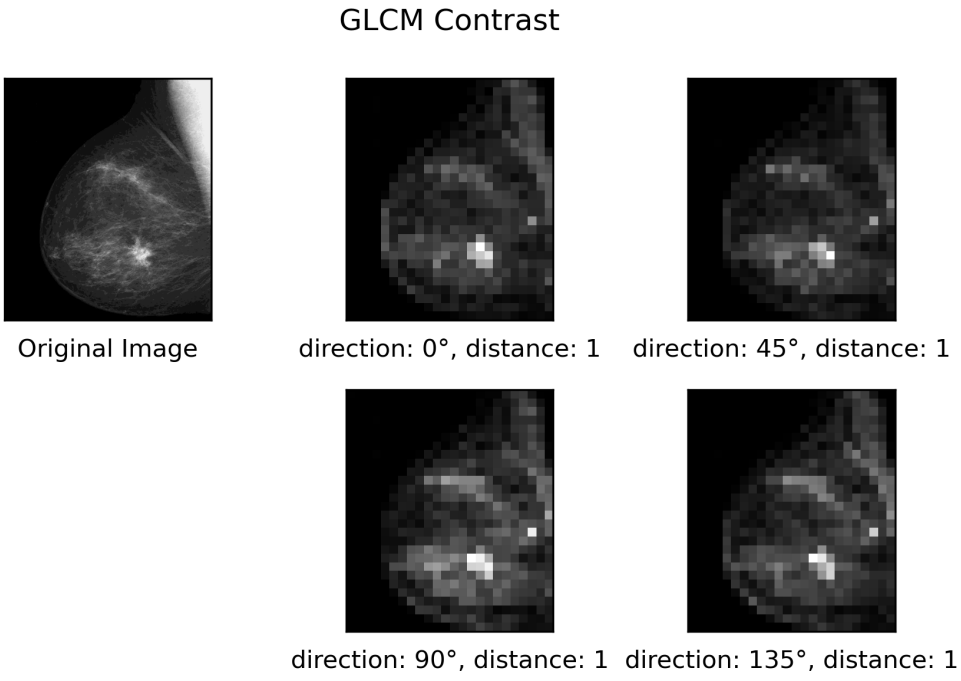

Figure 1.3: Contrast of GLCM calculated for the angels: 0°, 45°, 90° and 135°

1.3.3 Energy

Energy describes the uniformity of the image. In other words the lower changes of pixel values inside a block are, the higher values are calculated. Except of the background where the pixel values equal zero what cannot lead to any value by the used concept of summed up squares. In figure 1.4 can be seen that at the location of the tumor, where high changes are, the energy is calculated low.

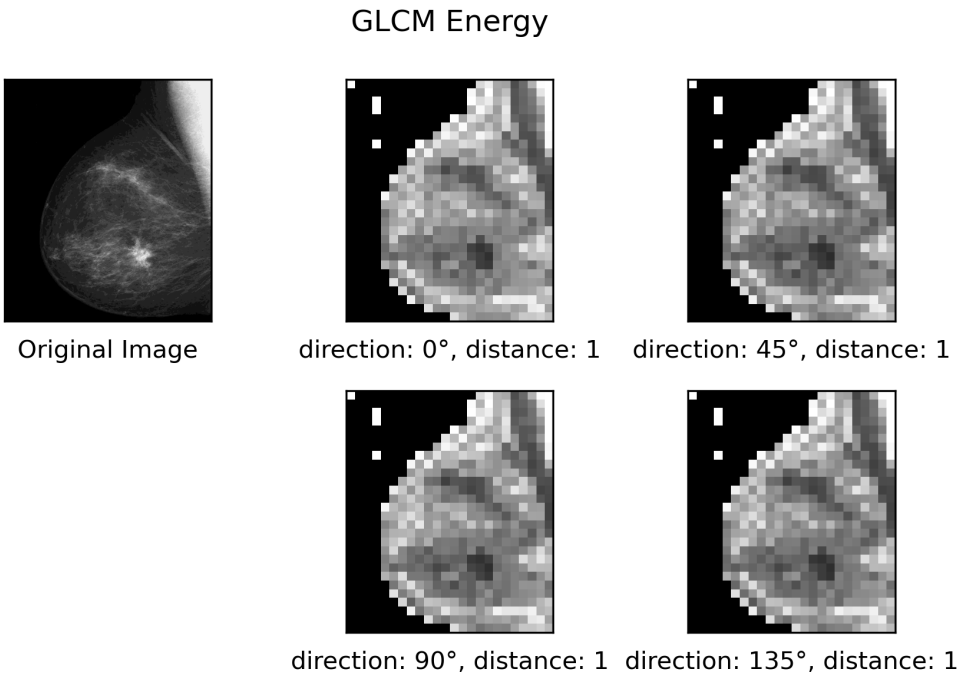

Figure 1.4: Energy of GLCM calculated for the angels: 0°, 45°, 90° and 135°

1.3.4 Homogeneity

Homogeneity describes distribution of pixels, which decreases with lower contrast. In figure 1.5 can be seen that at the location of the tumor, where we found a high contrast in figure 1.3, the homogeneity is calculated low.

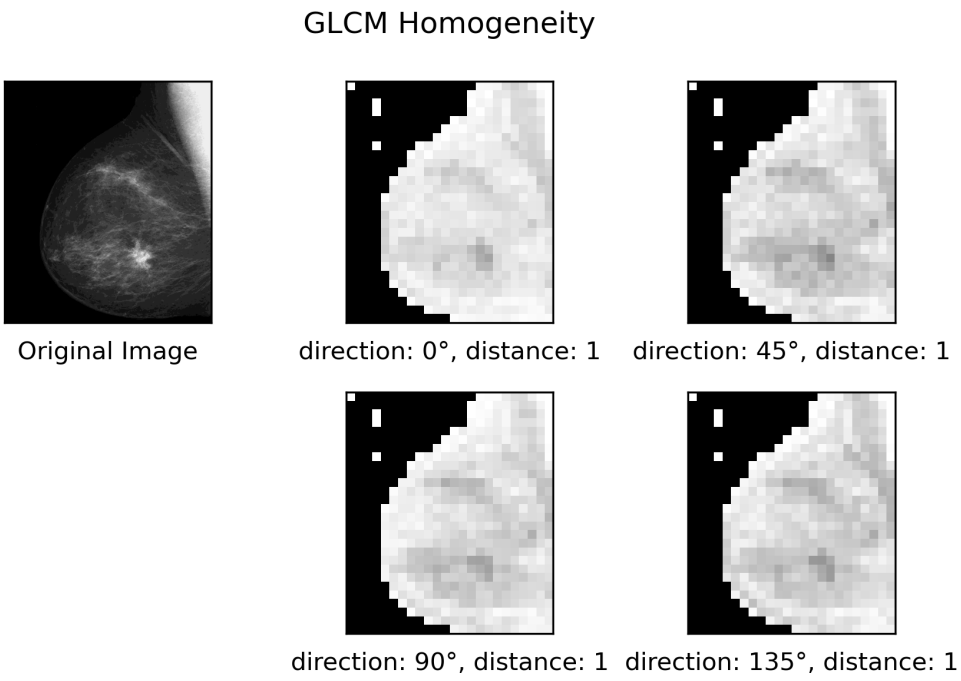

Figure 1.5: Homogeneity of GLCM calculated for the angels: 0°, 45°, 90° and 135°

1.4 Design Matrix

The design matrix displays all 16 texture descriptors as features with its belonging blocks as observations. Since there are 28x24 blocks and 4 descriptors with 4 angels each, the matrix has a size of 672 rows and 16 columns. Every feature has 672 observation.

1.5 Extra task: change distance in GLCM to 3

The distance is raised to 3 instead of 1 and the computed images are compared to each other.

1.5.1 Correlation

Increasing the distance to 3 instead of 1, the occurrence probability of having pixel pairs even decreases, like shown in figure 1.6.

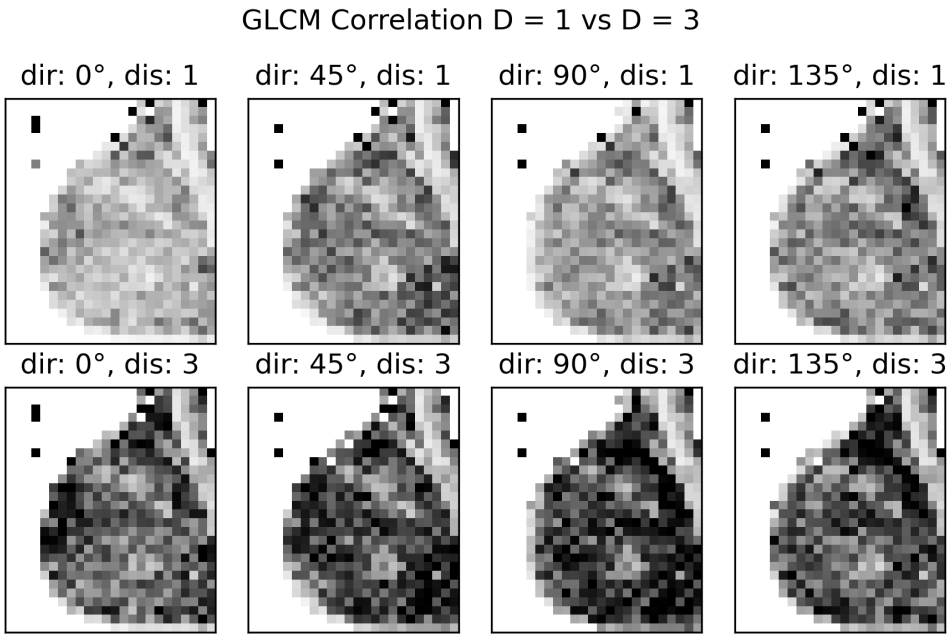

Figure 1.6: Correlation of GLCM with distance = 1 vs distance =3

1.5.2 Contrast

Increasing the distance to 3 instead of 1, lead to a better focused contrast calculation. In figure 1.7 can be seen, that the tumor can more likely be detected by the algorithm with higher distance.

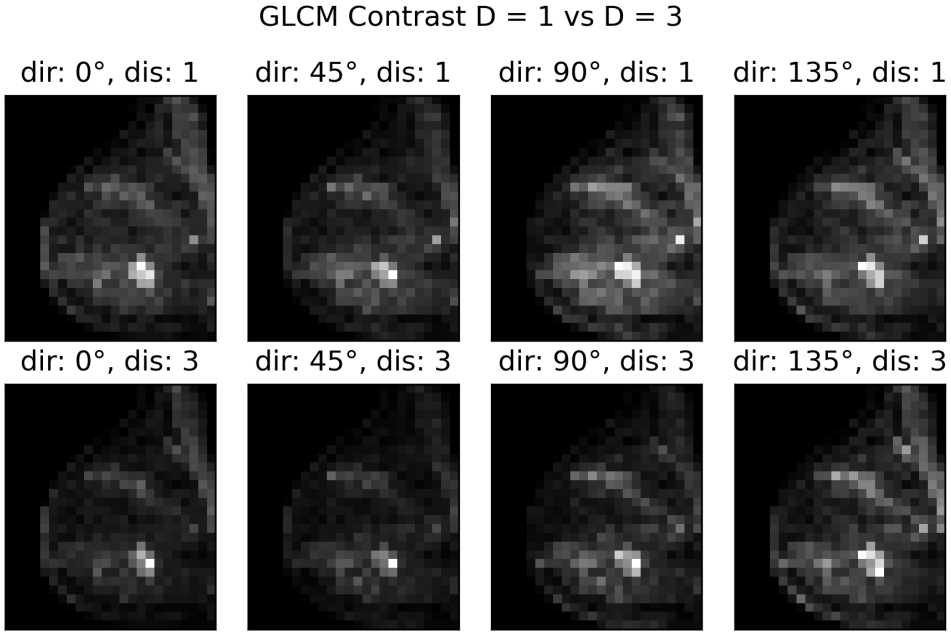

Figure 1.7: Contrast of GLCM with distance = 1 vs distance =3

1.5.3 Energy

As shown in figure 1.8 the increased distance does not lead to a striking difference. That could be seen as a possibility to increase performance of the algorithm, because of lower calculation effort.

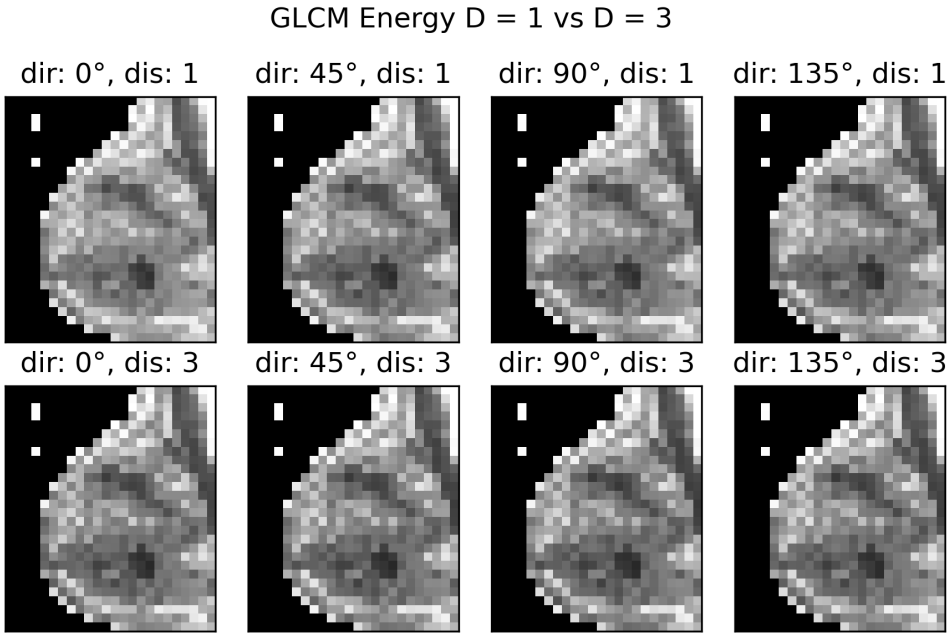

Figure 1.8: Energy of GLCM with distance = 1 vs distance =3

1.5.4 Homogeneity

Analog to the statement in subsection 1.3.4 figure 1.9 shows the correlation between homogeneity and contrast. The contrast with increased distance is more focused, so also the homogeneity is more focused in the location of the tumor.

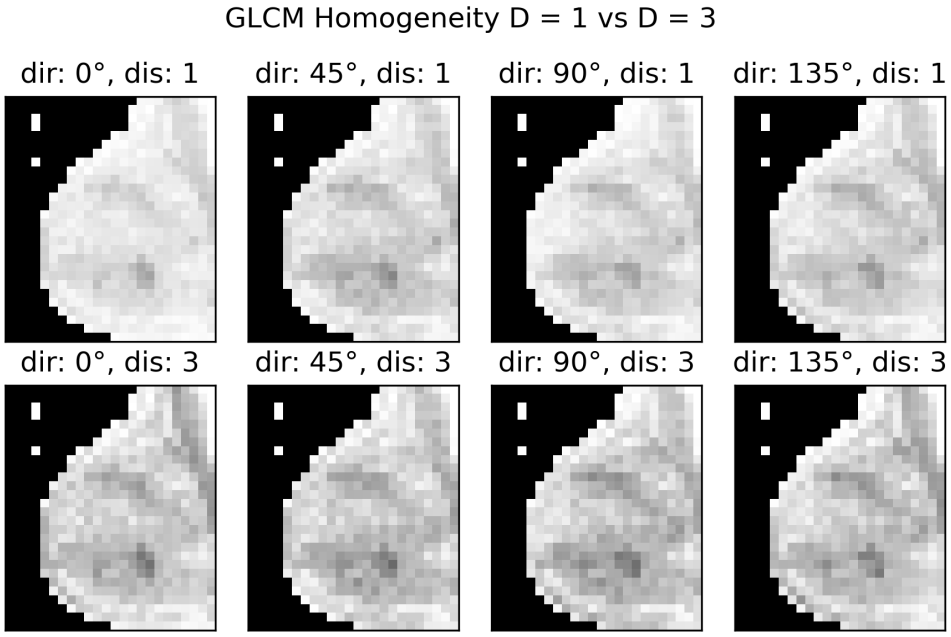

Figure 1.9: Homogeneity of GLCM with distance = 1 vs distance =3

2 Texture-based clustering

2.1 Clustering using k-means clustering

In this task the blocks represented by the design matrix are clustered using the k means clustering. To use the k-means clustering function of the scikit library the design matrix needs to has the size of (672,16). The clusters are set to 4 and the repetition for the initialization is set to 10.

2.2 Overlay to the original image

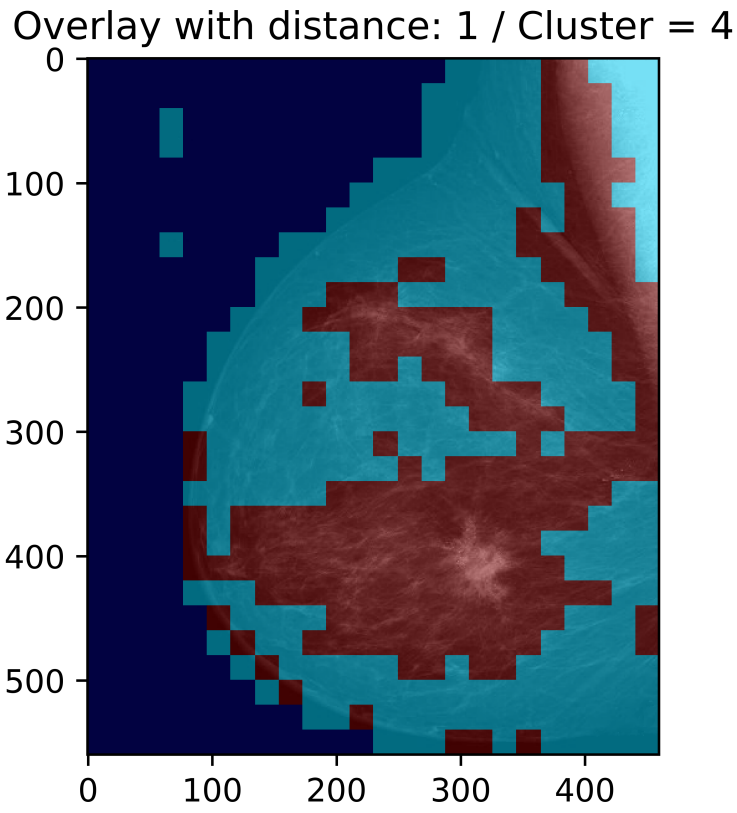

Figure 2.1: Overlay of the segments on the original image with distance = 1 and kmeans cluster = 4

In this part the blocks (20x20 pixels) corresponding to the four different labels should be visualized and afterwards the processed image is to be superimposed on the original image. All result images needs to be resized from 28x24 blocks to 560x480 pixels. The first result for a distance of 1 and a cluster of 4 can be seen in 2.1. Somehow a wrong label is detected by the kmeans.label_ function and the fist column is defined as an extra label. Therefore the fist column is manipulated to have the same label as the second column. However, this process results in the loss of a label as it can be recognized in the 2.1. There are only 3 colors instead of 4. The background of the image is dark blue, the breast in general is shown in light blue and the slightly denser tissue in red. The interesting part, the tumor or the node tissue is not labeled. Since there is a loss of one label, 2.2 shows the result for a kmeans algorithm with a cluster number of 5. The result shows 4 colors, here the background is displayed in light blue, the general breast in orange, the slightly denser tissue in dark blue and the node tissue is shown in red. Even better results are visualized in 2.3. Here the distance is set to 3 and the tumor tissue is even better localized.

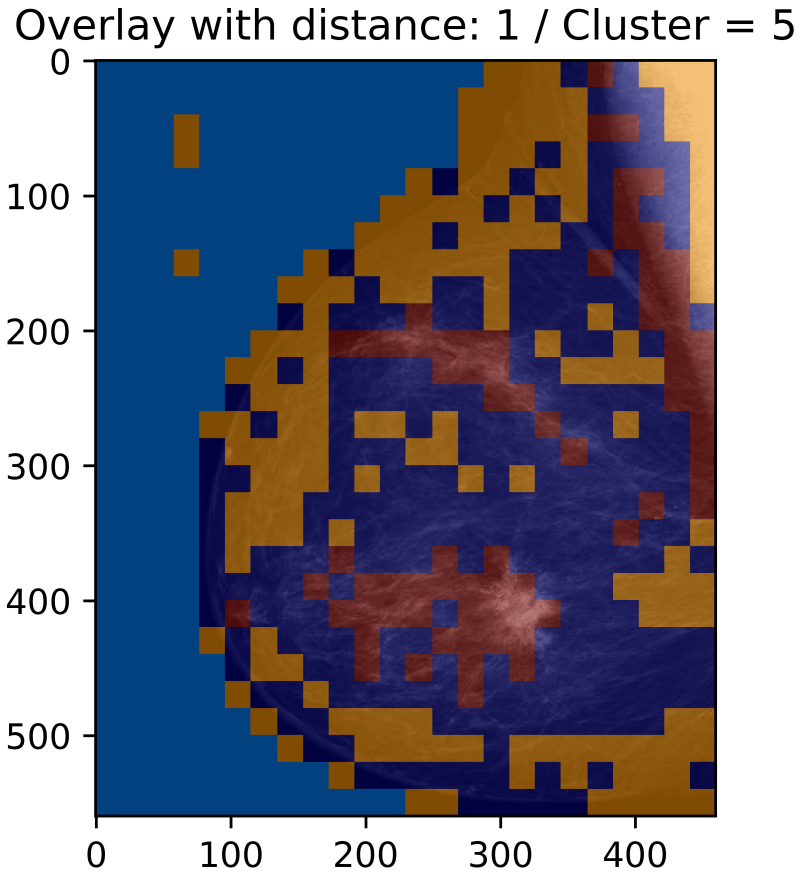

Figure 2.2: Overlay of the segments on the original image with distance = 1 and kmeans cluster = 5

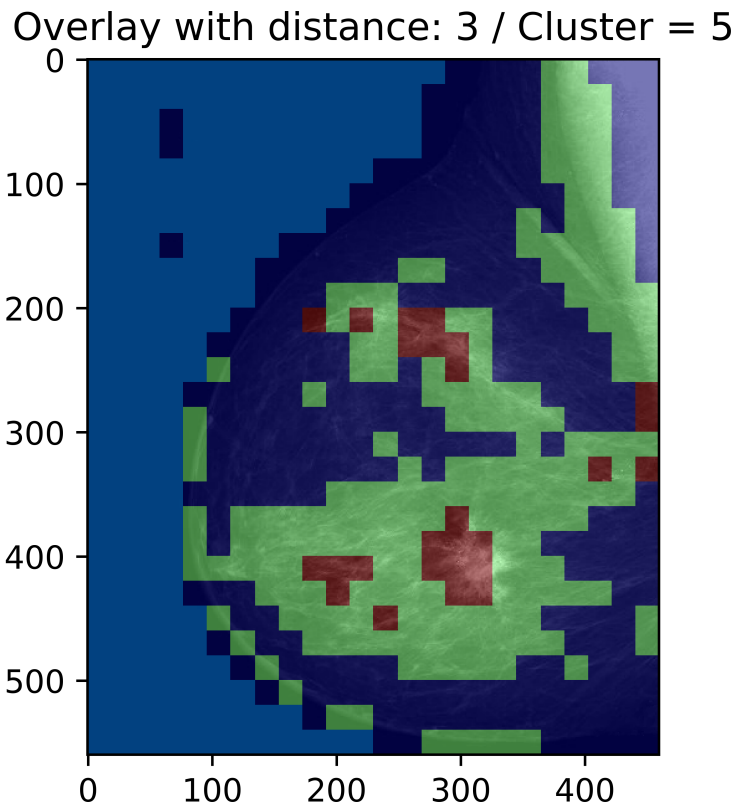

Figure 2.3: Overlay of the segments on the original image with distance = 3 and kmeans cluster = 5

List of Figures

1.1 Raw image from which several features gets extracted 1
1.2 Correlation of GLCM calculated for the angels: 0°, 45°, 90° and 135° . . 3
1.3 Contrast of GLCM calculated for the angels: 0°, 45°, 90° and 135° . . . 3
1.4 Energy of GLCM calculated for the angels: 0°, 45°, 90° and 135° 4
1.5 Homogeneity of GLCM calculated for the angels: 0°, 45°, 90° and 135° 5
1.6 Correlation of GLCM with distance = 1 vs distance =3 6
1.7 Contrast of GLCM with distance = 1 vs distance =3 6
1.8 Energy of GLCM with distance = 1 vs distance =3 7
1.9 Homogeneity of GLCM with distance = 1 vs distance =3 8

2.1 Overlay of the segments on the original image with distance = 1 and kmeans cluster = 4 9
2.2 Overlay of the segments on the original image with distance = 1 and kmeans cluster = 5 10
2.3 Overlay of the segments on the original image with distance = 3 and kmeans cluster = 5 11