Comparative Analysis of Lung Texture in High Resolution CT Images

Rafael Sêco
a2020131003@isec.pt

Henrique Oliveira
a2020153500@isec.pt

Aurora Guidubaldi
a2022108633@isec.pt

Department of Mathematics and Physics,
Coimbra Engineering School

Medical Image Processing,
Biomedical Engineering

Professor: Verónica Vasconcelos

Abstract

A Computed Tomography (CT) is a medical technique that uses x-rays to produce detailed images from the interior of the body.
During a CT scan, the patient must lie on a table that moves through the inside of a circular machine. A detector and a x-rays source moves around the patient to capture multiple images of the focusing area from different angles.
CT scans are non-invasive, although the patient might be exposed to a small amount of radiation.

1 Introduction

Computed Tomography is one of the most popular methods to obtain data that therefore can be transformed into images. In this project, it was used data from several high-resolution CT scans to analyse the texture between emphysema and normal patients' lungs.

1.1 Emphysema

Emphysema is a chronic, progressive, and irreversible lung condition that causes destruction of lung tissue and shortness of breath [1][2].
In CT scans, the presence of an emphysema condition can be identified as "holes", corresponding to zones of decreased pulmonary attenuation. The centrilobular emphysema is usually associated with destruction of alveolar walls in the central portion of the acinus and it is more often found on the upper long zones.

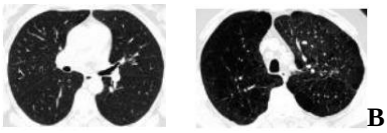

Figure 1 - Visual appearance of lung tissue on an axial CT for a normal patient (A) and an emphysema patient (B)

1.2 Objective

The goal of this project is to implement 4 methods in MATLAB to analyse the textures of ROIs (*Regions of Interest*) from the emphysema and normal patients' lungs.

As the objective of this work is not to obtain the images, but to analyse them, we were provided with a database of over than 4000 ROIs which had the labels "normal patient" and "patient with emphysema".

The first step taken was importing every ROI to MATLAB. The second step was using the four methods to obtain features with each method, and then analyse them to compare the texture from patients without emphysema and patients with emphysema.

2 Methods and features

The methods used in this project were the First Order Features, the Co-Occurrence matrix method, the Run-Length Primitives Matrix Method and the Grey Level Difference Method.

2.1 First Order Features

This method provides some quantitative features about the pixel from the selected area like the histogram, average, variance, symmetry, kurtosis, energy, and entropy.

Table 1 - First Order Features

Feature	Formula
Histogram	$h(k) = \frac{n_k}{MxN}$
Average	$\mu = \sum_{k=0}^{L-1} k h(k)$
Variance	$VAR = \sum_{k=0}^{L-1} (k-\mu)^2 h(k)$
Symmetry	$SKE = \frac{1}{\sigma^3} \sum_{i=0}^{L-1} h(i)(i-\mu)^3$
Kurtosis	$KUR = \frac{1}{\sigma^4} \sum_{i=0}^{L-1} h(i)(i-\mu)^4 - 3$
Energy	$Energy = \sum_{k=0}^{L-1} h(k)^2$
Entropy	$Entropy = -\sum_{k=0}^{L-1} h(k) \log_2(h(k))$

2.2 Co-Occurrence Matrix Method

The Co-Occurrence Matrix Method explores the distribution of the grey levels in a certain neighbourhood. The matrix is built under a certain direction and a certain angle. In the project, the Co-Occurrence matrix was built horizontally from the left to the right (***d=1*** *and* ***θ=0º***).

The descriptors obtained from this method can be calculated by the following expressions:

Table 2 - Co-Occurrence Matrix Features

Feature	Formula
Correlation	$\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{(i-m_r)(j-m_c)p_{ij}}{\sigma_r\sigma_c}$
Contrast	$\sum_{i=1}^{K}\sum_{j=1}^{K}(i-j)^2p_{ij}$
Uniformity	$\sum_{i=1}^{K}\sum_{j=1}^{K}p_{ij}^2$
Homogeneity	$\sum_{i=1}^{K}\sum_{j=1}^{K}\frac{p_{ij}}{1+\lvert i-j\rvert}$
Entropy	$-\sum_{i=1}^{K}\sum_{j=1}^{K}p_{ij}^2\,log_2\,p_{ij}$

2.3 Run-length Primitives Matrix Method

The Run-length Primitives Matrix Method [3] is based on the calculation of run-length primitives, which are a set of consecutive pixels with the same grey level, in a certain direction. In the project, an angle of ***θ=0º*** was used.

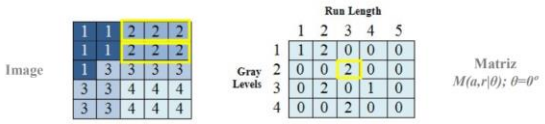

Figure 2 - Run-length Primitives Matrix

The descriptors obtained from this method can be calculated by the following expressions:

Table 3 - Co-Occurrence Matrix Method features

Feature	Formula
Short Run Emphasis	$SRE=\frac{1}{n_r}\sum_{a=1}^{L}\sum_{r=1}^{Nr}\frac{M(a,r)}{r^2}$
Long Run Emphasis	$LRE=\frac{1}{n_r}\sum_{a=1}^{L}\sum_{r=1}^{Nr}M(a,r)*r^2$
Grey Level Non-Uniformity	$GLNU=\frac{1}{n_r}\sum_{a=1}^{L}(\sum_{r=1}^{Nr}M(a,r))^2$
Run Length Non-Uniformity	$RLNU=\frac{1}{n_r}\sum_{r=1}^{Nr}(\sum_{a=1}^{L}M(a,r))^2$
Run Percentage	$RP=\frac{n_r}{n_p}$

2.4 Grey Level Difference Method

The Grey Level Difference Method is an image processing technique used to measure the texture of an image. This method was used to calculate the following descriptors:

Table 4 - Grey Level Difference Method Features

Feature	Formula
Angular Second Moment	$ASM_d=\sum_{k=0}^{L-1}H^2(k)$
Inverse Difference Moment	$IDM_d=\sum_{k=0}^{L-1}\frac{H(k)}{1+k^2}$
Correlation	$COR_d=\sum_{k=0}^{L-1}H(k)\frac{(k-\mu)}{\sigma}$
Variance	$VAR_d=\sum_{k=0}^{L-1}H(k)(k-\mu)^2$
Contrast	$CON_d=\sum_{k=0}^{L-1}H(k)k^2$

3 Results

In order to minimize the errors occurred when analysing each subROI, the group has decided to only analyse subROIs which were more than 99% inside the *Region Of Interest*. After this condition, the group was able to represent each subROI as it is shown in the following figure:

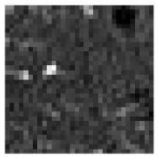

Figure 3 - Image representation of the first subROI being analysed

As this example was for a patient with an emphysema condition, the histogram obtained was the following:

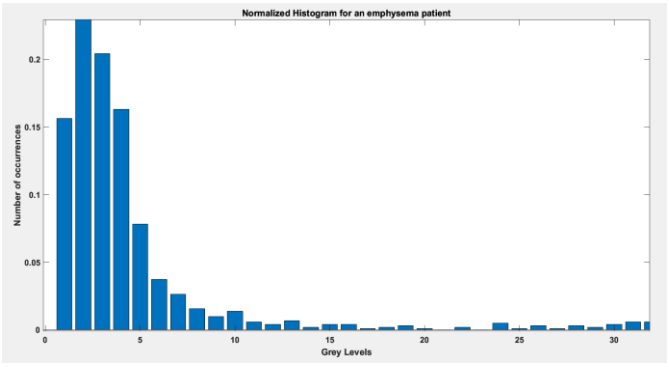

Figure 4 - Grey Level Normalized Histogram for the first patient analysed

It is possible to observe a higher grey level distribution on the left of the histogram.

The process was repeated for the rest of the subROIs available in the database, having some interesting results for the comparison of lung texture in emphysema and normal patients.

The following figure is a normalized histogram for a normal patient and it shows a higher distribution of higher grey levels among the histogram, when compared to a normalized histogram for an Emphysema patient:

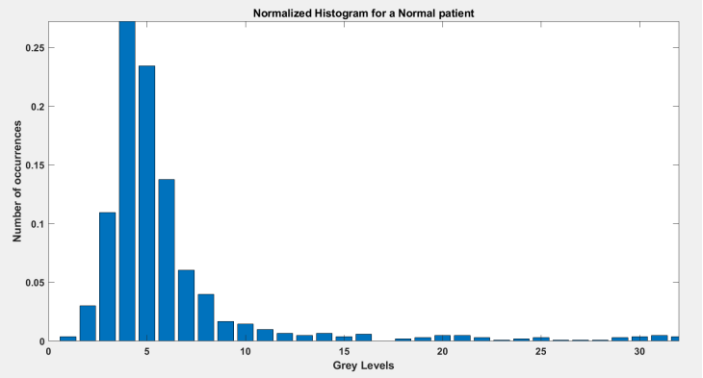

Figure 5 - Grey Level Normalized Histogram for a normal patient

3.1 First Order Features

Trough the features of this method, it is possible to verify that the *Media*, the *Variance*, and the *Entropy* are higher in patients without emphysema than they are in patients with emphysema. These values can be explained because the emphysema condition creates "holes" in the lungs of the patients, which turns the image darker. With this, the image values decrease, which implies a decrease in the mean and variance values. On the other hand, the *Entropy* is only slightly higher, which means that the distribution of grey levels is smaller in the images of patients with emphysema.

The *Symmetry* and the *Energy* are two features that measures the similarity between images and pixels. In these cases, they are slightly higher in patients with the disease, suggesting that the "holes" created by the emphysema turn the images more similar. In opposition we have the *Kurtosis* that are much bigger in the CT Scans of people with emphysema. This means that the intensity of the pixels in the images that are distributed around the mean is bigger.

The following graph bar displays the average value for each feature and its standard deviation.

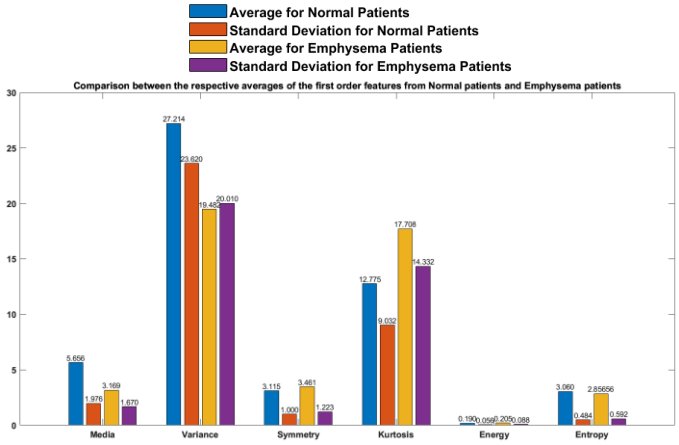

Figure 6 - Comparison between the respective averages of the first order features from Normal and Emphysema patients

The results can be also seen in a table form:

Table 5 - Comparison between the respective averages of the first order features from Normal and Emphysema patients

	Normal Patients		Emphysema Patients	
	Average	*Standard Deviation*	**Average**	*Standard Deviation*
Media	5.6559	1.9763	3.1688	1.6698
Variance	27.2143	23.6202	19.4815	20.0091
Symmetry	3.1146	1.0004	3.4615	1.2234
Kurtosis	12.7748	9.0321	17.7081	14.3321
Energy	0.1904	0.0565	0.2053	0.0880
Entropy	3.0595	0.4846	2.8566	0.5919

3.2 Co-Occurrence Matrix Method

This method was used to obtain a Co-Occurrence Matrix for each subROI and then extract its features. The features were stored in two different tables: one for emphysema and another for normal subROIs. The referred tables were also exported to Excel files.

The average and the standard deviation of each feature was calculated and displayed in the following graph:

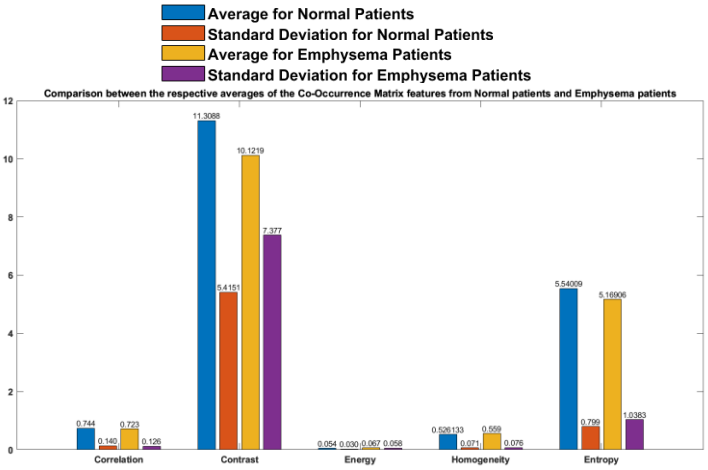

Figure 7 - Comparison between the respective averages of the Co-Occurrence Matrix features from Normal and Emphysema patients

Comparing the results obtained, it is possible to conclude that the *Correlation* is higher for normal patients and lower than 1, as it should. The *Contrast* is higher in normal cases of study, as expected, because these subROIs have a higher range of grey levels when compared to subROIs from patients with emphysema. Thus, the *Entropy* is also higher for normal patients, once there is a higher distribution of grey levels. The *Energy* (*Uniformity*) and the *Homogeneity* were expected to be higher for patients with emphysema because smaller grey levels are more consistent throughout all these subROIs and that is in conformity with the information represented both in the graph and in the table:

Table 6 - Comparison between the respective averages of the Co-Occurrence Matrix features from Normal and Emphysema patients

	Normal Patients		Emphysema Patients	
	Average	*Standard Deviation*	**Average**	*Standard Deviation*
Correlation	0.7439	0.1395	0.7230	0.1257
Contrast	11.3088	5.4151	10.1219	7.3770
Energy	0.0543	0.0302	0.0673	0.0582
Homogeneity	0.5261	0.0708	0.5593	0.0757
Entropy	5.5401	0.7990	5.1691	1.0383

3.3 Grey Level Run-Length Primitives Method

Using this method, it was possible to build a run-length matrix with 32 grey levels for each subROI analysed. The purpose was to extract the 5 principal features. Like the previous methods, the results were stored in two different tables and then exported to an Excel File. The next bar graph contains the average and the standard deviation for each feature:

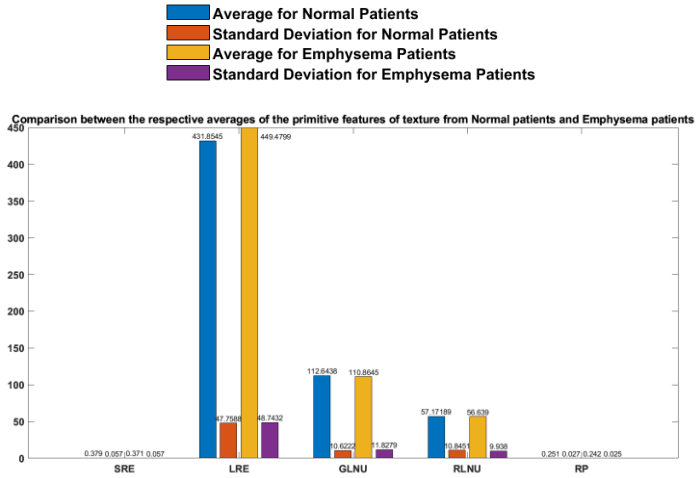

Figure 8- Comparison between the respective averages of the primitive features of texture from Normal and Emphysema patients

The results from the graph are also displayed in the next table:

Table 7 - Comparison between the respective averages of the primitive features of texture from Normal and Emphysema patients

	Normal Patients		Emphysema Patients	
	Average	*Standard Deviation*	**Average**	*Standard Deviation*
SRE	0.3793	0.0571	0.3707	0.0572
LRE	431.8545	47.7588	449.4799	48.7432
GLNU	112.6438	10.6222	110.8645	11.8279
RLNU	57.1719	10.8451	56.6390	9.9375
RP	0.2512	0.0265	0.2424	0.0251

By analysing both the graph and the table, it is possible to obtain very interesting results with this method.
The *Short Run Emphasis* refers as being a relative importance given to short-range spatial patterns in an image. In this case, it is higher for normal patients, but it is still a very small number. The *Long Run Emphasis* increases when long runs dominate, for example in homogenous areas. It is higher for emphysema patients.
The *Grey-Level Non-Uniformity* is a metric that increases when grey-level outliers dominate the histogram. So, it was expected to be lower for emphysema patients because smaller grey levels are more consistent.
The *Run-Length Non-Uniformity* metric is the opposite, but the average for this metric is higher for normal patients. Nevertheless, the standard deviation of this feature is also higher for normal patients, which means that is less accurate.
The *Run Percentage* metric increases if there are higher amounts of pixels with the same value followed by each other. If there are, more primitives in the matrix there will be related to the subROI being studied. This metric should be higher for emphysema patients for the exact same reason of the *Grey-Level Non-Uniformity* descriptor: smaller grey levels are more consistent. But in the case analysed, there is also a higher *Run Percentage* for normal patients with less accuracy (higher standard deviation).

3.4 Grey Level Difference Method

The Grey Level Difference Method [4] was used to obtain a grey level difference method matrix with an angle $\theta=0^{o}$ and a distance $d=1$ (from the left to the right), as the example of the next figure:

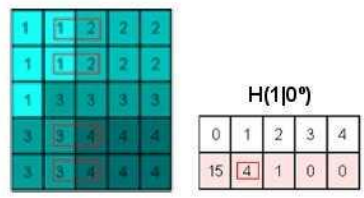

Figure 9 - Example of the Grey Level Difference Method

This method allows to describe the differences existing between each pixel of the image and the immediately adjacent ones. Similarly with the previous methods, the results were stored in two different tables and exported to an Excel file. The average and the standard deviation of each feature were calculated and then reported in the following bar graph:

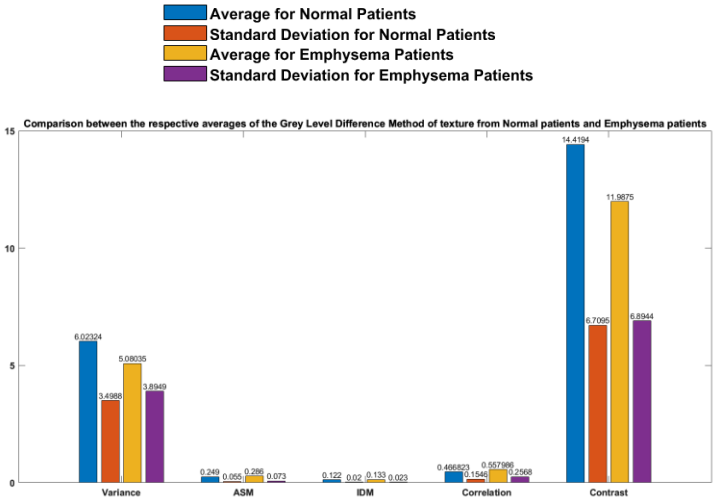

Figure 10 - Comparison between the respective averages of the Grey Level Difference Method of texture from Normal and Emphysema patients

The following table displays the results obtained:

Table 8 - Comparison between the respective averages of the Grey Level Difference Method of texture from Normal and Emphysema patients

	Normal Patients		Emphysema Patients	
	Average	*Standard Deviation*	**Average**	*Standard Deviation*
Variance	6.0232	3.4988	5.0803	3.8949
ASM	0.2488	0.0545	0.2856	0.0732
IDM	0.1217	0.0200	0.1329	0.0229
Correlation	0.4668	0.1546	0.5580	0.2567
Contrast	14.4194	6.7095	11.9875	6.8944

The *Variance* descriptor was expected to be higher for the normal patients because there is more discrepancy of grey levels.
The *Angular Second Moment* is a measure of the image's uniformity and images with higher ASM, as the emphysema patients, tends to be more uniform and less complex.
The *Inverse Difference Moment* is another metric to measure an image's uniformity and is compliant with the *Angular Second Moment*: its value is higher for Emphysema patients. Similarly to the Co-Occurrence Matrix Method, the *Contrast* descriptor is higher in normal cases of study, as expected, because these subROIs have a higher range of grey levels when compared to subROIs from patients with emphysema.

4 Conclusion

It can be concluded that is possible to use image texture methods to distinguish patients with emphysema from healthy patients based on the descriptors' values obtained using these methods.

Emphysema looks like a "hole" in a CT scan image, for this reason, scans with emphysema are characterized by less grey level variation than CT scans from patients without the disease. In fact, based on the values obtained with the Co-Occurrence Matrix Method and the Grey Level Difference Method, the images relating to patients with emphysema are characterized by greater homogeneity and lower contrast than those of normal patients. Furthermore, both the Co-Occurrence Matrix Method and the First Order Features method demonstrate that patients with emphysema are characterized by higher energy and lower entropy.

Using a general overview and based on the results obtained, it is possible to conclude that the Gray Level Difference Method turned out to be a good method as the results obtained from it are consistent with those obtained with the first two methods, but in any case, it is not the best method to apply to do texture analysis.

The Gray Level Run-Length Primitives Method has led to conflicting conclusions compared to those obtained by the other methods (for example, the result from the *RLNU* feature suggested that the images extracted from patients with emphysema are less uniform than those of healthy patients). So, between all of them, this turned out to be the worst method to use, it is therefore necessary to do a more in-depth analysis by varying the angle $\boldsymbol{\theta}$ to understand if, by varying the angle $\boldsymbol{\theta}$, the results improve.

The first two methods, on the other hand, are the best to use to distinguish patients with the disease from healthy ones and, among all, the features with the greatest significance are *Contrast* and *Kurtosis*, since, from them, can also be drawn conclusions relating to the others features, like the *Energy*, *Variance* and *Entropy*.

References

[1] Emphysema, Maio Clinic, https://www.mayoclinic.org/diseases-conditions/emphysema/symptoms-causes/syc-20355555

[2] Vasconcelos, Verónica, José Silvestre Silva, Luís Marques and João Barroso. "Statistical textural features for classification of lung emphysema in CT images: A comparative study." *5th Iberian Conference on Information Systems and Technologies* (2010): 1-5.

[3] Run Lengths (Biomedical Image Analysis), http://what-when-how.com/biomedical-image-analysis/run-lengths-biomedical-image-analysis/

[4] A. Rizal, R. Hidayat and H. A. Nugroho, "Modification of Grey Level Difference Matrix (GLDM) for Lung Sound Classification," *2018 4th International Conference on Science and Technology (ICST)*, 2018, pp. 1-5, doi: 10.1109/ICSTC.2018.8528650.