

ALMA MATER STUDIORUM · UNIVERSITY OF BOLOGNA

School of Science
Department of Physics and Astronomy
Master Degree in Physics

Automatic Pipeline for the Identification of Lung Lesions on CT Images of Patients Affected by COVID-19

Supervisor:
Prof. Gastone Castellani

Co-supervisor:
Dr.Nico Curti

Submitted by:
Riccardo Biondi

Academic Year 2019 /2020

Abstract

SARS-CoV-2 virus has widely spread all over the world since the beginning of 2020, and it is responsible of the COrina VIrus Disease (COVID). This disease affect lung areas and cause respiratory illness.

Patient affected by this pathology have shown a particular pattern of ground glass opacities(GGO) and consolidation(CS) in chest CT scans and that are made in relation both with the stage of the disease and its severity; moreover the study of these pattern on healed patient was used to check the actual recovery.

Considering this scenario, a segmentation of these areas is desirable. Up to now the gold standard for the identification and quantification of these areas is made by manual or semiautomatic segmentation; which are subjective and time consuming. So an automatic way for the identification is desired.

In this work I will present an automatic pipeline for the segmentation of GGO and CS. The pipeline achieve the segmentation by applying the color quantization as medical image segmentation technique, achieve a classification based on voxel intensity. Exploiting the properties of digital multichannel images, we was able to incorporates also other image properties beside the single voxel intensity.

The proposed approach, tested on three datasets, has shown to achieve good segmentation on chest CT with typical lesions.

Contents

Introduction

Since the end of 2019, COVID-19 has widely spread all over the world. Up to now the gold standard for the diagnosis of this disease are the reverse transcription-polymerase chain reaction (RT-PCR) and the gene sequencing of sputum, throat swab and lower respiratory tract secretion [1].

An initial prospective made by Huang and al. on chest CT scans [2] of patients affected by COVID-19, has shown the 98% of examined patients have bilateral patchy shadows or ground glass opacity (GGO) and consolidation(CS); moreover the severity, shape and involved percentage of lung were in relation with the stage of the disease [3]. In the end, other study have monitored the change on volume and shape of these features on healed patients [4] in order to monitoring their actual recovery. In Figure ?? are compared slices of an healthy control and a patient affected by COVID-19. We can clearly see the GGO and CS regions in the lung of the second lung image from left.



Figure 1: Ground Glass Opacity and Consolidation on chest CT scans of COVID-19 affected patients with different severity of the disease. From left to right we can observe an increasing of the involvement of the involved lung volume.

GGO and CS are not exclusive of COVID-19, but may be also caused by pulmonary edema, bacterial infection, other viral infection or alveolar haemorrhage [5]. The combination between CT scan information and other diagnostic techniques like the RT-PCR mentioned above, may help the diagnosis, the monitoring of the course of the disease and the checking of the recovery in healed patients; moreover the study of these patterns may help to understand the infection pathogenesis, which is not well known since COVID-19 is a new disease.

Identification and quantification of these lesions in chest CT scans is a fundamental task. Up to now the segmentation is made in a manual or semiautomatic way, which are time consuming and subjective, since involves the interaction with trained personnel; so an automatic and fast way for the identification of this features

is desired.

This thesis work, made in collaboration with the Department of Diagnostic and Preventive Medicine of the Poloclinico Sant'Orsola - Malpighi, aims the developing of an automatic pipeline for the identification of GGO and CS in COVID-19 affected patients. The developing was based and tested on chest CT scans provided Sant'Orsola, but also public repositories [6] [7] where used as benchmark.

We start the discussion by understanding what is CT image, its physical meaning and digital representation; so a brief review on the main image segmentation techniques will be presented, describing the main features of them. More details will be given on the techniques used for the actual implementation.

The discussion will continue by describing the main pipeline characteristics and the main pipeline structure. We will see how color quantization was used to achieve the segmentation and how the digital image properties were used in order to takes into account different image features. We also discuss how a preliminary lung segmentation will help the performances of the segmentation. After that we will continue the discussion by describe in details the actual pipeline implementation, going in deep on how each step is actually achieved. Also a brief description about the used frameworks will be given.

In the end we will discuss the segmentation results. The pipeline performances were checked trough different method, like visual comparison with other segmentation techniques, quantitative comparison with a ground truth segmentation and blind evaluation by experts. Also the segmentation achieved on healthy control was considered as benchmark.

Chapter 1

Image Segmentation techniques

Image segmentation consist in the partitioning of an image into non overlapping, inconsistent regions that are homogeneous respect to some characteristics such as intensity or texture [8]. Nowadays several non-invasive medical imaging techniques are available, such as Computed Tomography(CT), Magnetic Resonance Imaging (MRI) or X-Ray imaging, that provides a map of the subject anatomy. Image segmentation play a crucial role in many medical-imaging applications by automating or facilitate the delineation of anatomical structures and other regions of interest [8]. Manual segmentation is possible, but is time consuming and subject to operator variability; making the results difficult to reproduce [9], so automatic or semi-automatic methods are preferable.

A major difficulty of medical image segmentation is the high variability in medical images. First and foremost, the human anatomy itself shows major modes of variation. Furthermore many different modalities (X-ray, CT, MRI, etc.) are used to create medical images [10].

The results of segmentation can be used to perform feature extraction, that provides fundamental information about organs or lesion volumes, cell counting, etc. If the patient perform several analysis during time, image segmentation is a useful tool to monitor the evolution of particular lesions or tumors during a therapy.

In this chapter I will provide a brief introduction about medical images, focusing mainly on on Computed Tomography, which is the technique used tp acquire the images segmented in this work.

This part is followed by a discussion on the main image segmentation techniques. As before the focus will be mainly in these techniques that we have used in this work.

1.1 Medical Images

Digital images are represented by 3 dimensional tensor, where the three dimensions corresponds respectively to height, width and number of channel, where channel refers to the number of image components. For instance a Gray level image is represented with a $h \times w$ matrix, on the other hand, an RGB image is composed by three $h \times w$ matrices, each of them represent a different primary.

Each value of the tensor is in a range that change according to the image format. The most common are [0, 255] for 8-bit integer image, of [0, 1] for float. Also other formats are available, like 16-bit integers, widely used to represent medical images.

Medical images provides a map of the subject anatomy and so are computed

on a uniformly x-y-z spatial space, allowing a non invasive visualization of internal organs and tissues. At each point the data is represented by a 16-bit integer. The meaning of the data change according to the image acquisition modality(CT, MRI, PET, etc.).

Computed Tomography(CT) is a medical imaging techniques which aims to reproduce cross-section images and the 3D anatomy of the examined subject, each data represent the capability of the corresponding volume to attenuate an x-ray beam. In order to be able to match results from different scans the beam attenuation is measured in Hounsfield Units(HU) :

$$CT - number = k \times \frac{\mu - \mu_{H_2O}}{\mu_{H_2O}} \quad (1.1)$$

Where μ is the linear attenuation coefficient of the tissue, μ_{H_2O} is the linear attenuation coefficient of the water, took as a reference. , and k is a constant which can be 1000 or 1024 according to manufacturer scan. The linear attenuation coefficient of the air is considered as 0, so the corresponding CT number is -1000 ; for the bones, that have a density double than water, the CT number is 1000 .

The resulting image tensor have a 16-bit depth and each voxel value is in relation with the capability to attenuate an x-ray beam of the tissue in object, which is different between tissues.

1.2 Review on Image Segmentation Methods

During the years, several segmentation methods have been developed based on a lot of different approaches. These methods can be categorized in several way, for example we can divide them into *supervised* or *unsupervised* if they requires or not a set of training data, or can be classified according to the used information type, like *Pixel classification methods*, which use only information about pixel intensity, or *Boundary following* methods, which use edge information, etc. In this section I will provide a brief review on the main segmentation methods, organized in the same way as in [8] that divides the methods in 8 categories:

1. Thresholding,
2. Region growing,
3. Classifiers,
4. Clustering,
5. Markov Random Fields models,
6. Artificial Neural Networks,
7. Deformable Models,
8. Atlas guided approaches.

1.2.1 Thresholding

Thresholding approach is very simple and basically segments a scalar image by creating a binary partitioning of image intensities [8]. It can be applied on an image to distinguish regions with contrasting intensities and thus differentiate between tissue regions represented within the image [9]. Figure ?? show an histogram of a scalar image with two classes, threshold based approach attempts to determine an intensity value, called *threshold* which separate the desired classes [8]. To achieve the segmentation we can group all the pixels with intensity higher than the threshold in one class and all the remaining in the other class.

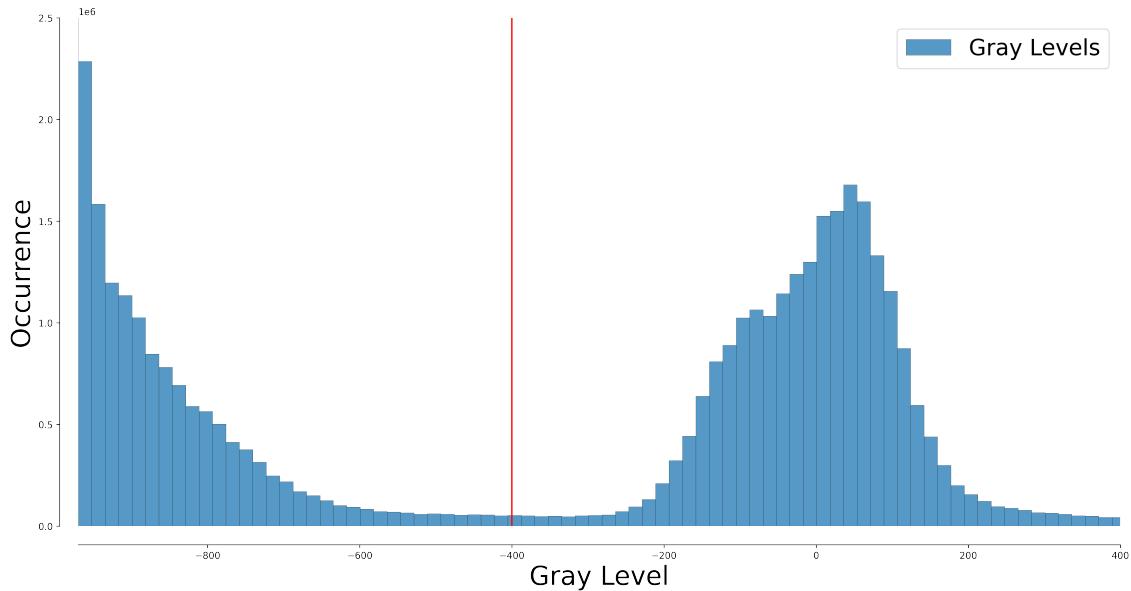


Figure 1.1: Histogram of a GL image with two well delineated regions. The threshold value(red line) was set visually at -400 HU

The threshold value is usually setting by visual assessment, but can also be automatized by algorithm like otsu one.

Sometimes may happen that more than two classes are present in the image, so we can set more than one threshold values in order to achieve this multi-class segmentation, also in this case there are algorithms to automatized this process, like an extension of the previous one called *multi otsu threshold*.

This is a simple but very effective approach to segment images when different structures have an high contrast in intensities. Threshold doesn't takes into account the spatial characteristic if the image, so it is sensitive to noise and intensity inhomogeneity, that corrupt the image histogram of the image and making difficult the separation [8]. To overcome these difficulties several variation of threshold have been proposed based on local intensities and connectivity.

Threshold is usually used as initial step in sequence of image processing operations, followed by other segmentation technique that improve the segmentation quality. Since threshold use only intensity information, can be considered a pixel classification technique.

1.2.2 Region Growing Approach

Region growing approach allows to extract connected regions from an image. This algorithm start at seed location in the image(usually manually selected) and check the adjacent pixels against a predefined homogeneity criterion [9], based on intensity, and/or edges. If the pixels met the criterion, they are added to the region. A continuos application of the rule allow the region to grow.

Like thresholding, region growing is used in combination with other image segmentation operations, and usually allows the delineation of small and simple structures such as tumor and lesions [8].

Regions growing can also be sensitive to noise so extracted regions may have holes or even become disconnected. May also happen that separate rion becomes connected due to partial volume effect.

When we use this approach we have to consider that for each region we want to segment a seed must be planted. There are some algorithm, related to region growing, that does not require a seed point, like split and merge one. Split and merge operates in a recursive fashion. The first step is to check the pixel intensity homogeneity, if they are not homogeneous, the region is splitted into two equal sized sub-regions. This step leads to an oversegmentation, so a merging step is performed, which merge together adjacent regions with similar intensities [9].

1.2.3 Deformable Model

Deformable Model use an artificial, closed, contour/surface able to expand or contract over time and conforme to a specific image feature [9]. This approach is physically motivated model-based thechnique for the detection of region boundaries [8]. The curve/surface is placed near the desidered boundary and it is deformed by the action of internal and external forces that act iteratively. The external forces are usually derived from the image.

This approach has the capability to directly generate closed parametric curves or surfaces from images and an also incorporate smoothness constraint that providesrobustness to noise and spurioous edges.

However this approach requires a manual interaction to place the appropriate set of parameters.

1.2.4 Markov Random Field

Markov Random Field(MRF) is not a proper segmentation method but its a statistical model that's used within segmentation methods that model the spatial interaction between neighbouring pixels. It's often incorporated in clustering algorithms such as K-means with a Bayesian prior probability.

This model is used because most pixels belong to the same class as their neighbouring pixels, this means that any anatomical structure that consist of only one pixel has a very low probability of occurring [8].

A difficulty of this model is that it is very sensitive to the parameters that controls the strenght of the spatial interactions. An other MRF disavantage is that requires computationally intensive algorithms. However, despite these disavantages, MRF are widely used to model segmentation classes and intensity inhomogeneities [8].

1.2.5 Classifiers Approach

Classifiers approaches use statistical pattern recognition techniques to segment images by using a mixture model that assume each pixels belonging to one of a known set of classes [9]. To assign each pixel to the corresponding classes, use the so called *feature space*, which is the space of any function of the image like intensity. An example of 1D feature space is image histogram.

The feature of each pixel form a pattern that is classified by assign a probability measure for the inclusion of each pixel in each class [9].

This approach assume a prior knowledge about the total number in the image and the probability of occurrence of each class. Generally this quantity aren't known, so we need a set of training data to use as reference.

There are different techniques which use this approach:

- **k-Nearest Neighborhood** : each pixel is classified in the same class as the training data with the closest intensity;
- **Maximum likelihood or Bayesian** : Assume that pixel intensities are independent samples from a mixture of probability distributions and the classification is obtained by assign each pixel to the class with the highest posterior probability.

This approach requires a structure to segment with distinct and quantifiable features. It is computational efficient and can be applied to multichannel images. This approach doesn't consider a spatial modelling and need a manual interaction to obtain the training data that must be several since the use of the same training set for a large number of scans can lead to biased results.

1.2.6 Clustering

Clustering approach is similar to classifiers one but in an unsupervised fashion, so doesn't require a training dataset. Clustering iteratively alternate between segmenting the image and characterizing the properties of each class. In this way we can say that clustering approach train itself by using the data available information. We can identify 3 main clustering algorithms:

- **k-means clustering:** that iteratively compute a mean intensity for each class and segments the image by classifying each pixel in the class with the closest mean;
- **Fuzzy C-means:** this algorithm generalize the K-means clustering in order to achieve soft- segmentation;
- **Expectation Maximization:** use the same clustering principle as k-means by assuming that the pixel follows a Gaussian mixture model. It iterates between posterior probability and compute the the Maximum Likelihood estimates for the means, covariances and mixing coefficients of the mixture model.

This approach doesn't require training data, but suffer to an high sensitivity to the initial parameters and do not incorporates spatial model, so it is a pixel classification technique [8].

The most used algorithm for clustering is the k-means clustering, which seek to assign each point to a particular cluster in a way that minimize the average square distance between points in the same cluster [11]. A vector representing the mean is used to describe each cluster, so this technique is described as a centroid model [12]. The labeling is performed by assign to each point to the cluster with the nearest centroid. Each point is assigned to the cluster with the nearest mean.

Given an integer k and a set of n data points from \mathbb{R}^d , the k-means clustering seek to find k centers that minimize a potential function given by the sum of squares:

$$\Phi = \sum_{x \in S} \min \|x - c\|^2 \quad (1.2)$$

Where $S \subset \mathbb{R}^d$ is a set of points. In this work \mathbb{R}^d will be the colors space and S is the space of color of each voxel.

The steps of the algorithm are the following:

1. Select the value of k as initial centroids
2. Form k cluster by allocating every point to its most nearest centroid
3. Recalculate the centroid for each cluster until the centroid does not change.

Arthur and Vassilvitskii [11] have pointed that this algorithm is not accurate and can produce arbitrarily bad clusters. So they have developed a popular algorithm, the "k-means++" which improves the clustering accuracy by made an accurate choice of the initial cluster centers.

They pointed out that the bad clustering is caused to the fact that $\frac{\Phi}{\Phi_{opt}}$ is unbounded even if the number of clusters and points are fixed, where Φ_{opt} is the potential function in the optimal centroids case. They have proposed a variant for the choosing of the centroids, instead of chose the centroids randomly, the weight the initial points according to the distance square ($D(x)^2$) from the closest center already chosen. So the final algorithm is equal to the k-means except for the initial centroids selection that is made as follows:

1. Take one center c_1 , chosen uniformly at random from S .
2. Take a new center c_i , choosing $x \in S$ with probability $\frac{D(x)^2}{\sum_{x \in S} D(x)^2}$
3. Repeat the step 2 till k centers are choose
4. Proceed like a classical k-means clustering.

They have proved that this approach leads to better results in less time. For more details refer to [11]

1.2.7 Color Quantization

Color quantization is the process of reducing the number of colors in a digital image. The main objective of quantization process is that significant information should be preserved while reducing the number of colors in an image, in other word quantization process shouldn't cause significant information loss in the image. Color

quantization, accepted as a pre-processing application, is used to reduce the number of colors in images with minimum distortion such that the reproduced image should be very close to the original image visually, as in Figure ??.

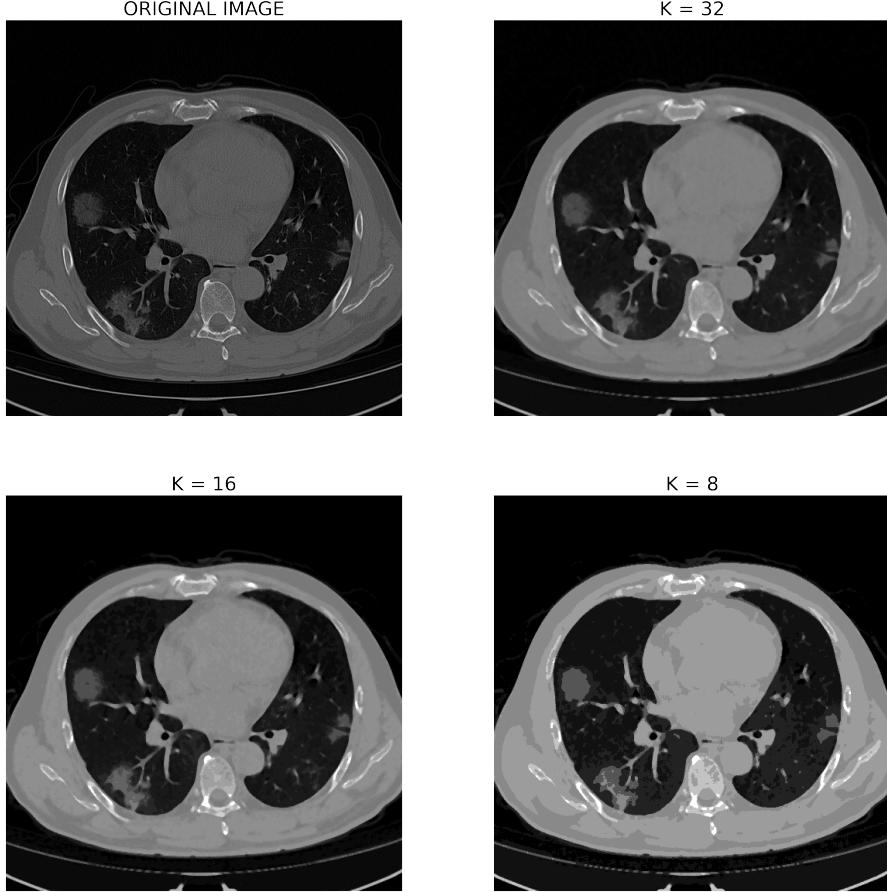


Figure 1.2: *Color quantized GL image. We observe the original image, a 32 color image which look similar to the original one, a 16 colors image and 8 colors image. Tissue are grouped by color similarity. Reduce the color to the number of cluster will*

Color quantization play an important role in many filed of applications such as segmentation, compression, color texture analysis, watermarking, text localization/detection, non photorealistic rendering and content-based retrieval [13].

Color quantization may be used for image segmentation. Use that for this purpose means to reduce the number of colors to the number of the different objects we seek to segment. This means that is is based on the assumption that to each different object class in the image is assigned an unique characteristic color. This is the case of image segmentation, in which each image color represent a particular characteristic of the tissue displayed(i.e in x-ray represent μ). To perform this technique, different algorithms may be used to group the colors, like clustering algorithm or the principal component analysis.

1.2.8 U-Net

Artificial Neural Networks are formed by using artificial neurons derived from physiological models [9]. Neural Networks are made by nodes that simulate a biological

learning. Each node of the network it is able to perform an elementary operation. For imagery analysis are usually used Convolutional Neural Networks(CNN), also known as shift invariant or space invariant artificial neural networks (SIANN).

In biological image processing usually is used the so called U-Net architecture, which is a kind of convolutional neural network which allows to overcome the two main drawbacks of this kind of networks :

- Needs of a huge size fo training data
- Size of the considered network.

That because, for medical and biological segmentation purposes, training dataset with a huge size are not available.

Convolutional Networks usually are used nn classification tasks, which requires only one label. For biological and medical purposes, the segmentation should include localization and a class label should be assigned on each voxel.

To achieve this purposes, in 2015 for the ISB cell tracking challenge, Olaf Ronneberger, Philipp Fischer, and Thomas Brox have developed this kind of network [14]. This kind of network is a modification and simplification of a fully convolutional neural network, making it suitable to works with few training samples. The whole structure is divided into two main parts:

- Contraction path(*encoder*) : sequence of convolutional and pooling layers, which aim to extract features and reduce the input dimensionality.
- Expansion Path(*decode*) : second set of convolutional and up-sampling layers, to reconstruct the feature map size and the segmentation mask, which aims to process the extracted features

Decode path tends to lose some of the igher level features that the encoder learned: Using shortcut connection, the output of the encoding layers are directly passed to the decoder layer, preserving the important features [15].

As I've said the network is composed by two path: the contractive path and the extractive path. The contracting path is a typical convolutional network that consists of repeated application of convolutions, each followed by a rectified linear unit (ReLU) and a max pooling operation. During the contraction, the spatial information is reduced while feature information is increased. The expansive pathway combines the feature and spatial information through a sequence of up-convolutions and concatenations with high-resolution features from the contracting path [14].

The contractive path consist in the application of 3×3 convolutions, each of them followed by a rectified linear unit(ReLU). and a 2×2 pooling operation. very step in the expansive path consists of an upsampling of the feature map followed by a 2×2 convolution (“up-convolution”) that halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3×3 convolutions, each followed by a ReLU. The cropping is necessary due to the loss of border pixels in every convolution. At the final layer a 1×1 convolution is used to map each 64-component feature vector to the desired number of classes

One of the important modification of this network is that in the upsampling part have also a large number of feature channels, which allow the network to propagate

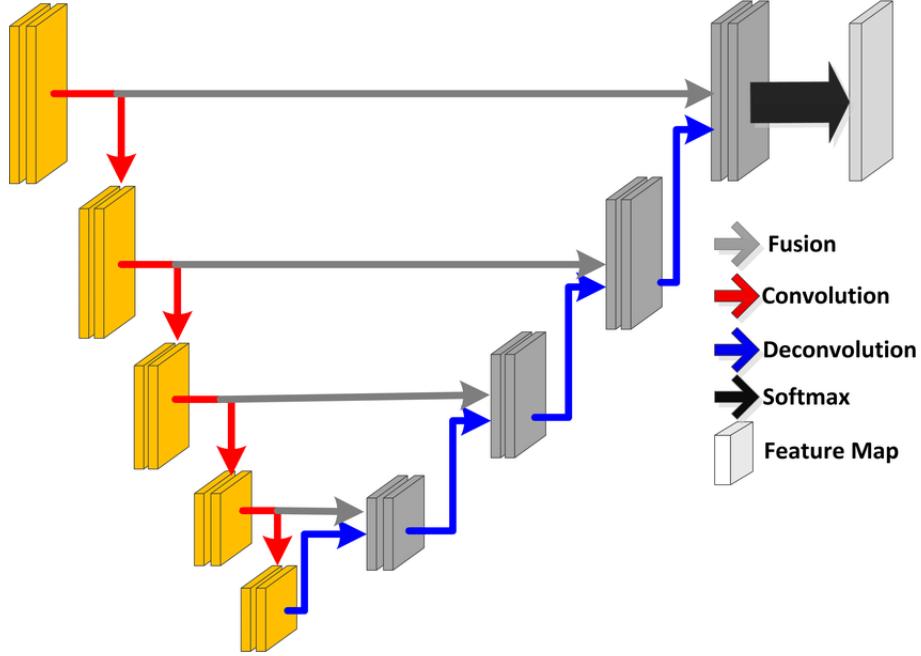


Figure 1.3: *UNEt* network architecture: We can see the *U*-shape made by symmetry between expansion and contraction path. The gray arrows indicates the shortcut used to prevent information loss

context information to higher resolution layers, as a consequence the expansive path is symmetric to the contractive one, making the *U* shape, as we can see in the structure displayed in Figure ???. In a *U*-Net the segmentation map only contains the pixels for which the full context is available [14].

In order to work with few training data, this network makes a huge data augmentation, by applying an elastic deformation to the training images, that allow the network to learn invariance to such kind of transformation. .MainSegmentation

Chapter 2

Lung Lesion Segmentation

The aim of this works of thesis is to develop a pipeline for identification of GGO and CS areas in chest CT scans of patients affected by COVID-19, with following characteristics:

- **Fully Automated:** to remove the dependency from any external operator, and so the subjectivity to its experience.
- **Fast:** in order to compete with certified software and to provides a segmentation in few minutes.

Before stat to speak about the the pipeline, is useful to define the ares we wish to identify.

Austin in [16] has defined the GGO and CS as :

"Hazy increased attenuation of lung, but with preservation of bronchial and vascular margins; caused by partial filling of air spaces, interstitial thickening, partial collapse of alveoli, normal expiration, or increased capillary blood volume, which is different from consolidation in which bronchovascular margins are obscured."



Figure 2.1: Chest CT image of a patient affected by COVID-19. We can see GGO and CS on the right lung. We can appreciate how the different structures, like lesions, are identified by similar color, so the basic segmentation idea was to group voxel by color similarity.

An example of these areas is in Figure ??, in which chest CT scan of patient affected by COVID-19 are visualized. We can notice that the different structures are characterized by similar gray level: the basic idea was to use the color quantization for medical image segmentation, grouping voxels based on color similarity, and assigning to each tissue a characteristic color. This can be done since in CT scan exists a relationship between the tissue in the voxels and the Gray Levels used to display it, given by the Hounsfiend Units(eq ??): the colors are proportional to HU, which are defined as a linear transformation of the linear attenuation coefficient(μ).

We can consider different properties beside the single voxel intensity. As we can see, lesion areas involves many closest voxels. It is interesting to incorporates also neighborhood voxel information in the color quantization. Moreover, contrast between sick and healthy areas may change according to the severity of the disease. As a consequence it is interesting to incorporates also different gamma of the image, in order to enhance these regions.

To takes into account these multiple information, I've build a multichannel image in which each channel corresponds to a different image property. To exploites these properties I've applied different filters to the raw image tensor.

In this way I've build a color space in which each primary color corresponds to a different image property.

Once I've build the color space, we I've to find the characteristic color of each tissue under study, which is represented by a centroids in the color space. In order to perform this task and to achieve the centroids estimation, a simple k-means clustering was used, since it provides a good balance between segmentation performance and computational efficiency.

K-means clustering requires a prior knowledge about the number of cluster, which in our case is given by the anatomical structure of the lung. I have consider a different cluster for each anatomical structure. Once I have estimated the centroids for each tissue, I have used them for the actual segmentation, by assign each voxel to the cluster of the closest centroids

Before each of these steps we need a preliminary phase that aim to isolate the lung regions in order to exclude the extra lung areas and reduce false positives and motion artifacts. In the end the pipeline structure is divided in three main blocks as we can see in Figure ???. Each block will perform a different step of the segmentation:

- **Pre-Processing and lung extraction:** Preliminary step, it involves the registration of HU, isolation of lung regions and removal of bronchial structures and motion artifacts.
- **Training,** that compute the set of centroids used for segmentation. It involves the clustering of a training dataset.
- **Labeling :** that assign each voxel to the cluster with the closest centroids.

The training step is the one which allows the estimation of the centroids. Once the centroids are estimated this step is no more necessary. In the end the segmentation pipeline results in only two steps: **lung extraction** and **labeling**. The final structure is summarized in Figure ?? in which we can observe the flowchart of each step with an image that shows the partial results.

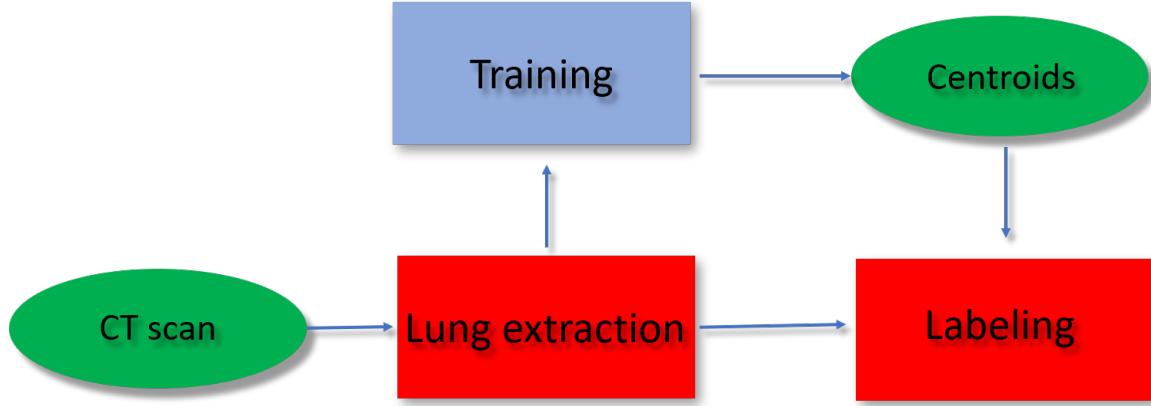


Figure 2.2: Flow chart of the main structure of the developed pipeline. The training process, which allows the estimation of the centroids, is performed only one time.

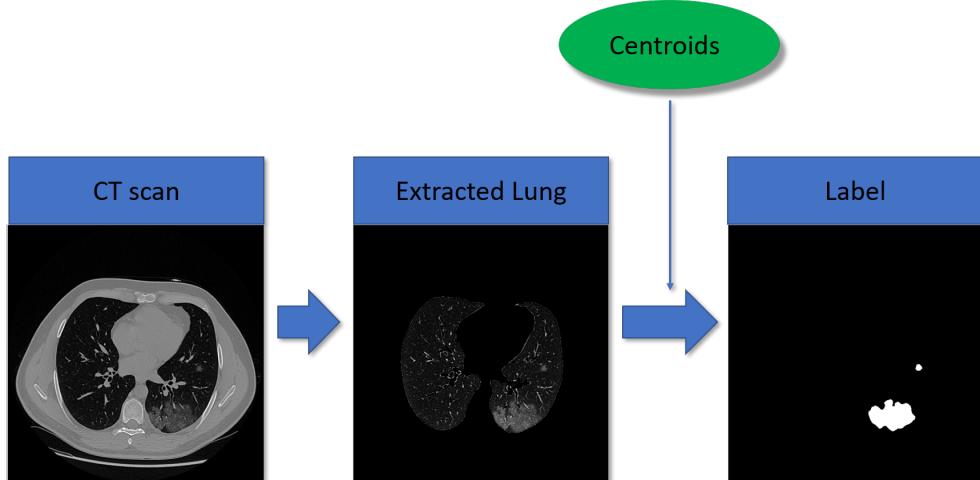


Figure 2.3: Actual segmentation step, from left to right we can see the input image stack, the isolated lung regions and the final label. To perform the labeling a set of pre-computed centroids was used.

2.0.1 Description

As we have seen, the pipeline workflow involves three main steps. In this section I will discuss the general structure of the pipeline, more details about the actual implementation will be given in the next chapter. To perform the color quantization I've found the characteristic color (centroids in the color space) of each tissue and use them for the actual segmentation, dividing the pipeline in two main steps.

Pre Processing and Lung Extraction

This preliminary step is performed before both training and labeling and involves the managing of the HU, the isolation of lung regions and the removal of the main bronchial structures.

The registration of the HU on a common space is necessary to overcome the issues that may raise from the different padding values and multiplicative constant for HU computation (equation ??) used by the different manufacturer of the CT scans. The k constant in the HU definition (equation ??) may change according to

the scan manufacturer or scan model; moreover, during the scan acquisition, all the regions outside the CT tube aren't sampled, so to obtain a square $N \times N$ image for each slice some padding values are added, which different values according to the scan manufacturer. So a first registration was performed.

Lung segmentation is a pivotal pre-processing step in many image analysis such as identification and classification of lung pathologies [14]. The lung isolation allow us to fund a mask for the lung regions, and thus excluding all the body regions, the CT tube and the extra-lung organs like intestine and heart, avoiding the formation of false positives.

Automatic lung segmentation algorithm are typically developed and tested on limited datasets and usually over a limited spectrum of visual variability by containing mainly cases without severe pathologies [14]. Rule based approach, like thresholding, region growing, ect, usually fails for CT scans of patients with severe Interstitial Lung Disease (ILD), as we can see in Figure ???. To achieve the lung segmentation I've used a pre-trained U-Net [14] [17].

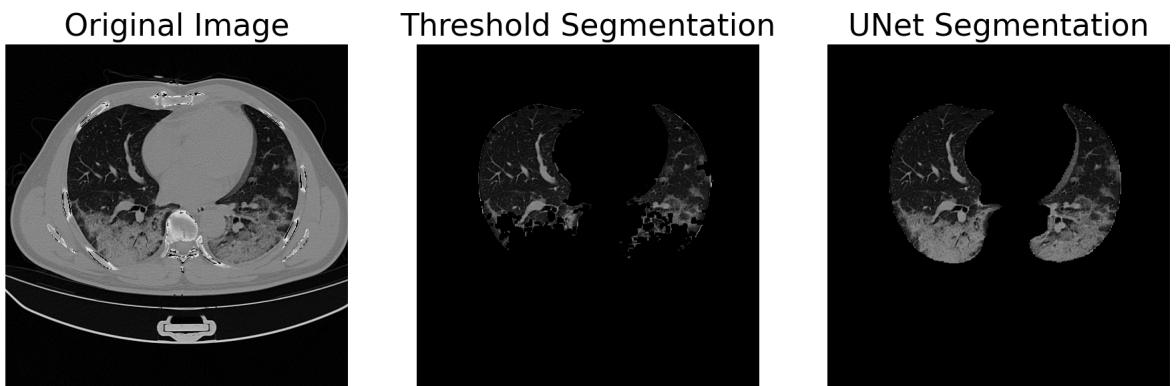


Figure 2.4: From left to right the original CT scan of a patient with severe ILD, the lung segmented by threshold and it's connected components, and lung segmentation achieved by the pre-trained U-Net. We can clearly see the missing areas in the first segmentation, correctly identified in the U-Net one.

This kind of segmentation includes in the lung region also motion artifacts and bronchial structures, which we will see are the principal causes of false positives. To achieve a better segmentation a refinement process is performed, which aims to remove the main bronchial structures from the selected lung regions. Bronchi have an elongated shape respect to the other structure which usually are rounded: the basic idea was to use this kind of information. In order to perform this task, I've computed the covariant matrix of the derivative in a neighborhood and the corresponding eigenvalues. If a particular region has an elongated shape, one of the eigenvalues (corresponding to the eigenvector in the direction of the structure) will have an higher value, otherwise both eigenvalues have a lower value. So we have applied this filter on each slice of the scans and took the maximum eigenvalues.

In Figure ???, I've displayed the image after the lung segmentation by the neural network, and the corresponding eigenvalues map. As we can see the higher values of the map corresponds to the edges of the main bronchial structures. To create the mask for these structures a simple threshold on the map was taken. Since the main bronchial structures are large, this process is able to remove only part of the edges,

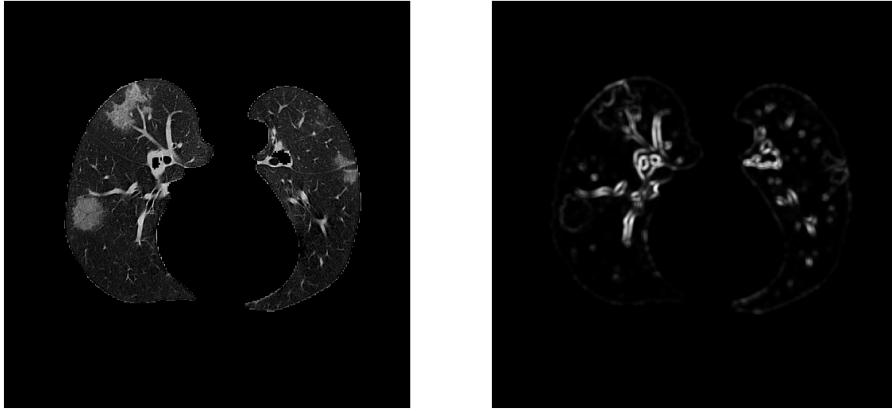


Figure 2.5: From left to right: lung regions selected by the U-Net; ,maximum eigenvalues map of the lung. As we can see the U-Net does not exclude the bronchial structure from the lung, on the other hand, the maximum eigenvalues map delineates very well these regions. We have used this map to remove the unwanted bronchial regions.

but the inner structure is preserved. In order to refine the segmentation, this process is repeated a second time, allowing a more accurate exclusion of the structures.

Training

This step involves the estimation of the centroids for each tissue. To achieve this purpose I have chose to perform a clustering using the k-means algorithm.

I have to takes into account that the k-means clustering requires a balanced representation for each cluster. As we can see in Figure ??, the main part of the image is composed to background. THis cluster is overrepresented. Moreover we can observe that the amount of involved lung volume may change according to the severity of the disease: the first scan presents a low amount pf GGO and CS; se second one have a larger region involves.

To overcome these issue, I have simply removed the background from the segmentation and carefully selected the scans of the training set to ensure an balanced representation of each cluster.

In summary, the implementation of this step involves the building of the multi-channel image, which allow us to takes into account also the neighborhood information, the managing of the over represented clusters and the actual centroids estimation.

Labeling

This step involves the actual segmentation. The script which perform it requires as inputs the CT scans after the lung extraction, and the previously estimated centroids. This block of the pipeline simply assigns each voxel to the cluster corresponding to the nearest centroids then select the only one corresponding to GGO and CS areas. In this way we are performing a pixel classification by assigning the regions to a particular labels according only to their intensities information: this allow us to group on the same cluster objects that are spatially disconnected as often happen in the medical imaging field.

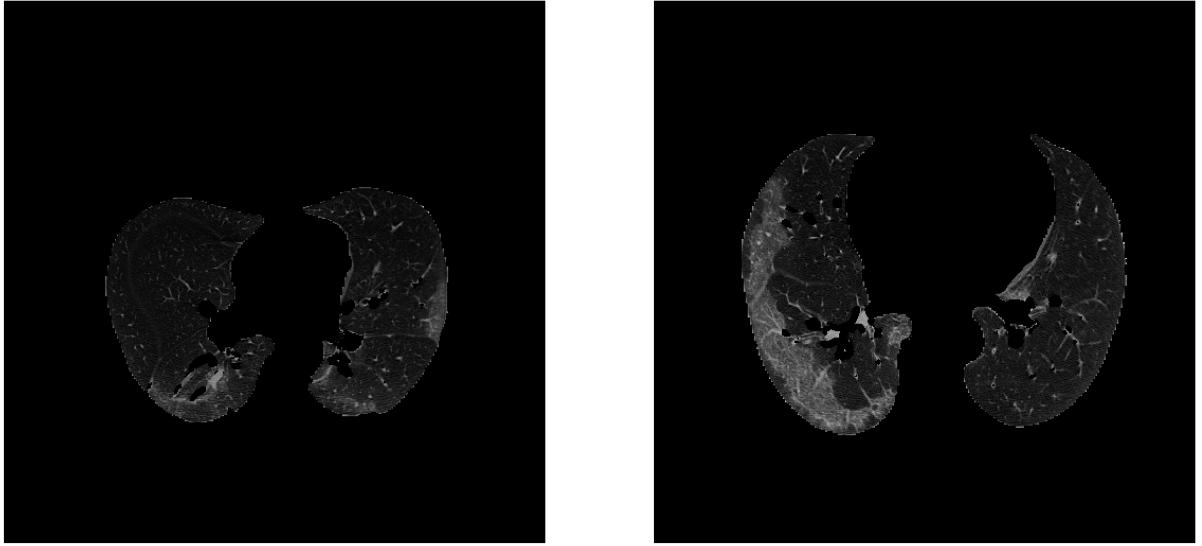


Figure 2.6: Lung regions with different ground glass areas. We can notice that the main part of the image is composed only by the background. From left to right respectively a scan with small lesion areas and one with a large lesion.

The distance between voxel colors and each centroid is defined as the euclidean distance:

$$d(x_j, c_i) = \sqrt{(x_j - c_i)^2}$$

Where x_j is the color vector for the j_{th} voxel and c_i is the i_{th} centroid.

2.1 Implementation

In this section I will discuss in detail the pipeline implementation, focusing on how each step is achieved and implemented.

The whole code is open source and available on GitHub [18] and the pipeline installation is automatically tested on both Windows and Linux by using AppveyorCI and TravisCI. The installation is managed by setup.py, which provides also the full list of dependencies. The code documentation was generated by using sphinx and its available online (<https://covid-19-ggo-segmentation.readthedocs.io/en/latest/?badge=latest>).

The pipeline provides scripts to perform lung segmentation and labeling on a single scan. It provides also an already trained set of centroids. The pipeline allows also to train other set of centroids by with the TRAIN script. This new set can be used instead of the default one.

Till now I've used the (silent) hypothesis that the segmentation involves only one scan. In order to automatize the segmentation on several CT scans, powershell and bash script are provided.

The pipeline operates only on the image array, so no spatial informations are took into account. However, using SimpleITK, I've taken care on the preservation of these informations, in order to allows also quantitative measures.

2.1.1 Lung Extraction

This script aims to achieve the HU registration, the lung segmentation and the bronchial removal.

The lung mask was created by applying the suitable method from [17].

In order to achieve the first step, first of all all the values less than the air value are setted to this value, after that the air value is subtracted from all the voxels. These operation are performed on the image tensor by using NUMPY ndarray method.

Listing 2.1: HU registering function

```

1 import numpy as np
2
3 def shif_and_crop(tensor) :
4
5     tensor[tensor < -1024] = tensor[tensor > -1024].min()
6     tensor = tensor - tensor.min()
7
8     return tensor
9
10

```

More inreresting is the estimation of the eignevalues map. To achieve this purpose I've implemented a function based on CV2.CORNEREIGENVALSANDVEC in OPENCV. This function takes as input an 8-bit gray scale image, so we have to rescale the GL value of the obtained image tensor. Since I don't want to lose information about the rescaled values, I've simply made a copy of this object.

In the end the computing of the eigenvalues map is made by the following function:

Listing 2.2: HU registering function

```

1
2 import cv2
3 import numpy as np
4 from functools import partial
5
6 def max_eigenvalues_map(tensor, block_size, kernel_size) :
7
8     func = partial(cv2.cornerEigenValsAndVecs,
9                   blockSize = block_size,
10                  ksize = kernel_size)
11     res = np.array(list(*list(map(func, tensor))))
12     res = res.transpose(1, 0, 2, 3)
13     res = res[:, :, :, :2]
14     max_eigenvals = np.max(res, axis = 3)
15
16     return max_eigenvals
17
18

```

Once the required parameter were setted, the OPENCV function is iterate over the axial slices. The following transposition and selection of tensor elements, are needed since the function returns also the eigenvectors, in which we are not interested in.

At the end of the whole procedure, the resulting tensor is saved into a medical

image formats. In this step I have paid attention to preserve also spatial informations.

2.1.2 Training

This step consist in the estimation of the centroids of the color space. May be really time consuming, but it is performed only once. To achieved the estimation of centroids, a k-means clustering of the multichannel images of several CT scan from different patients is performed. In order to achieve a uniform representation of each cluster, the scans included in the training set were carefully selected from the 3 datasets used in this work. The main rule used for the selection is that in each scan must be present a huge amount of infection areas and also a well representation of artifacts, in order to takes into account all the possible features. The achievement of the training task involves two main steps :

1. **Preparation of images** : involves the building of the multi channel images, and the registration in a common space;
2. **Clustering** :

Preparation of Images

This step involves the preparation of images, with the building of the multi channel image that incorporates neighbouring and edges information as well as the registration in a common space and the managing of an allocation memory problem.

The used image is composed by 4 channel built as follows:

- Pure image after Contrast Limited Adaptive Histogram Equalization (CLAHE),whi a block sizeof 10×10 pixels
- Image after a median blurring with kernel size equal to 11 pixes
- Image after a gamma correction with $\gamma = 1.5$
- Standard filtered image with a kernel of size 3 pixels

In order to compute the different image filter, I've vectorize the corresponding OPENCV functions by iteration over the axial direction. Since these function requires as input an 8-bit gray scale image, I've to rescale the input tensor.

In Figure ?? I have displayed the 4 different channels of the image. Each channel allow us to consider different information.

The histogram equalized image and the gamma corrected allows to take into account information about the single voxel. The histogram equalization is applied in order to enhance the image contrast by improving the GL usage. For each slice the histogram is equalized considering only a 10×10 area, in order to take care of the over-amplification of the contrast.

The gamma correction is a non linear operation and is used to decode the luminance, and to made in evidence the less evident lesions. Both of these filters involves the single voxel.

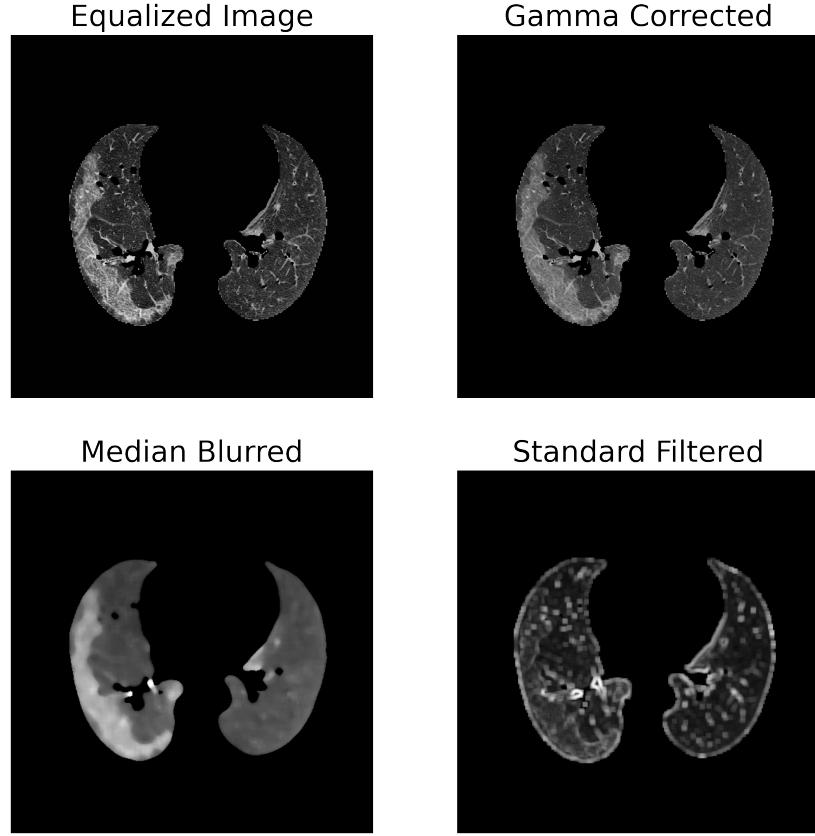


Figure 2.7: *Channels of the image. From left to right and from top to bottom the image after the histogram equalization, the gamma correction, the median blurring and the std filtering. These channels allow us to consider information about single voxels, neighbouring voxels and their variability. The histogram equalization is applied over small 10×10 areas to avoid over-amplification of the contrast. The gamma correction was performed using $\gamma = 1.5$, and the median and std filter kernel sizes are respectively 11 and 3. Each image is normalized according to mean and standard deviation of the whole scan.*

The median blurring allow us to consider also the information of the neighborhood voxels, allowing the reduction of the outliers. The usage of the filter is justified since the lesions involves several closest voxels.

The last channel used is the image after the application of a local standard deviation filter which consist in the replacement of each pixel value with the standard deviation of its neighborhood, help us to distinguish the bronchial structures and motion artifacts not removed in the previous step. Each image channel is normalized according to means and standard deviation of the whole scan.

Clustering

After the construction of the multichannel image of for each input series, all the images are shuffled and divided into several sub-samples.

The clustering was performed independently on each sub-sample and only the centroids set which minimize the intra cluster variance is taken. That because since the results of the k-means clustering depends on the initial choice centroids. Even

if the k-means++ allows an optimal way to chose the initial centroids, also may happen that during the minimization of the potential function the algorithm found a local minimum rather than the global one. The different clustering allow us to estimate more than one possible solution and estimate chose the best one. Moreover the creation of several sub-samples is made since the creation of a single, huge array with several images is not always possible, since requires a huge quantity of memory to be allocated, so we have chose to divide all the images into several sub-samples and chose the optimal ones.

As I have said before, the k-means clustering requires a balance in the different cluster representation. The problem of the under-represented clusters is overcome by a proper choose of the training set. To manage the over-represented cluster (the background) and the clustering of the multiple subsamples, I have defined the KMEANS_ON_SUBSAMPLES function. Implemented as follows.

Listing 2.3: kmenas_on_subsamples

```
1 import cv2
2 import numpy as np
3 from tqdm import tqdm
4
5
6 def kmeans_on_subsamples(imgs, n_centroids, stopping_criteria,
7                         centr_init, weight = None) :
8
9     if weight is not None :
10        vector = np.asarray([ el[w != 0]
11                             for el, w in zip(imgs, weight)], dtype = np.
12        ndarray)
13    else :
14        vector = np.asarray([el.reshape((-1, el[-1])) for el in imgs], dtype=np.ndarray)
15
16    centroids = []
17    for el in tqdm(vector) :
18
19        _, _, centr = cv2.kmeans(el.astype(np.float32), n_centroids,
20        , None, stopping_criteria, 10, centr_init)
21        centroids.append(centr)
22
23    return np.asarray(centroids, dtype= np.float32)
24
```

As we can see the clustering of the different subsamples is managed simply iterating over the subsamples vector and storing the partial results in a list. The background removal is managed by passing a weight tensor. The value of this tensor must be zero for the voxel representing the background, and one otherwise. This allows to build the sample vector removing the background values.

An other problem may be the estimation of the correct number of clusters since k-means clustering requires a prior knowledge on the number of clusters. In our case the anatomical knowledge about the lung may help, since we can consider one cluster for each anatomical structure. I have found that 5 clusters are a viable choice, and the considered structures are the following:

- Lung Parenchima;

- Edges;
- Vessel surrounding bronchial structures;
- Ground Glass Opacities and consolidation;
- Bronchi.

The whole step is summarized in the pseudocode in Figure ??.

Algorithm 1: Pseudo-code for the training script

Data: CT scans with Extracted lung

Result: Centroid matrix

```

foreach scan  $\in$  input_scans do
    | read the scan
    | sample  $\leftarrow$  image_array
end

/* prepare subsamples */ 
sample  $\leftarrow$  build_multichannel(sample)
sample  $\leftarrow$  shuffle(sample)
subsamples  $\leftarrow$  split(sample)
/* start the first clustering */ 
foreach Sub  $\in$  subsamples do
    | center  $\leftarrow$  kmeans(sub, number of centroids)
    | centroid_vector, internal_variance  $\leftarrow$  append(center)
end

/* Refinement */ 
centroid_matrix  $\leftarrow$  centroid_vector(min(internal_variance))

```

2.1.3 Labeling

This is the last step of the pipeline, which involves the actual segmentation. This task is performed by simply assign each voxels to the cluster corresponding to the nearest centroids, in this way an hard segmentation is achieved. This step is performed by the function IMLABELING implemented as follows:

Listing 2.4: kmenas_on_subsamples

```

1 import numpy as np
2
3 def imlabeling(image, centroids, weight = None) :
4
5
6     if weight is not None :
7         distances = np.asarray([np.linalg.norm(image[weight != 0] - c,
8             axis = 1) for c in centroids])
9

```

```

10     weight[weight != 0] = np.argmin(distances, axis = 0)
11     return weight
12 else :
13     distances = np.asarray([np.linalg.norm(image - c, axis = 3) for
14         c in centroids])
15     labels = np.argmin(distances, axis = 0)
16     return labels
17
18

```

The script takes as input the CT scan after the lung extraction and it build the multichannel image as described before. After that it assigns each voxel to the cluster of nearest centroids, which is the one that minimize the distance :

$$cluster = \arg \min_S \sum_{i=1}^k \sum_S \|x - \mu_i\| \quad (2.1)$$

where x is the color vector of the voxel and μ is the i th centroid. During this process the background is automatically assigned to the 0 label, by passing a mask which assume False on the volxel background and True for the other one. To summarize the process, the pseudocode of the script is reported in algorithm ?? I've tested this algorithm on three different dataset, the results are described in the next chapter.

Algorithm 2: Pseudo-code for the labeling script

Data: CT scan to label, centroids

Result: GGO label

image←build_multi_channel

/* Compute distances and found the minimum */

foreach $c \in centroids$ **do**

| distances← $\|image - c\|^2$

end

labels← $\arg \min (distances)$

2.2 Optimization

During each step of the pipeline we have to set different parameters, like the kernel size for median and std filter, as well as the number of centroids to use fro the actual segmentation. In this section I will birefly describe how each one of these parameters was optimized, in order to obtain the best segmentation performances.

2.2.1 Estimation of the Number of Clusters

The designed algorithm for the centroids estimation is the k-means clustering that requires a prior knowledge about the number of clusters to use. This is very important since a bad choice will badly affect the whole segmentation results. In order to

chose the proper number of clusters, I've consider two different sources of information: the anatomical knowledge about the lung and the internal variability of the lung.

From anatomical knowledge about the lung, we can derive 5 clusters, corresponding to:

- Lung Parenchima;
- Edges;
- Vessel surrounding bronchial structures;
- Ground Glass Opacities and consolidation;
- Bronchi.

Notice that the background of the image is not considered as a cluster since it is removed from the segmentation for the reasons explained before. In order to verify that this number of clusters is the best one, I have considered the internal cluster variability.

Clustering techniques try to group the data into different clusters in order to maximize the difference between points in different clusters and to maximize the similarity within each cluster. If the number of centroids is less then the clusters one, the similarity within each cluster is low. Increasing the number of centroids, will reduce reduce the internal variability till 0 (if number of clusters is equal to the number of points). This means that after a certain point the diminishing of the internal variability is no more significant, since do not correspond to the good number of clusters but only to the increasing of their number.

To found the correct number of clusters we are seek to for a number of clusters which still provides a small amount of internal variability.

To achieve this purpose, the clustering was repeated several times increasing the number of clusters and for each iteration the internal variability was measured by the sum of squares estimate error (SSE) :

$$SSE = \sum (x_i - c_j)^2 \quad (2.2)$$

Once this task is completed, the results was printed in Figure ??.

The optimal number of cluster is the one that corresponds to the elbow of the curve. It is difficult to find this feature from visually, I took the numerical value of the elbow considering the point which maximize the distance between the right joining the first and the last point. From this analysis I have found that the optimal number of clusters is 5, which corresponds to the same that I have found considering the lung anatomy.

2.2.2 Kernel Size Optimization

During the building of the multi channel image, I had to compute different image features, that requires the setting of different parameters, like median or standard filter kernel sizes. In order to achieve the best segmentation, I have performed an optimization step that aims to find the parameters that allows to obtain the best segmentation.

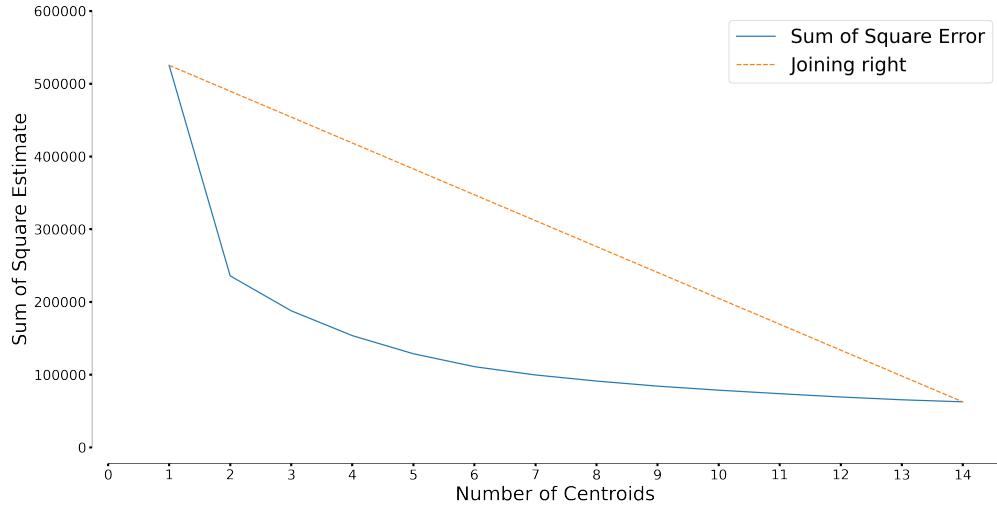


Figure 2.8: Intra-cluster variance vs the number fo clusters(blue); right joining firts and last point(orange). We can se the elbow shape of the graph; the elbow location in the one which maximize the distance between the right joining line and the elbow curve.

Notice that this process is not necessary and a good segmentation can be achieved also by setting this parameters manually.

In order to perform the optimization, I have used SCIKIT-OPTIMIZE [19], more specifically the GC_MINIMIZE METHOD..

This method seeks to perform a Bayesian optimization by using a Gaussian process to approximate an objective function. The function values are assumed to follow a multivariate Gaussian. The covariance of the function values are given by a GP kernel between the parameters. Then a smart choice to choose the next parameter to evaluate can be made by the acquisition function over the Gaussian prior which is much quicker to evaluate [19].

In the end an objective function is minimized. I have used the objective function defined in

I have decided to use the IoU instead of accuracy since the number of pixels concerning the labeled object would be very few against the number of pixels related to the background. Thus, the label would be a matrix with a large amount of zeros (background) and only few ones (object). In this case the standard metric functions have to consider an unbalanced number of samples; so the solution was to use the IoU which measures the ration between the Intersection and Union of the output labels and the binary ground truth [15]:

Algorithm 3: Parameter Optimization Algorithm

Data: Test scans, Ground Truth

Function `objective(parameters, ref_labels, CT_scan):`

```

    labels ← segment(CT_scan)
    iou = IoU(labels, ref_labels)
    return 1 - iou
  
```

`best_parameters ← gc_minimize(objective, n_calls, n_random_init)`

This process allows to optimize the parameters in order to obtain the best results

as possible. As reference labels I have used the ones valuated as gold standard from two experts radiologists whith more than 5 years of experience.

This procedure allows only to tune the parameters for a better segmentation, but the learning process remains unsupervised.

Chapter 3

Results

The developed pipeline was trained and tested on three datasets (Sant'Orsola, ZENOZO and MOSMED), described in the first section. After that I will discuss the segmentation performances. The main method for the evaluation were a quantitative comparison whit a gold standard segmentation, made and validate by 5 experts with at least 2 years of experience. In collaboration with Sant'Orsola, the pipeline results were compared with some semiautomatic segmentation by a blind evaluation made by experts. As a control also segmentation on healthy patient were made, in order to ensure that no lesion were detected. This kind of segmentation was useful also to find the main source of errors.

3.1 DataSet Description

This section is dedicated to the description of the datasets used for the developing and test of the pipeline. The description includes general image characteristics and some metadata. If within the datasets are provided also some ground truth manual segmented labels, also the segmentation masks are described.

3.1.1 Sant'Orsola

Sant'Orsola data was the ones mainly considered in this work. It consist into 83 anonymized CT scans from 83 different patients affected by COVID-19 and 8 scans from healthy controls. Within these scans, also manual annotation were provided. These annotations were obtained with a semi-automatic approach. The built of each annotation may requires several hours.

The series are distributed as follows:

Property	Value
Number of Scans	83
Distribution by sex(M/F/O)	66.3/33.7/0
Distribution by age(min/median/max)	35/60/89

For 5 scans were provided also other semi-automatic segmentations. These segmentation were obtained by refining the initial segmentation of a certified software. The building of these labels has required several days. In the end this results are validated by 5 experts with at least 2 years of experience. These 5 segmentation represents the gold standard and are used as ground truth.

3.1.2 MOSMED

MosMed is a dataset which contains 1110 anonymized CT scan of human lung from both patients affected by COVID-19 in several stages of the disease, and healthy controls. A small subset of this scans is labeled. The scans are obtained between 1st March and 25th of April 2020 by different Russian hospitals. This dataset was born with educational and AI developing purposes. The studies are divided into 5 categories, from healthy patients to the most severe cases. Each scan of the dataset is saved in *.nifti* format and during the conversion from the original dicom series only 1 image every 10 was preserved. The resulting dataset have the following characteristics:

Property	value
Number of Scans	1110
Distribution by sex(M/F/O)	42/56/2
Distribution by age(min/median/max)	18/47/87
Number of studies in each category	254/648/125/45/2

The CT scans are organized into 5 categories, depending on the percentage of the involved lung volume :

Class	Description
CT-0	Normal lung tissues
CT-1	presence of GGO, lung parenchima involved less than 25%
CT-2	GGO, involvement of lung parenchima in 25 – 50%
CT-3	GGO and consolidation, involvement of lung parenchima in 50 – 75%
CT-4	GGO, consolidation and reticular changes, lung parenchima involved more than 75%

Of these five categories only 50 annotations are available, mostly involving only the patients of CT-1 groups. Scans have been annotated by the experts of Research and Practical Clinical Center for Diagnostics and Telemedicine Technologies of the Moscow Health Care Department.

3.1.3 ZENODO

This dataset consists into 20 CT scans of patients affected by COVID-19, labeled by two expert radiologists and verified by third one. The anatomic structures labeled are the left and right lung and the infections regions. Each file is in nifti format and no metadata were available.

Unfortunately only half of the scans are in HU, the remaining are in 8-bit gray scale, which is not suitable to verify the pipeline since requires it as input an image in HU.

3.2 Accuracy

In this section I will discuss the pipeline segmentation compared with the annotations. The pipeline was trained over 10 CT scans, carefully selected from the three available datasets. This ensure a balanced cluster representation. Using this set of

centroids I have segmented the CT scans of the 3 available datasets, and match the achieved results with the annotations, when available.

The automatic pipeline was run on the servers of the Department of Physics and Astronomy (DIFA), and it's able to achieve a segmentation in less than 2 minutes.

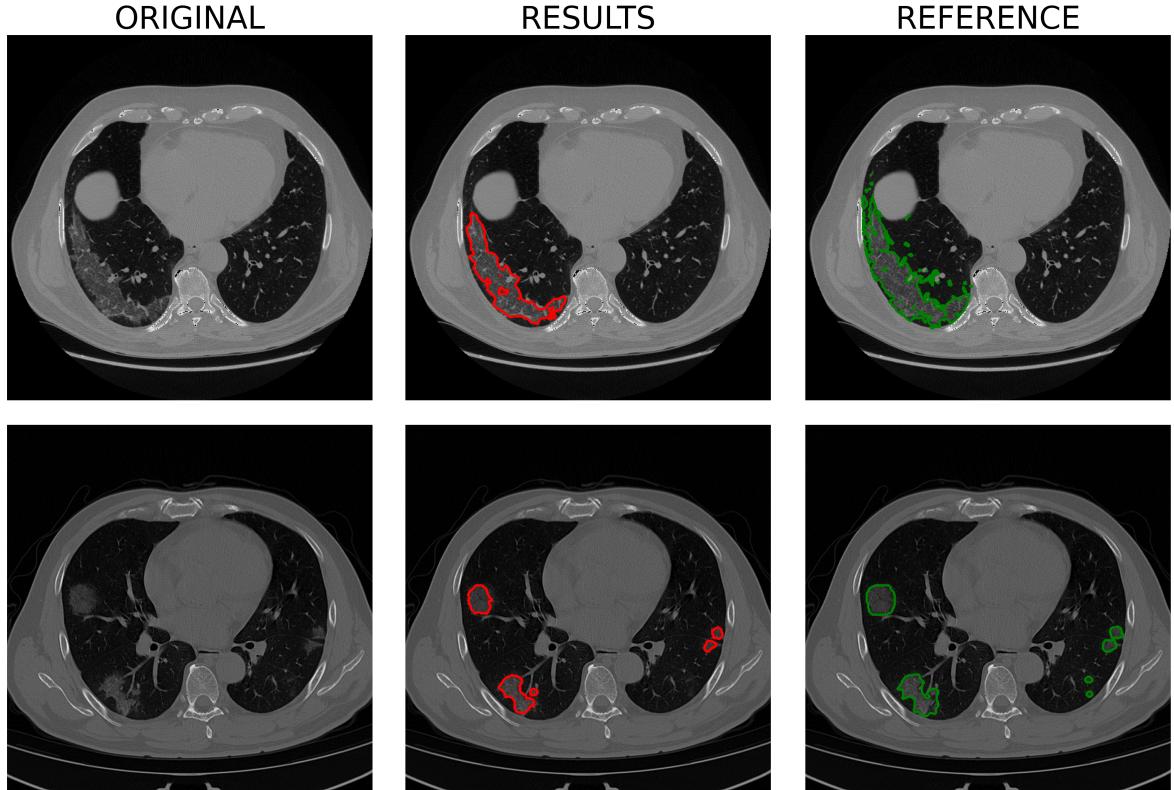


Figure 3.1: Comparison between the achieved segmentation (red) and the reference (green). We clearly see that the GGO and CS areas are well identified and segmented.

In Figure ?? I've reported a comparison between the achieved results and the reference segmentation. We can observe that GGO and CS areas are correctly segmented.

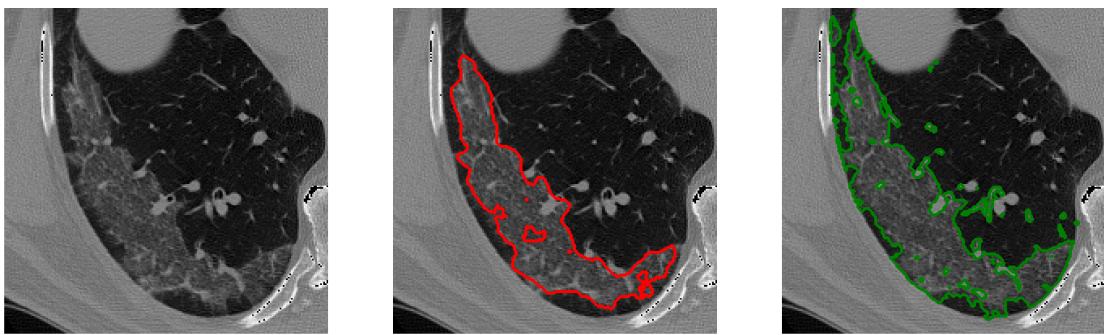


Figure 3.2: Focus on lesion area. From left to right we can observe the original image, the predicted lesion areas and the reference one. We can observe that the annotation identify a larger area. The annotation identify also some spots outside the lesions which seems to be healthy tissue

In Figure ?? I have reported a zoom on the identified lesions areas. As we can see both the pipeline segmentation and the annotations correctly identify the areas of interest. We can observe that in the semi-automatic method there are some spot which do not seem to belong to a lesion area. Moreover the automatic segmentation seems to identify a lesions with less areas.

In Figure ?? I've reported a 3D rendering of the lung, with the identified lesions. In pink we can see the manual annotation, in green the segmentation achieved by the pipeline. From these images is more clear that the total estimated volume is different between the two methods. In particular the annotations incorporates a larger volume than the pipeline segmentation.

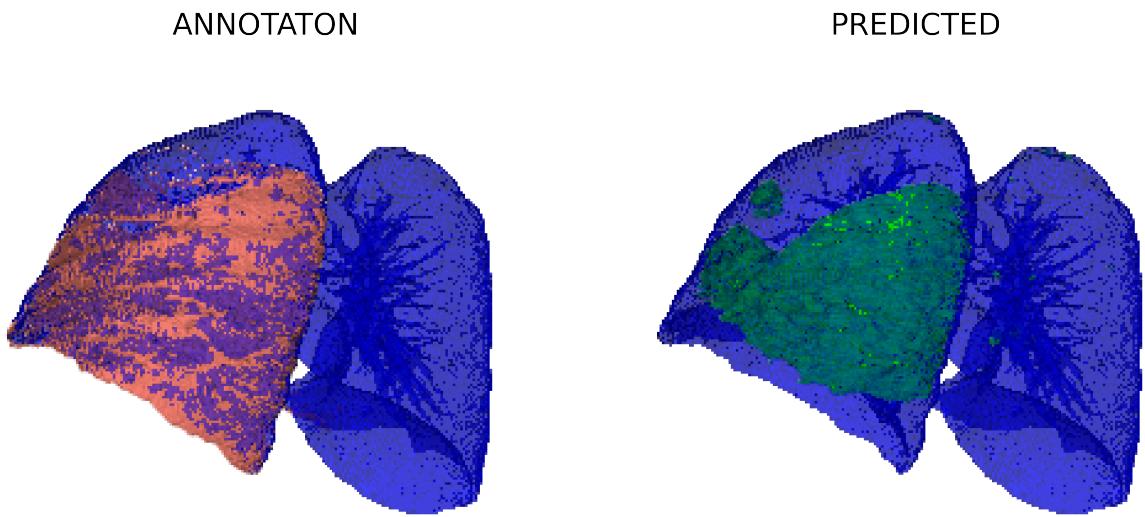


Figure 3.3: 3D representation of lesions. From left to right the manual annotation (pink) and the pipeline segmentation (green). We can observe that the segmented size of the area is different.

In order to assess the quality of the achieved segmentation, I've used two different method.

First of all I've compared both the pipeline segmentation and the annotations with a gold standard. In this way I was able to measure the proportion of the areas correctly identified.

After that the segmentation was submitted to five experts in order to made a blind evaluation, comparing the pipeline segmentation with the reference one.

In the section below I will describe these kind of measures.

3.2.1 Gold Standard

In order to obtain quantitative measures about the performances of the developed pipeline, I have compared the results with the manual annotation provided by the Department of Diagnostic and Preventive Medicine of the Poloclinico Sant'Orsola - Malpighi.

The comparison was made using sensitivity and specificity scores.

Sensitivity refers to the ability to correctly detect ill areas. It is defined as the number of voxel classified as lesions among all the lesions. in other world is defined as the number of true positives over the total number of lesions voxels (True POsitives + False negatives) :

$$Sensitivity : \frac{TruePositive}{TruePositive + FalseNegatives} \quad (3.1)$$

specificity relates to the ability to correctly reject healthy areas. Is defined as the number of rejected pixels(True Negative) agains the total numner of healthy areas (True Negative + False Positives) :

$$spcificity : \frac{TrueNegative}{TrueNegative + FalsePositives} \quad (3.2)$$

As ground truth I have considered the 5 gold standard segmentations.
The results of these measures are displayed in table ??

	Predicted		Annotation	
	Sensitivity	Specificity	Sensitivity	Specificity
Patient 1	0.412	~ 1.00	0.676	0.999
Patient 2	0.399	~ 1.00	0.698	0.995
Patient 3	0.570	~ 1.00	0.653	0.999
Patient 4	0.512	~ 1.00	0.325	0.999
Patient 5	0.628	~ 1.00	0.974	0.999

Table 3.1: Sensitivity and Specificity for the pipeline segmentation and annotation. As a ground truth was used a sem-automatic segmentation made and evaluated by 5 experts with at least 2 years of experience.

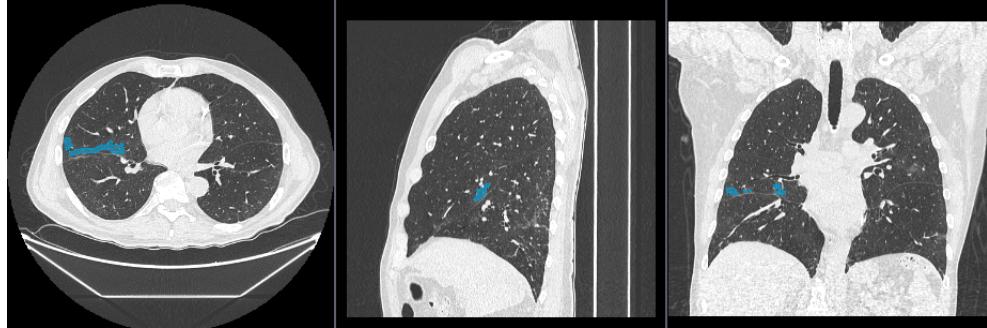
The first thing we can notice is that the segmentation performances achieved by the semi-automatic method seems to have better quality perfomances, achieving a better capability to detect lesion areas. However we have to point out that the semi-automatic method requires several hours and the involved of trained experts. The proposed pipeline requires only few minutes and no trained personnel.

We have also to point out that the achieved semiautomatic segmentation is subject on the experience of the operator. Only few patients per days may be segmented without lack of performances, since after several subsequent segmentation the level of concentration will drop.

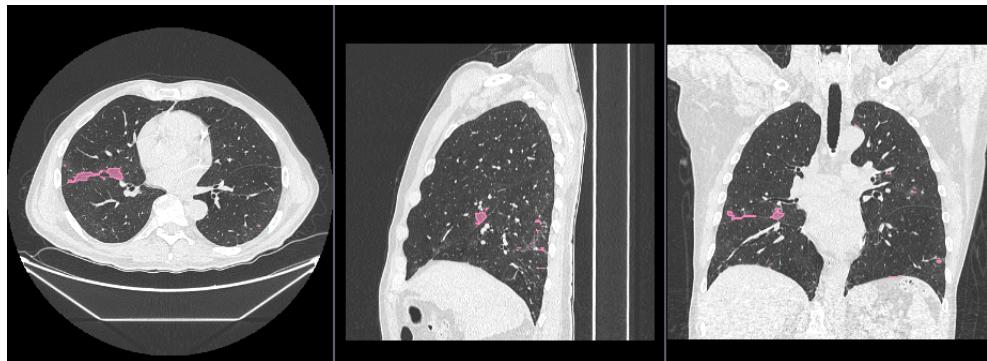
We have also to point out that the developed pipeline may segment several patient without lack of performances. The training personnel may segment only one or two patient per days,

The specificity values allow us to think that the developed pipeline doesn't produce an high number of false positives. The sensitivity may suggest that some lesion areas may be missed. We will see that each area is identified, however there is an underestimation of the total lesion volume.

In Figure ?? I've reported a comparison between the ground truth (blue) and the pipeline segmentation (pink) for the first patient in axial, sagittal and coronal view. We can see how the main lesion areas are correctly identified, however a low underestimation of the total lesion volume is presents, toghether with the presence



(a) Ground truth for Axial, sagittal and coronal view of the first patient



(b) Predicted lesions areas for Axial, sagittal and coronal view of the first patient

Figure 3.4: Comparison between the gold standard segmentation(blue) and the pipeline results(pink) for axial, sagittal and coronal view of a patient with a low involvement of lung parenchima. We can see that te main lesion areas are identified, even if an underestimation of the total volume is present together with some small misclassified points.

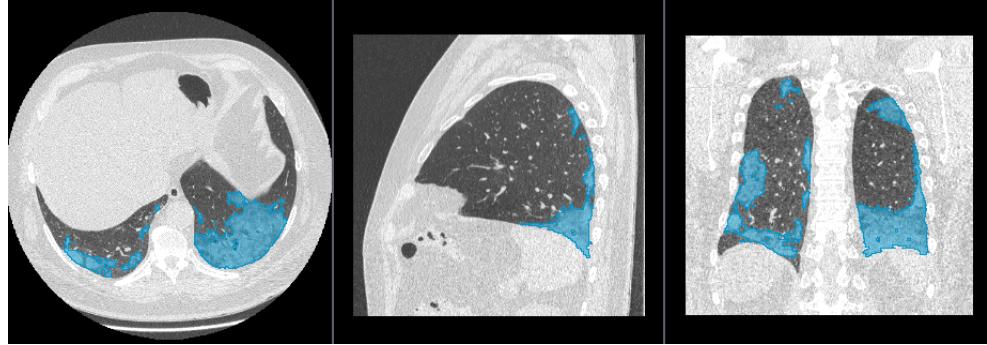
of false positives in the lowest back regions. This is a particular case, since the involvement of lung volume is very low, as well as the contrast between lesions and healthy lung.

In Figure ?? I've reported a comparison between the ground truth (blue) and the pipeline segmentation (pink) for the third patient in axial, sagittal and coronal view. The patient presents an high involvement of lung regions. The lesion areas, which in this case presents an high contrast with the lung regions, are correctly selected. In this case no misclassified points are detected. This suggest that the pipeline is able to correctly identify the lesion areas.

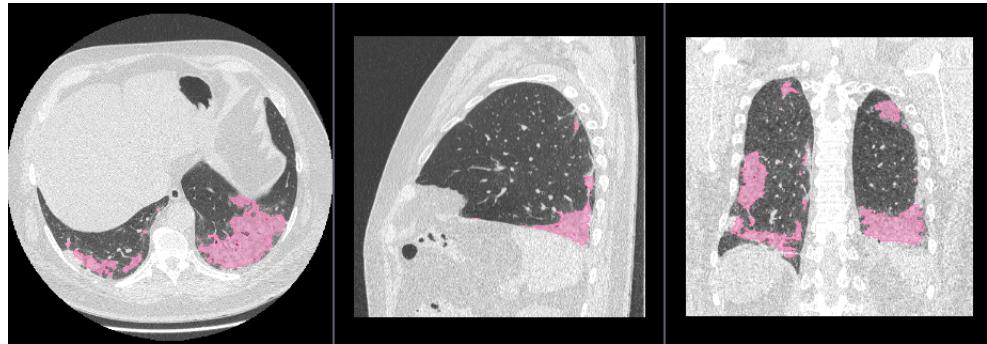
3.2.2 Expert Evaluation

An other measure of the quality of the segmentation used in this work has been blind evaluation. Five expert with at least 2 years of experience have compared the pipeline segmentation and the reference labels provided within the dataset (obtained by semiautomatic technique).

To perform the evaluation we have randomly select 40 patients from the three datasets and organized the scans as follows: for each patient we have displayed, slice by slice, two images: the scans with the pipeline segmentation (the one to test), and the semi-automatic labels provided within the dataset (control) as in Figure ??.



(a) Ground truth for Axial, sagittal and coronal view of the third patient



(b) Predicted lesions areas for Axial, sagittal and coronal view of the third patient

Figure 3.5: Comparison between the gold standard segmentation(blue) and the pipeline results(pink) for axial, sagittal and coronal view for a patient with a large involvement of lung parenchima. We can see how the lesion areas are correctly identified

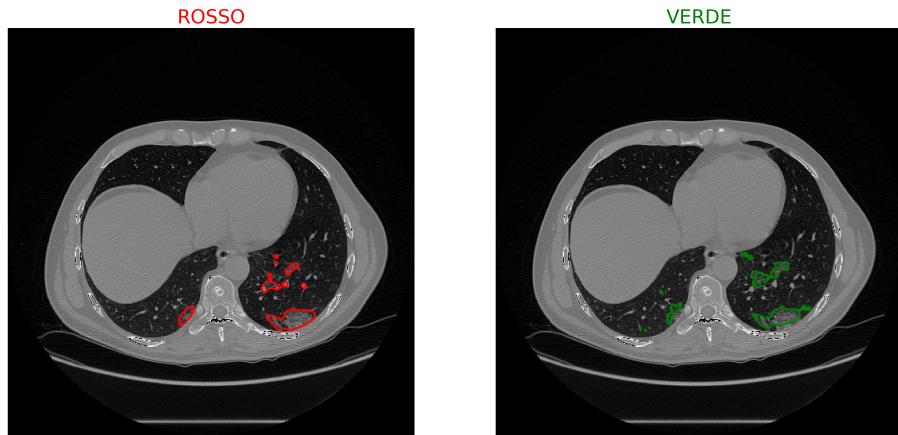


Figure 3.6: Example of comparison submitted to the experts. The two images report the segmentation of the same slices achieved with two different techniques: Manual and pipeline. Was asked to the expert to decide which one is better, without knowing which technique is used to obtain the segmentation.

The scans, organized in this way, was evaluated by experts, to which has been asked to decide, slice by slice, which segmentation is better or if the quality is equal.

The core of this method is that the expert don't know which scan corresponds to a specific segmentation technique, so this will lead to an ideally unbiased evaluation.

In order to ensure that, the order of segmentation was shuffled between patients.

Conclusions

In this work I've developed a fully automated pipeline for the identification of lesions (GGO and CS) in chest CT scans of patient affected by COVID-19. The whole pipeline achieves the segmentation by using the color quantization technique, that allows to exploit the color similarity between voxels belonging from the same tissue. The multichannel nature of digital images has allowed also to consider other properties besides the single voxel intensity: also neighboring voxel information are takes into account.

As a preliminary step, a lung segmentation is performed by using a pre trained U-Net, followed by a bronchial removal, which aims to reduce the rate of false positives. To achieve the segmentation, we have found the characteristic color of each tissue by performing a k-means clustering in the color space. The actual segmentation is achieved by assigning each voxel to the cluster corresponding to the nearest color.

The pipeline, tested on three different datasets, has provided good segmentations for the typical lesion cases with an accurate identification of GGO and CS. The proposed algorithm has also shown some limitations due to the eventual presence of motion artifacts, caused by heartbeat and respiratory cycle; this behaviour was observed also on segmentation made on healthy controls, which have shown some misclassified points in these areas.

In the end we have developed a fully automated pipeline which achieve a good segmentation for the most typical cases.

Color quantization has shown to be a suitable approach to face this kind of problems, and some improvements, like removal of motion artifacts are still possible.

Bibliography

- [1] Zhao Fu and et al. “CT features of COVID-19 patients with two consecutive negative RT-PCR tests after treatment.” In: *Scientific reports* (July 2020). DOI: [10.1038/s41598-020-68509-x](https://doi.org/10.1038/s41598-020-68509-x).
- [2] Huang C, Wang Y, and et al. “Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China”. In: *Lancet* (Feb. 2020). DOI: [10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5).
- [3] Adam Bernheim et al. “Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to Duration of Infection”. In: *Radiology* 295.3 (2020). PMID: 32077789, p. 200463. DOI: [10.1148/radiol.2020200463](https://doi.org/10.1148/radiol.2020200463). eprint: <https://doi.org/10.1148/radiol.2020200463>. URL: <https://doi.org/10.1148/radiol.2020200463>.
- [4] Tao Ai et al. “Correlation of Chest CT and RT-PCR Testing for Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases”. In: *Radiology* 296.2 (2020). PMID: 32101510, E32–E40. DOI: [10.1148/radiol.2020200642](https://doi.org/10.1148/radiol.2020200642). eprint: <https://doi.org/10.1148/radiol.2020200642>. URL: <https://doi.org/10.1148/radiol.2020200642>.
- [5] Collins J Stern E J. “Ground-glass opacity at CT: the ABCs”. In: *American Journal of Roentgenology* 169 (1997), pp. 355–366. DOI: [doi:10.2214/ajr.169.2.9242736](https://doi.org/10.2214/ajr.169.2.9242736).
- [6] Ma Jun et al. *COVID-19 CT Lung and Infection Segmentation Dataset*. Version Verson 1.0. Zenodo, Apr. 2020. DOI: [10.5281/zenodo.3757476](https://doi.org/10.5281/zenodo.3757476). URL: <https://doi.org/10.5281/zenodo.3757476>.
- [7] N.A. Vladzymyrskyy A.V. Ledikhova N.V. Gombolevskiy V.A. Blokhin I.A. Gelezhe P.B. Gonchar A.V. Morozov S.P. Andreychenko A.E. Pavlov and Chernina V.Y. *MosMedData: Chest CT Scans With COVID-19 Related Findings*. Version Verson 1.0. 2020. URL: <https://mosmed.ai/>.
- [8] Dzung L. Pham, Chenyang Xu, and Jerry L. Prince. “Current Methods in Medical Image Segmentation”. In: *Annual Review of Biomedical Engineering* 2.1 (2000). PMID: 11701515, pp. 315–337. DOI: [10.1146/annurev.bioeng.2.1.315](https://doi.org/10.1146/annurev.bioeng.2.1.315). eprint: <https://doi.org/10.1146/annurev.bioeng.2.1.315>. URL: <https://doi.org/10.1146/annurev.bioeng.2.1.315>.
- [9] D. Withey and Z. J. Koles. “A Review of Medical Image Segmentation: Methods and Available Software”. In: 2008.

- [10] Dr J P Chaudhari Pooja V. Supe Prof. K. S. Bhagat. “Image Segmentation and Classification for Medical Image Processing”. In: *International Journal on Future Revolution in Computer Science & Communication Engineering* 5.1 (), pp. 45–52.
- [11] David Arthur and Sergei Vassilvitskii. “K-Means++: The Advantages of Careful Seeding”. In: vol. 8. Jan. 2007, pp. 1027–1035. DOI: 10.1145/1283383.1283494.
- [12] Laurence Morissette and Sylvain Chartier. “The k-means clustering technique: General considerations and implementation in Mathematica”. In: *Tutorials in Quantitative Methods for Psychology* 9 (Feb. 2013), pp. 15–24. DOI: 10.20982/tqmp.09.1.p015.
- [13] Ozturk Celal, Hancer Emrah, and Karaboga Dervis. “Color Image Quantization: A Short Review and an Application with Artificial Bee Colony Algorithm”. In: *Informatica* 25.3 (2014), pp. 485–503. ISSN: 0868-4952. DOI: 10.15388/Informatica.2014.25.
- [14] Johannes Prayer Forian Pan Jeanny Röhrich Sebastian Prosch Helmut Langs Georg Hofmanninger. “Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem”. In: *European Radiology Experimental* 4 (2020). ISSN: 1. DOI: 10.1186/s41747-020-00173-2. URL: <https://doi.org/10.1186/s41747-020-00173-2>.
- [15] Nico Curti. *Implementation and optimization of algorithms in Biomedical Big Data Analytics*. <https://nico-curti2.gitbook.io/phd-thesis/>. 2019.
- [16] J. Austin et al. “Glossary of terms for CT of the lungs: Recommendations of the Nomenclature Committee of the Fleischner Society”. In: *Radiology* 200 (Sept. 1996), pp. 327–31. DOI: 10.1148/radiology.200.2.8685321.
- [17] Hoa Nguyen Johannes Hofmanninger. *Automated lung segmentation in CT under presence of severe pathologies*. <https://github.com/JoHof/lungmask>. 2020.
- [18] Biondi Riccardo et al. *COVID-19 Lung Segmentation*. <https://github.com/RiccardoBiondi/segmentation>. 2020.
- [19] Tim Head et al. *scikit-optimize/scikit-optimize*. Version v0.8.1. Sept. 2020. DOI: 10.5281/zenodo.4014775.