Alma Mater Studiorum · University of Bologna

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

Implementation of an automated pipeline for predicting the response to neo-adjuvant chemo-rediotherapy of colorectal cancer

Supervisor:
Prof. Gastone Castellani

Co-supervisor:
Dr. Nico Curti

Submitted by:
Giuseppe Filitto

Academic Year 2020/2021

Abstract

Colorectal cancer is a malignant neoplasm of the large intestine resulting from the uncontrolled proliferation of one of the cells making up the colorectal tract.
In Western countries, colorectal cancer is the second largest malignant tumor after that of the breast in women and the third after that of the lung and prostate in men. Risk factors for this kind of cancer include colon polyps, long-standing ulcerative colitis, diabetes II and genetic history (HNPCC or Lynch syndrome). In order to get information about diagnosis, therapy evaluation on colorectal cancer, analysis on radiological images can be performed through the application of dedicated algorithms.
In this scenario, the correct and fast identification of the cancer regions is a fundamental task. Up to now this task is performed using manual or semi-automatic techniques, which are time-consuming and subjected to the operator expertise.
The aim of this project is to provide an automated pipeline to predict the response to neo-adjuvant chemo-radiotherapy of patients affect by colorectal cancer.

. . . To my family and Nicole

Contents

Introduction

Colorectal cancer is a malignant neoplasm of the large intestine resulting from the uncontrolled proliferation of the cells making up the colorectal tract. Colorectal cancer is the second malignant tumor per number of deaths after the lung cancer and the third for number of new cases after the breast and lung cancer[1].
Among the risk factors for this kind of cancer, non hereditary could range from colon polyps to long-standing ulcerative colitis, from Crohn's disease to old age. Also genetic history (HNPCC or Lynch syndrome) and nutritional factors as diabetes II can increase the probability of develop cancer [2]. Preventive measures for colorectal cancer include physical activity, reducing the consumption of processed meat and alcohol, and avoiding smoking[3].

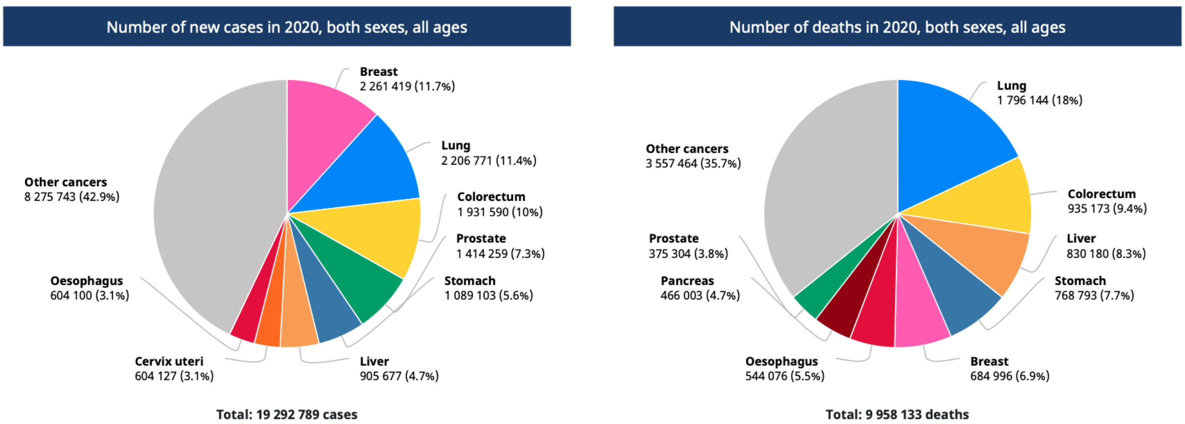

Figure 1: World's cancer cases and deaths. From [1]

Screening and diagnosis methods for colorectal cancer can be based on different techniques. The gold standard in medical routines is colonoscopy which is an invasive technique[4]. Among medical imaging techniques, Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) are the most used[2]. In particular Magnetic Resonance

Imaging (MRI) is used for pre-operative predictions and for the evaluation of the neo-adjuvant therapy of patients affected by colorectal cancer [2].

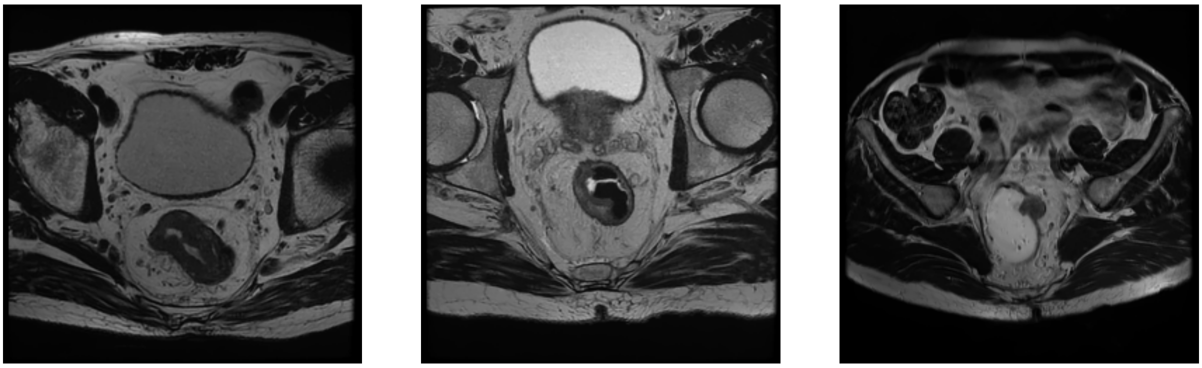

Figure 2: Example of MRI scans of patients affected by colorectal cancer. From Sant'Orsola original Dataset.

In order to get information about diagnosis, therapy evaluation, stage of colorectal cancer, analysis on radiological images can be performed through the application of dedicated algorithms.
In this scenario, the correct and fast identification of the cancer regions is a fundamental task. Up to now, this segmentation task is performed using manual or semi-automatic techniques, which are time-consuming (requiring hours per day) and subjected to the operator expertise since it requires the interaction with trained specialists[2, 4]. Moreover, due to the highly sensitivity to the operator expertise, the obtained results cannot be reproduced[5]. To overcome these issues, an automatic and fast way is required.
The aim of this project is to provide an automated pipeline to predict the response to neo-adjuvant chemo-radiotherapy of patients affect by colorectal cancer. The work is based and tested on MRI scans provided by IRCCS Sant'Orsola-Malpighi Policlinic.
The discussion will start focusing on medical digital images to understand their properties and features. After that, an overview on the main segmentation method for the identification of the cancer regions will be given. Then, the main pipeline characteristics and structure will be described. In particular, how the segmentation was achieved using a Convolutional Neural Network and how extracting and processing medical image features. Finally, the result will be shown and discussed.

Chapter 1

Medical Digital Images

A medical digital image is the representation of the anatomical (or functional) structure of the patient composed by a finite number of picture elements called *pixels*. Each pixel is a discrete numeric representation for its intensity or gray-level, that is an output coming from its two-dimensional function $f(x, y)$ fed as input by its spatial coordinates denoted with (x, y) on the x-axis and y-axis, respectively [6].
A digital image can be processed by computers, this process is called *digital image processing*. It is useful to divide the mentioned process into two main categories: the methods whose output and input data are images (*image processing*) and the methods whose input data can be images and the output data are attributes extracted from the images themselves (*image analysis*).

1.1 General Properties

The physical meaning of the image data depends on the performed image modality. For example Computed Tomography (CT) and Magnetic Resonance Imaging (MRI), give structural information about the anatomy of the patient. Other techniques, such as Positron Emission Tomography (PET) or Functional Magnetic Resonance Imaging (fMRI) give information about the functional properties of the patient's target organs. However, we can distinguish some general characteristics of digital images:

Pixel depth is the number of bits used to encode the values of each pixel and it is related to the memory space used to store the amount of the encoded information[7]. Higher the number of bits, higher the information stored but also more memory space is required[7]. A group of 8 bits is called byte and represent the smallest quantity that can be stored in the memory of a computer. For example, if an image has a pixel depth of

16 or 12 bits the computer will always store two bytes per pixel[7]. With a pixel depth of 8 bits it is possible to codify and store integer numbers between 0 and 255 ($2^8 - 1$). There are also two formats for the encoding in binary of floating-point numbers: single precision 32-bit and the double precision 64-bit.

Pixel data represent numerical values of the pixels are stored according to the data type. Radiological images like CT and MR store 16 bits for each pixel as integers. Image data may also be of complex type even if this data type is not common and can be bypassed by storing the real and imaginary parts as separate images. For example, complex data is provided by arrays that in MRI store acquired data before the reconstruction (the so called k-space) or after the reconstruction if you choose to save both magnitude and phase images[7].

Metadata are information that describe the image stored usually at the beginning of the file as a header[7]. In the case of medical images, metadata have an important role due to the nature of the images itself. For example, a magnetic resonance image will have parameters related to the pulse sequence used, timing information, number of acquisitions. More, a PET image will have information about the radiopharmaceutical injected and the weight of the patient. Medical image metadata can also include information about the patient.

1.1.1 Medical Image Formats

Image file formats provide a standard way to store information of an image in a computer file[8]. Medical image file formats can be divided in two categories. The first is formats intended to standardize the images generated by diagnostic modalities. The second is formats born with the aim to facilitate and strengthen post-processing analysis[7].

DICOM is the acronym of Digital Imaging and COmmunications in medicine. It is not only a file format but also a network communication protocol[7]. However here, we will discuss DICOM only as a medical image format.
DICOM file format establishes that the pixel data cannot be separated from the metadata[7]. In other words, metadata and pixel data are merged in a unique file. The header contains the description of the entire procedure used to generate the image in terms of acquisition protocol and scanning parameters[7]. It also contains patient information such as name, gender, age, weight, and height. For these reasons, the DICOM header is modality-dependent and varies in size. In practice, the header allows the image to be *self-descriptive.*

1.2 Spatial Domain Filtering

Filtering is a technique for modifying or enhancing an image. The term *spatial domain* refers to the plane of the image itself, where the related processing methods are based on the direct manipulation of the pixels. Among the various categories of spatial processing there are *intensity transformations* and *spatial filtering*. The former operate on single pixels while the latter on every pixel's neighborhood.
Mathematically, we can express this processes as follow:

$$g(x, y) = T[f(x, y)] \tag{1.1}$$

where $f(x, y)$ is the input image, $g(x, y)$ the output image and T is an operator defined on f around a point (x, y). The operation on the point located in (x, y) usually involves the application of a matrix called *mask* or *kernel*. It must have $M \times N$ dimensions with M and N odd, in order to make the center of the mask coincide with the pixel in question, and occupy a small section of the image. The application of the above-mentioned mask (or kernel) on an image is called *spatial filtering*.

1.2.1 Spatial Filter

A spatial filter consists in a (usually square) region called *mask* and a pre-defined operation applied on pixels of the region covered by the mask[9]. Filtering creates a new pixel with the same coordinates as the center of the neighborhood whose value is the result of the operation. For each (x, y) of the image, the filter transform $g(x, y)$ is the linear combination of the mask coefficient $w(s, t)$ and the pixels of the image affected by the mask itself.
In general, we can write:

$$g(x, y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s, t) f(x + s, y + t) \tag{1.2}$$

1.2.2 Correlation and Convolution

Spatial filtering is a correlation or convolution process. Correlation is the process of moving a *mask* or *kernel* over the image and computing the sum of products at each location[9]. The mechanics of convolution are the same, except that the filter is first rotated by 180° degree. In other words, correlation or convolution are filter shift functions. The correlation and convolution of a filter $w(x, y)$ of size $m \times n$ with an image $f(x, y)$ can be written as follow:

$$Correlation : w(s,t) \times f(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t) \tag{1.3}$$

$$Convolution : w(s,t) \circledast f(x,y) = \sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x-s, y-t) \tag{1.4}$$

1.2.3 Smoothing Filters

Smoothing filters are used for blurring and for noise reduction[9]. This is used in removal of small details and bridging of small gaps in lines or curves. Smoothing spatial filters include *linear filters* and *nonlinear filters*[9].
The general implementation for filtering an $M \times N$ image with a weighted averaging filter of size $m \times n$ is given by:

$$g(x,y) = \frac{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t) f(x+s, y+t)}{\sum_{s=-a}^{a} \sum_{t=-b}^{b} w(s,t)} \tag{1.5}$$

where $m = 2a + 1$ and $n = 2b + 1$.

Linear filtering is based on the *mean filter* [10]. The mean filter is a simple sliding spatial filter that replaces the center value in the mask region with the average of all the neighboring pixel values including itself. These filters are also called *low pass filters* since the process of averaging drastically lowers high frequencies. The mask or kernel is a square. Larger kernels of size 5×5 or 7×7 produces more denoising but make the image more blurred[10]. A common mean filter can be described by a 3×3 matrix with all elements equal to 1, so that the output pixel corresponds to a value of:

$$R = \frac{1}{9} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \mathbf{z} = \frac{1}{9} \sum_{i=1}^{9} z_i \tag{1.6}$$

or using a weighted mean filter:

$$R' = \frac{1}{16} \begin{pmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{pmatrix} \mathbf{z} \tag{1.7}$$

Non-Linear filtering is based on the *median filter*[10]. The median filter principle is similar to the mean filter. A 3×3, 5×5, 7×7 kernel or mask is scanned over pixel matrix of the entire image. The median of the pixel values in the mask region is calculated, and the center pixel of the mask region is replaced with the calculated median value[10]. Mathematically:

$$g(p) = median\{f(p), where\, p \in N_8(p)\} \tag{1.8}$$

where $g(p)$ is the median pixel value, $f(p)$ all pixel values under mask, and $N_8(p)$ 8-neighborhood of pixel p.
This filter is particularly effective in the presence of *impulse noise* (or *salt-and-pepper noise*)[9].

Notes: Adaptive filters are commonly used in image processing to enhance or restore data by removing noise without significantly blurring the structures in the image[11]. This means to not smoothing the areas of the image in which there is a large jump in intensity values (i.e. when there is an *edge*) and at the same time applying the filter to lower the noise. In this case, the local variance will be evaluated in relation to the variance of the noise that occurs.
Mathematically:

$$\hat{f}(x,y) = f(x,y) - \frac{\sigma^2_{noise}}{\sigma^2_{local}}[f(x,y) - m_{local}] \tag{1.9}$$

where if $\sigma^2_{noise} = \sigma^2_{local}$ the mean local value is associated to the output, while if $\sigma^2_{noise} \ll \sigma^2_{local}$ the function is not altered.
The same concept can be used for the median filter, defining the Adaptive Median Filter algorithm: it is first defined a portion of the image S_{xy} of variable size until a pre-defined size S_{xyMAX}; then it is calculated from this area the maximum intensity value (z_{max}), the minimum value (z_{min}) and the median value (z_{median}). In the first step of the algorithm, the aim is to see if the median value coincides with the intensity extremes present in the area, that is, if inside the mask there is actually an area with discontinuities due to noise. In the second step of the algorithm, the aim is to check if the point in the center of the area, z_{xy}, is to be modified with the median value or to be left unchanged. If the maximum window size is reached, then simply the median value is returned.

1.3 Segmentation

Image segmentation consists of the partitioning of an image into non-overlapping consistent regions that are homogeneous respect to some characteristics, such as intensity or texture[8]. The results of segmentation can be used to perform feature extraction, that provides fundamental information about organs or lesion volumes, to monitor the evolution of a particular disease and/or to evaluate the effects of therapeutical treatments etc... Therefore, segmentation plays a crucial role for clinitians in identifying diseases such as tumors. Segmentation, depending on the technique, can be manual, semi-manual or automatic:

Manual is still the most reliable and precise method but it is time-consuming, highly operator-dependent and subject to operator expertise[2].

Semi-Manual is a faster method compared to the manual one and it is based on the traditional image processing methods such as thresholding and clustering. However, despite the time savings it is operator-dependent[2].

Automatic is the faster method compared to the other ones and it is not operator-dependent. However, the implementation of the algorithms is harder to perform[2].

1.3.1 Methods

During the year several segmentation methods have been developed[8]. There are several ways to classify these methods. For example, depending if they require or not a training set of data, they can be classified into *supervised* or *unsupervised* methods. More, they can be classified depending on the information type they use, like *Pixel classification* methods, which use only information about pixel intensity, or *Boundary following* methods which use edge information etc...[8].
Among the most common ones:

Thresholding is a very simple and common approach to segmentation. This method is applied on the *histogram* of the image. The histogram of a digital image with intensity levels L in the range $[0,\ L-1]$, is a discrete function $h(l_k) = n_k$ where l_k is the k-th intensity value and n_k is the number of pixels with intensity l_k.
Thresholding consists in binarizing an image through an (if) clause on the intensity value of each point after having determined a threshold value $T \in [0,\ L-1]$. The threshold value T is usually chosen by visual assessment on the image histogram but it can be automatize by algorithms like the *Otsu algorithm.* One drawback of this method is that some parts of the image can belong to the same class even if they belong to different

objects. In fact, thresholding does not take into account the spatial characteristics of the image. Moreover, it is sensitive to noise and intensity inhomogeneity that corrupt the image histogram and make difficult the classification of pixels[8].

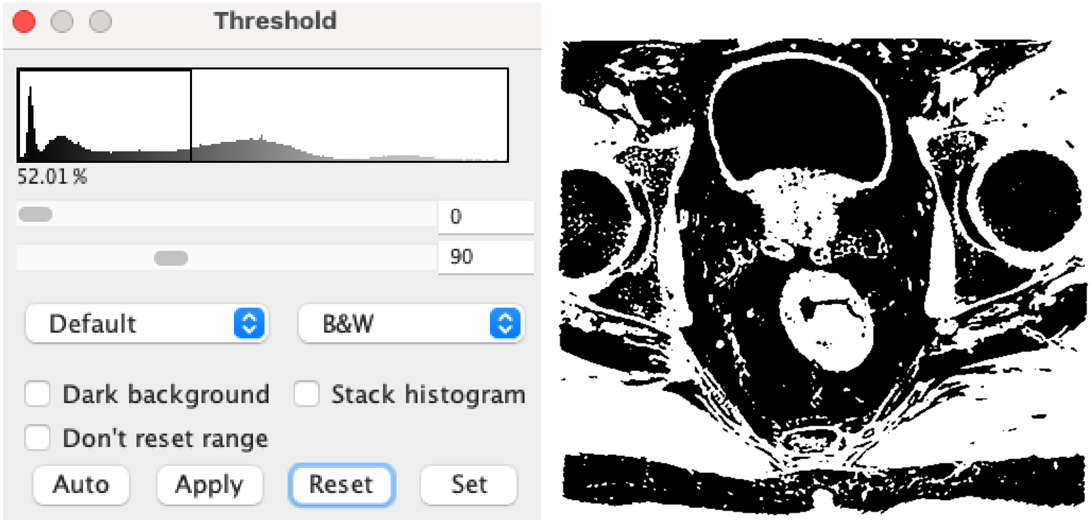

Figure 1.1: Example of thresholding segmentation using Fiji software[12].
Left) Image Histogram. *Right)* Result of thresholding.

Artificial Neural Networks are computational architectures derived from neural physiological models[13]. Artificial Neural Networks (ANNs) have evolved into a broad family of techniques. For visual analysis are usually used Convolutional Neural Networks (CNNs) based on *convolution kernels* or *filters* that slide along input data to extract feature maps[14]. Several architectures have been developed over the years, for different tasks and fields of application. In bio-medical image processing, the so-called U-Net[15], is one of the most common architecture. U-Net is a kind of CNN which allows overcoming the requirement of many training data[8, 15]. However a better explanation of ANNs will be provided in the following chapter.

1.4 Radiomics

Radiomics consists in methods that, using data-characterization algorithms, extracts from medical images, a large number of features which have the potential to uncover disease characteristics that fail to be appreciated by the naked eye[16]. The main objective of radiomics is to assist the subjective interpretation of the clinitians with an objective prediction of all those data invisible to the radiologist, transforming medical images into data, defined as *biomarkers*[2]. In the new era of precision medicine, radiomics is an emerging translational research field that aims to find associations between qualitative and quantitative information extracted from clinical images and clinical data to support decision making process.

1.4.1 Radiomic Features

Radiomic features can be divided into five groups[16, 2]:

- size and shape based–features like descriptors of the image intensity histogram, gray-level co-occurrence matrix (GLCM);
- run length matrix (RLM);
- size zone matrix (SZM);
- neighborhood gray tone difference matrix (NGTDM) derived textures, textures extracted from filtered images;
- fractal features.

1.4.2 Possible Purposes Of Radiomics

The possible applications of radiomics are based on a very wide range, from the prediction of clinical outcomes to the oncological diagnosis. In this subsection, a brief overview of some general possible purposes will be given.

Prediction of clinical outcomes: Radiomic features may be useful for predicting patient survival and describing intratumoral heterogeneity as demonstrated in a study by Aerts et al. [17]. More, the usefulness of radiomics for predicting the immunotherapy response of patients with non-small cell lung cancer (NSCLC) using pretreatment CT and PET-CT images has been demonstrated by other studies[2].

Prediction of metastases: Radiomic features can also predict the metastatic potential of tumors. For example, many radiomic features were identified as predictors of distant metastasis of lung adenocarcinoma in a study by Coroller et al.[18]. They concluded that radiomic features may be useful in identifying patients at high risk of developing distant metastases, guiding clinicians in choosing the most effective treatment for individual patients.

Genetic evaluation of cancer: The biological mechanisms of colorectal cancer were studied for the construction of different imaging models. In particular, It has been showed that radiomic features can be associated with some biological genes[2].

Prediction of physiological events: Another possible application of radiomics analysis is the prediction of physiological events. Indeed, radiomics can be applied for the characterization and investigation of complex physiological events such as brain activity, which is usually studied with specific imaging techniques such as functional magnetic resonance "fMRI"[2].

Chapter 2

Artificial Neural Networks

2.1 Convolutional Neural Networks

2.1.1 U-Net

Chapter 3

Pipeline

Chapter 4

Results

Conclusions

Bibliography

[1] International Agency for Research on Cancer. *Rectum-fact-sheet.* 2020. URL: https://gco.iarc.fr/today/data/factsheets/cancers/9-Rectum-fact-sheet.pdf.

[2] Jia Cheng Yuan. "Analisi radiomica per predire la risposta alla chemio-radioterapia neoadiuvante nel carcinoma del colon retto". Tesi di laurea. Alma Mater Studiorum Università di Bologna, 2019/2020.

[3] Jemal A Siegel RL Miller KD. "Cancer statistics, 2019". In: *CA Cancer J Clin* 69(1).7-34 (2019).

[4] Jovana Panic. "Colorectal cancer segmentation on MRI images using Convolutional Neural Networks". Tesi di Laurea Magistrale. Politecnico di torino, 2018/2019.

[5] Trebeschi et al. "Deep Learning for Fully-Automated Localization and Segmentation of Rectal Cancer on Multiparametric MR". In: *Scientific Reports* 7.1 (2017), p. 5301. DOI: 10.1038/s41598-017-05728-9. URL: https://doi.org/10.1038/s41598-017-05728-9.

[6] Rafael C. Gonzalez and Richard E. Woods. *Digital image processing.* 2018.

[7] Michele Larobina and Loredana Murino. "Medical image file formats". In: *Journal of digital imaging* 27.2 (Apr. 2014), pp. 200–206. DOI: 10.1007/s10278-013-9657-9. URL: https://pubmed.ncbi.nlm.nih.gov/24338090.

[8] Riccardo Biondi. "Implementation of an Authomated Pipeline for the Identification of Ground Glass Opacities in Chest CT Scans of Patient Affected by COVID-19". MA thesis. riccardo.biondi7@unibo.it: Alma Mater Studiorum - Università di Bologna, School of Science, Dec. 2020.

[9] *Processing in Spatial Domain – Spatial Filtering.* URL: https://ceng.eskisehir.edu.tr/serkangunal/BIM472/odev/BIM472%20-%2004%20-%20Processing%20in%20Spatial%20Domain%20(Part%202).pdf.

[10] Bhumika Gupta and Shailendra Singh Negi. "Image Denoising with Linear and Non-Linear Filters: A REVIEW". In: *International Journal of Computer Science Issues (IJCSI)* 10.6 (Nov. 2013), pp. 149–154. URL: https://www.proquest.com/scholarly-journals/image-denoising-with-linear-non-filters-review/docview/1498210097/se-2?accountid=9652.

[11] Carl-Fredrik Westin, Hans Knutsson, and Ron Kikinis. "Chapter 2 - Adaptive Image Filtering". In: *Handbook of Medical Image Processing and Analysis (Second Edition)*. Ed. by ISAAC N. BANKMAN. Second Edition. Burlington: Academic Press, 2009, pp. 19–33.

[12] Johannes Schindelin et al. "Fiji: an open-source platform for biological-image analysis". In: *Nature Methods* 9.7 (2012), pp. 676–682. DOI: 10.1038/nmeth.2019. URL: https://doi.org/10.1038/nmeth.2019.

[13] 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. eprint: https://doi.org/10.1146/annurev.bioeng.2.1.315. URL: https://doi.org/10.1146/annurev.bioeng.2.1.315.

[14] *convolutional neural network*. URL: https://en.wikipedia.org/wiki/Convolutional_neural_network#Image_recognition.

[15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation". In: *CoRR* abs/1505.04597 (2015). arXiv: 1505.04597. URL: http://arxiv.org/abs/1505.04597.

[16] *Radiomics*. URL: https://en.wikipedia.org/wiki/Radiomics.

[17] Michal R. Tomaszewski and Robert J. Gillies. "The Biological Meaning of Radiomic Features". In: *Radiology* 298.3 (2021). PMID: 33399513, pp. 505–516. DOI: 10.1148/radiol.2021202553. eprint: https://doi.org/10.1148/radiol.2021202553. URL: https://doi.org/10.1148/radiol.2021202553.

[18] Thibaud P Coroller et al. "CT-based radiomic signature predicts distant metastasis in lung adenocarcinoma." In: *Radiother Oncol* 114.3 (Mar. 2015), pp. 345–350.