TMS016 Spatial Statistics and Image Analysis

Chalmers University of Technology

Project report 3

Distinguishing tablet components in micro-CT images using image segmentation.

Project group number 10

Emma Stavås
stavase@student.chalmers.se
Martina H. Billstein
marbil@student.chalmers.se
Devosmita Chatterjee
devcha@student.chalmers.se

May 17, 2019

Abstract

Not done...Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut purus elit, vestibulum ut, placerat ac, adipiscing vitae, felis. Curabitur dictum gravida mauris. Nam arcu libero, nonummy eget, consectetuer id, vulputate a, magna. Donec vehicula augue eu neque. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Mauris ut leo. Cras viverra metus rhoncus sem. Nulla et lectus vestibulum urna fringilla ultrices. Phasellus eu tellus sit amet tortor gravida placerat. Integer sapien est, iaculis in, pretium quis, viverra ac, nunc. Praesent eget sem vel leo ultrices bibendum. Aenean faucibus. Morbi dolor nulla, malesuada eu, pulvinar at, mollis ac, nulla. Curabitur auctor semper nulla. Donec varius orci eget risus. Duis nibh mi, congue eu, accumsan eleifend, sagittis quis, diam. Duis eget orci sit amet orci dignissim rutrum.

Nam dui ligula, fringilla a, euismod sodales, sollicitudin vel, wisi. Morbi auctor lorem non justo. Nam lacus libero, pretium at, lobortis vitae, ultricies et, tellus. Donec aliquet, tortor sed accumsan bibendum, erat ligula aliquet magna, vitae ornare odio metus a mi. Morbi ac orci et nisl hendrerit mollis. Suspendisse ut massa. Cras nec ante. Pellentesque a nulla. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Aliquam tincidunt urna. Nulla ullamcorper vestibulum turpis. Pellentesque cursus luctus mauris.

Contents

1 Introduction

Image segmentation is a method for detecting different groups of pixels in an image with the same features, e.g. texture, color or intensity, and separating them from one another [1]. This can for example be useful when creating maps from satellite images, where roads and rivers should be separated from a background of forests and cities. Image segmentation is also used within the medical field to separate different tissues and bones in images from MR-scans.

In this project, image segmentation will be used to determine the location and spread of various components in tablets developed and produced by AstraZeneca. The segmentation will be done on several micro-CT images of tablets at different cross-sections. Our goal is to analyze a few different images of tablets and compare them to one another. The images will be subject to preprocessing, segmentation and analysis. The share of each component will also be calculated and compared to theoretical values provided by AstraZeneca. The tablets have been produced with three different methods, namely the Direct Compression method, Rolling compaction method and Twin-screw granulation method. This study compares the resulting images of the tables produced from those methods, but will not cover the different mechanisms of the methods and how that might affect the components in the tablets.

2 Data

Micro-CT images of tablets were given from AstraZeneca. Four tablets were picked out for analysis; two made with the Direct Compression (DC) method, and one picture each from the Rolling compaction (RC) and Twin-screw granulation (TSG) methods. For each tablet, five images were selected. The images are cross sections taken consecutively in the middle of the tablets. Each image is of the size 4868×4868 pixels. An example of what the micro-CT images look like is presented in figure 1. The image in the figure is made with the DC method and will be used as a first test to establish how the segmentation of the other images will be performed.

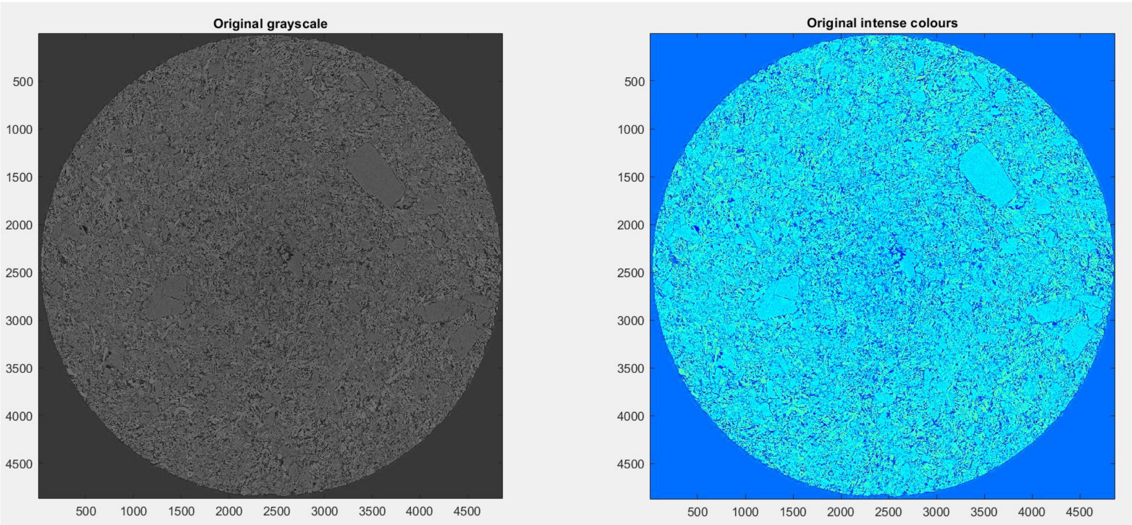

Figure 1: *Example of an original micro-CT image from a tablet produced by the DC method. This image is used as a test to establish methods for further segmentations. The left image has colormap gray and the right image has colormap jet from MATLAB.*

A list of the different types of components and their expected percentage share can be seen in table 1. Additionally, the tablets may also include air bubbles.

Table 1: *The fraction of different components in the tablets.*

Ingredient	Volume fraction
Drug substance	25%
Filler 1	33%
Filler 2	33%
Binder	3%
Disintegrant	4%
Lubricant	2%

An image from each of the tablets that will be studied in this report are seen in figure 2. Comparing the images from different tablets suggest that some differences in the composition of the tablets can be expected depending on what production method was used.

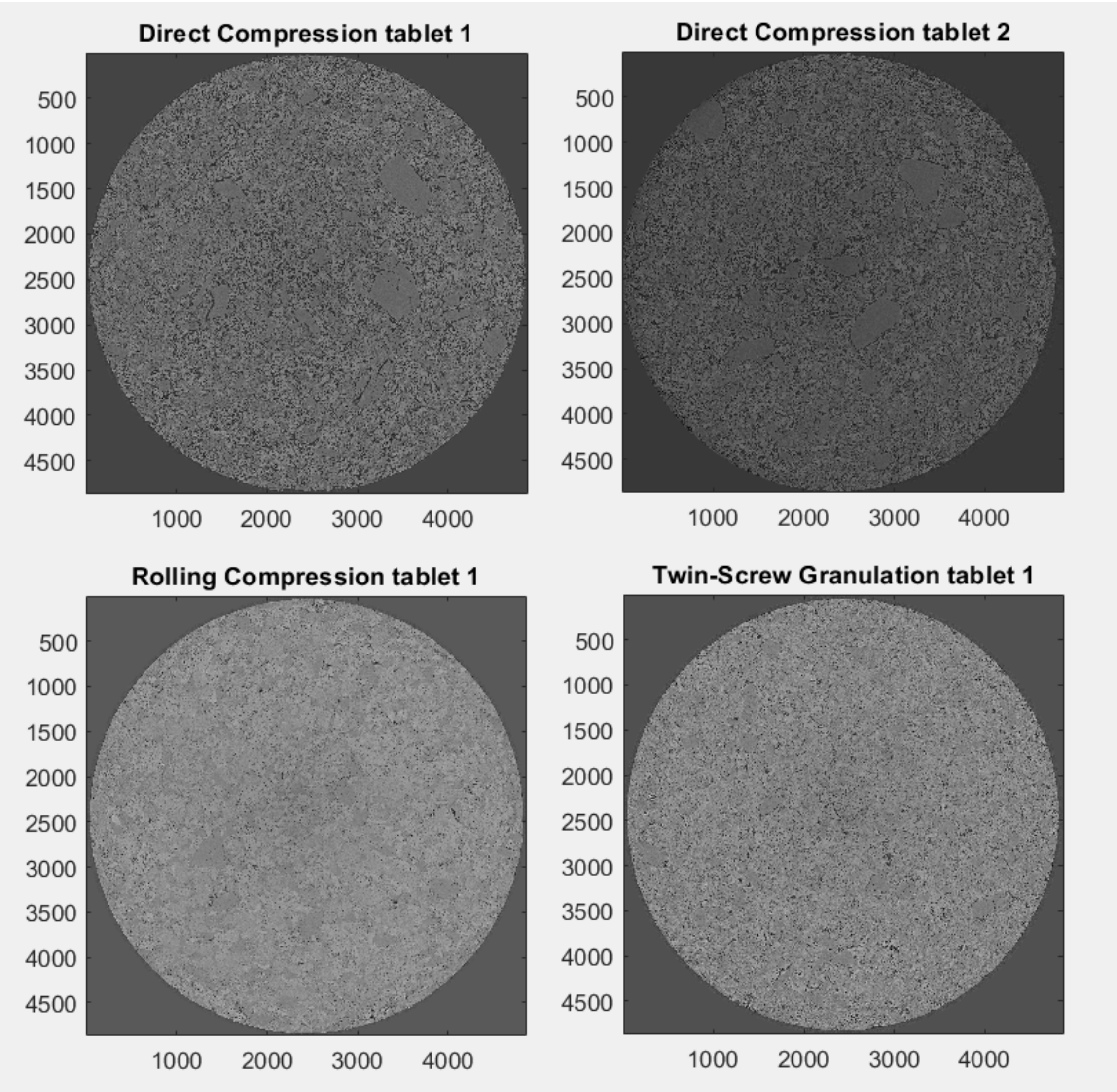

Figure 2: *The tablets studied in this report with MATLAB colormap gray.*

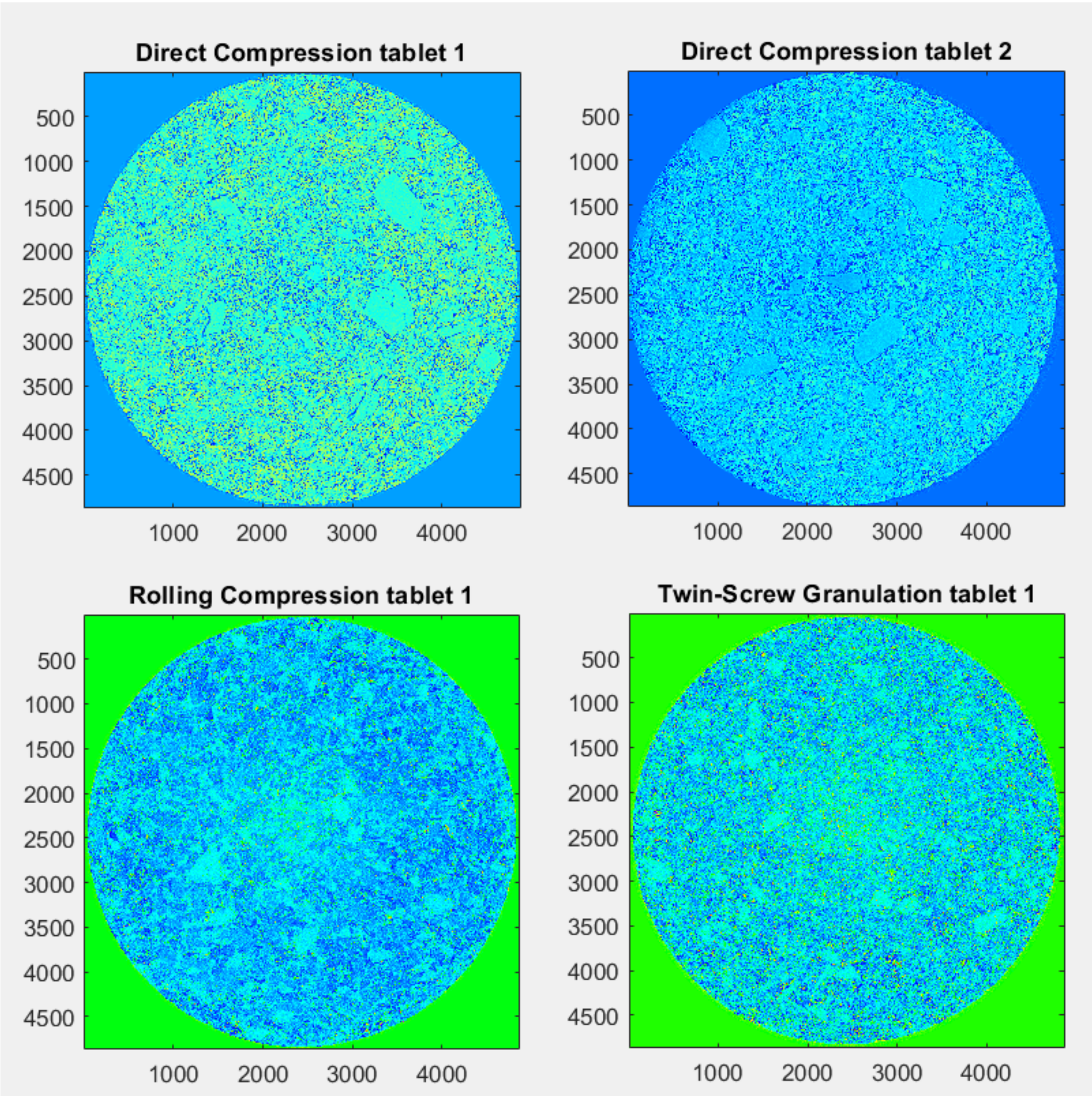

Figure 3: *The tablets studied in this report with MATLAB colormaps jet or hsv.*

3 Methods

3.1 Preprocessing of images:

The micro-CT images analysed in this report were originally in .tif-format and in grayscale. Before any tests were performed the pixel values were scaled to range between 0 and 1. The images did not appear to include any significant noise, and thus denoising was not expected to be required. Before performing segmentation and fitting a Generalised Mixture Model (GMM) to the images, the background had to be removed. Not removing the background would leave a large segment of the image with almost the same pixel values with close to zero variance, which is not supported by the algorithm used for fitting the GMM. The background was removed with the use of edge detection. Edges were first detected with two different thresholds which would later yield two different filters. A higher threshold would result in a filter that kept less of the background (Low background filter) and a lower threshold would result in a filter that kept more of the background (High background filter). The edges that were detected were then dilated in order to get a continuous border around the tablet, within which any holes were finally filled. The steps for creating the two filters in the case with the test image is visualised by the binary images in figure 4.

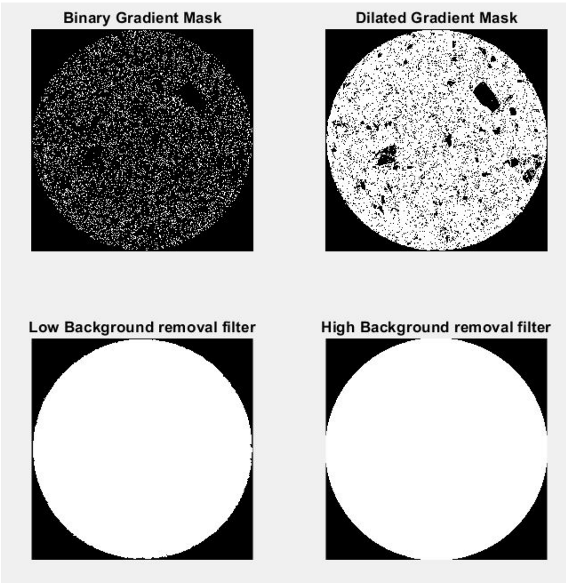

Figure 4: *Filters created to remove background from test image.*

Applying the filters (one at a time) to the test image created to new edited images where all pixels excluded by the filter were set to zero. In figure 5 the difference between the two filters can be seen when comparing the border of the tablet. Only pixel values within the blue area were kept.

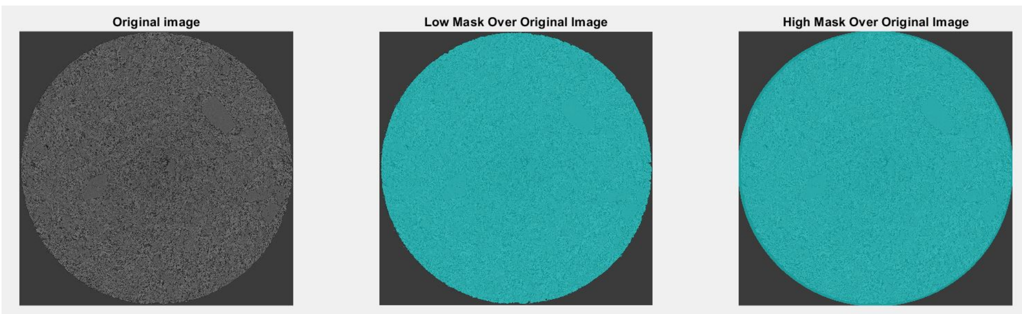

Figure 5: *Filters over original test image*

Histograms of the original test image and the two edited (filtered) images are seen in figure 6. As seen in the plots, there is a clear peak in the original image which corresponds to the dark background. In the filtered images, this peak is removed, although remains can still be detected in the histogram of the image where more background has been kept (high filter).

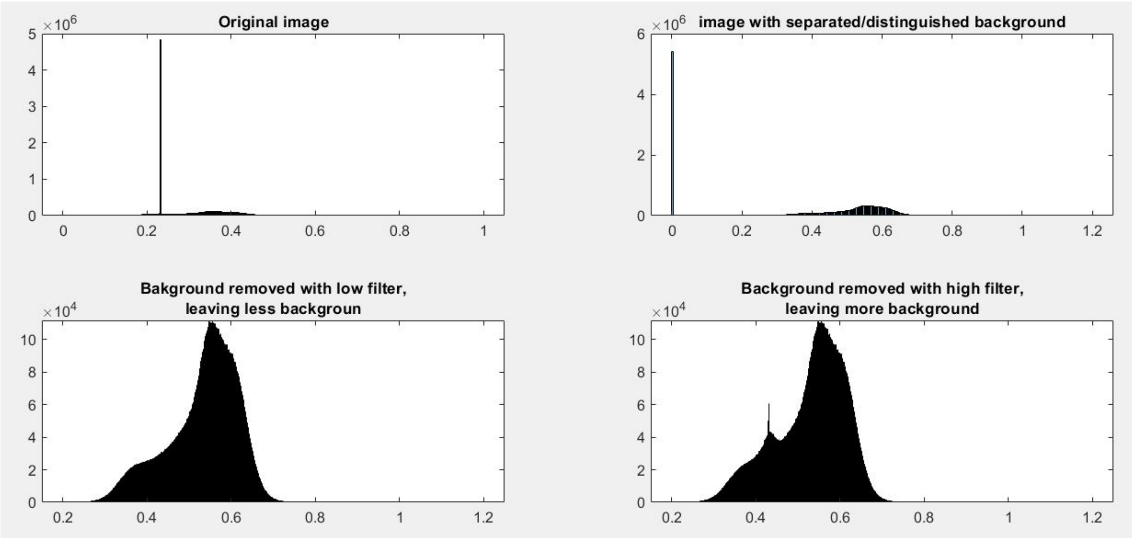

Figure 6: *Histogram of original image and edited versions*

Based on the histograms it also appears as if at least 3 gaussian models would be needed to fit the data. For the filtered image where more background had been kept, 4 models appeared to be necessary.

3.2 GMM-segmentation

Before a GMM could be fitted, the 2-dimensional image (n by m by d=1) was reshaped and stacked as a (n*m) by d=1 variable. MATLAB functions provided during the course (*normmix_sgd* and *normmix_classify*) where used with the number of iterations set to 20 and stepsize set to 1. The number of iterations were altered in some cases if the model didn't converge within the set number. In figure 7, the iteration process of the case where the test image filtered with a high filter and number of segments, K, set to 3 can be seen to converge as expected within 20 iterations.

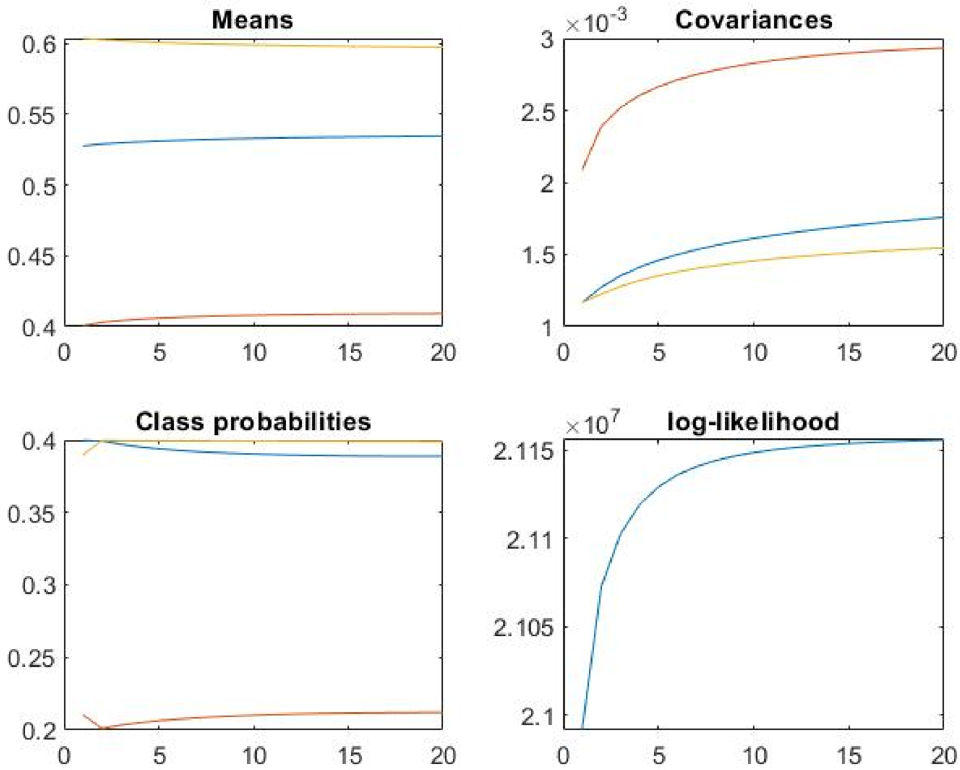

Figure 7: *Iteration of GMM model with K=3*

The results from segmenting the test image at different Ks and with different filters indicated that the best option was to use either the high filter with K=4 or the low filter with K=3 (Results for test image is attached in appendix I). One image from each of the studied tablets were segmented with other settings as well, in order to see that they yielded similar results, then the four remaining images for each tablet were segmented only for the two best settings (High filter K=4 and Low filter K=3). The shares for each segment in the analysed images were computed in each case.

4 Results

Here we will present some of the segmented images (at least one for each tablet) as well as the computed shares of each component in the different tablets. All segmented images will not be presented since there will be too many, but some will be attached in an appendix. A mean of the five segmented images for each tablet (with K=3) will be used to estimate the shares for the three largest components, which should be close to the theoretical values. For one of the images from one tablet, the results from segmenting with different values of K will be presented, as well as the results from segmenting with different filters applied to the original image.

5 Discussion

Here we will compare the computed shares for each tablet to the theoretical values, as well as commenting and comparing the spread of the components in the different tablets. Only looking at figure 3, it appears as if the tablets made with DC have one component which tends to cluster and form large pieces while the other two components are more homogeneously spread out in smaller pieces in the tablet. Compared to the DC-tablets, the image of the RC tablet appear to have more homogeneously spread out components. The component that formed large pieces in the DC-tablets appear to have been split into smaller pieces, but the other components appear to have formed larger pieces. It almost appears as if there are only two components in the RC-tablet, as if two components have "melted together". The TSG-tablet looks like a cross of the DC- and RC-tablets. We will also discuss the limits with this study, for example how we might have improved the segmentation if we had more time/knowledge (Analysing more tablets with more images of each, applying other filters or using other segmentation-algorithms s.a mean-shift). We will comment on the problems we encountered and how they were solved (such as not being able to detect the components with very low shares and thus setting K=3 in order to only find the three largest components).

6 Conclusions

Here our final conclusions from the discussion will be clearly stated, such as what method we found to be the best for segmenting the images and what steps we found necessary to yield good results. The results will be briefly stated, presenting what estimated shares we found in the 3 different types of tablets and what differences/conclusions we could draw regarding the three production methods based on the segmented images.

References

[1] Siddharth Singh Chouhan, Ajay Kaul, and Uday Pratap Singh. Image segmentation using computational intelligence techniques: Review. *Archives of Computational Methods in Engineering*, 26(3):533–596, Jul 2019. ISSN 1886-1784. doi: 10.1007/s11831-018-9257-4. URL `https://doi.org/10.1007/s11831-018-9257-4`.

7 Appendix I

7.1 Appendix I - Results from segmenting test image

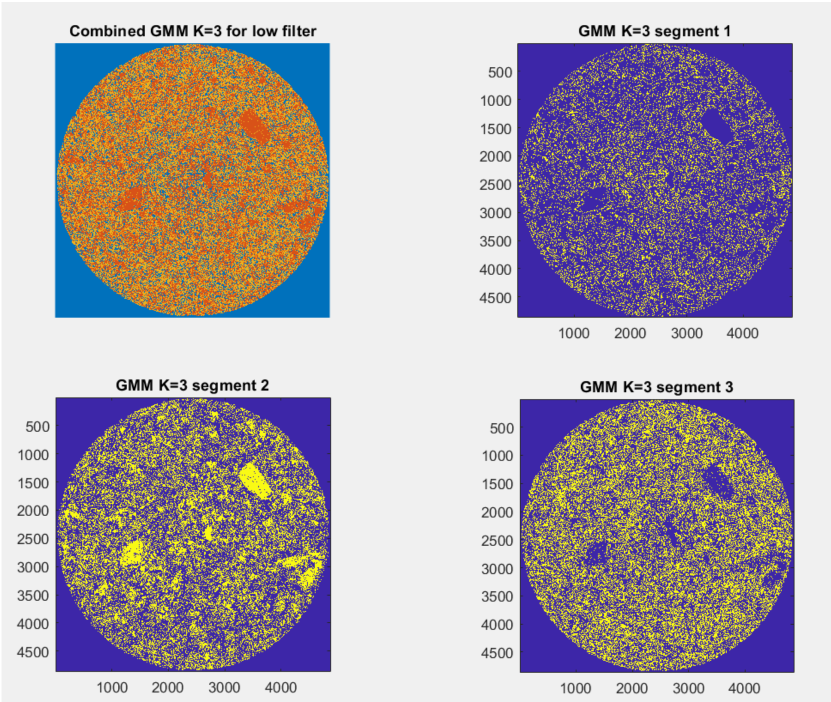

Figure 8: *GMM-segmentation of test image with K=3 and low filter*

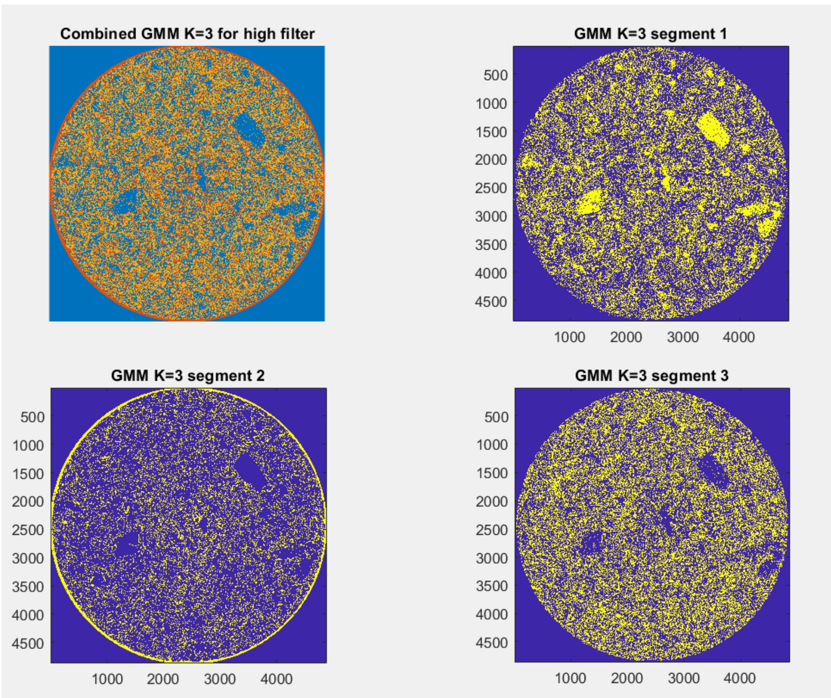

Figure 9: *GMM-segmentation of test image with K=3 and high filter*

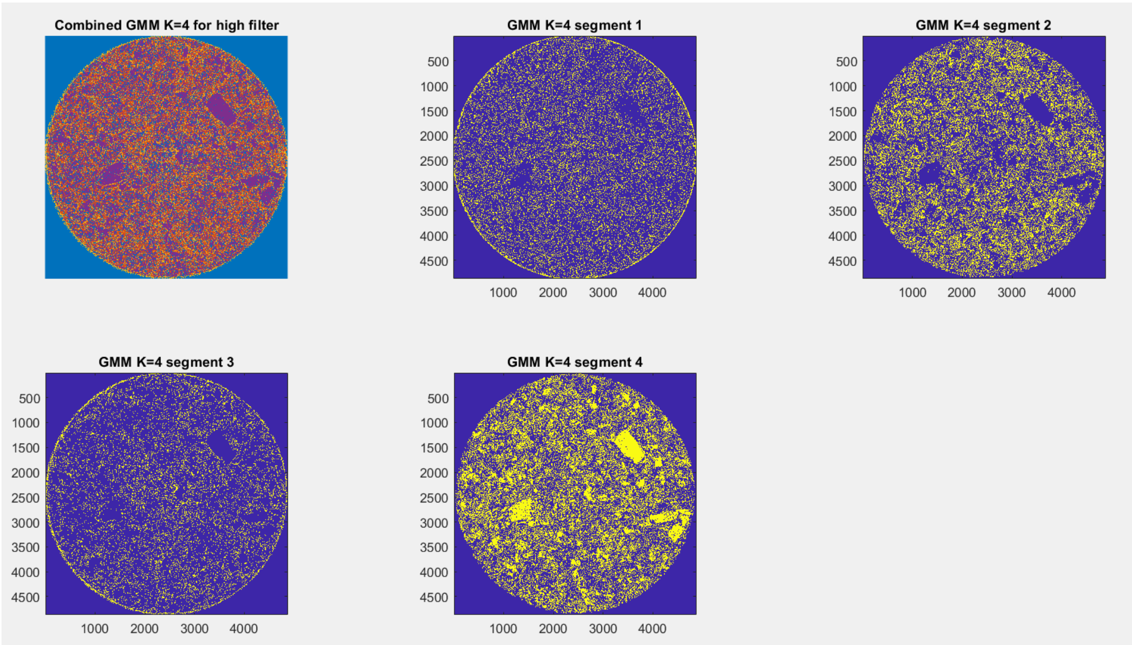

Figure 10: *GMM-segmentation of test image with K=4 and high filter*

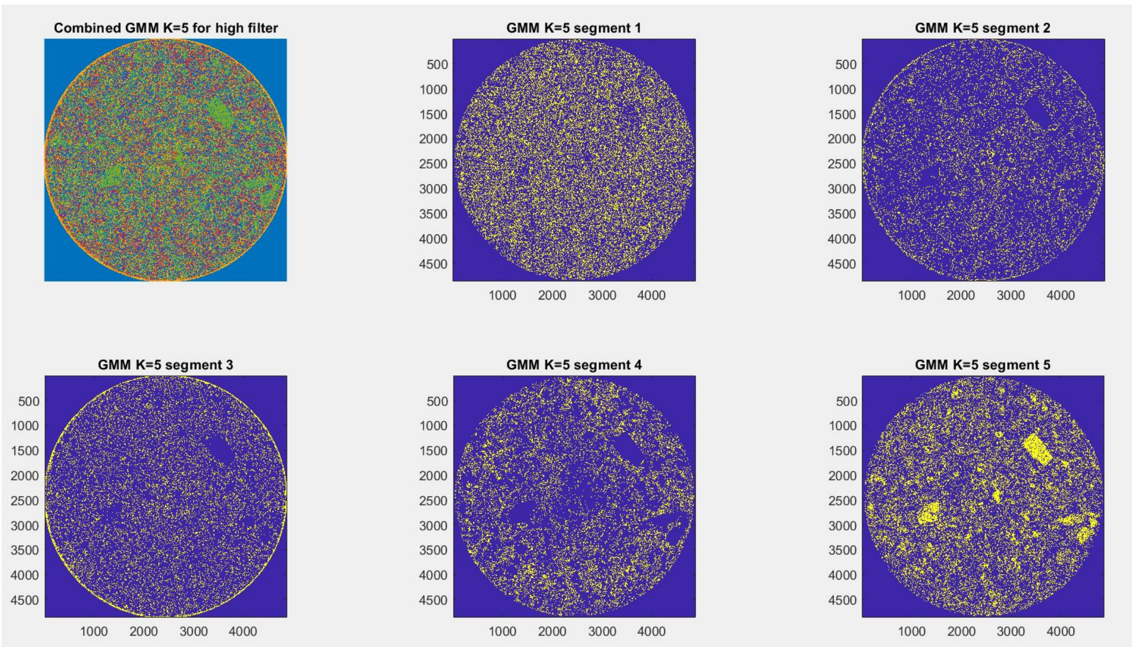

Figure 11: *GMM-segmentation of test image with K=5 and high filter*