

ESCOLA DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM GERONTOLOGIA BIOMÉDICA
MESTRADO EM GERONTOLOGIA BIOMÉDICA

CHRISTIAN MATTJIE DE OLIVEIRA

**EXPLORANDO A ABORDAGEM RADIÔMICA PARA A IDENTIFICAÇÃO DE
COVID-19 EM TOMOGRAFIA COMPUTADORIZADA PULMONAR**

Porto Alegre
2022

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul



Pontifícia Universidade Católica do Rio Grande do Sul
Pró-Reitoria de Graduação
Av. Ipiranga, 6681 - Prédio 1 - 3º. andar
Porto Alegre - RS - Brasil
Fone: (51) 3320-3500 - Fax: (51) 3339-1564
E-mail: prograd@pucrs.br
Site: www.pucrs.br

**PONTIFICAL CATHOLIC UNIVERSITY OF RIO GRANDE DO SUL
SCHOOL OF MEDICINE
BIOMEDICAL GERONTOLOGY POST GRADUATE PROGRAM**

EXPLORING THE RADIOMICS APPROACH FOR COVID-19 IDENTIFICATION IN LUNG COMPUTED TOMOGRAPHY

CHRISTIAN MATTJIE DE OLIVEIRA

Dissertation submitted to the Pontifical Catholic University of Rio Grande do Sul in partial fulfillment of the requirements for the degree of Master in Biomedical Gerontology.

Line of research in Biological Aspects of Aging.

Advisor: Prof. Ana Maria Marques da Silva

**Porto Alegre
2022**

Ficha Catalográfica

D278e de Oliveira, Christian Mattjie

Exploring the radiomics approach for COVID-19 identification in lung computed tomography / Christian Mattjie de Oliveira. – 2022.

84.

Dissertação (Mestrado) – Programa de Pós-Graduação em Gerontologia Biomédica, PUCRS.

Orientadora: Profa. Dra. Ana Maria Marques da Silva.

1. Tomografia Computadorizada. 2. COVID-19. 3. Radiômica. 4. Discretização. 5. Importância de Atributos. I. Marques da Silva, Ana Maria. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS
com os dados fornecidos pelo(a) autor(a).

Bibliotecária responsável: Clarissa Jesinska Selbach CRB-10/2051

CHRISTIAN MATTJIE DE OLIVEIRA

EXPLORING THE RADIOMICS APPROACH FOR COVID-19
IDENTIFICATION IN LUNG COMPUTED TOMOGRAPHY

Dissertation submitted to the Pontifical
Catholic University of Rio Grande do
Sul in partial fulfillment of the
requirements for the degree of Master in
Biomedical Gerontology.

Line of research in Biological Aspects of
Aging.

Aproved at: _____ of _____ of _____.

EXAMINATION BOARD

Prof. Dr. Daniel Rodrigo Marinowic

Prof. Dra. Ana Cláudia Patrocínio

PORTO ALEGRE

2022

ACKNOWLEDGMENTS

I would like to thank my family, friends, and partners for all the support they have given me during this period and throughout my life. I would also like to thank the professors of the Graduate Program in Biomedical Gerontology, especially my advisor Prof^a Ana Maria Marques da Silva, for their dedication through this pandemic since they did not measure efforts to teach and orient us.

My sincere thanks to the Coordination for the Improvement of Personnel (CAPES) for the full scholarship that allowed my exclusive dedication to this master's degree - Funding Code 001.

EXPLORANDO A ABORDAGEM RADIÔMICA PARA A IDENTIFICAÇÃO DE COVID-19 EM TOMOGRAFIA COMPUTADORIZADA PULMONAR

RESUMO

O surto de pneumonia de COVID-19 causou transtornos globais e foi declarado uma pandemia pela Organização Mundial da Saúde em 13 de março de 2020. Os exames radiológicos do tórax, como radiografias do tórax ou tomografias computadorizadas, têm um papel vital no diagnóstico da COVID-19. Vários estudos propuseram o uso de modelos de classificação utilizando características radiômicas extraídas dos pulmões em imagens radiológicas, principalmente para o diagnóstico e avaliação da gravidade da COVID-19. Entretanto, poucos desses estudos exploram como os parâmetros de extração de características, como a discretização, impactam as características extraídas. Portanto, este estudo visa implementar modelos para identificar a COVID-19 através da assinatura radiômica enquanto investiga diferentes parâmetros de pré-processamento e discretização. O conjunto de dados utilizado foi de 180 (128 COVID e 52 não COVID) tomografias de tórax realizadas no Hospital São Lucas da PUCRS que foram divididas em conjuntos de treinamento (50%), validação (25%) e teste (25%). Realizamos segmentação dos pulmões, aplicamos diversos filtros e discretizamos a imagem com 6 tamanhos diferentes de bin: 1, 5, 10, 25, 50, e 75. As características foram extraídas de todos os filtros aplicados e tamanhos de bin. Os atributos Wavelet e não-wavelet foram fundidos em 36 combinações de tamanhos de bin com 1774 atributos para cada pulmão. Um modelo de classificação foi treinado com cada combinação de características e os três melhores modelos foram escolhidos para a otimização. Identificamos algumas de nossas limitações e utilizamos quatro estratégias alternativas para tentar superá-las: SMOTE, subamostragem, seleção de atributos e somente utilizar atributos da imagem original. O melhor desempenho foi alcançado pelo modelo SMOTE NW25-1 com um AUC de 0,800. Os três melhores modelos para cada uma destas estratégias alternativas também foram otimizados. Dos 15 modelos otimizados, os seis melhores foram selecionados para análise da importância dos atributos. Os filtros laplaciano da gaussiana e wavelet foram os que geraram os atributos mais relevantes. Nossos resultados indicam que os tamanhos menores de bin, em uma faixa de 1 a 25, podem ser mais investigados para extração de características na imagem original e na maior parte dos filtros. Os filtros laplaciano da gaussiana e wavelet podem ter melhor desempenho com bins ainda menores, em uma faixa de 1 a 10.

Palavras-chave: Tomografia computadorizada, COVID-19, Radiômica, Discretização, Importância de atributos

EXPLORING THE RADIOMICS APPROACH FOR COVID-19 IDENTIFICATION IN LUNG COMPUTED TOMOGRAPHY

ABSTRACT

The COVID-19 pneumonia outbreak has caused global turmoil and was declared a pandemic by the World Health Organization on March 13, 2020. Chest radiological examinations, such as chest X-rays or CT scans, play a vital role in the diagnosis of COVID-19. Several studies have proposed the use of classification models using radiomic features extracted from the lungs in radiological images, mainly for COVID-19 diagnosis and severity assessment. However, few of these studies explore how feature extraction parameters, such as discretization, impact the extracted features. Therefore, this study aims to implement models for identifying COVID-19 through the radiomic signature while investigating different preprocessing and discretization parameters. Our dataset was composed by 180 (128 COVID and 52 non-COVID) chest CT scans performed at Hospital São Lucas da PUCRS which were divided into training (50%), validation (25%), and test (25%) sets. We performed lung segmentation, applied several filters, and discretized the image with 6 different bin sizes: 1, 5, 10, 25, 50, and 75. Features were extracted from all applied filters and bin sizes. Wavelet and non-wavelet features were merged into 36 combinations of bin sizes with 1774 features for each lung. A classification model was trained with each combination of features and the best three models were chosen for the optimization. We identified some of our limitations and used four alternative strategies to try to overcome them: SMOTE, undersampling, feature selection, and only using features from the original image. The best performance was achieved by SMOTE NW25-1 model with an AUC of 0.800. The best three models for each of these alternative strategies were also optimized. Of the 15 optimized models, the six best were selected for feature importance analysis. The laplacian of gaussian and wavelet filters were the ones that generated the most relevant features. Our results indicate that smaller bin sizes, in a range from 1 to 25 may be further investigated for feature extraction in the original image and most filters. Laplacian of gaussian and wavelet filters may perform better with even smaller bin sizes, with a range from 1 to 10.

Keywords: Computed tomography, COVID-19, Radiomics, Discretization, Feature importance

LIST OF ACRONYMS

AUC – Area Under the Curve

CT – Computed Tomography

COVID-19 – Corona Virus Disease 2019

FBS – Fixed Bin Size

FBN – Fixed Bin Number

GLCM – Gray Level Co-occurrence Matrix

GLDM – Gray Level Dependence Matrix

GLRLM – Gray Level Run Length Matrix

GLSZM – Gray Level Size Zone Matrix

LOG – Laplacian of Gaussian

MRI – Magnetic Resonance Imaging

PACS – Picture Archiving and Communication System

SMOTE – Synthetic Minority Over-sampling Technique

ROC – Receiver Operator Characteristic

CONTENTS

1	INTRODUCTION	10
2	BACKGROUND	12
2.1	COVID-19	12
2.1.1	Symptoms, transmission capacity and lethality	12
2.1.2	Vacines and variants	12
2.2	Computed Tomography	13
2.3	Radiological characteristics of COVID-19 on CT	14
2.4	Radiomics	15
2.5	Discretization	16
2.5.1	Discretization methods in radiomics review	17
2.6	State of the art on COVID-19 radiomics	17
3	OBJECTIVES	30
3.1	Main objective	30
3.2	Specific goals	30
4	METHOD	31
4.1	Study design	31
4.2	Sample	31
4.2.1	Inclusion and exclusion criteria	31
4.2.2	Lung CT Dataset	32
4.3	Image Processing	33
4.3.1	Segmentation	33
4.3.2	Filtering	33
4.3.3	Discretization	34
4.4	Feature extraction	35
4.5	Classification	36
4.6	Model Performance Evaluation	38
5	RESULTS AND DISCUSSION	39
5.1	Image Processing	39
5.2	Training and Optimization of Classification Models	40

5.3	Exploring the Limitations	42
5.3.1	Dealing with Imbalanced Dataset	42
5.3.2	Reducing the Number of Features	43
5.3.3	Performance on validation and test sets	45
5.4	Preprocessing Impact and Feature Importance	46
5.5	Discretization Impact	48
5.6	Limitations	50
6	CONCLUSION	51
	REFERENCES	52
	ANNEX A – Consubstantiated opinion of the research ethics committee	65
	ANNEX B – Abbreviations for features shown in feature importance	70
	ANNEX C – The impact of discretization in radiomics: what we know so far	72

1. INTRODUCTION

In December 2019, a group of patients with atypical pneumonia of unknown cause was associated with bat meat consumed at an exotic animal meat market in Wuhan, Hubei, China. A new type of beta coronavirus could be identified through unbiased sequencing of patient samples. The COVID-19 pneumonia outbreak, caused by the coronavirus strain SARS-Cov-2 (coronavirus severe acute respiratory syndrome 2), caused global turmoil and was declared a pandemic by the World Health Organization (WHO) on March 13, 2020. By December 23, 2021, more than 270 million cases had been confirmed worldwide.

The most used test to diagnose the COVID-19 is reverse transcription followed by polymerase chain reaction (RT-PCR). The test, however, has low specificity when performed before the onset of symptoms and in asymptomatic cases, which leads to a large number of false-negative results [Kucirka et al., 2020]. Therefore, as a way to help make a reliable diagnosis, several studies use the radiological features of chest X-ray images of the lungs of suspected COVID-19 individuals [Xie et al., 2020b].

Chest radiological examinations, such as chest X-ray or CT scans, play a vital role in the diagnosis of COVID-19 [Hu and Wang, 2020]. CT is the standard method for evaluating lung changes, even in the early stage of the disease when the patient has few or no symptoms [Rubin et al., 2020]. In addition, CT is used to monitor the disease evolution. From the images generated, it is possible to assess whether the lung lesions are mild or severe [Qiu et al., 2021], which is of great importance for the prognosis, severity analysis, and treatment of patients infected with COVID-19. Observation of the radiological lung patterns can reveal different types of pulmonary diseases. These patterns can be described based on the disease and the affected tissue region. Knowledge of the disease-related patterns is essential for the differentiation and monitoring of lung diseases. For example, studies have shown that COVID-19 induces abnormal pneumonia that leads to a bilateral, peripheral, opaque ground-glass pattern [Shi et al., 2020].

Computational methods to quantitatively analyze medical images have been used for decades to detect pathological patterns, aid diagnosis, and predict treatment outcomes. A new field called radiomics emerged in the last years, integrating medical image processing techniques with machine learning methods [Lambin et al., 2012] [Aerts et al., 2014]. Radiomics is a field that aims to extract a large number of quantitative features from medical images using image processing and data characterization algorithms. It has the potential to identify disease patterns that are difficult to locate or even invisible to the eye, mining information from multidimensional data [Kocak et al., 2019] [Idris and Hacking, 2017].

The computerized analysis is used to quantify specific features of an image numerically. Quantitative analysis of morphology, intensity, and texture features is helpful in diagnosis, and prognosis [Yoon et al., 2019]. The image features are used to classify between

malign and benign lesions or to identify pathological patterns and predict the severity of diseases. Several models for classification and prediction of pneumonia based on COVID-19 using binary or multi-class classification techniques have been developed to aid diagnosis using CT and X-ray images [Apostolopoulos et al., 2020] [Cardobi et al., 2021] [Lin et al., 2021] [Shi et al., 2022]. In addition, other studies have attempted to identify radiomic patterns and features related to COVID-19 [Chen et al., 2021a] [Huang et al., 2021b] [Chao et al., 2021]. However, there has not been much progress in investigating the impact of preprocessing methods in extracting features to identify the disease.

Discretization, also named quantization or rebinning, is a standard procedure in radiomic feature extraction and consists of grouping nearby intensity levels according to predetermined intervals. The process is necessary to reduce the influence of noise in the extraction process so that radiomic features can be calculated efficiently and practically.

There is no ideal universal method and/or discretization value for extracting imaging features. Several studies have already investigated the impact on feature extraction with different discretization methods and in different imaging modalities (CT, magnetic resonance imaging, and positron emission tomography) for different tasks, mainly in oncology [de Farias et al., 2021] [Wang et al., 2019] [Desseroit et al., 2016] [Kolinger et al., 2021] [Molina et al., 2016]. These studies reinforce the importance of defining the method according to the imaging modality and medical task. For example, radiomic analysis of cancer is usually restricted to the tumor region, while the analysis of lung diseases usually considers the whole organ. Due to differences in the region of interest (ROI) size and types of tissues present, discretization methods can have positive and negative effects on the analysis.

Many radiomic COVID-19 studies use a single method, and discretization value [Guiot et al., 2020] [Gülbay et al., 2021] [Huang et al., 2021b] which is essentially the software default procedure, or do not even mention this process. Our study investigated the impact of different discretization procedures on the extraction of radiomic features to identify COVID-19 patients using lung CT images.

2. BACKGROUND

2.1 COVID-19

2.1.1 Symptoms, transmission capacity and lethality

Among the main symptoms of COVID-19 are fever, dry cough, fatigue, and dyspnea followed by less common symptoms such as body aches, sore throat, diarrhea, conjunctivitis, headache, loss of taste and smell, rash on the skin, and discoloration of fingers or toes [World Health Organization, 2022a] [Tian et al., 2020]. Rarer but severe symptoms include shortness of breath, chest pain or pressure, and difficulty speaking and moving around.

The first symptoms of the disease take approximately 5 to 6 days to appear after contact with the virus; however, this period can be extended up to 14 days after infection [World Health Organization, 2022a]. COVID-19 is most transmissible when people show symptoms [Chen et al., 2020a]; however, there are several cases where individuals are asymptomatic.

The transmission capacity of SARS-Cov-2 is high. A study conducted in China indicated, through epidemiological investigations, that the average number of people infected per one individual carrying the virus is 1.81, even using some containment strategies [Liu et al., 2020]. This number can vary according to social isolation measures. In Brazil, a study showed that transmission rates in Brazilian metropolises were between 1.29 and 2.01 during the first 80 days since the first case in the country [Sousa et al., 2020]. Since the transmission also occurs among asymptomatic individuals, it is challenging to contain the virus [Liu et al., 2020].

A study carried out by the Emergency Response Epidemiology Team for Coronavirus Pneumonia, from the beginning of the pandemic, involving 72,314 patients infected with the virus, showed that 80.9% of cases are considered “mild,” not requiring hospitalization [Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020]. The remainder of those infected develop difficulty breathing and need hospital care. The lethality of the virus depends on epidemiological, social, and economic factors, but this study, with a significant sample, comes in at 2.3% [Novel Coronavirus Pneumonia Emergency Response Epidemiology Team, 2020]. As of February 13, 2022, the lethality of COVID-19 in Brazil was 2.3% [Dong et al., 2020] and 1.4% worldwide, and it has dropped considerably in the last year due to mass vaccination worldwide.

2.1.2 Vaccines and variants

Vaccines are essential to prevent and control infectious disease outbreaks. Due to the number of infected and dead caused by COVID-19, in May of 2020, the 73rd World

Health Assembly recognized the role of mass immunization as a global-public-health goal for preventing, containing, and stopping transmission of SARS-CoV-2 [World Health Organization, 2020]. The immunization process reduces the risks of contracting the disease by using the body's natural defenses to create resistance [World Health Organization, 2022c]. Therefore, an effective vaccine helps to protect the individuals in two ways; the first one is direct protection, where groups are vaccinated with the primary purpose of preventing the disease's harsher symptoms on the individuals vaccinated. The second way is indirect protection, where those in contact with vaccinated individuals, even if unvaccinated, have a smaller chance of contracting the virus, reducing overall transmission.

Unfortunately, being vaccinated does not entirely protect individuals and does not mean people can stop taking precaution measures. All viruses can change over time, including SARS-CoV-2. Most of those changes have little to no impact on their properties. Nonetheless, some changes significantly affect the virus, such as how easily it spreads, the associated disease severity, the performance of vaccines, etc. Today there are five variants considered of concern by the WHO: Alpha, Beta, Gamma, Delta, and Omicron [World Health Organization, 2022b]. The Omicron variant is the latest one, discovered in November 2020, and it spreads more quickly than the original virus and the Delta variant. It is expected that anyone contaminated with Omicron can spread the virus even when vaccinated or without presenting symptoms [Centers Disease Control Prevention, 2022].

Regardless of the total of 10,095,615,243 vaccine doses that have been administered until February 7, 2022, the COVID-19 pandemic is far from over. A growing anti-vaccine movement can cause significant losses for the global population. This happens because a set of unvaccinated subjects provides a receptacle full of opportunities for the virus to mutate and spread [Goldman, 2021]. In the face of this scenario, a variant can emerge that is resistant to the vaccine. Thus far, the known variants can still be controlled by the existent vaccines, but this is not an unchangeable reality. In fact, the Delta variant is exhibiting an increased frequency of infections among the vaccinated [Farinholt et al., 2021]. Therefore, studies regarding the diagnosis and prognosis of COVID-19 must continue to be explored and improved.

2.2 Computed Tomography

Computed Tomography (CT) is an imaging exam that uses ionizing radiation to produce internal images of the human body. The X-ray tube emits the radiation, which performs a rotational movement around the patient. The radiation beam is transformed as a function of its interaction with the patient's tissues. All projections are measured by a set of detectors diametrically opposite to the X-ray tube. Data from all detectors generate sinograms, which are used to reconstruct internal slices of the body [Goldman, 2007].

The resulting tomographic images are cross-sections that allow visualization of the internal structures of the patient's body, reducing the anatomical overlap compared to a

conventional X-ray image projection. Denser regions attenuate the incident X-ray beams to a greater extent and, as in other radiological techniques, are represented in the image as a lighter region. In contrast, darker regions represent low density tissues. Several parameters (slice thickness, field of view, voltage, current-time, etc.) are adjusted in the acquisition, influencing the X-ray beam and, consequently, the CT image [Mahesh, 2013].

The grayscale used on CT reconstructed images is the Hounsfield scale, where each pixel has a value in Hounsfield Units (HU) associated with the X-ray attenuation coefficients of the tissues. The HU scale varies from -1024 to 3071 (12 bits), and its primary reference is the attenuation coefficient of water, as shown in equation 2.1.

$$\text{HU Value} = 1000 \times \frac{\mu - \mu_{H_2O}}{\mu_{H_2O} - \mu_{air}} \quad (2.1)$$

Therefore, gray values in CT images have physical meaning, and specific values may be associated with specific tissues and organs [Little, 2015]. For example, -1000 HU is the value associated with air, while 0 is associated with water in all CT scanners. Other examples from tissues on the body include healthy lungs with HU values between -900 to -700, muscle with 50 HU, fat with -100, intravenous contrast with 300, and bone with HU values over 1000.

2.3 Radiological characteristics of COVID-19 on CT

CT images can show indications of lung diseases, such as COVID-19, even before the onset of symptoms [Li and Xia, 2020]. The main chest CT findings of patients with COVID-19 are ground-glass opacities, crazy-paving patterns, consolidations, and linear opacities [Fu et al., 2020a] [Ding et al., 2020] [Chung et al., 2020]. In most cases, the disease affects more than one lung lobe and evolves into bilateral pneumonia [Chung et al., 2020] [Fatima et al., 2020]. Lower lung lobes have shown to be the most affected [Fu et al., 2020a] [Ding et al., 2020].

[Fatima et al., 2020] study presents the radiological findings according to the stages of the disease 2.1:

- Ultra-early stage: Before the onset of symptoms, with negative laboratory tests. CT images show multiple bilateral and subpleural scattered ground-glass opacities.
- Early stage: 1-3 days after symptom onset. CT images show multiple bilateral ground-glass opacities. Irregular interlobular septa begin to develop.
- Rapidly progressing stage: 3-7 days after onset of symptoms. CT findings include subpleural posterior consolidations, scattered air bronchograms, and overlapping irregular septa.
- Consolidation stage: Second week after symptom onset. CT images show a decrease in the size and density of consolidations.




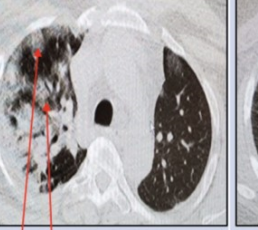

Ultra-Early Stage	Early Stage	Rapid progression Stage	Consolidation Stage	Dissipation Stage
				
CT scan demonstrates Bilateral, subpleural, multiple scattered ground glass opacities.	CT scan shows multiple, bilateral ground glass opacities. Irregular, interlobular septa begin to develop.	CT findings include subpleural, posterior consolidations, dispersed air bronchograms along with superimposed irregular septa.	There is a decrease in size and density of consolidations.	CT scan shows patchy consolidation, reticular opacities (strip-like opacities), bronchial and interlobular septal thickening.

Figure 2.1: COVID-19 radiological characteristics in chest CT for each disease stage. Adapted from Fatima et al. [2020].

- Dissipation stage: Usually occurs between 2 and 3 weeks after symptom onset. CT images show irregular consolidation, reticular (strip-like) opacities, bronchial and interlobular septal thickening.

2.4 Radiomics

Due to medicine digitalization, medical imaging data is now more available than ever. Together with the increased availability of artificial intelligence (AI) methods, the field of medical imaging is evolving exponentially. The new field of radiomics is one of the results of these developments [Tang, 2020].

Radiomics is a relatively new word, meaning the extraction of large numbers of quantitative features from medical images [Kocak et al., 2019]. This process effectively converts images into mineable data and rises in contrast to the traditional practice of only looking at medical images as pictures intended only for visual interpretation [Gillies et al., 2016]. Radiomic features can be related to intensity, shape, size, texture, among other information present in the images [Gillies et al., 2016]. Furthermore, radiomic features can be selected in advance (handcrafted features) or determined by the algorithm itself (deep learning features). Models have been proposed in recent years by analyzing these radiomic features to aid in diagnostic, prognostic, and predictive decision-making.

Radiomics is mostly used with CT, PET and MR, but can also be applied for other imaging modalities, including radiographs and ultrasound [Wilson and Devaraj, 2017] [Mu et al., 2020] [Zhang et al., 2017] [Chang et al., 2021] [Pang et al., 2021]. Due to the benefits of different modalities, it is possible to make an integrated cross-modality approach. This kind of approach enables the use of additional imaging information, resulting in more relevant features and, consequently, a powerful characterization [van Timmeren et al., 2020]. Radiomic features carry information regarding the pathology of interest, complementary to

other patient data, such as clinical and genomic ones. Consequently, when combined with pertinent data, radiomics can extract accurate and robust information, potentially offering and improving a clinical decision system [Lambin et al., 2017].

Although it is possible to use radiomics for several conditions, it is mostly used and developed in the field of oncology [Bera et al., 2021] [Liu et al., 2019] [Amini et al., 2021] [Pang et al., 2021]. It has quickly advanced toward clinical application enabling more precise diagnostics and treatment of patients. This rapid advance can be associated with the support offered from the National Cancer Institute, Quantitative Image Network, and other initiatives from the National Cancer Institute Cancer Imaging Program [Gillies et al., 2016].

The radiomics field combines multiple computational technologies, and the approaches used are usually focused on clinical tasks since radiomic data carry information regarding tumor biology. Recent technological advances have benefited clinical tasks such as cancer treatment and diagnosis. The application of radiomic studies in oncology is commonly diagnosis, which involves dividing samples according to disease or disease types, or predicting outcomes, which consists in separating the patients into risk groups [Shur et al., 2021]. Therefore, radiomics can even be used as a virtual biopsy without the need for invasive methods. Also, it can be helpful in disease monitoring since it offers information regarding disease progression.

2.5 Discretization

Discretization groups nearby pixel intensity values in the images according to pre-determined intervals (bins). There is no ideal and universal discretization method for extracting imaging features [van Timmeren et al., 2020] [van Griethuysen et al., 2017a]. This process is mostly performed in two ways.

The first is to define a fixed number of bins (Fixed Bin Number - FBN), where the width of the interval is defined as the range of image gray or intensity values divided by the number of defined intervals. The range is defined by each image's minimum and maximum intensity values or gray levels. For example, in our study, the range is defined as the difference between the minimum and maximum HU values of the lung in CT images. However, not all CT scans have the same maximum and minimum gray levels, especially after segmentation. In addition, different scans would have different bin widths, and the discretized values would not be directly related between scans. Thus, the FBN method would imply giving up the biological information carried with HU values for CT scans since each one would be discretized according to its value range.

The second most common form of image discretization is to determine a specific bin width value (Fixed bin size - FBS). With this method, the number of intervals is determined by the range of gray levels divided by the width of each bin. FBS gives direct control over the range of values represented in each bin [van Timmeren et al., 2020]. It is also guar-

anteed that the same image contrast will be used to extract all features. In CT images, where gray levels are defined in HU, and the values are absolute and directly correlate to organ and tissue's physical characteristics, the FBS method may achieve better performance.

Each imaging modality has its physical formation principle, and therefore, the image gray levels are specifically determined. In this sense, radiomics has particularities for each modality. Processing methods that perform well in one specific image modality will not necessarily perform well in others. Therefore, besides the discretization procedure, it is necessary to define the parameters of each method – the number of bins for the FBN and the bin width for the FBS - which also have different influences according to the imaging modality.

Radiomic algorithms and libraries recommend that the discretization method should be determined according to the imaging modality, body part examined, and task. Still, many radiomic studies use the default values defined by libraries (like FBS with bin width 25 - Pyradiomics library default).

2.5.1 Discretization methods in radiomics review

A review study was conducted to identify the discretization methods with the best performance for different imaging modalities and medical tasks in radiomic studies. Currently, there is not enough information to define optimal values for all modalities and tasks. Thus, the best-performing methods and values identified in the review could be used as a basis for choosing and investigating discretization methods in future works.

The manuscript was submitted to the IEEE Journal of Biomedical and Health Informatics on February 03, 2022. The journal has a JCR impact factor of 5.772 (2020). The manuscript can be found at the end of this dissertation, in Annex C.

2.6 State of the art on COVID-19 radiomics

The applications of radiomics in the diagnosis and prognosis of pulmonary diseases have been investigated in recent years. Motivated by the pandemic, studies have investigated the feasibility of using radiomic models for different COVID-related tasks. A literature review was conducted using the Pubmed and Scopus databases seeking to establish the state of the art in the use of CT-based radiomics in the study of COVID-19. The terms used in the search are described below.

Pubmed: (COVID-19) or (coronavirus)) AND ((computed tomography) or (CT)) AND (coronavirus) OR (COVID-19)) AND (radiomics OR radiomic AND features)

Scopus: TITLE-ABS-KEY ((covid-19) OR (coronavirus)) AND TITLE-ABS-KEY ((computed AND tomography) OR (ct)) AND TITLE-ABS-KEY (radiomics OR radiomic AND features)

We found 76 publications in Pubmed and 86 in Scopus (studies included in the databases until 07/21/2020). After disregarding duplicates, 98 studies were considered for this review. Of these, 32 papers were disregarded (18 through abstract and 14 through full-text review) because they did not meet the requirement of using chest CT radiomic features for COVID-19 diagnosis or prognosis related to medical tasks. Two papers were also disregarded because they were not available for access. Thus, 64 studies were reviewed to determine the state of the art in using radiomics applied to COVID-19 on chest CT.

Table 2.1 presents a summary of all the studies evaluated in this state of the art, with data regarding the purpose of the study, datasets, discretization methods, filters, features, and segmentation method used as well as the evaluation metrics and results for each of the studies.

Table 2.1: State of the art on chest CT radiomics for COVID-related tasks

Ref.	Aim	Sample	Disc. Method	Filters	Extracted Features	Extraction Software	Segmentation	Metrics	Uses CD?
Cardobi et al. [2021]	Diag.	115	FBS: 25	Original, Wavelet, LoG, Exponential, Logarithm, Gradient, LBP3D.	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Whole Lung	AUC: 0.83	No
Caruso et al. [2021]	Diag.	120	-	Original, LoG	1st order	TexRAD	-	Sensitivity: 60%; specificity: 80%	Yes
Chen et al. [2021b]	Diag.	134	FBS: 25	Original, Wavelet, LoG	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM	Pyradiomics	Lesion	AUC: 0.925; sensitivity: 0.816; specificity: 0.816	Yes
Chen et al. [2021c]	Diag.	185	-	Original	Shape, 1st order, GLCM, GLRLM,	IBEX	Lesion	AUC: 0.915	Yes
Pizzi et al. [2021]	Diag.	58	-	Original, Wavelet, LoG, Square, Squareroot, Logarithm, Exponential, Gradient, LBP2D, LBP3D	Shape, 1st order, GLCM, GLRLM, NGTDM, GLDM, GLSZM	Pyradiomics	Lesion	AUC: 0.868; sensitivity: 93%; specificity: 75%;	No
Di et al. [2021]	Diag.	3330 imgs	-	Original	1st order, GLCM	-	Lobe and Lesion	Accuracy: 0.8979	No
Fang et al. [2020a]	Diag.	75	-	Original	Shape, 1st order, GLCM, GLRLM	MATLAB in-house developed	Whole Lung	AUC: 0.826	No

Fang et al. [2020b]	Diag.	239	-	Original	Shape, 1st order, GLCM, GLSZM, GLDM, NGTDM, GLRLM	uAI-Discover NCP R001	-	AUC: 0.995	Yes
Feng et al. [2020]	Diag.	350	-	Original, Laplacian sharpening, Gaussian, Shot noise	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Lesion	Sensitivity: 0.941; specificity: 0.981; accuracy: 0.962	Yes
Guiot et al. [2020]	Diag.	2022	FBS: 25	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, NGTDM, GLDM	RadiomiX	Whole Lung	AUC: 0.882; accuracy: 85.18%; sensitivity: 69.52%; specificity: 91.63%	Yes
Gülbay et al. [2021]	Diag.	134	FBN: 64	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	-	Lesion	AUC: 0.907; accuracy: 83%; sensitivity: 79.5%; specificity: 85.7%	No
Huang et al. [2021b]	Diag.	154	FBS: 25	Original, LoG, LBP	Wavelet, Shape, 1st order, GLCM, GLRLM, GLSZM, NGTDM	Pyradiomics	Lesion	AUC: 0.959; sensitivity: 89.9%; specificity: 90.7%; accuracy: 90.3%	Yes
Kang et al. [2021]	Diag.	170	-	Original	Shape, 1st order, etc	-	Whole Lung and Lesion	Diag-accuracy: 91.2%; SA-accuracy: 95%	Yes
Lin et al. [2021]	Diag.	319	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM	Pyradiomics	Lesion	AUC: 0.911	No
Liu et al. [2021]	Diag.	146	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	-	Lesion	AUC: 0,98; sensitivity: 0,94; specificity: 0,93	Yes

Li et al. [2021b]	Diag.	550	FBS: 25	Wavelet, LoG	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM	Pyradiomics	Whole Lung	F-score - GGO: 93.84%; cord: 92.37%; solid: 95.47%; subsolid: 84.42%	Yes
Peng et al. [2022]	Diag.	145	FBS: 25	Original, Wavelet	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Lobe and Lesion	AUC: 0.977; sensitivity: 0.944; specificity: 0.870; accuracy: 0.915	Yes
Rezaei et al. [2021]	Diag.	278	FBN: 64	Original	Shape, 1st order, GLCM, GLRLM, GLDM	-	Whole Lung	AUC: 0.997	No
Santone et al. [2021]	Diag.	35	-	Original	1st order, GLDM, GLRLM, GLSZM	Pyradiomics	Whole Image	Accuracy: 0.83	No
Shi et al. [2021b]	Diag.	2685	FBN:30	Original	1st order, GLCM, GLRLM, GLSZM, NGTDM	-	Lesion	Sensitivity: 90.7%; speci- ficity: 87.2%; accuracy: 89.4%	No
Soleymani et al. [2021]	Diag.	106	FBN: 128	Original	GLCM, GLRLM	S-IBEX	Lesion	-	No
Stefano and Comelli [2021]	Diag.	199	-	Original	-	-	Parenchyma	DSC: 74.83 ± 11.18%; sensitivity: 76.50% ± 16.79%	No

Sun et al. [2020]	Diag.	2522	-	Original	1st order, GLRLM, NGTDM	GLCM, GLSZM,	-	Lesion	AUC: 96.35%; accuracy: 91.79%; sensitivity: 93.05%; specificity: 89.95%	No
Tabatabaei et al. [2021]	Diag.	66	FBS: 64	Original	Shape, 1st order, GLRLM, NGTDM	GLCM, GLSZM, GLDM,	Pyradiomics	Lesion	AUC: 0.97; sensitivity: 89%; precision: 90%; F1score: 89%; accuracy: 89%	No
Tan et al. [2020]	Diag.	219	-	Original, Wavelet, Square, Squareroot, Logarithm, Exponential, Gradient, LBP2D, LBP3D	Shape, 1st order, GLCM, GLRLM, GLSZM, NGTDM		Pyradiomics	Healthy Tissue	AUC: moderate&severe: 0.95; moderate&control: 0.98; severe&control: 0.95	Yes
Velichko et al. [2022]	Diag.	417	-	Original, Wavelet, Gabor, LBP3D	GLCM, Gabor, Laws texture, Lapped		-	Lesion	AUC: 0.994; accuracy: 0.98; sensitivity: 0.98; specificity: 0.98	No
Wang et al. [2020]	Diag.	266	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, NGTDM, GLDM		Pyradiomics	Whole Lung and Lesion	AUC: 0.87; sensitivity: 73,5%; specificity: 81,8%	No
Wang et al. [2021a]	Diag.	218	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, NGTDM, GLDM		Pyradiomics	Lesion	AUC: 0.81; sensitivity: 72.2%; specificity: 75.1%	Yes
Xie et al. [2021]	Diag.	301	-	Original, Wavelet	1st order, GLRLM, NGTDM	GLCM, GLSZM,	-	Lesion	AUC: 0.98	No

Xin et al. [2021]	Diag.	103	-	Original		Shape, 1st order, GLDM, GLSZM	Python	Lesion	AUC: 0.990	Yes
Yang et al. [2021]	Diag.	180	-	Original		Shape, 1st order, GLCM, GLRLM, GLSZM, NGTDM	LIFEx	Lesion	AUC: 0.940; accuracy: 89.83%; sensitivity: 94.22%; specificity: 85.44%	No
Zeng et al. [2020]	Diag.	78	-	Original		Shape, 1st order, GLCM, GLZLM, GLDM, GLRLM	LIFEx	Lesion	AUC: 0.87	Yes
Zhang et al. [2021c]	Diag.	386	-	Original		Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Whole Lung	COVID - AUC: 0.922; sensitivity: 0.879; specificity: 0.900	No
Zhao et al. [2021]	Diag.	112	-	Original		1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	-	Whole Lung	AUC: 0.9470; sensitivity: 0.9670; specificity: 0.9270	No
Zhu et al. [2021]	Diag.	2522	-	Original		1st order, GLCM, GLRLM, GLSZM, NGTDM	-	Lesion	AUC: 91.32; accuracy: 91.31; sensitivity: 91.62; specificity: 91.01	No
Cai et al. [2022]	SA	520	-	Original		Shape, 1st order,	-	Lesion	AUC of each stage: 0.965; 0.958; 0.998; 0.975	Yes
Huang et al. [2021a]	Other	162	-	Original, LoG	Wavelet,	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Lung Kit	Whole Lung and Lesion	AUC: 0.875 - 0.837; sensitivity: 0.920 - 0.680; specificity: 0.826 - 0.913	No

Wang et al. [2021c]	Other	50	FBN: 8, 16, 32, 64, 128, 256	Original	1st order, GLRLM, NGTDM	GLCM, GLSZM, -	Lesion	AUC: 0.773; accuracy: 0.726	Yes	
Zhang et al. [2021a]	Other	107	-	Original, LoG, SquareRoot, Exponential; Logarithm	Wavelet, Square, Exponential; Logarithm	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Lesion	AUC: 0.857; sensitivity: 87.5%; specificity: 70.7%	No
Arru et al. [2021]	OP	221	-	Original	Shape, 1st order, GLCM, GLRLM, NGTDM, GLDM	-	Whole Lung and Lesion	SA-AUC: 0.88; ICU admission - AUC: 0.82;	No	
Bartolucci et al. [2021]	OP	115	-	Original	1st order, GLDM, GLSZM, GLCM, GLRLM, NGTDM	3DSlicer	Lesion	AUC: 0.82	Yes	
Cai et al. [2020]	OP	203	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, NGTDM, GLDM	Lung intelligence Kit	Lobe and Lesion	AUC: 0.812; sensitivity: 0,625; specificity: 0,600	Yes	
Chao et al. [2021]	OP	295	-	Original, LoG, Squareroot, Exponential	Wavelet, Square, Logarithm, Exponential	Shape, 1st order, GLCM, GLDM, GLRLM, GLSZM, NGTDM	Pyradiomics	Lung Lobes	AUC: 0.884; sensitivity: 96.1%	Yes
Chen et al. [2021a]	OP	40	-	Original, Logarithm, Squareroot, Exponential	Wavelet, Square, Exponential	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM	-	Lesion	AUC: 0.88; C-index: 0.85	Yes

Fu et al. [2020b]	OP	64	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM,	Quantitative Analysis Kit	Whole Lung	AUC: 0.833; sensitivity: 80.95%; specificity: 74.42%,	Yes
Ke et al. [2021]	OP	96	-	Original, Wavelet	Shape, 1st order, "texture"	Lung Intelligence Kit	Lesion	C-index: 0.907	Yes
Shiri et al. [2021]	OP	152	FBN: 64	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Whole Lung and Lesion	AUC: 0.95; accuracy: 0.88; sensitivity: 0.88; specificity: 0.89	Yes
Wang et al. [2021b]	OP	400	-	Original, LoG	Wavelet, 1st order	Pyradiomics	Whole Lung	AUC: 0.876	Yes
Wu et al. [2020]	OP	429	FBS: 25	Original	Shape. 1st order, GLCM, GLRLM, NGTDM	Pyradiomics	Whole Lung	AUC of early and late stages: 0.862; 0,976	Yes
Yue et al. [2020]	OP	31	-	Original, Wavelet	Shape, 1st order, "texture"	Pyradiomics	Lobe and Lesion	AUC: 0.97; sensitivity: 1.0; specificity: 0.89	Yes
Berta et al. [2021]	SA	60	FBS: 3, 5 e 10	Original	1st order	-	Lobe and Lesion	WAVE.f: 84%; WAVE.th: 75%;	No
Chen et al. [2020b]	SA	86	-	Original	Shape, 1st order, GLRLM, GLDM, GLSZM, GLCM, NGTDM	Pyradiomics	Lesion	AUC: 0.91; sensitivity: 0.82; specificity: 0.86	Yes
Li et al. [2020]	SA	217	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM	Pyradiomics	Whole Lung	AUC: 0.861	Yes
Li et al. [2021a]	SA	316	-	Original, Wavelet	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	-	Lesion	AUC: 0.696	Yes

Purkayastha et al. [2021]	SA	981	FBN: 8, 16, 32, 64	Original	Shape, 1st order, GLCM, GLRLM, GLSZM	-	Lesion	AUC: 3days: 0.897; 5days: 0.933; 7days: 0.927	Yes
Qiu et al. [2021]	SA	1160	le- sions	Original	1st Order, GLCM, GLRLM, LoG, contourlet, ACM, absolute gradient, autoregression, GLDM	-	Lesion	AUC:0.87	No
Shi et al. [2022]	SA	260	-	Original	-	-	Lesion	AUC: 0.978	Yes
Tang et al. [2021]	SA	118	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, NGTDM, GLDM	Pyradiomics	Whole Lung and Lesion	AUC: 0,98; acuracy: 0.89; TPR:0.910; TNR:0.858	Yes
Wei et al. [2020]	SA	81	-	Original	1st order, GLCM, GLSZM, GLRLM	Pyradiomics	Lung Lobes	AUC: 0.95; accuracy: 0.91; sensitivity: 0.81; specificity: 0.95	Yes
Xie et al. [2020a]	SA	150	FBN: 128	Original	Shape, 1st order, GLCM, GLRLM, GLSZM	Artificial intelligence Kit	Lesion	AUC: 0.905; accuracy: 89.5 %; sensitivity: 83.3 %; specificity: 90.0 %	Yes
Xiong et al. [2021]	SA	219	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Whole Lung	AUC of each stage: 0.619; 0.519; 0.478; CTscore: 0.548	No
Xu et al. [2021a]	SA	3024	-	Original, Wavelet, LoG, Square, Logarithm, Gradient; LBP3D	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Lesion	AUC: 0.919	Yes

Xu et al. [2021b]	SA	284	-	Original, Wavelet, Logarithm, Exponential, Gradient, Square, Squareroot, LBP2D, LBP3D	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM, NGTDM	Pyradiomics	Lesion	AUC of each stage: 0.97; 0.86; 0.83; 0.89	No
Zhang et al. [2021b]	SA	509	-	Original	Shape, 1st order, GLCM, GLRLM, GLSZM, GLDM	Pyradiomics	Lesion	AUC: 0.75	Yes

Diag., OP and SA refer to diagnosis, outcome prediction and severity assessment, respectively. LoG refer to Laplacian of Gaussian filter.

Our review shows the wide variety of dataset sizes in the studies. The studies use from 31 [Yue et al., 2020] up to 3024 [Xu et al., 2021a] patients' data. Most of the studies' datasets are in between 100 and 300 patients (31 studies), 13 studies have less than 100 patients in their dataset, and 20 studies have more than 300 patients. It is acknowledged that different discretization methods and values can affect the image features extraction. Despite that, it was observed that from the 64 studies, only 17 discuss the discretization method, and only 3 of them explore more than one discretization parameter [Purkayastha et al., 2021], [Berta et al., 2021], [Wang et al., 2021c]. These three studies used only one discretization method with different parameters. They improved their performance on the classification models using all the extracted features, regardless of the discretization parameters used. Therefore, they didn't recommend specific methods or parameters for discretization.

Regarding the segmentation step in the studies, only one of the studies did not use it [Santone et al., 2021] and one didn't even mention this step [Caruso et al., 2021]. Computer-aided methods using non-segmented images may lead to biases [Rizzo et al., 2018] [Shi et al., 2021a], since the model may associate elements from outside the lungs, which are not related to the disease. Limiting the region of interest guarantees that the extracted features come from the radiological information present in the lungs. In Maguolo study [Maguolo and Nanni, 2021], they put black squares on the middle of chest radiographs (completely erasing the lungs) from four different datasets, then trained a model to differentiate the original dataset of each sample. Surprisingly, the model could determine the original dataset with an AUC of over 0.9 with squares of different sizes. This shows that COVID-19 datasets need to be standardized since using multiple of the available datasets may lead models to differentiate inner characteristics of each one instead of differentiating COVID-19 from other pathologies. The study also reinforces the importance of segmentation, as removing the information from outside the lungs guarantees that the model will not take it into account.

Different methods of segmentation were used in the studies. Thirty studies segmented only the lesion region, 12 studies segmented the whole lung, and 3 studies used lobe segmentation. Another 13 studies did lesion segmentation and lobe or lung segmentation as well. Some studies used parenchyma segmentation and one only segmented the healthy tissue of the lungs [Tan et al., 2020].

Regarding image preprocessing or filtering in radiomics, 40 studies only used the original image. The most common preprocessing used were wavelets and Laplacian filters. One or both are used in 20 studies. Regarding the software or libraries used for feature extraction, 16 of the 64 studies did not specify the software or library. Besides, 28 studies used the python library Pyradiomics, and the other 20 studies used different software, including in-house developed ones. Furthermore, two studies did not specify what radiomic features were used. One study used only first-order features, and the other 61 articles also used GLCM and other descriptors. Regarding the metrics used for model's performance, all the

studies presented AUC, specificity, and sensibility higher than 0.60. Most of them achieved a result between 0.80 and 0.95. The articles found were subdivided into four categories: those whose purpose was the diagnosis of COVID-19, those that did a severity assessment, the ones that aimed at making an outcome prediction, and others who did not fit any of these categories. The vast majority, 55%, is regarding diagnosis. 23% are related to severity assessment and 17% to study outcome prediction. Of the three studies that did not fit these categories, one studies the prediction of re-positive cases [Wang et al., 2021c], one is regarding rapid progression prediction [Zhang et al., 2021a]. The last one studies residual lung lesion prediction [Huang et al., 2021a].

Regarding the inclusion of clinical data in the radiomics, 59% of the studies use clinical factors in their analysis. Approximately 90% of the studies in the outcome category, 10 out of the 11, use clinical data. Regarding severity assessment, the results are similar. Of the 15 studies, only 4 of them do not use clinical characteristics in their evaluation [Berta et al., 2021], [Qiu et al., 2021], [Xiong et al., 2021] and [Xu et al., 2021b]. The combination with clinical characteristics was expected, considering that the severity and the outcome can be reinforced with clinical variables. Of the 35 studies regarding COVID-19 diagnosis, 16 of them used clinical data in their evaluation, and, from the 'other' category [Wang et al., 2021c], and [Cai et al., 2022] included clinical data in their analyses.

3. OBJECTIVES

3.1 Main objective

Implement models for identifying COVID-19 disease through the radiomic signature of CT lung images investigating different preprocessing and discretization parameters.

3.2 Specific goals

- Evaluate the performance of COVID identification models on lung CT images and explore their limitations.
- Evaluate the importance of the radiomic features on the classification models.
- Investigate the impact of preprocessing methods on the performance of the classification models.
- Investigate the impact of discretization methods on the performance of the classification models.

4. METHOD

4.1 Study design

A retrospective, cross-sectional study was carried out at the Hospital São Lucas (HSL) of the Pontifical Catholic University of Rio Grande do Sul (PUCRS), located in Porto Alegre. The study was approved by the PUCRS Research Ethics Committee (CEP) under protocol CAAE 30791720.5.0000.5336, report number 4.850.213. The consubstantiated decision of the CEP can be found in Annex A.

4.2 Sample

The data of this study were collected from patients who underwent lung CT scans at HSL between December 16, 2020, and February 08, 2021. 442 CT scans were collected from the Picture Archiving and Communication System (PACS). The test used to distinguish between positive and negative COVID-19 patients was the Reverse Transcription-Polymerase Chain Reaction (RT-PCR) test. The following criteria of inclusion and exclusion in the groups are described below.

4.2.1 Inclusion and exclusion criteria

Patients were included in the COVID group if there was 1 (one) week or less between the positive RT-PCR and the CT scan, with no other negative RT-PCR results within 10 (ten) days of the CT scan date. Patients were included in the non-COVID group if there were 3 days or less between the negative RT-PCR result and the CT scan, with no other positive RT-PCR results within 60 days of the scan date. From the 442 CT exams collected, 233 were discarded because they had no RT-PCR data in the lab system.

Initially, the COVID group was formed with 151 CT scans. 11 CT scans were excluded for having more than 7 days between the scan and the RT-PCR, 7 were excluded as they were from patients with another CT scan already included (only the CT scan closest to the RT-PCR was considered), one was excluded for being incomplete, and 4 were excluded for having RT-PCR with a negative result with less than one week of the examination.

Initially, the non-COVID group had 58 CT scans. 3 CT scans were excluded for having more than 3 days of the negative RT-PCR, one was excluded for having a metallic artifact, and 2 were excluded as they were from patients with a CT scan already included. As for the COVID group, only the CT scan closest to the RT-PCR was considered.

Considering both inclusion and exclusion criteria, at the end the COVID group was composed of 128 CT scans, and the non-COVID group was formed of 52 CT scans. The

complete criteria for each group are shown in figure 4.1. In addition, the patients' age and sex distribution of both groups are shown in 4.1.

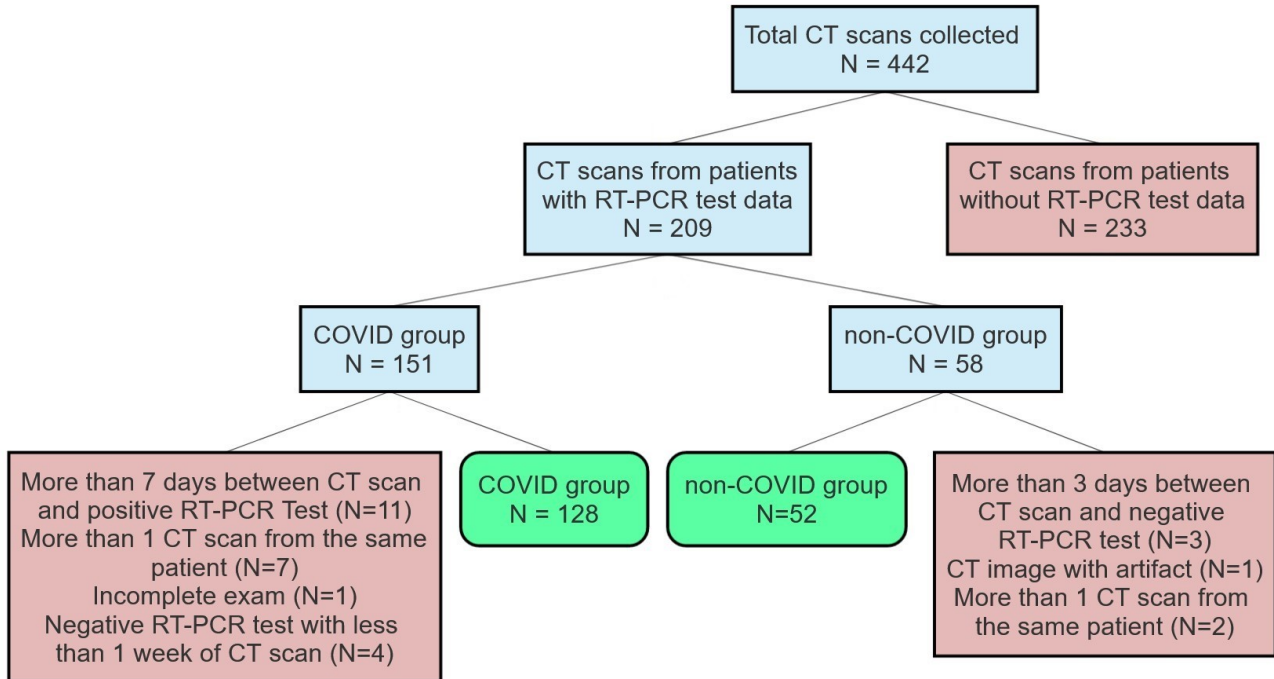


Figure 4.1: COVID and non-COVID group formation and criteria

4.2.2 Lung CT Dataset

A total of 180 lung CT scans were used in this study, one from each patient. All scans were acquired in a Siemens CT, model Emotion 16 from the year 2007. The acquisition protocol was the same for all patients, with 110 kV tube voltage, current modulation and 2 mm slice thickness. The *Hounsfield* values are in the range of -1024 to 3071 (12 bits). In case of multiple CT-scans for the same patient, only the closest to the corresponding RT-PCR test was used. We divided the dataset randomly and stratified it into training (50%), validation (25%), and test (25%) sets. No augmented data were used. The classification model only sees the images from the test set at the last performance evaluation.

Table 4.1: Demographic data of the patients in each group

	COVID group	non-COVID group
Age (years)	64 ± 16	60 ± 20
Sex (male)	59	22
Sex (female)	69	30
Total	128	52

4.3 Image Processing

The following subsections describe the image processing pipeline applied to the lung CT images before the feature extraction.

4.3.1 Segmentation

To extract the radiomic features only from the organ of interest, we performed segmentation of the lungs using the U-net R231CovidWeb model. The model is identical to the previous U-net R-231 [Hofmanninger et al., 2020], but was trained with additional images from patients with COVID-19 pneumonia. The segmentation model has already been used in other COVID-19 studies [Paluru et al., 2021] [Heidarian et al., 2021] [de Moura et al., 2022]. The left and right lungs were individually segmented to further extract features according to the laterality.

4.3.2 Filtering

The Pyradiomics library contains built-in filters that can be applied after the discretization process. The idea behind using these filters is to highlight possible radiomic patterns, which can help in the classification differentiating between COVID and non-COVID groups. Features were extracted from the original images and also using each of the filters embedded in the Pyradiomics library. The filters applied in this work are listed below, along with the specific parameters.

- Laplacian of Gaussian (LoG): applies a 3D Laplacian of Gaussian filter to the input image and yields a derived image for each sigma value specified. A Laplacian of Gaussian image is obtained by convolving the image with the second derivative (Laplacian) of a Gaussian kernel. The Gaussian kernel is used to smooth the image and is defined as [Johnson et al., 2013]:

$$G(x, y, z, \sigma) = \frac{1}{(\sigma\sqrt{2\pi})^3} e^{-\frac{x^2+y^2+z^2}{2\sigma^2}} \quad (4.1)$$

The Gaussian kernel is convolved by the laplacian kernel $\nabla^2 G(x, y, z)$, which is sensitive to areas with rapidly changing intensities, enhancing edges. The width of the filter in the Gaussian kernel is determined by σ and can be used to emphasize more fine (low σ values) or coarse (high σ values) textures. σ values used were 1, 3 and 5 and, therefore, three different LoG filtered images were generated.

- Square: calculates the square of each of the image gray intensities, and then rescales them linearly to the initial range [van Griethuysen et al., 2017b]. After applying the filter, negative values in the original image will be made negative again.

- Square Root: calculates the square root of each of the image gray intensities, and then rescales them linearly to the initial range [van Griethuysen et al., 2017b]. After applying the filter, negative values in the original image will be made negative again.
- Logarithm: calculates the logarithm of each of the image gray intensities, adds 1, and then rescales them linearly to the initial range [van Griethuysen et al., 2017b]. After applying the filter, negative values in the original image will be made negative again.
- Exponential: calculates the exponential of each of the image gray intensities, where the filtered intensity is $e^{\text{original gray intensity}}$, and then rescales them linearly to the initial range [van Griethuysen et al., 2017b]. After applying the filter, negative values in the original image will be made negative again.
- Gradient: Computes the gradient of an image using directional derivatives. Convolutions with first-order derivative operators are used to compute the directional derivative at each pixel position. The derivative operators are a 1×3 and 3×1 matrix with the values -1, 0 and 1. The filtered intensity is equal to the square root of the sum of the squares of the convolutions, and is computed as [Johnson et al., 2013]:

$$G(x, y) = \sqrt{[f(x + 1, y) - f(x - 1, y)]^2 + [f(x, y + 1) - f(x, y - 1)]^2} \quad (4.2)$$

Where $f(x - 1, y)$, $f(x + 1, y)$, $f(x, y - 1)$, and $f(x, y + 1)$ refer to the gray intensities of the pixels right before and after the computed one in the x and y axis.

- LBP3D: Using spherical harmonics, compute and return the Local Binary Pattern (LBP) in 3D as described by Banerjee et al. [2013]. The icosphere radius and subdivision were set to 1, and 2 levels were used, generating two different LBP3D filtered images.
- Wavelet: Computes the wavelet transforms of the image. We use Coiflets as they are Pyradiomics default and were already investigated in many radiomic studies, including COVID-19 ones [Xu et al., 2021a] [Xu et al., 2021b] [Chao et al., 2021]. We only use one level of wavelet; therefore, eight different images are generated according to the Low (L) and High (H) pass filters in each direction (x, y, z): LLL, HLL, LHL, LLH, HHL, HLH, LHH, HHH.

4.3.3 Discretization

This study deals with CT images, whose HU values carry biological information regarding different tissues of the human body. Therefore, the Fixed Bin Size (FBS) discretization method was chosen to not lose the biological information in the discretization process instead of FBN. Using smaller bin widths leads to less biological information being lost in

the discretization process and means dealing with more noise. Larger bin widths can be used to homogenize the image by reducing or removing the influence of noise, but with the possibility of erasing relevant information from the radiomic signature. For example, in CT images, the use of bin size equal to 1 HU effectively means not performing any discretization process.

We investigated 6 different bin widths to find the value that highlights the differences in the radiomic signature of COVID and no-COVID lung CT images.

The value of 25 is the default bin size in the Pyradiomics library, except for wavelet-based features, which the default value is 10. We evaluated the classification model performance using bin sizes of 1, 5, 10, 25, 50, and 75 for both wavelets and non-wavelets features.

4.4 Feature extraction

The Pyradiomics library (version 3.0.1) was used to extract first and second-order features from each of the lungs of the patients in the COVID and non-COVID groups [van Griethuysen et al., 2017b]. We extracted features from six different classes listed below:

- First order (18 features): These features describe the intensity distribution of the voxels using simple and commonly used metrics. First, a pixel value shift of 1024 was applied, adding 1024 to the gray values to calculate the Energy, Total Energy, and Average Square Root features. This process prevents negative values from calculating these features and is recommended in the Pyradiomics library documentation [van Griethuysen et al., 2017b].
- 3D Shape (14 attributes): This group of features describes the size and shape of the lung by analyzing the mask generated in the segmentation process. These attributes are independent of the distribution of gray levels and therefore are not calculated for any discretization methods and applied filters.
- Gray Level Co-occurrence Matrix – GLCM (24 features): This matrix contains information regarding the gray-level spatial distribution, considering the relationship between pixel pairs and the frequency of each intensity within a 26-connected neighborhood (for 3D images).
- Gray Level Run Length Matrix – GLRLM (16 features): This matrix quantifies gray level runs, which are the length of consecutive pixels that have the same gray level intensity.
- Gray Level Size Zone Matrix - GLSZM (16 features): This matrix quantifies gray level zones in the image. These zones are defined as the number of connected voxels with the same gray level intensity. Voxels are considered connected if they are within a 26-connected neighborhood.

- Gray Level Dependence Matrix – GLDM (14 features): This matrix quantifies gray level dependencies in the image. Gray level dependency is defined as the number of connected voxels where the difference in pixel intensity is less than α . We used the default α value of 1.

Considering the original image and all filters, 1070 non-wavelet features and 704 wavelet features were extracted for each bin size in each lung.

4.5 Classification

Prior to classification, min-max normalization was applied to each feature. This process scales and translates each feature individually such that it is in the given range on the training set. The minimum value of each feature is set to 0, while the maximum is set to 1, and intermediate values are proportionally scaled. This step is needed to ensure the model will not give higher importance to features only because they have higher values or higher ranges between their minimum and maximum values.

We performed all classifications using the optimized distributed gradient boosting library XGBoost (Version 0.9) [Chen and Guestrin, 2016] in Google Colaboratory [Bisong, 2019]. We performed the classification using the training and validation sets evenly stratified into a 5-fold cross-validation. We ran this process 10 times and used the mean F1-score to select the best models for hyperparameter optimization, as shown in figure 4.2.

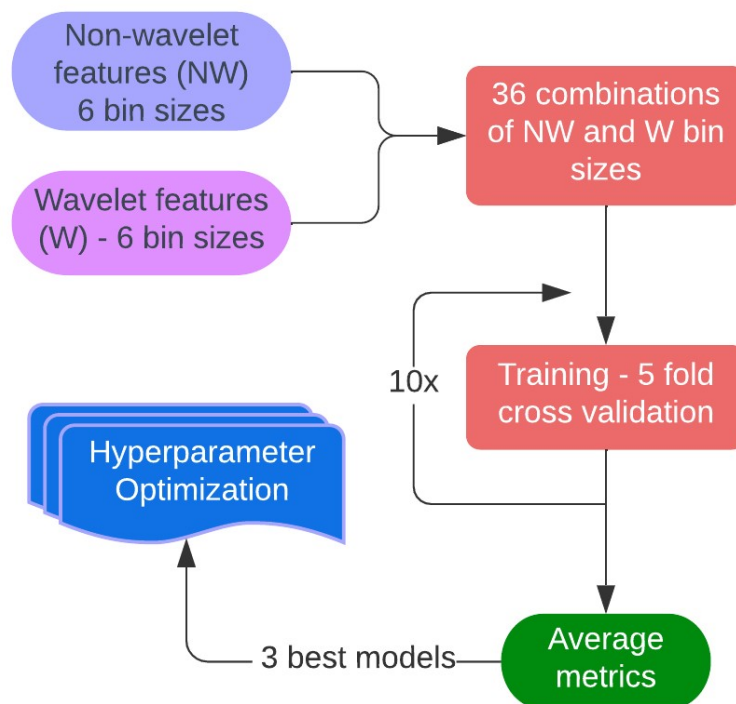


Figure 4.2: Workflow of the model selection process.

Feature extraction with the FBS discretization method is commonly performed with a single bin size and a specific bin size for the wavelet features. Even though using all

our data (6 bin sizes) for classification could have a better performance, it wouldn't be a realistic or reproducible method for differentiating COVID-19. Besides, it would require high computational power. Instead, we unified the features into 36 combinations of non-wavelet features (NW) and wavelet features (W) (6 sizes for non-wavelet and wavelet features each). With both lungs, there were a total of 3548 features included in each combination.

We optimized the hyperparameters using the Hyperopt library (Version 0.1.2) developed by Bergstra et al. [2015]. This method is based on an objective function, similar to deep learning models. In our 2 labels (COVID and non-COVID) classification problem, the selected objective function is binary logistic, and the booster is gbtrees. The hyperparameters are then changed in a guided way, aiming to minimize the loss function. This process is a lot faster than the traditional grid-search while achieving a lot better results than random search [Putatunda and Rama, 2018].

The optimization happens as follows: We train the model again in a 5-fold cross-validation using only the training data. Then, the model's Area Under the Curve (AUC) of the receiver operating characteristic (ROC) is evaluated in the validation set, and the hyperparameters are altered according to the Hyperopt optimization. After that, the optimization procedure starts again using the new values for the hyperparameters. This process is repeated 3000 times and the model with the highest AUC is saved for further evaluation, as shown in figure 4.3.

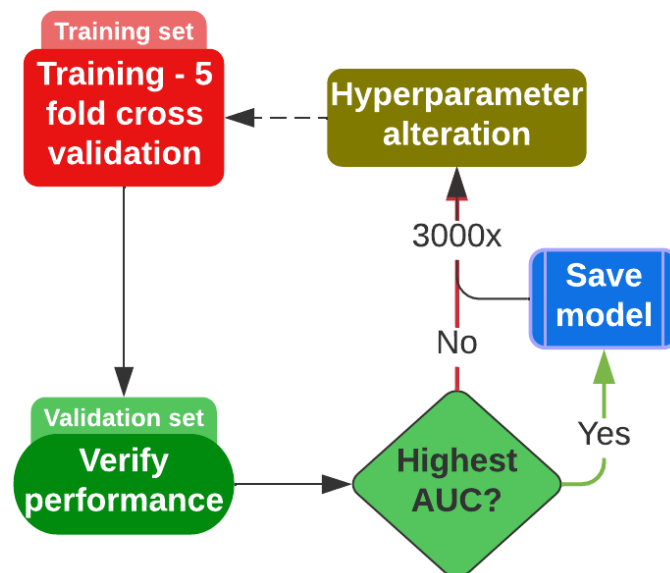


Figure 4.3: Workflow of the hyperparameters optimization process.

Optimization parameters were set as:

- Number of Hyperopt probes (`n_hyperopt_probes`) = 3000
- Hyperopt algorithm (`hyperopt_algo`) = `tpe.suggest`
- Number of boosting rounds (`num_boost_round`) = 250

- Holdout size (holdout_size) = 0.10

The optimized hyperparameters were:

- max_depth
- alpha
- lambda
- reg_lambda
- min_child_weight
- learning_rate
- colsample_bytree
- num_leaves
- n_estimators.

Finally, we evaluate the best models' performance on the test dataset and identify the relevant features for classification and their importance using the normalized gain. The gain implies the relative contribution of each feature to the model and is calculated by taking each feature's contribution for each tree in the model [Chen and Guestrin, 2016]. A higher value implies it is more important for generating a prediction. The normalized gain is the gain of each feature divided by the sum of the gain of all features. Features that weren't used have a gain equal to 0. Therefore, normalized gain shows how much each feature contributed to the classification relative to the total contributions of all features.

4.6 Model Performance Evaluation

We use the following metrics for model performance evaluation: accuracy, sensitivity(recall), precision, F1-score, and AUC of the ROC. Metrics are calculated as follows [Powers, 2020]:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4.3)$$

$$\text{Recall} = \text{Sensitivity} = \frac{TP}{TP + FN} \quad (4.4)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (4.5)$$

$$F1 \text{ Score} = 2 \times \frac{\text{Precision} \times \text{Sensitivity}}{\text{Precision} + \text{Sensitivity}} \quad (4.6)$$

Where TP, FP, TN, and FN refers to true positive, false positive, true negative, and false negative, respectively.

5. RESULTS AND DISCUSSION

5.1 Image Processing

This study used automatic segmentation of the lung CT images, which were independently analyzed. The segmentation was performed by the U-net R231CovidWeb model. Filters were applied to further enhance the differences between the radiomic signature of the COVID and non-COVID groups. Features were extracted from the original image, and all filter applied images. Filters used were exponential, gradient, Laplacian of Gaussian (LoG), logarithm, square, square root, wavelet, and local binary pattern 3D (LBP3D). All images used in this study were carefully assessed to avoid biases, like the ones shown in Maguolo and Nanni [2021] study. Figures 5.1 and 5.2 shows the processed images before feature extraction for all filters used.

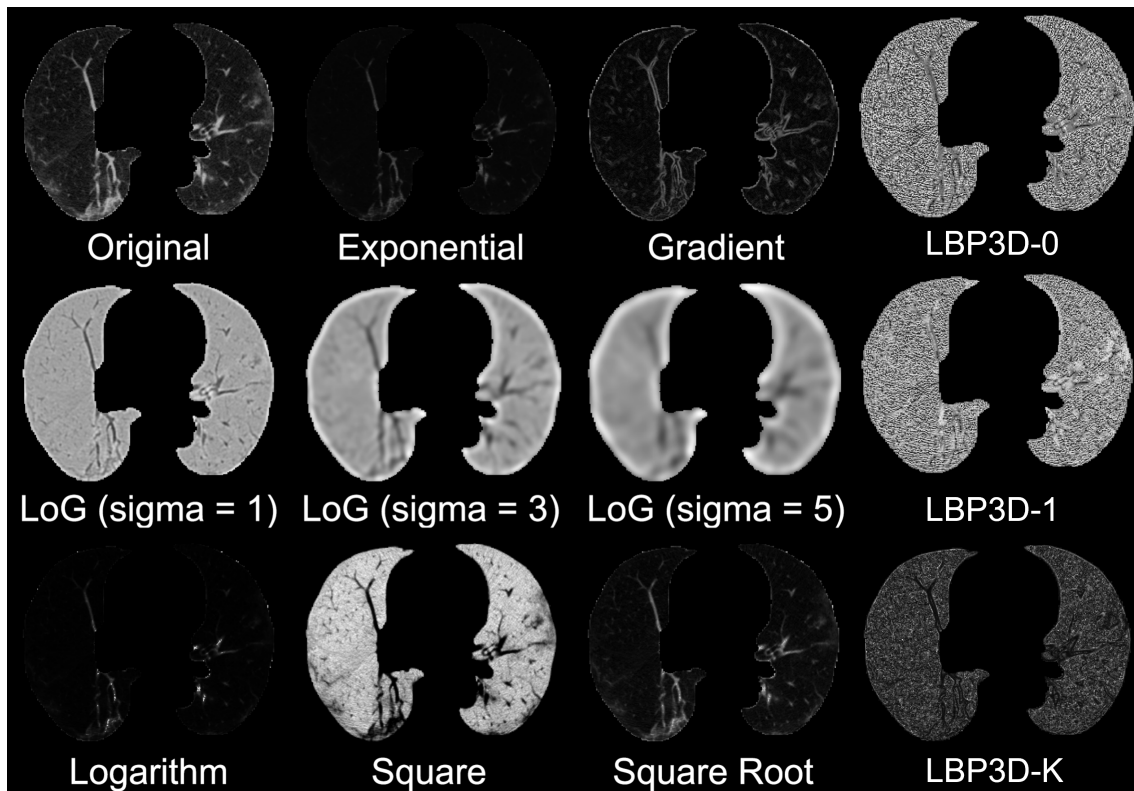


Figure 5.1: Segmented original and filtered images prior to feature extraction

The exponential filter shows a darkened image, with some white spots on the border. These indicate the highest gray values, which are probably tissues from outside the lungs that weren't correctly segmented. This filter greatly increases the gray level intensity through exponential operations, and then rescales them back to the original range. Therefore, these distinct tissues ended up with a lot higher values than the rest of the image, de-emphasizing the differences on the lung tissues.

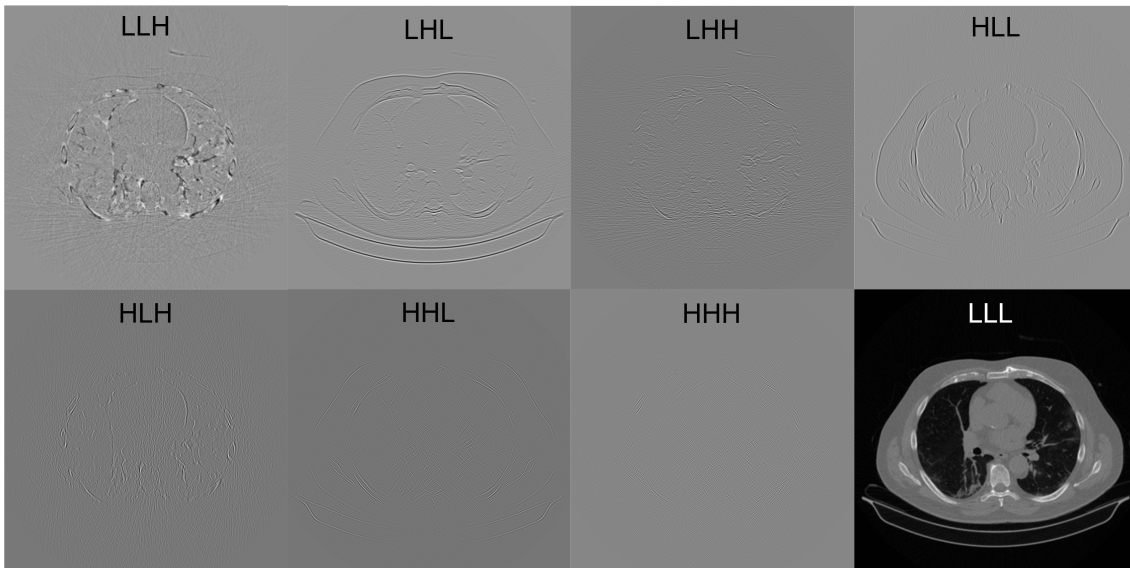


Figure 5.2: Wavelet filtered images prior to feature extraction

As expected, the gradient filter highlighted the edges, especially the blood vessels edges on the lungs. However, this filter does not highlight opaque areas, which might be relevant for the identification of COVID-19. The LoG filter with σ of 1 also highlighted the blood vessels. Using a σ of 3 and 5 resulted in a proportionally smoother image.

The logarithm filter highlighted the main blood vessels while darkening most of the smaller ones. The square filter highlighted all blood vessels on the image as well as the opaque zone, showing a good capability of highlighting visual lung characteristics. The square root filter resulted in a slightly darker version of the original image.

5.2 Training and Optimization of Classification Models

Table 5.1 shows the performance of the classification models in the training dataset in 5-fold cross-validation with the average F1-score metric before optimization, using all combinations of wavelet features, non-wavelet features, and bin sizes (1, 5, 10, 25, 50, 75).

Models will be further referenced as “NW(bin size for non-wavelet features) – W(bin size for wavelet features)”. e.g., NW5-W25 is the model with a bin size of 5 HU for non-wavelet features and 25 HU for wavelet features. The three models with the highest mean F1-scores on both groups (on all data) were selected for optimization using the Hyperopt library [Bergstra et al., 2015]. The models with the best F1 scores were NW1-W25, NW5-W5, and NW5-W25.

Table 5.2 shows the performance metrics for the classification models in the validation set after the optimization process using the macro average and weighted average. The macro average gives a simple average of the performance on both classes, while the weighted average proportionally scales the metrics from each class according to its number of samples.

Table 5.1: Mean F1-score results for classification in 5-fold cross-validation for all 36 combinations of bin sizes

Bin Size		non-wavelet features (NW)						
		1	5	10	25	50	75	
wavelet features (W)	1	COVID	0.837	0.838	0.840	0.827	0.838	0.847
		non-COVID	0.405	0.409	0.408	0.357	0.350	0.320
		All data	0.621	0.624	0.624	0.592	0.594	0.583
	5	COVID	0.823	0.831	0.823	0.831	0.829	0.841
		non-COVID	0.371	0.378	0.346	0.399	0.367	0.331
		All data	0.597	0.604	0.585	0.615	0.598	0.586
	10	COVID	0.823	0.837	0.811	0.825	0.837	0.848
		non-COVID	0.340	0.382	0.309	0.381	0.341	0.331
		All data	0.582	0.609	0.560	0.603	0.589	0.590
	25	COVID	0.837	0.830	0.819	0.831	0.825	0.842
		non-COVID	0.405	0.386	0.370	0.382	0.350	0.283
		All data	0.621	0.608	0.595	0.607	0.588	0.562
	50	COVID	0.818	0.822	0.809	0.828	0.824	0.838
		non-COVID	0.342	0.337	0.337	0.393	0.361	0.262
		All data	0.580	0.579	0.573	0.611	0.593	0.550
	75	COVID	0.823	0.827	0.807	0.824	0.815	0.831
		non-COVID	0.353	0.349	0.303	0.351	0.333	0.273
		All data	0.588	0.588	0.555	0.588	0.574	0.552

Table 5.2: Performance of the optimized models on the validation set

Model	Metric	COVID	Non-COVID	Macro avg	Weighted avg
1-25	Precision	0.838	0.875	0.856	0.849
	Recall	0.969	0.538	0.754	0.844
	F1-score	0.899	0.667	0.783	0.832
	Accuracy	-	-	0.844	-
	AUC	0.916	-	-	-
5-1	Precision	0.821	1.000	0.910	0.872
	Recall	1.000	0.462	0.731	0.844
	F1-score	0.901	0.632	0.766	0.823
	Accuracy	-	-	0.844	-
	AUC	0.966	-	-	-
10-1	Precision	0.762	1.000	0.881	0.831
	Recall	1.000	0.231	0.615	0.778
	F1-score	0.865	0.375	0.620	0.723
	Accuracy	-	-	0.778	-
	AUC	0.813	-	-	-
N on validation set		32	13		

All models had a better performance classifying COVID images rather than non-COVID. This was expected since the COVID group is more than twice as large as the non-COVID, which may bias the model, as it learns to classify much more samples as the predominant class. This is more evident in the optimized NW5-W1 and NW10-W1 models,

where the precision is 1 (or 100%) for the non-COVID class, while the recall is low (0.46 for the NW5-W1 model and 0.23 for the NW10-W1 model). This means that all samples classified as non-COVID are indeed from this class, but only a part of all non-COVID samples are correctly classified. The recall of 1 with not so high precision (0.82 for the NW5-W1 model and 0.76 for the NW10-W1 model) for the COVID class on both models corroborates this behavior, as it shows that the model correctly classified all ground-truth COVID samples while also classifying other non-COVID samples as COVID. Despite having the same accuracy as the NW5-W1, the NW1-W25 model proved to be more robust, with a higher average F1 score. Still, this model also had a low recall for the non-COVID class.

Two hypotheses were built for the low performance of the classification models in the validation set. The first hypothesis is that the imbalanced dataset is the primary source of the low performance since the higher number of the input COVID samples in the training process led the model to classify almost all inputs as from the COVID group. The second hypothesis is that we used too many features for a small dataset, which led to useless information that could not be precisely filtered on the training process. Results show up to this point, regarding our first models, will be further referenced as Raw models.

5.3 Exploring the Limitations

5.3.1 Dealing with Imbalanced Dataset

Regarding the first hypothesis, we tried to address the problem in two different ways. The first was to use data augmentation directly on the extracted features using Synthetic Minority Over-sampling Technique (SMOTE). The second was to under-sample our COVID group until it had the same number of samples as the non-COVID group. We didn't try to use extra data augmentation on the original images, like flipping and rotation considering most of the extracted features, like the first order, GLCM, and GLSZM, are independent of these variations [Parekh and Jacobs, 2016].

SMOTE [Chawla et al., 2002] is an oversampling method that has great applications in artificial intelligence. It consists of creating synthetic examples to deal with classes with fewer samples. The synthetic examples are created by operating in the "feature space" rather than "data space". It calculates the difference between the feature vector of one input and its neighbors multiplied by a random number between 0 and 1. Finally the value is added to the feature vector under consideration. This process is applied to the extracted features on the training dataset (in our case, the inputs are the extracted features), creating new feature values so that both COVID and non-COVID groups had the same number of samples. However, using SMOTE with non-wavelet features extracted with bin size 75 led to features with higher values than the acceptable on SMOTE library (64-bit). Due to this limitation, non-wavelet features extracted with bin size 75 were not used in this analysis.

Results using the SMOTE technique are shown in table 5.3. Given their performance, the selected models for optimization were NW1-W5, NW5-W5, and NW25-W1.

Table 5.3: Mean F1-score results for classification in 5-fold cross-validation for 30 combinations of bin sizes using SMOTE

		Bin Size	non-wavelet features (NW)				
			1	5	10	25	50
wavelet features (W)	1	COVID	0.818	0.817	0.821	0.826	0.817
		non-COVID	0.414	0.423	0.427	0.430	0.403
		All data	0.616	0.620	0.624	0.628	0.610
	5	COVID	0.830	0.819	0.795	0.814	0.807
		non-COVID	0.445	0.445	0.394	0.441	0.378
		All data	0.638	0.632	0.594	0.627	0.593
	10	COVID	0.809	0.814	0.796	0.805	0.798
		non-COVID	0.381	0.409	0.377	0.396	0.382
		All data	0.595	0.611	0.586	0.600	0.590
	25	COVID	0.812	0.802	0.810	0.811	0.810
		non-COVID	0.405	0.427	0.417	0.433	0.407
		All data	0.609	0.615	0.613	0.622	0.609
	50	COVID	0.814	0.797	0.798	0.804	0.806
		non-COVID	0.401	0.411	0.362	0.400	0.368
		All data	0.608	0.604	0.580	0.602	0.587
	75	COVID	0.810	0.795	0.788	0.803	0.805
		non-COVID	0.390	0.351	0.346	0.368	0.383
		All data	0.600	0.573	0.567	0.585	0.594

To undersample our COVID group, we simply removed random samples from it until both COVID and non-COVID groups had the same number of instances. Results using this approach are shown in table 5.4. Given their performance, the selected models for optimization were NW5-W10, NW5-W25, and NW1-W75.

5.3.2 Reducing the Number of Features

Following our second hypothesis, that we used too many features for a small dataset, we came up with two alternative methods. The first was to perform feature selection before classification. The second was to use the features extracted from the original segmented image without any additional filters.

We performed feature selection by checking the correlation of each extracted feature with all the other ones. Features with a correlation higher than 0.9 were discarded. Of the original 3548 features used in each model, a minimum of 925 to a maximum of 1064 were dropped (according to the combination of bin sizes), and 2484 to 2623 features were used to train the models. Results using feature selection are shown in table 5.5. Given their performance, the selected models for optimization were NW1-W1, NW25-W25, and NW50-W1.

Table 5.4: Mean F1-score results for classification in 5-fold cross-validation for all combinations of bin sizes undersampling the COVID class.

		Bin Size	non-wavelet features (NW)					
			1	5	10	25	50	75
wavelet features (W)	1	COVID	0.690	0.686	0.680	0.698	0.687	0.668
		non-COVID	0.464	0.459	0.444	0.482	0.443	0.456
		All data	0.577	0.572	0.562	0.590	0.565	0.562
	5	COVID	0.700	0.707	0.680	0.708	0.680	0.678
		non-COVID	0.472	0.486	0.467	0.488	0.432	0.460
		All data	0.586	0.597	0.573	0.598	0.556	0.569
	10	COVID	0.712	0.705	0.677	0.676	0.689	0.669
		non-COVID	0.489	0.499	0.472	0.459	0.445	0.453
		All data	0.600	0.602	0.575	0.567	0.567	0.561
	25	COVID	0.699	0.715	0.684	0.696	0.705	0.675
		non-COVID	0.479	0.500	0.473	0.496	0.470	0.474
		All data	0.589	0.608	0.579	0.596	0.587	0.575
	50	COVID	0.696	0.685	0.684	0.681	0.660	0.663
		non-COVID	0.485	0.472	0.461	0.477	0.442	0.436
		All data	0.590	0.578	0.573	0.579	0.551	0.549
	75	COVID	0.703	0.675	0.674	0.663	0.680	0.663
		non-COVID	0.534	0.442	0.452	0.461	0.465	0.424
		All data	0.619	0.559	0.563	0.562	0.573	0.543

Table 5.5: Mean F1-score results for classification in 5-fold cross-validation for all combinations of bin sizes excluding highly correlated features

		Bin Size	non-wavelet features (NW)					
			1	5	10	25	50	75
wavelet features (W)	1	COVID	0.841	0.837	0.831	0.827	0.841	0.828
		non-COVID	0.381	0.370	0.324	0.380	0.382	0.349
		All data	0.611	0.604	0.578	0.604	0.611	0.589
	5	COVID	0.826	0.834	0.816	0.824	0.826	0.824
		non-COVID	0.345	0.379	0.361	0.333	0.365	0.336
		All data	0.585	0.606	0.589	0.578	0.595	0.580
	10	COVID	0.823	0.826	0.826	0.823	0.836	0.819
		non-COVID	0.326	0.365	0.337	0.359	0.354	0.300
		All data	0.574	0.595	0.582	0.591	0.595	0.560
	25	COVID	0.826	0.831	0.825	0.840	0.830	0.827
		non-COVID	0.364	0.369	0.364	0.420	0.366	0.326
		All data	0.595	0.600	0.595	0.630	0.598	0.577
	50	COVID	0.823	0.829	0.817	0.823	0.834	0.810
		non-COVID	0.323	0.345	0.294	0.364	0.372	0.285
		All data	0.573	0.587	0.555	0.593	0.603	0.548
	75	COVID	0.815	0.824	0.815	0.823	0.828	0.818
		non-COVID	0.307	0.313	0.295	0.355	0.350	0.282
		All data	0.561	0.569	0.555	0.589	0.589	0.550

Regarding our second strategy to deal with too many features per sample, results of classification using only the features extracted from the original image (without any applied

filters) are shown in table 5.6. Given their performance, the selected models for optimization were NW1-W1, NW5-W25, and NW25-W25. This strategy will be further referenced as “Original features”.

Table 5.6: Mean F1-score results for classification in 5-fold cross-validation for 6 different bin sizes using only the features extracted from the original image

Bin Size	1	5	10	25	50	75
COVID	0.831	0.837	0.832	0.829	0.831	0.825
non-COVID	0.403	0.448	0.358	0.406	0.367	0.406
All data	0.617	0.642	0.595	0.617	0.599	0.616

5.3.3 Performance on validation and test sets

The performance of the classification models are shown in table 5.7. The AUC on the validation set was used to optimize the model, and thus, the performance was higher than in the test set for all models. The performance on the test set shows the capability of the model to generalize to new data. Higher metrics on the validation set with lower performance on the test set indicate overfitting, where the model learned too many intrinsic characteristics of the training and validation sets and can't be applied to other datasets. A clear example of this modeling error happened on the NW5-W1 model, where the AUC was 0.966 on the validation set but only 0.581 on the test set. Considering 0.5 would be a random selection, it is a significantly lower performance than on the validation set. Other models indicating overfitting were NW1-W5 using SMOTE and NW50-W1 using feature selection.

Table 5.7: Performance of all optimized models on the validation and test sets

	NW-W	Validation				Test			
		COVID	F1-score N-COVID	Avg	AUC	COVID	F1-score N-COVID	Avg	AUC
First models (raw)	1-25	0.899	0.667	0.783	0.916	0.861	0.444	0.653	0.678
	5-1	0.901	0.632	0.766	0.966	0.849	0.353	0.601	0.581
	10-1	0.865	0.375	0.620	0.813	0.849	0.353	0.601	0.731
SMOTE	1-5	0.841	0.476	0.658	0.851	0.789	0.211	0.500	0.538
	5-5	0.873	0.526	0.700	0.916	0.857	0.500	0.679	0.745
	25-1	0.870	0.571	0.720	0.875	0.849	0.353	0.601	0.800
Under-sample	1-75	0.714	0.692	0.703	0.794	0.630	0.444	0.537	0.615
	5-10	0.696	0.774	0.735	0.846	0.444	0.444	0.444	0.642
	5-25	0.500	0.706	0.603	0.912	0.542	0.476	0.509	0.649
Feature selection	1-1	0.866	0.609	0.737	0.846	0.829	0.400	0.614	0.671
	25-25	0.845	0.421	0.633	0.892	0.833	0.333	0.583	0.632
	50-1	0.882	0.636	0.759	0.857	0.833	0.333	0.583	0.542
Only original features	1	0.886	0.600	0.743	0.901	0.812	0.381	0.596	0.660
	5	0.870	0.571	0.720	0.906	0.824	0.455	0.639	0.692
	25	0.901	0.632	0.766	0.947	0.833	0.333	0.583	0.637

The SMOTE strategy to deal with imbalanced data had the most significant impact on the NW5-W5 model. In the test set, it achieved an F1-score of 0.5 on the non-COVID samples, which is still low but higher than all other optimized models. Despite their low performance on non-COVID samples, SMOTE NW25-W1 model had the higher AUC in the test set (0.8). Under-sampling the COVID class improved the performance of the non-COVID class when compared to the other models. However, the F1-score for the non-COVID class was still low (0.444 - 0.476). Also, it significantly lowered the metrics on the COVID class, leading to a lower average F1-score on all three models, with a maximum of 0.537 on the NW5-W25 model.

Using feature selection did not have a significant impact on the performance compared to our first models. All three models had an F1-score of 0.4 or lower for the non-COVID class. The NW1-W1 was the best model using feature selection, achieving an AUC of 0.671. Only using features extracted from the original image resulted in one more robust model (NW5-W25), which reached the third-best average F1-score (0.639) and the fourth-best AUC (0.692).

Of the optimized models, the ones with an AUC higher than 0.65 and F1-score higher than 0.6 were selected for feature importance analysis. The ROC curve for the six best models is shown in figure 5.3.

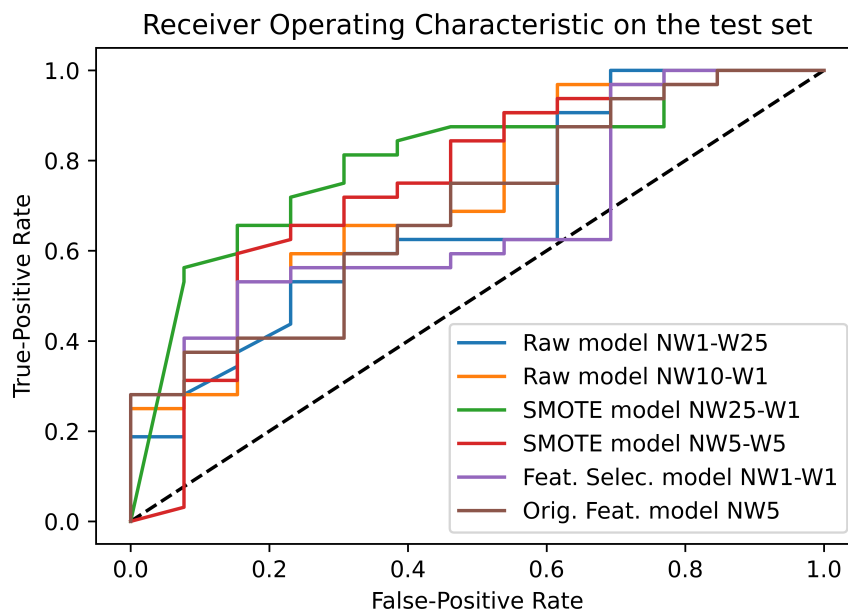


Figure 5.3: Receiver Operator Characteristic curve on the test set

5.4 Preprocessing Impact and Feature Importance

The 10 most important features for each of the 6 best models are shown in figure 5.4 along with their normalized gain. The feature gain is proportional to how much it

improved the models' performance. Abbreviations for each of the presented features are shown in Annex B. Features that were mostly used for classification were extracted from the

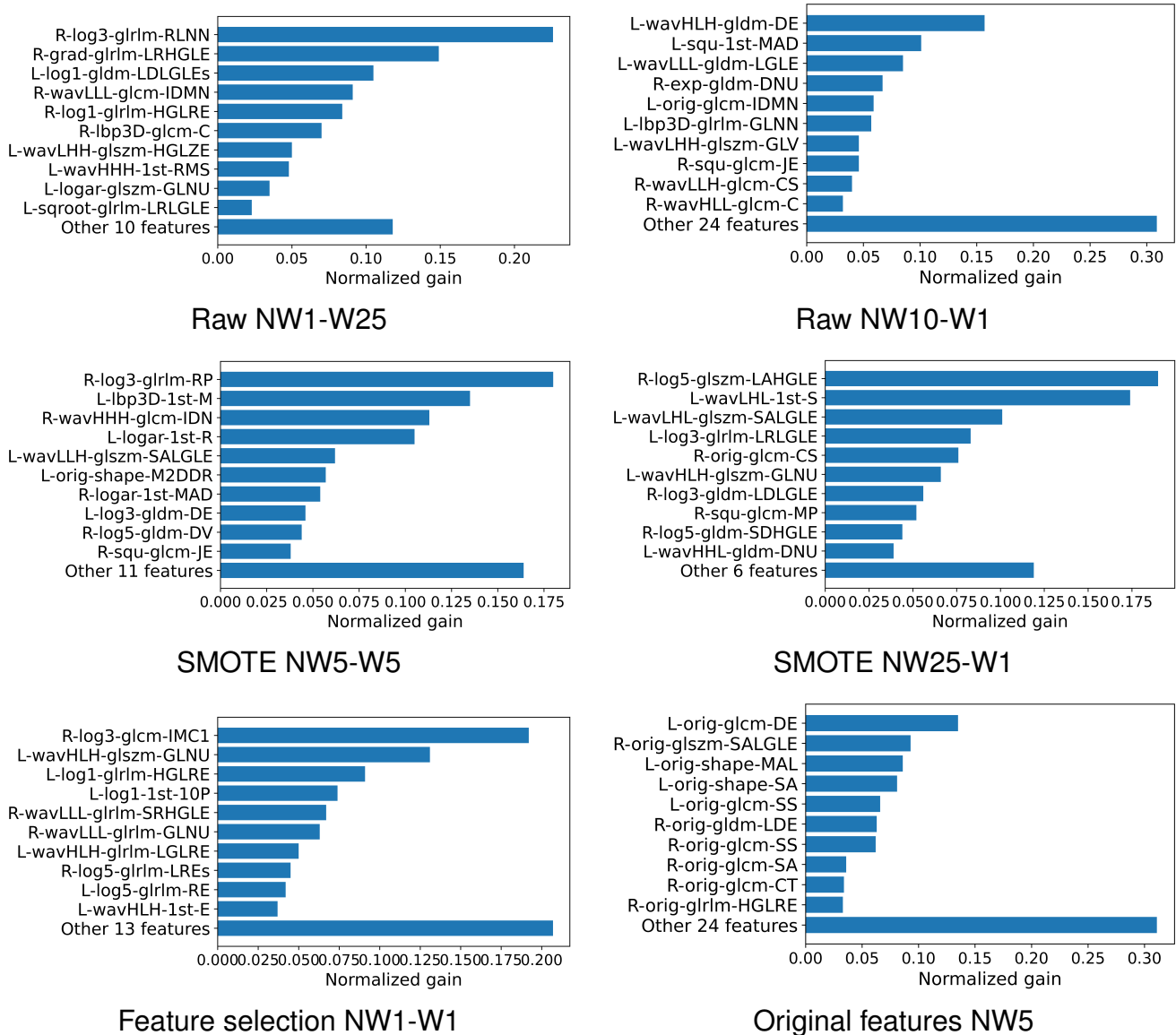


Figure 5.4: Feature importance for the models with the highest performance on the test set images generated with wavelet and laplacian filters.

From the classification models that used all feature groups (all five except the Original features NW5), four of five have a feature from the right lung using the Laplacian filter as their most important feature, with a σ value of 3 in three models and 5 in one model (SMOTE NW25-W1). On Raw NW1-W25 model, the most important feature was Run Length Non-Uniformity Normalized with 22.6% normalized gain, on SMOTE NW5-W5 it was Run Percentage with 18.0%, on the NW1-W1 model using feature selection it was Informational Measure of Correlation 1 with 19.2%, and on SMOTE NW25-W1 it was Large Area High Gray Level Emphasis with 19.0%. Even though the most important features were different in each model, the Laplacian filter may highlight the differences between COVID and

non-COVID radiomic signatures. [Chao et al., 2021] study also found that multiple features derived from the Laplacian filter had high importance on the outcome prediction of COVID-19 patients, as did wavelet features.

Wavelet features were important for classification in all models that used them. The most important feature on the Raw NW10-W1 model was the Dependence Entropy extracted using one of the wavelet filters, with a 15.7% normalized gain. [Chen et al., 2021a] study also found that wavelet features have a high importance for COVID-19 prognosis prediction, as from their 15 selected features with a higher correlation, 8 were derived from the wavelet filter. However, they used segmented lesions while we used whole lung segmentation, and they did not explore Laplacian filters. [Huang et al., 2021b] build a nomogram to discriminate COVID-19 and influenza pneumonia using seven radiomic features extracted from lung lesions, where three features are derived from the wavelet filters.

Other filters that had a few important features for classification were logarithm, local binary pattern 3D, square, square root, exponential, and gradient. However, due to the low number of features selected, the low performance of our model, and the randomness of the training and optimization process, we cannot infer those features extracted using these filters are relevant to differentiate COVID and non-COVID radiomic signatures.

Only one feature extracted from the original image was among the ten most important ones in three of the five final models (not considering Original features NW5). This indicates that filtering images may be better for differentiating COVID and non-COVID than using the original images. Regarding our model that only used features extracted from the original image without filtering, the most important ones were mainly from GLCM, GLSZM, and shape feature groups. Therefore, if trying to only use features extracted from the original images, we recommend a bin size of 5 or similar and further exploration of GLCM, GLSZM, and shape feature groups. However, it should be noted that shape is highly dependent on the segmentation method.

5.5 Discretization Impact

To our knowledge, this is the first study that sought to evaluate discretization methods for radiomic feature extraction on whole lung segmented images for the identification of COVID-19. Even though the efficiency of our models wasn't particularly high, we can see patterns in our chosen models and their performances regarding the discretization.

Models using non-wavelet features extracted with bin size 75 had, generally, lower performance on all models prior to optimization when compared to other bin sizes with the same strategies. Consequently, none of the models chosen for optimization used non-wavelet features extracted with a bin size 75. For non-wavelet features with a bin size 50, the scenario was similar. Even though one model was selected for optimization (NW50-W1 using feature selection), it showed low performance on the test set, with an AUC of 0.542.

Regarding wavelet features, the only optimized model that used a bin size of 50 or 75 was the NW1-W75 under-sampling the COVID class. However, it also had low performance on the test set, with an F1-score of 0.537.

These results indicate that larger bin sizes may not be ideal for feature extraction using the whole lungs, at least when identifying the COVID radiomic signature. This may happen due to higher bin sizes erasing relevant information that could be used to differentiate COVID and non-COVID but may also apply to other lung diseases.

The bin size of 25 for non-wavelet features and a bin size of 1 for wavelet features were used in the model with the highest AUC on the test set (SMOTE NW25-W1). Figure 5.4 shows that the most important features for classification also came from the wavelet and laplacian filters. Wavelet features with a bin size of 1 were also used in the other two of the best models (Feature Selection NW1-W1 and Raw NW10-W1) and showed a high importance in all of them. The only model that used wavelet features extracted with higher bin sizes (Raw NW1-W25) had lower importance for wavelet features than the other models. These results indicate that wavelet filters with a small bin size may be relevant in differentiating COVID and non-COVID radiomic signatures. Therefore, wavelet features may be less susceptible to noise than other ones. The same may apply to features extracted from laplacian filtered images using larger σ values (as 3 and 5, which led to more relevant features), since the filter smooths out the image, and may eliminate noise impact. This allows smaller bin sizes or a higher number of gray levels to be used.

The model with the highest average F1-score on the test set used non-wavelet features and wavelet features extracted with a bin size of 5 (SMOTE NW5-W5). The most important features on this model were extracted using the original image and multiple filters, including laplacian of gaussian, local binary pattern 3D, wavelet, logarithm, and square.

Our results indicate that bin sizes between 1 and 25 should be further explored for non-wavelet features. It is possible that specific discretization strategies might perform better for some filters than others, as is the case for the wavelet filter. The LoG is one of the filters that may perform better than others with smaller bin sizes, as this filter already reduces noise. Therefore, features from the laplacian filter might be further investigated with a bin size between 1 and 10, while other filters may have a better performance using higher values. For wavelet features, we also recommend further investigation for smaller bin sizes, between 1 and 10, as features extracted using these parameters had higher importance. Other wavelet families may also be explored, as we only used Coiflets in this study. A previous study regarding lung cancer [Desseroit et al., 2016] found that image discretization has an important role in feature repeatability, and they also found that an ideal smaller bin size of 10 HU was the best bin for their purpose. However, they do not mention filtering and use lesion segmentation.

5.6 Limitations

Despite our efforts, our study has some limitations. First, we did not have access to clinical information, which could strongly improve our COVID identification model and was used in several studies analyzed on our state of the art. Also, our datasets were not particularly large, which led to models with a lower generalization capability to forecast unseen data. Furthermore, we used imbalanced datasets since we did not have access to more non-COVID samples using the same acquisition protocol. Further studies should try to use larger balanced datasets and use clinical data and radiomic features.

6. CONCLUSION

We implemented several models for COVID-19 identification using features extracted from multiple filtered lung-segmented images with different discretization parameters. Models were trained with combinations of non-wavelet and wavelet features using six different bin sizes (1, 5, 10, 25, 50, 75). The models with the highest performance were selected for hyperparameter optimization. We identified the limitations of our models (imbalanced data and too many features per sample) and proposed strategies to overcome them. Models were trained again using SMOTE, under-sampling, feature selection, and only using features extracted from the original image. The best models using each of these strategies also had their hyperparameters optimized. All optimized models were evaluated in the test set and the best model (SMOTE NW25-W1) reached an AUC of 0.800.

The six best-performing models on the test set were chosen for feature importance analysis. Their F1-score varied between 0.601 and 0.679 and their AUC between 0.671 and 0.800. The filters that led to the most important features for classification were identified: wavelet and laplacian of gaussian. We recommend further exploration of these filters for the identification of COVID-19 in lung images, as they may highlight differences in the radiomic signature of the lungs.

Regarding discretization methods, our results indicate that features extracted with bin sizes 50 and 75 aren't ideal for discriminating the COVID and non-COVID groups. These larger bin sizes may erase relevant information of the COVID-19 radiomic signature. We recommend further investigation of smaller bin sizes, between 1 and 25 for feature extraction. Multi-parameter feature extraction may improve COVID identification performance, using different discretization for different filters. Regarding the wavelet and laplacian of gaussian filters, we suggest bin sizes between 1 and 10, as these filters already partially remove noise and showed a good performance with smaller bin sizes.

We suggest further research on the COVID-19 radiomic signature of the lungs include clinical data on their analysis and use larger and balanced datasets so that the models have a higher generability.

REFERENCES

- Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Carvalho, S., Bussink, J., Monshouwer, R., Haibe-Kains, B., Rietveld, D., Hoebers, F., Rietbergen, M. M., Leemans, C. R., Dekker, A., Quackenbush, J., Gillies, R. J., and Lambin, P. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature Communications*, 5(1).
- Amini, M., Nazari, M., Shiri, I., Hajianfar, G., Deevband, M. R., Abdollahi, H., Arabi, H., Rahmim, A., and Zaidi, H. (2021). Multi-level multi-modality (PET and CT) fusion radiomics: prognostic modeling for non-small cell lung carcinoma. *Physics in Medicine & Biology*, 66(20):205017.
- Apostolopoulos, I. D., Aznaouridis, S. I., and Tzani, M. A. (2020). Extracting possibly representative COVID-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases. *Journal of Medical and Biological Engineering*, 40(3):462–469.
- Arru, C., Ebrahimian, S., Falaschi, Z., Hansen, J. V., Pasche, A., Lyhne, M. D., Zimmermann, M., Durlak, F., Mitschke, M., Carriero, A., Nielsen-Kudsk, J. E., Kalra, M. K., and Saba, L. (2021). Comparison of deep learning, radiomics and subjective assessment of chest CT findings in SARS-CoV-2 pneumonia. *Clinical Imaging*, 80:58–66.
- Banerjee, J., Moelker, A., Niessen, W. J., and van Walsum, T. (2013). 3d LBP-based rotationally invariant region description. In *Computer Vision - ACCV 2012 Workshops*, pages 26–37. Springer Berlin Heidelberg.
- Bartolucci, M., Benelli, M., Betti, M., Bicchi, S., Fedeli, L., Giannelli, F., Aquilini, D., Baldini, A., Consales, G., Natale, M. E. D., Lotti, P., Vannucchi, L., Trezzi, M., Mazzoni, L. N., Santini, S., Carpi, R., Matarrese, D., Bernardi, L., Mascacchi, M., Cavigli, E., Bindi, A., Cozzi, D., Miele, V., Busoni, S., Taddeucci, A., Allesscia, G., Zini, C., Dedola, G. L., Mazzocchi, S., Pozzessere, C., and and, A. V. (2021). The incremental value of computed tomography of COVID-19 pneumonia in predicting ICU admission. *Scientific Reports*, 11(1).
- Bera, K., Braman, N., Gupta, A., Velcheti, V., and Madabhushi, A. (2021). Predicting cancer outcomes with radiomics and artificial intelligence in radiology. *Nature Reviews Clinical Oncology*, 19(2):132–146.
- Bergstra, J., Komer, B., Eliasmith, C., Yamins, D., and Cox, D. D. (2015). Hyperopt: a python library for model selection and hyperparameter optimization. *Computational Science & Discovery*, 8(1):014008.

- Berta, L., Mattia, C. D., Rizzetto, F., Carrazza, S., Colombo, P., Fumagalli, R., Langer, T., Lizio, D., Vanzulli, A., and Torresin, A. (2021). A patient-specific approach for quantitative and automatic analysis of computed tomography images in lung disease: Application to COVID-19 patients. *Physica Medica*, 82:28–39.
- Bisong, E. (2019). Google colab. In *Building Machine Learning and Deep Learning Models on Google Cloud Platform*, pages 59–64. Apress.
- Cai, Q., Du, S.-Y., Gao, S., Huang, G.-L., Zhang, Z., Li, S., Wang, X., Li, P.-L., Lv, P., Hou, G., and Zhang, L.-N. (2020). A model based on CT radiomic features for predicting RT-PCR becoming negative in coronavirus disease 2019 (COVID-19) patients. *BMC Medical Imaging*, 20(1).
- Cai, S., Chen, Y., Zhao, S., He, D., Li, Y., Xiong, N., Li, Z., and Hu, S. (2022). Dynamic 3d radiomics analysis using artificial intelligence to assess the stage of COVID-19 on CT images. *European Radiology*.
- Cardobi, N., Benetti, G., Cardano, G., Arena, C., Micheletto, C., Cavedon, C., and Montemezzi, S. (2021). CT radiomic models to distinguish COVID-19 pneumonia from other interstitial pneumonias. *La radiologia medica*, 126(8):1037–1043.
- Caruso, D., Pucciarelli, F., Zerunian, M., Ganeshan, B., Santis, D. D., Polici, M., Rucci, C., Polidori, T., Guido, G., Bracci, B., Benvenga, A., Barbato, L., and Laghi, A. (2021). Chest CT texture-based radiomics analysis in differentiating COVID-19 from other interstitial pneumonia. *La radiologia medica*, 126(11):1415–1424.
- Centers Disease Control Prevention (2022). What you need to know about variants. <https://www.cdc.gov/coronavirus/2019-ncov/variants/about-variants.html>. Accessed: 2022-02-12.
- Chang, Y.-C., Stoyanova, R., Danilova, S., Pretell-Mazzini, J. A., Kerr, D. A., Wilky, B. A., and Subhawong, T. (2021). Radiomics on radiography predicts giant cell tumor histologic response to denosumab. *Skeletal Radiology*, 50(9):1881–1887.
- Chao, H., Fang, X., Zhang, J., Homayounieh, F., Arru, C. D., Digumarthy, S. R., Babaei, R., Mobin, H. K., Mohseni, I., Saba, L., Carriero, A., Falaschi, Z., Pasche, A., Wang, G., Kalra, M. K., and Yan, P. (2021). Integrative analysis for COVID-19 patient outcome prediction. *Medical Image Analysis*, 67:101844.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16:321–357.
- Chen, H., Zeng, M., Wang, X., Su, L., Xia, Y., Yang, Q., and Liu, D. (2021a). A CT-based radiomics nomogram for predicting prognosis of coronavirus disease 2019 (COVID-19) radiomics nomogram predicting COVID-19. *The British Journal of Radiology*, 94(1117):20200634.

- Chen, H. J., Mao, L., Chen, Y., Yuan, L., Wang, F., Li, X., Cai, Q., Qiu, J., and Chen, F. (2021b). Machine learning-based CT radiomics model distinguishes COVID-19 from non-COVID-19 pneumonia. *BMC Infectious Diseases*, 21(1).
- Chen, N., Zhou, M., Dong, X., Qu, J., Gong, F., Han, Y., Qiu, Y., Wang, J., Liu, Y., Wei, Y., Xia, J., Yu, T., Zhang, X., and Zhang, L. (2020a). Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in wuhan, china: a descriptive study. *The Lancet*, 395(10223):507–513.
- Chen, T. and Guestrin, C. (2016). Xgboost: A scalable tree boosting system. *CoRR*, abs/1603.02754.
- Chen, Y., Wang, Y., Zhang, Y., Zhang, N., Zhao, S., Zeng, H., Deng, W., Huang, Z., Liu, S., and Song, B. (2020b). A quantitative and radiomics approach to monitoring ARDS in COVID-19 patients based on chest CT: a retrospective cohort study. *International Journal of Medical Sciences*, 17(12):1773–1782.
- Chen, Z., Li, X., Li, J., Zhang, S., Zhou, P., Yu, X., Ren, Y., Wang, J., Zhang, L., Li, Y., Wu, B., Hou, Y., Zhang, K., Tang, R., Liu, Y., Ding, Z., Yang, B., Deng, Q., Lin, Q., Nie, K., Cai, Z., Ma, S., and Kuang, Y. (2021c). A COVID-19 risk score combining chest CT radiomics and clinical characteristics to differentiate COVID-19 pneumonia from other viral pneumonias. *Aging*, 13(7):9186–9224.
- Chung, M., Bernheim, A., Mei, X., Zhang, N., Huang, M., Zeng, X., Cui, J., Xu, W., Yang, Y., Fayad, Z. A., Jacobi, A., Li, K., Li, S., and Shan, H. (2020). CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology*, 295(1):202–207.
- de Farias, E. C., di Noia, C., Han, C., Sala, E., Castelli, M., and Rundo, L. (2021). Impact of GAN-based lesion-focused medical image super-resolution on the robustness of radiomic features. *Scientific Reports*, 11(1).
- de Moura, L. V., Mattjie, C., Dartora, C. M., Barros, R. C., and da Silva, A. M. M. (2022). Explainable machine learning for COVID-19 pneumonia classification with texture-based features extraction in chest radiography. *Frontiers in Digital Health*, 3.
- Desseroit, M.-C., Tixier, F., Weber, W. A., Siegel, B. A., Rest, C. C. L., Visvikis, D., and Hatt, M. (2016). Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of non–small cell lung cancer tumors: A repeatability analysis in a prospective multicenter cohort. *Journal of Nuclear Medicine*, 58(3):406–411.
- Di, D., Shi, F., Yan, F., Xia, L., Mo, Z., Ding, Z., Shan, F., Song, B., Li, S., Wei, Y., Shao, Y., Han, M., Gao, Y., Sui, H., Gao, Y., and Shen, D. (2021). Hypergraph learning for identification of COVID-19 with CT imaging. *Medical Image Analysis*, 68:101910.

- Ding, X., Xu, J., Zhou, J., and Long, Q. (2020). Chest CT findings of COVID-19 pneumonia by duration of symptoms. *European Journal of Radiology*, 127:109009.
- Dong, E., Du, H., and Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, 20(5):533–534.
- Fang, M., He, B., Li, L., Dong, D., Yang, X., Li, C., Meng, L., Zhong, L., Li, H., Li, H., and Tian, J. (2020a). CT radiomics can help screen the coronavirus disease 2019 (COVID-19): a preliminary study. *Science China Information Sciences*, 63(7).
- Fang, X., Li, X., Bian, Y., Ji, X., and Lu, J. (2020b). Radiomics nomogram for the prediction of 2019 novel coronavirus pneumonia caused by SARS-CoV-2. *European Radiology*, 30(12):6888–6901.
- Farinholt, T., Doddapaneni, H., Qin, X., Menon, V., Meng, Q., Metcalf, G., Chao, H., Gingras, M.-C., Avadhanula, V., Farinholt, P., Agrawal, C., Muzny, D. M., Piedra, P. A., Gibbs, R. A., and Petrosino, J. (2021). Transmission event of SARS-CoV-2 delta variant reveals multiple vaccine breakthrough infections. *BMC Medicine*, 19(1).
- Fatima, S., Ratnani, I., Husain, M., and Surani, S. (2020). Radiological findings in patients with COVID-19. *Cureus*.
- Feng, J., Guo, Y., Wang, S., Shi, F., Wei, Y., He, Y., Zeng, P., Liu, J., Wang, W., Lin, L., Yang, Q., Li, C., and Liu, X. (2020). Differentiation between COVID -19 and bacterial pneumonia using radiomics of chest computed tomography and clinical features. *International Journal of Imaging Systems and Technology*, 31(1):47–58.
- Fu, F., Lou, J., Xi, D., Bai, Y., Ma, G., Zhao, B., Liu, D., Bao, G., Lei, Z., and Wang, M. (2020a). Chest computed tomography findings of coronavirus disease 2019 (COVID-19) pneumonia. *European Radiology*, 30(10):5489–5498.
- Fu, L., Li, Y., Cheng, A., Pang, P., and Shu, Z. (2020b). A novel machine learning-derived radiomic signature of the whole lung differentiates stable from progressive COVID-19 infection. *Journal of Thoracic Imaging*, Publish Ahead of Print.
- Gillies, R. J., Kinahan, P. E., and Hricak, H. (2016). Radiomics: Images are more than pictures, they are data. *Radiology*, 278(2):563–577.
- Goldman, E. (2021). How the unvaccinated threaten the vaccinated for COVID-19: A darwinian perspective. *Proceedings of the National Academy of Sciences*, 118(39):e2114279118.
- Goldman, L. W. (2007). Principles of CT and CT technology. *Journal of Nuclear Medicine Technology*, 35(3):115–128.

Guiot, J., Vaidyanathan, A., Deprez, L., Zerka, F., Danthine, D., Frix, A.-N., Thys, M., Henket, M., Canivet, G., Mathieu, S., Eftaxia, E., Lambin, P., Tsoutzidis, N., Miraglio, B., Walsh, S., Moutschen, M., Louis, R., Meunier, P., Vos, W., Leijenaar, R. T. H., and Lovinfosse, P. (2020). Development and validation of an automated radiomic CT signature for detecting COVID-19. *Diagnostics*, 11(1):41.

Gülbay, M., Özbay, B. O., Mendi, B. A. R., Baştuğ, A., and Bodur, H. (2021). A CT radiomics analysis of COVID-19-related ground-glass opacities and consolidation: Is it valuable in a differential diagnosis with other atypical pneumonias? *PLOS ONE*, 16(3):e0246582.

Heidarian, S., Afshar, P., Enshaei, N., Naderkhani, F., Rafiee, M. J., Fard, F. B., Samimi, K., Atashzar, S. F., Oikonomou, A., Plataniotis, K. N., and Mohammadi, A. (2021). COVID-FACT: A fully-automated capsule network-based framework for identification of COVID-19 cases from chest CT scans. *Frontiers in Artificial Intelligence*, 4.

Hofmanninger, J., Prayer, F., Pan, J., Röhrich, S., Prosch, H., and Langs, G. (2020). Automatic lung segmentation in routine imaging is primarily a data diversity problem, not a methodology problem. *European Radiology Experimental*, 4(1).

Hu, L. and Wang, C. (2020). Radiological role in the detection, diagnosis and monitoring for the coronavirus disease 2019 (covid-19). *European Review for Medical and Pharmacological Sciences*, 24(8):4523–4528.

Huang, J., Wu, F., Chen, L., Yu, J., Sun, W., Nie, Z., Liu, H., Yang, F., and Zheng, C. (2021a). CT-based radiomics helps to predict residual lung lesions in COVID-19 patients at three months after discharge. *Diagnostics*, 11(10):1814.

Huang, Y., Zhang, Z., Liu, S., Li, X., Yang, Y., Ma, J., Li, Z., Zhou, J., Jiang, Y., and He, B. (2021b). CT-based radiomics combined with signs: a valuable tool to help radiologist discriminate COVID-19 and influenza pneumonia. *BMC Medical Imaging*, 21(1).

Idris, M. and Hacking, C. (2017). Radiomics. <https://radiopaedia.org/articles/56602>. Accessed: 2021-11-25.

Johnson, H. J., McCormick, M., Ibáñez, L., and Consortium, T. I. S. (2013). *The ITK Software Guide*. Kitware, Inc., third edition. *In press*.

Kang, M., Hong, K. S., Chikontwe, P., Luna, M., Jang, J. G., Park, J., Shin, K.-C., Park, S. H., and Ahn, J. H. (2021). Quantitative assessment of chest CT patterns in COVID-19 and bacterial pneumonia patients: a deep learning perspective. *Journal of Korean Medical Science*, 36(5).

Ke, Z., Li, L., Wang, L., Liu, H., Lu, X., Zeng, F., and Zha, Y. (2021). Radiomics analysis enables fatal outcome prediction for hospitalized patients with coronavirus disease 2019 (COVID-19). *Acta Radiologica*, 63(3):319–327.

Kocak, B., , Durmaz, E. S., Ates, E., Kilickesmez, O., , and and (2019). Radiomics with artificial intelligence: a practical guide for beginners. *Diagnostic and Interventional Radiology*, 25(6):485–495.

Kolinger, G. D., García, D. V., Kramer, G. M., Frings, V., Zwezerijnen, G. J. C., Smit, E. F., Langen, A. J. D., Buvat, I., and Boellaard, R. (2021). Effects of tracer uptake time in non-small cell lung cancer 18f-FDG PET radiomics. *Journal of Nuclear Medicine*, page jnumed.121.262660.

Kucirka, L. M., Lauer, S. A., Laeyendecker, O., Boon, D., and Lessler, J. (2020). Variation in false-negative rate of reverse transcriptase polymerase chain reaction–based SARS-CoV-2 tests by time since exposure. *Annals of Internal Medicine*, 173(4):262–267.

Lambin, P., Leijenaar, R. T., Deist, T. M., Peerlings, J., de Jong, E. E., van Timmeren, J., Sanduleanu, S., Larue, R. T., Even, A. J., Jochems, A., van Wijk, Y., Woodruff, H., van Soest, J., Lustberg, T., Roelofs, E., van Elmpt, W., Dekker, A., Mottaghy, F. M., Wildberger, J. E., and Walsh, S. (2017). Radiomics: the bridge between medical imaging and personalized medicine. *Nature Reviews Clinical Oncology*, 14(12):749–762.

Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G., Granton, P., Zegers, C. M., Gillies, R., Boellard, R., Dekker, A., and Aerts, H. J. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer*, 48(4):441–446.

Li, C., Dong, D., Li, L., Gong, W., Li, X., Bai, Y., Wang, M., Hu, Z., Zha, Y., and Tian, J. (2020). Classification of severe and critical covid-19 using deep learning and radiomics. *IEEE Journal of Biomedical and Health Informatics*, 24(12):3585–3594.

Li, L., Wang, L., Zeng, F., Peng, G., Ke, Z., Liu, H., and Zha, Y. (2021a). Development and multicenter validation of a CT-based radiomics signature for predicting severe COVID-19 pneumonia. *European Radiology*, 31(10):7901–7912.

Li, W., Cao, Y., Yu, K., Cai, Y., Huang, F., Yang, M., and Xie, W. (2021b). Pulmonary lesion subtypes recognition of COVID-19 from radiomics data with three-dimensional texture characterization in computed tomography images. *BioMedical Engineering OnLine*, 20(1).

Li, Y. and Xia, L. (2020). Coronavirus disease 2019 (COVID-19): Role of chest CT in diagnosis and management. *American Journal of Roentgenology*, 214(6):1280–1286.

Lin, L., Liu, J., Deng, Q., Li, N., Pan, J., Sun, H., and Quan, S. (2021). Radiomics is effective for distinguishing coronavirus disease 2019 pneumonia from influenza virus pneumonia. *Frontiers in Public Health*, 9.

Little, B. P. (2015). Approach to chest computed tomography. *Clinics in Chest Medicine*, 36(2):127–145.

Liu, H., Ren, H., Wu, Z., Xu, H., Zhang, S., Li, J., Hou, L., Chi, R., Zheng, H., Chen, Y., Duan, S., Li, H., Xie, Z., and Wang, D. (2021). CT radiomics facilitates more accurate diagnosis of COVID-19 pneumonia: compared with CO-RADS. *Journal of Translational Medicine*, 19(1).

Liu, L., Lei, X., Xiao, X., Yang, J., Li, J., Ji, M., Du, W., Tan, H., Zhu, J., Li, B., Jin, Z., Liu, W., Wu, J., and Liu, Z. (2020). Epidemiological and clinical characteristics of patients with coronavirus disease-2019 in shiyan city, china. *Frontiers in Cellular and Infection Microbiology*, 10.

Liu, Z., Wang, S., Dong, D., Wei, J., Fang, C., Zhou, X., Sun, K., Li, L., Li, B., Wang, M., and Tian, J. (2019). The applications of radiomics in precision diagnosis and treatment of oncology: Opportunities and challenges. *Theranostics*, 9(5):1303–1322.

Maguolo, G. and Nanni, L. (2021). A critic evaluation of methods for COVID-19 automatic detection from x-ray images. *Information Fusion*, 76:1–7.

Mahesh, M. (2013). The essential physics of medical imaging, third edition. *Medical Physics*, 40(7):077301.

Molina, D., Pérez-Beteta, J., Martínez-González, A., Martino, J., Velásquez, C., Arana, E., and Pérez-García, V. M. (2016). Influence of gray level and space discretization on brain tumor heterogeneity measures obtained from magnetic resonance images. *Computers in Biology and Medicine*, 78:49–57.

Mu, W., Jiang, L., Zhang, J., Shi, Y., Gray, J. E., Tunali, I., Gao, C., Sun, Y., Tian, J., Zhao, X., Sun, X., Gillies, R. J., and Schabath, M. B. (2020). Non-invasive decision support for NSCLC treatment using PET/CT radiomics. *Nature Communications*, 11(1).

Novel Coronavirus Pneumonia Emergency Response Epidemiology Team (2020). The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) — china, 2020. *China CDC Weekly*, 2(8):113–122.

Paluru, N., Dayal, A., Jenssen, H. B., Sakinis, T., Cenkeramaddi, L. R., Prakash, J., and Yalavarthy, P. K. (2021). Anam-net: Anamorphic depth embedding-based lightweight cnn for segmentation of anomalies in covid-19 chest ct images. *IEEE Transactions on Neural Networks and Learning Systems*, 32(3):932–946.

Pang, T., Wong, J. H. D., Ng, W. L., and Chan, C. S. (2021). Semi-supervised GAN-based radiomics model for data augmentation in breast ultrasound mass classification. *Computer Methods and Programs in Biomedicine*, 203:106018.

Parekh, V. and Jacobs, M. A. (2016). Radiomics: a new application from established techniques. *Expert Review of Precision Medicine and Drug Development*, 1(2):207–226.

Peng, S., Pan, L., Guo, Y., Gong, B., Huang, X., Liu, S., Huang, J., Pu, H., and Zeng, J. (2022). Quantitative CT imaging features for COVID-19 evaluation: The ability to differentiate COVID-19 from non- COVID-19 (highly suspected) pneumonia patients during the epidemic period. *PLOS ONE*, 17(1):e0256194.

Pizzi, A. D., Chiarelli, A. M., Chiacchiaretta, P., Valdesi, C., Croce, P., Mastrodicasa, D., Villani, M., Trebeschi, S., Serafini, F. L., Rosa, C., Cocco, G., Luberti, R., Conte, S., Maz-zamurro, L., Mereu, M., Patea, R. L., Panara, V., Marinari, S., Vecchiet, J., and Caulo, M. (2021). Radiomics-based machine learning differentiates “ground-glass” opacities due to COVID-19 from acute non-COVID-19 lung disease. *Scientific Reports*, 11(1).

Powers, D. M. W. (2020). Evaluation: from precision, recall and f-measure to roc, informed-ness, markedness and correlation. *CoRR*, abs/2010.16061.

Purkayastha, S., Xiao, Y., Jiao, Z., Thepumnoeysuk, R., Halsey, K., Wu, J., Tran, T. M. L., Hsieh, B., Choi, J. W., Wang, D., Vallières, M., Wang, R., Collins, S., Feng, X., Feldman, M., Zhang, P. J., Atalay, M., Sebro, R., Yang, L., Fan, Y., hua Liao, W., and Bai, H. X. (2021). Machine learning-based prediction of COVID-19 severity and progression to critical illness using CT imaging and clinical data. *Korean Journal of Radiology*, 22(7):1213.

Putatunda, S. and Rama, K. (2018). A comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of XGBoost. In *Proceedings of the 2018 International Conference on Signal Processing and Machine Learning - SPML '18*. ACM Press.

Qiu, J., Peng, S., Yin, J., Wang, J., Jiang, J., Li, Z., Song, H., and Zhang, W. (2021). A radiomics signature to quantitatively analyze COVID-19-infected pulmonary lesions. *Inter-disciplinary Sciences: Computational Life Sciences*, 13(1):61–72.

Rezaeiyo, S. M., Abedi-Firouzjah, R., Ghorvei, M., and Sarnameh, S. (2021). Screening of COVID-19 based on the extracted radiomics features from chest CT images. *Journal of X-Ray Science and Technology*, 29(2):229–243.

Rizzo, S., Botta, F., Raimondi, S., Origgi, D., Fanciullo, C., Morganti, A. G., and Bellomi, M. (2018). Radiomics: the facts and the challenges of image analysis. *European Radiology Experimental*, 2(1).

Rubin, G. D., Ryerson, C. J., Haramati, L. B., Sverzellati, N., Kanne, J. P., Raouf, S., Schluger, N. W., Volpi, A., Yim, J.-J., Martin, I. B., Anderson, D. J., Kong, C., Altes, T., Bush, A., Desai, S. R., Goldin, J., Goo, J. M., Humbert, M., Inoue, Y., Kauczor, H.-U., Luo, F., Mazzone, P. J., Prokop, M., Remy-Jardin, M., Richeldi, L., Schaefer-Prokop, C. M., Tomiyama, N., Wells, A. U., and Leung, A. N. (2020). The role of chest imaging in patient management during the COVID-19 pandemic. *Chest*, 158(1):106–116.

- Santone, A., Belfiore, M. P., Mercaldo, F., Varriano, G., and Brunese, L. (2021). On the adoption of radiomics and formal methods for COVID-19 coronavirus diagnosis. *Diagnostics*, 11(2):293.
- Shi, F., Wang, J., Shi, J., Wu, Z., Wang, Q., Tang, Z., He, K., Shi, Y., and Shen, D. (2021a). Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for COVID-19. *IEEE Reviews in Biomedical Engineering*, 14:4–15.
- Shi, F., Xia, L., Shan, F., Song, B., Wu, D., Wei, Y., Yuan, H., Jiang, H., He, Y., Gao, Y., Sui, H., and Shen, D. (2021b). Large-scale screening to distinguish between COVID-19 and community-acquired pneumonia using infection size-aware classification. *Physics in Medicine & Biology*, 66(6):065031.
- Shi, H., Han, X., Jiang, N., Cao, Y., Alwalid, O., Gu, J., Fan, Y., and Zheng, C. (2020). Radiological findings from 81 patients with COVID-19 pneumonia in wuhan, china: a descriptive study. *The Lancet Infectious Diseases*, 20(4):425–434.
- Shi, H., Xu, Z., Cheng, G., Ji, H., He, L., Zhu, J., Hu, H., Xie, Z., Ao, W., and Wang, J. (2022). CT-based radiomic nomogram for predicting the severity of patients with COVID-19. *European Journal of Medical Research*, 27(1).
- Shiri, I., Sorouri, M., Geramifar, P., Nazari, M., Abdollahi, M., Salimi, Y., Khosravi, B., Askari, D., Aghaghazvini, L., Hajianfar, G., Kasaeian, A., Abdollahi, H., Arabi, H., Rahmim, A., Radmard, A. R., and Zaidi, H. (2021). Machine learning-based prognostic modeling using clinical data and quantitative radiomic features from chest CT images in COVID-19 patients. *Computers in Biology and Medicine*, 132:104304.
- Shur, J. D., Doran, S. J., Kumar, S., ap Dafydd, D., Downey, K., O'Connor, J. P. B., Papanikolaou, N., Messiou, C., Koh, D.-M., and Orton, M. R. (2021). Radiomics in oncology: A practical guide. *RadioGraphics*, 41(6):1717–1732.
- Soleymani, Y., Jahanshahi, A. R., Hefzi, M., Ghaziani, M. F., Pourfarshid, A., and Khezerloo, D. (2021). Evaluation of textural-based radiomics features for differentiation of COVID-19 pneumonia from non-COVID pneumonia. *Egyptian Journal of Radiology and Nuclear Medicine*, 52(1).
- Sousa, G. J. B., Garces, T. S., Cestari, V. R. F., Moreira, T. M. M., Florêncio, R. S., and Pereira, M. L. D. (2020). Estimation and prediction of COVID-19 cases in brazilian metropolises. *Revista Latino-Americana de Enfermagem*, 28.
- Stefano, A. and Comelli, A. (2021). Customized efficient neural network for COVID-19 infected region identification in CT images. *Journal of Imaging*, 7(8):131.

Sun, L., Mo, Z., Yan, F., Xia, L., Shan, F., Ding, Z., Song, B., Gao, W., Shao, W., Shi, F., Yuan, H., Jiang, H., Wu, D., Wei, Y., Gao, Y., Sui, H., Zhang, D., and Shen, D. (2020). Adaptive feature selection guided deep forest for COVID-19 classification with chest CT. *IEEE Journal of Biomedical and Health Informatics*, 24(10):2798–2805.

Tabatabaei, M., Tasorian, B., Goyal, M., Moini, A., and Sotoudeh, H. (2021). Feasibility of radiomics to differentiate coronavirus disease 2019 (covid-19) from h1n1 influenza pneumonia on chest computed tomography: A proof of concept. *Iranian Journal of Medical Sciences*, 46(6).

Tan, H.-B., Xiong, F., Jiang, Y.-L., Huang, W.-C., Wang, Y., Li, H.-H., You, T., Fu, T.-T., Lu, R., and Peng, B.-W. (2020). The study of automatic machine learning base on radiomics of non-focus area in the first chest CT of different clinical types of COVID-19 pneumonia. *Scientific Reports*, 10(1).

Tang, X. (2020). The role of artificial intelligence in medical imaging research. *British Journal of Radiology*, 2(1):20190031.

Tang, Z., Zhao, W., Xie, X., Zhong, Z., Shi, F., Ma, T., Liu, J., and Shen, D. (2021). Severity assessment of COVID-19 using CT image features and laboratory indices. *Physics in Medicine & Biology*, 66(3):035015.

Tian, S., Hu, N., Lou, J., Chen, K., Kang, X., Xiang, Z., Chen, H., Wang, D., Liu, N., Liu, D., Chen, G., Zhang, Y., Li, D., Li, J., Lian, H., Niu, S., Zhang, L., and Zhang, J. (2020). Characteristics of COVID-19 infection in beijing. *Journal of Infection*, 80(4):401–406.

van Griethuysen, J. J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G., Fillion-Robin, J.-C., Pieper, S., and Aerts, H. J. (2017a). Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, 77(21):e104–e107.

van Griethuysen, J. J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G., Fillion-Robin, J.-C., Pieper, S., and Aerts, H. J. (2017b). Computational radiomics system to decode the radiographic phenotype. *Cancer Research*, 77(21):e104–e107.

van Timmeren, J. E., Cester, D., Tanadini-Lang, S., Alkadhi, H., and Baessler, B. (2020). Radiomics in medical imaging—“how-to” guide and critical reflection. *Insights into Imaging*, 11(1).

Velichko, E., Shariaty, F., Orooji, M., Pavlov, V., Pervunina, T., Zavjalov, S., Khazaei, R., and Radmard, A. R. (2022). Development of computer-aided model to differentiate COVID-19 from pulmonary edema in lung CT scan: EDECOVID-net. *Computers in Biology and Medicine*, 141:105172.

- Wang, H., Wang, L., Lee, E. H., Zheng, J., Zhang, W., Halabi, S., Liu, C., Deng, K., Song, J., and Yeom, K. W. (2020). Decoding COVID-19 pneumonia: comparison of deep learning and radiomics CT image signatures. *European Journal of Nuclear Medicine and Molecular Imaging*, 48(5):1478–1486.
- Wang, H. Y. C., Donovan, E. M., Nisbet, A., South, C. P., Alobaidli, S., Ezhil, V., Phillips, I., Prakash, V., Ferreira, M., Webster, P., and Evans, P. M. (2019). The stability of imaging biomarkers in radiomics: a framework for evaluation. *Physics in Medicine & Biology*, 64(16):165012.
- Wang, L., Kelly, B., Lee, E. H., Wang, H., Zheng, J., Zhang, W., Halabi, S., Liu, J., Tian, Y., Han, B., Huang, C., Yeom, K. W., Deng, K., and Song, J. (2021a). Multi-classifier-based identification of COVID-19 from chest computed tomography using generalizable and interpretable radiomics features. *European Journal of Radiology*, 136:109552.
- Wang, S., Dong, D., Li, L., Li, H., Bai, Y., Hu, Y., Huang, Y., Yu, X., Liu, S., Qiu, X., Lu, L., Wang, M., Zha, Y., and Tian, J. (2021b). A deep learning radiomics model to identify poor outcome in COVID-19 patients with underlying health conditions: A multicenter study. *IEEE Journal of Biomedical and Health Informatics*, 25(7):2353–2362.
- Wang, X.-H., Xu, X., Ao, Z., Duan, J., Han, X., Tang, X., Fu, Y.-F., Wu, X.-S., Wang, X., Zhu, L., Zeng, W., and Guo, S. (2021c). Elaboration of a radiomics strategy for the prediction of the re-positive cases in the discharged patients with COVID-19. *Frontiers in Medicine*, 8.
- Wei, W., wen Hu, X., Cheng, Q., ming Zhao, Y., and qiong Ge, Y. (2020). Identification of common and severe COVID-19: the value of CT texture analysis and correlation with clinical characteristics. *European Radiology*, 30(12):6788–6796.
- Wilson, R. and Devaraj, A. (2017). Radiomics of pulmonary nodules and lung cancer. *Translational Lung Cancer Research*, 6(1):86–91.
- World Health Organization (2020). 73rd world health assembly decisions. <https://www.who.int/news/item/07-08-2020-73rd-world-health-assembly-decisions>. Accessed: 2022-02-12.
- World Health Organization (2022a). Coronavirus disease (covid-19) symptoms. <https://www.who.int/health-topics/coronavirus>. Accessed: 2022-01-05.
- World Health Organization (2022b). Tracking sars-cov-2 variants. <https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/>. Accessed: 2022-02-12.
- World Health Organization (2022c). Vaccines and immunization. <https://www.who.int/health-topics/vaccines-and-immunization>. Accessed: 2022-02-12.
- Wu, Q., Wang, S., Li, L., Wu, Q., Qian, W., Hu, Y., Li, L., Zhou, X., Ma, H., Li, H., Wang, M., Qiu, X., Zha, Y., and Tian, J. (2020). Radiomics analysis of computed tomography helps predict poor prognostic outcome in COVID-19. *Theranostics*, 10(16):7231–7244.

- Xie, C., Ng, M.-Y., Ding, J., Leung, S. T., Lo, C. S. Y., Wong, H. Y. F., and Vardhanabhuti, V. (2020a). Discrimination of pulmonary ground-glass opacity changes in COVID-19 and non-COVID-19 patients using CT radiomics analysis. *European Journal of Radiology Open*, 7:100271.
- Xie, X., Zhong, Z., Zhao, W., Zheng, C., Wang, F., and Liu, J. (2020b). Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: Relationship to negative RT-PCR testing. *Radiology*, 296(2):E41–E45.
- Xie, Z., Sun, H., Wang, J., Xu, H., Li, S., Zhao, C., Gao, Y., Wang, X., Zhao, T., Duan, S., Hu, C., and Ao, W. (2021). A novel CT-based radiomics in the distinction of severity of coronavirus disease 2019 (COVID-19) pneumonia. *BMC Infectious Diseases*, 21(1).
- Xin, X., Mo, R., Shao, M., Yang, W., Li, D., Zhang, Y., Wang, H., Liu, B., Tian, S., Chen, W., Wu, J., Zhu, B., Zhou, K., Du, C., and Zhang, B. (2021). A human-computer collaboration for COVID-19 differentiation: combining a radiomics model with deep learning and human auditing. *Annals of Palliative Medicine*, 10(7):7329–7339.
- Xiong, F., Wang, Y., You, T., Han Li, H., Ting Fu, T., Tan, H., Huang, W., and Jiang, Y. (2021). The clinical classification of patients with COVID-19 pneumonia was predicted by radiomics using chest CT. *Medicine*, 100(12):e25307.
- Xu, Q., Zhan, X., Zhou, Z., Li, Y., Xie, P., Zhang, S., Li, X., Yu, Y., Zhou, C., Zhang, L., Gevaert, O., and Lu, G. (2021a). AI-based analysis of CT images for rapid triage of COVID-19 patients. *npj Digital Medicine*, 4(1).
- Xu, Z., Zhao, L., Yang, G., Ren, Y., Wu, J., Xia, Y., Yang, X., Cao, M., Zhang, G., Peng, T., Zhao, J., Yang, H., Hu, J., and Du, J. (2021b). Severity assessment of COVID-19 using a CT-based radiomics model. *Stem Cells International*, 2021:1–10.
- Yang, N., Liu, F., Li, C., Xiao, W., Xie, S., Yuan, S., Zuo, W., Ma, X., and Jiang, G. (2021). Diagnostic classification of coronavirus disease 2019 (COVID-19) and other pneumonias using radiomics features in CT chest images. *Scientific Reports*, 11(1).
- Yoon, Y., Hwang, T., Choi, H., and Lee, H. (2019). Classification of radiographic lung pattern based on texture analysis and machine learning. *Journal of Veterinary Science*, 20(4).
- Yue, H., Yu, Q., Liu, C., Huang, Y., Jiang, Z., Shao, C., Zhang, H., Ma, B., Wang, Y., Xie, G., Zhang, H., Li, X., Kang, N., Meng, X., Huang, S., Xu, D., Lei, J., Huang, H., Yang, J., Ji, J., Pan, H., Zou, S., Ju, S., and Qi, X. (2020). Machine learning-based CT radiomics method for predicting hospital stay in patients with pneumonia associated with SARS-CoV-2 infection: a multicenter study. *Annals of Translational Medicine*, 8(14):859–859.

Zeng, Q.-Q., Zheng, K. I., Chen, J., Jiang, Z.-H., Tian, T., Wang, X.-B., Ma, H.-L., Pan, K.-H., Yang, Y.-J., Chen, Y.-P., and Zheng, M.-H. (2020). Radiomics-based model for accurately distinguishing between severe acute respiratory syndrome associated coronavirus 2 (SARS-CoV-2) and influenza a infected pneumonia. *MedComm*, 1(2):240–248.

Zhang, B., Tian, J., Dong, D., Gu, D., Dong, Y., Zhang, L., Lian, Z., Liu, J., Luo, X., Pei, S., Mo, X., Huang, W., Ouyang, F., Guo, B., Liang, L., Chen, W., Liang, C., and Zhang, S. (2017). Radiomics features of multiparametric MRI as novel prognostic factors in advanced nasopharyngeal carcinoma. *Clinical Cancer Research*, 23(15):4259–4269.

Zhang, B., yi-di-li Ni-jia Ti, M., Yan, R., An, N., Chen, L., Liu, S., Chen, L., Chen, Q., Li, M., Chen, Z., You, J., Dong, Y., Xiong, Z., and Zhang, S. (2021a). CT-based radiomics for predicting the rapid progression of coronavirus disease 2019 (COVID-19) pneumonia lesions. *The British Journal of Radiology*, 94(1122):20201007.

Zhang, M., Zeng, X., Huang, C., Liu, J., Liu, X., Xie, X., and Wang, R. (2021b). An AI-based radiomics nomogram for disease prognosis in patients with COVID-19 pneumonia using initial CT images and clinical indicators. *International Journal of Medical Informatics*, 154:104545.

Zhang, X., Wang, D., Shao, J., Tian, S., Tan, W., Ma, Y., Xu, Q., Ma, X., Li, D., Chai, J., Wang, D., Liu, W., Lin, L., Wu, J., Xia, C., and Zhang, Z. (2021c). A deep learning integrated radiomics model for identification of coronavirus disease 2019 using computed tomography. *Scientific Reports*, 11(1).

Zhao, C., Xu, Y., He, Z., Tang, J., Zhang, Y., Han, J., Shi, Y., and Zhou, W. (2021). Lung segmentation and automatic detection of COVID-19 using radiomic features from chest CT images. *Pattern Recognition*, 119:108071.

Zhu, Q., Ye, H., Sun, L., Li, Z., Wang, R., Shi, F., Shen, D., and Zhang, D. (2021). GACDN: generative adversarial feature completion and diagnosis network for COVID-19. *BMC Medical Imaging*, 21(1).

**ANNEX A – Consubstantiated opinion of the research ethics
committee**

PONTIFÍCIA UNIVERSIDADE
CATÓLICA DO RIO GRANDE
DO SUL - PUC/RS



PARECER CONSUBSTANCIADO DO CEP

DADOS DA EMENDA

Título da Pesquisa: Análise da assinatura radiômica em imagens de tomografia computadorizada na pneumonia por COVID-19

Pesquisador: Ana Maria Marques da Silva

Área Temática:

Versão: 3

CAAE: 30791720.5.0000.5336

Instituição Proponente: UNIAO BRASILEIRA DE EDUCACAO E ASSISTENCIA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 4.850.213

Apresentação do Projeto:

O pesquisador principal Ana Maria Marques da Silva, responsável pelo projeto com número de CAAE 30791720.5.0000.5336 e Título: Análise da assinatura radiômica em imagens de tomografia computadorizada na pneumonia por COVID-19 encaminhou ao CEP-PUCRS emenda contendo os seguintes documentos:

- 1) ProjetoRadiomicaCOVIDemenda1.docx
- 2) ProjetoRadiomicaCOVIDemenda1.pdf
- 3) ProjetoRadiomicaCOVIDemenda1comalteracoes.docx
- 4) TCUEmenda1p1de2.pdf
- 5) TCUEmenda1p2de2.pdf
- 6) AnuenciaUFU.pdf
- 7) LattesPesquisadores.docx
- 8) OrcamentoEmenda1.pdf
- 9) FolhadeRostoEmenda1.pdf

Objetivo da Pesquisa:

O pesquisador principal Ana Maria Marques da Silva, responsável pelo projeto com número de CAAE 30791720.5.0000.5336 e Título: Análise da assinatura radiômica em imagens de tomografia computadorizada na pneumonia por COVID-19 encaminhou ao CEP-PUCRS emenda contendo os

Endereço: Av.Ipiranga, 6681, prédio 50, sala 703

Bairro: Partenon

CEP: 90.619-900

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3320-3345

Fax: (51)3320-3345

E-mail: cep@pucrs.br

PONTIFÍCIA UNIVERSIDADE
CATÓLICA DO RIO GRANDE
DO SUL - PUC/RS



Continuação do Parecer: 4.850.213

seguintes documentos:

- 1) ProjetoRadiomicaCOVIDemenda1.docx
- 2) ProjetoRadiomicaCOVIDemenda1.pdf
- 3) ProjetoRadiomicaCOVIDemenda1comalteracoes.docx
- 4) TCUDEmenda1p1de2.pdf
- 5) TCUDEmenda1p2de2.pdf
- 6) AnuenciaUFU.pdf
- 7) LattesPesquisadores.docx
- 8) OrcamentoEmenda1.pdf
- 9) FolhadeRostoEmenda1.pdf

Avaliação dos Riscos e Benefícios:

O pesquisador principal Ana Maria Marques da Silva, responsável pelo projeto com número de CAAE 30791720.5.0000.5336 e Título: Análise da assinatura radiômica em imagens de tomografia computadorizada na pneumonia por COVID-19 encaminhou ao CEP-PUCRS emenda contendo os seguintes documentos:

- 1) ProjetoRadiomicaCOVIDemenda1.docx
- 2) ProjetoRadiomicaCOVIDemenda1.pdf
- 3) ProjetoRadiomicaCOVIDemenda1comalteracoes.docx
- 4) TCUDEmenda1p1de2.pdf
- 5) TCUDEmenda1p2de2.pdf
- 6) AnuenciaUFU.pdf
- 7) LattesPesquisadores.docx
- 8) OrcamentoEmenda1.pdf
- 9) FolhadeRostoEmenda1.pdf

Comentários e Considerações sobre a Pesquisa:

O pesquisador principal Ana Maria Marques da Silva, responsável pelo projeto com número de CAAE 30791720.5.0000.5336 e Título: Análise da assinatura radiômica em imagens de tomografia computadorizada na pneumonia por COVID-19 encaminhou ao CEP-PUCRS emenda contendo os seguintes documentos:

- 1) ProjetoRadiomicaCOVIDemenda1.docx
- 2) ProjetoRadiomicaCOVIDemenda1.pdf
- 3) ProjetoRadiomicaCOVIDemenda1comalteracoes.docx
- 4) TCUDEmenda1p1de2.pdf

Endereço: Av.Ipiranga, 6681, prédio 50, sala 703

Bairro: Partenon

CEP: 90.619-900

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3320-3345

Fax: (51)3320-3345

E-mail: cep@pucrs.br

**PONTIFÍCIA UNIVERSIDADE
CATÓLICA DO RIO GRANDE
DO SUL - PUC/RS**



Continuação do Parecer: 4.850.213

- 5) TCUDEmenda1p2de2.pdf
- 6) AnuenciaUFU.pdf
- 7) LattesPesquisadores.docx
- 8) OrcamentoEmenda1.pdf
- 9) FolhadeRostoEmenda1.pdf

Considerações sobre os Termos de apresentação obrigatória:

Todos os termos foram apresentados.

Conclusões ou Pendências e Lista de Inadequações:

Não há pendências.

Considerações Finais a critério do CEP:

Diante do exposto, o CEP-PUCRS, de acordo com suas atribuições definidas na Resolução CNS n° 466 de 2012 e a Norma Operacional n° 001 de 2013 do CNS, manifesta-se pela aprovação da emenda ao projeto de pesquisa Análise da assinatura radiômica em imagens de tomografia computadorizada na pneumonia por COVID-19 proposto pela pesquisadora Ana Maria Marques da Silva com número de CAAE 30791720.5.0000.5336.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_1770290_E1.pdf	25/06/2021 14:28:30		Aceito
Outros	CartaRespostaPendenciaEmenda1.pdf	25/06/2021 14:27:51	Ana Maria Marques da Silva	Aceito
Folha de Rosto	folhaDeRostoEmenda1.pdf	25/06/2021 14:18:18	Ana Maria Marques da Silva	Aceito
Outros	LattesPesquisadores.docx	07/06/2021 20:52:32	Ana Maria Marques da Silva	Aceito
Orçamento	OrcamentoEmenda1.pdf	07/06/2021 20:49:37	Ana Maria Marques da Silva	Aceito
Outros	CartaApresentacaoEmenda1.docx	07/06/2021 20:49:04	Ana Maria Marques da Silva	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCUDEmenda1p1de2.pdf	07/06/2021 20:47:03	Ana Maria Marques da Silva	Aceito
TCLE / Termos de Assentimento / Justificativa de	TCUDEmenda1p2de2.pdf	07/06/2021 20:46:48	Ana Maria Marques da Silva	Aceito

Endereço: Av.Ipiranga, 6681, prédio 50, sala 703

Bairro: Partenon

CEP: 90.619-900

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3320-3345

Fax: (51)3320-3345

E-mail: cep@puhrs.br

PONTIFÍCIA UNIVERSIDADE
CATÓLICA DO RIO GRANDE
DO SUL - PUC/RS



Continuação do Parecer: 4.850.213

Ausência	TCUDEMenda1p2de2.pdf	07/06/2021 20:46:48	Ana Maria Marques da Silva	Aceito
Brochura Pesquisa	ProjetoRadiomicaCOVIDemenda1.pdf	07/06/2021 20:45:34	Ana Maria Marques da Silva	Aceito
Projeto Detalhado / Brochura Investigador	ProjetoRadiomicaCOVIDemenda1.docx	07/06/2021 20:45:04	Ana Maria Marques da Silva	Aceito
Declaração de Instituição e Infraestrutura	AnuenciaUFU.pdf	07/06/2021 20:44:37	Ana Maria Marques da Silva	Aceito
Projeto Detalhado / Brochura Investigador	ProjetoRadiomicaCOVIDemenda1comalteracoes.docx	07/06/2021 20:43:49	Ana Maria Marques da Silva	Aceito
Solicitação Assinada pelo Pesquisador Responsável	CartaApresentacaoEmenda1.pdf	07/06/2021 20:43:34	Ana Maria Marques da Silva	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

PORTO ALEGRE, 15 de Julho de 2021

Assinado por:

Paulo Vinicius Sporleder de Souza
(Coordenador(a))

Endereço: Av.Ipiranga, 6681, prédio 50, sala 703

Bairro: Partenon

CEP: 90.619-900

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3320-3345

Fax: (51)3320-3345

E-mail: cep@pucrs.br

ANNEX B – Abbreviations for features shown in feature importance

Table B.1: Abbreviations for the most important features on the test set

Abbreviation	Lung side - Filter - Feature group - Feature
L-lbp3D-1st-M	Left - local binary pattern 3D - first order - Median
L-lbp3D-glrlm-GLNN	Left - local binary pattern 3D - GLCM - Inverse Difference Moment Normalized
L-log1-1st-10P	Left - LoG(1mm) - first order - 10% percentile
L-log1-gldm-LDLGLEs	Left - LoG(1mm) - GLDM - Large Dependence Low Gray Level Emphasis
L-log1-glrlm-HGLRE	Left - LoG(1mm) - GLRLM - High Gray Level Run Emphasis
L-log3-gldm-DE	Left - LoG(3mm) - GLDM - Dependence Entropy
L-log3-glrlm-LRLGLE	Left - LoG(3mm) - GLRLM - Long Run Low Gray Level Emphasis
L-log5-glrlm-RE	Left - LoG(5mm) - GLRLM - Run entropy
L-logar-1st-R	Left - logarithm - first order - Range
L-logar-glszm-GLNU	Left - logarithm - GLSZM - Gray Level Non-Uniformity
L-orig-glcm-DE	Left - original - GLCM - Difference Entropy
L-orig-glcm-IDMN	Left - original - GLCM - Inverse Difference Moment Normalized
L-orig-glcm-SS	Left - original - GLCM - Sum of Squares
L-orig-shape-M2DDR	Left - original - shape - Maximum 2D Diameter (Row)
L-orig-shape-MAL	Left - original - shape - Minor Axis Length
L-orig-shape-SA	Left - original - shape - Surface Area
L-sqroot-glrlm-LRLGLE	Left - square root - GLRLM - Long Run Low Gray Level Emphasis
L-squ-1st-MAD	Left - square - first order - Mean Absolute Deviation
L-wavHHH-1st-RMS	Left - wavelet (HHH) - first order - Root Mean Squared
L-wavHHL-gldm-DNU	Left - wavelet (HHL) - GLDM - Dependence Non-Uniformity
L-wavHLH-1st-E	Left - wavelet (HLH) - first order - Energy
L-wavHLH-gldm-DE	Left - wavelet (HLH) - GLDM - Dependence Entropy
L-wavHLH-glrlm-LGLRE	Left - wavelet (HLH) - GLRLM - Low Gray Level Run Emphasis
L-wavHLH-glszm-GLNU	Left - wavelet (HLH) - GLSZM - Gray Level Non-Uniformity
L-wavLHH-glszm-GLV	Left - wavelet (LHH) - GLSZM - Gray Level Variance
L-wavLHH-glszm-HGLZE	Left - wavelet (LHH) - GLSZM - High Gray Level Zone Emphasis
L-wavLHL-1st-S	Left - wavelet (LHL) - first order - Skewness
L-wavLHL-glszm-SALGLE	Left - wavelet (LHL) - GLSZM - Small Area Low Gray Level Emphasis
L-wavLLL-gldm-LGLE	Left - wavelet (LLL) - GLDM - Low Gray Level Emphasis
R-exp-gldm-DNU	Right - exponential - GLDM - Dependence Non-Uniformity
R-grad-glrlm-LRHGLE	Right - gradient - GLRLM - Long Run High Gray Level Emphasis
R-lbp3D-glcm-C	Right - local binary pattern 3D - GLCM - Contrast
R-log1-glrlm-HGLRE	Right - laplacian ofgaussian (1mm) -GLRLM - High Gray Level Run Emphasis
R-log3-glcm-IMC1	Right - LoG(3mm) - GLCM - Informational Measure of Correlation 1
R-log3-gldm-LDLGLE	Right - LoG(3mm) - GLDM - Large Dependence Low Gray Level Emphasis
R-log3-glrlm-RLNN	Right - LoG(3mm) - GLRLM - Run Length Non-Uniformity Normalized
R-log3-glrlm-RP	Right - LoG(3mm) - GLRLM - Run Percentage
R-log5-gldm-DV	Right - LoG(5mm) - GLDM - Dependence Variance
R-log5-gldm-SDHGLE	Right - LoG(5mm) - GLDM - Small Dependence High Gray Level Emphasis
R-log5-glrlm-LREs	Right - LoG(5mm) - GLRLM - Long Run Emphasis
R-log5-glszm-LAHGLE	Right - LoG(5mm) - GLSZM - Large Area High Gray Level Emphasis
R-logar-1st-MAD	Right - logarithm - first order - Mean Absolute Deviation
R-orig-glcm-CS	Right - original - GLCM - Cluster Shade
R-orig-glcm-CT	Right - original - GLCM - Cluster Tendency
R-orig-glcm-SA	Right - original - GLCM - Sum Average
R-orig-glcm-SS	Right - original - GLCM - Sum of Squares
R-orig-gldm-LDE	Right - original - GLDM - Large Dependence Emphasis
R-orig-glrlm-HGLRE	Right - original - GLRLM - High Gray Level Run Emphasis
R-orig-glszm-SALGLE	Right - original - GLSZM - Small Area Low Gray Level Emphasis
R-squ-glcm-JE	Right - square - GLCM - Joint Energy
R-squ-glcm-MP	Right - square - GLCM - Maximum Probability
R-wavHHH-glcm-IDN	Right - wavelet (HHH) - GLCM - Inverse Difference Normalized
R-wavHLL-glcm-C	Right - wavelet (HLL) - GLCM - Contrast
R-wavLLH-glcm-CS	Right - wavelet (LHH) - GLCM - Cluster Shade
R-wavLLL-glcm-IDMN	Right - wavelet (LLL) - GLCM - Inverse Difference Moment Normalized
R-wavLLL-glrlm-GLNU	Right -wavelet (LLL) - GLRLM - GrayLevelNonUniformity
R-wavLLL-glrlm-SRHGLE	Right -wavelet (LLL) - GLRLM - Short Run High Gray Level Emphasis

ANNEX C – The impact of discretization in radiomics: what we know so far

The impact of discretization in radiomics: what we know so far

Christian Mattjie, Giordana Salvi de Souza, Rafaela Cappelari Ravazio, Luis Vinicius de Moura, Ana Maria Marques da Silva

Abstract—Discretization is a standard pre-processing step in radiomic features extraction. The most common discretization methods are resampling the intensity values into a fixed number of bins or according to fixed bin size. This review aims to summarize our current knowledge on how discretization impacts radiomic features extraction on distinct medical imaging modalities and medical tasks. A literature survey was conducted following the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines. A total of 42 eligible original articles were included in this scoping review. The results are presented in sections regarding discretization methods, feature groups, feature extraction methods, and a summary for each medical imaging modality (computed tomography, positron emission tomography, magnetic resonance imaging, ultrasound, single photon emission tomography, and electronic portal imaging). There is significant variability in the best discretization method, even for the same imaging modality and organ, tissue or disease evaluated. We recommend the best methods and values for organs and diseases. Finally, we reinforce there is an agreement among the reviewed studies that different quantization methods have a high impact on the radiomic features.

Index Terms—Medical imaging; Quantization; Rebinning; Feature Extraction;

I. BACKGROUND

Due to the use of digital technologies as tools for diagnosis, and treatment in the last decades, medical data is now widely available. Digital medical imaging offers real opportunities to improve medical outcomes and enhance efficiency in healthcare. Powered by artificial intelligence (AI) methods to facilitate early detection, allow for improved diagnosis, and lead to treatment predictions, the field of medical imaging is evolving exponentially. The emerging field of radiomics, also called imaging analytics, is one of the results of combining AI e medical imaging data [1]. Radiomics consists of extracting a large number of quantitative descriptors (features) from medical images using image processing and data characterization algorithms. The process effectively converts images into high-dimensional mineable data and rises in contrast to the traditional clinical practice of interpreting medical images as pictures qualitatively [2].

Radiomics has the potential to identify disease characteristics that are difficult to locate or even invisible to the human eye [3][4]. Computer-based texture analyses are used to quantify numerically specific features of an image. The quantitative analysis

Manuscript received 02/03/2022, This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001 and Comissão Nacional de Energia Nuclear of Brazil (CNEN)

C. Mattjie, G. S. de Souza, R. C. Ravazio, and A. M. Marques da Silva are with the Graduate Program in Biomedical Gerontology, School of Medicine, Pontifical Catholic University of Rio Grande do Sul, Porto Alegre, Brasil and with Medical Image Computing Laboratory, School of Technology, PUCRS, Porto Alegre, Brazil (email: christian.mattjie@acad.pucrs.br, giordana.souza@edu.pucrs.br, rafaela.ravazio@edu.pucrs.br, ana.marques@pucrs.br).

L. V. de Moura is with Medical Image Computing Laboratory, School of Technology, PUCRS, Porto Alegre, Brazil (email: luis.moura@edu.pucrs.br).

of morphology, intensity, and texture features is helpful in the diagnosis and prognosis of several diseases and pathologies. Texture analysis can be further categorized into structural, model-based, transformational, and statistics-based [5].

Discretization, also known as data binning or quantization, is a standard procedure before medical image features extraction. It simplifies and compresses data by reducing the number of possible intensity values or gray levels represented in the data. For medical images, discretization consists of clustering the original gray-levels or voxel intensity values to a specific interval, called bin. Two parameters are required in data binning: the bin number (or the number of bins) and the bin size (or bin width). An 8-bit image can be represented with $2^8 = 256$ bins when the bin size is equal to 1, or it can be discretized to a smaller number of grey levels, such as 128 bins, when the bin size is 2. This process is required to compute radiomic features efficiently. For example, a 12-bit medical image with 1024 intensity values will generate a $2^{1024} \times 2^{1024} \times 2^{1024}$ gray level co-occurrence matrix (GLCM) for texture features extraction, which might be time-consuming to mathematical computing [6]. The most common discretization methods are to resample intensity values into a fixed bin number (FBN) or according to a fixed bin size (FBS), also called in the literature as fixed bin width (FBW). Both methods are dependent on the range of intensity values of the region of interest (ROI) [2]. The intensity range can be preserved from the original image or determined according to the limited range inside the ROI. Fixing the bin number sets a specific number of intensity or gray levels, where each bin size is determined according to the range of intensity values divided by the number of bins. Similarly, setting a fixed bin size determines the number of intensity levels through the full range of intensities divided by the bin size. Different discretization methods and values affect feature extraction [7]. For example, a small bin number (larger bin size) homogenizes the original image and may smooth out important information related to the radiomic signature [7]. However, using too many bins (smaller bin size) may include noise, strongly affecting feature extraction, especially in second-order radiomic features, such as gray level run length matrix (GLRLM) and gray level size zone matrix (GLSZM). Few studies explore how different discretization methods impact radiomic features extraction in medical imaging analysis.

There is no optimal discretization procedure for all medical image modalities. The choice depends on the medical task and the specific imaging modality. There are several studies with positron emission tomography (PET) investigating the impact of discretization procedures for lung lesion detection [8], pediatric cancer [9], and non-small cell lung cancer treatment [10] [11] [12]. Magnetic resonance (MR) imaging researchers have been actively studying the discretization procedures before the feature extraction, related to lachrymal gland tumors and breast lesions [13], glioblastoma [14] and prostate studies [15]. Other imaging modalities, like computed tomography (CT), ultrasound (US), single-photon emission computed tomography (SPECT), and projection images in electronic portal imaging devices (EPID), face specific challenges in image discretization. However, most studies

use the default values defined by the radiomic libraries, even though their documentation remarks the importance of exploring bin sizes that make sense for the specific type of data and task.

This review aims to summarize our current knowledge on how discretization impacts radiomic features extraction on distinct imaging modalities and medical tasks.

II. METHODS

This scoping review was conducted within the guidelines of the scoping review methodology framework and recommendations guidelines of the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) Checklist [16].

A. Search, eligibility criteria

Articles published between January 1, 2012 and January 13, 2022 with any study design were searched through the following databases: PubMed, Science Direct and Scopus. We used the combination of the following terms: discretization, bin width, bin size, quantization, binning, radiomics, medical, imaging and clinical. The complete search strategy for each database is presented in the Appendix. Reviews, abstracts and non-English articles were not considered.

Studies were included if they meet the following inclusion criteria: (1) radiomics studies related to medical imaging for diagnosis or therapy purposes; (2) at least two discretization methods or bin values used; (3) recommendation of one or more discretization methods according to the medical imaging modality and task.

B. Study selection

The duplicated studies were removed using the Zotero software (Rosenzweig Center for History and New Media, USA). First, two reviewers (C.M. and G.S.S.) independently screened the titles and abstracts applying the inclusion/exclusion criteria. Then, the discrepancies were resolved by a third reviewer (A.M.M.S.). The decision rule for consensus was a simple majority. Then, the reviewers retrieved the full-text papers of the selected abstracts and, subsequently, performed an independent second-step selection based on the full-text article.

The following relevant data were extracted into an evidence table from each study to summarize and compare the studies and findings: reference (first author and publication year), study designation, population size, type of imaging modality, software used for radiomic features extraction, number of radiomic features extracted, radiomic features class extracted, number of radiomics features selected, prediction category, and key findings the study. Two authors (C.M. and G.S.S.) analysed and interpreted the data in agreement. The data gathered in the evidence table were used to define the main research themes for this scoping review.

C. Risk of bias

No quality scoring for study selection was applied. Articles were read by the two reviewers in duplicate, using the third reviewer only in cases of doubt. Data extracted from studies that did not meet the criteria were excluded.

III. RESULTS AND DISCUSSION

Figure 1 shows the selection process of the studies for this scoping review, according to the PRISMA-P flow chart. A total of 42 eligible original articles were included in this review to summarize their characteristics and radiomics approach implementation.

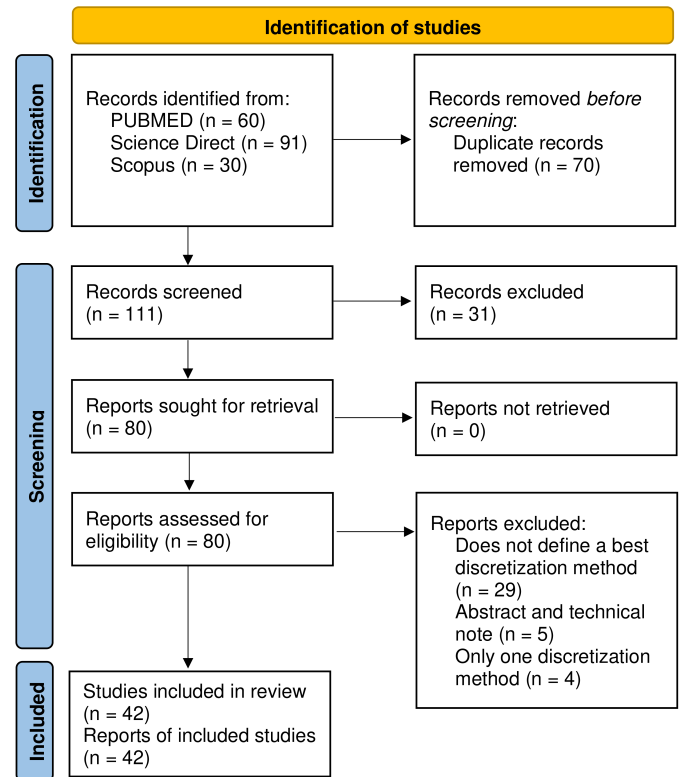


Fig. 1. PRISMA-P flow chart: articles selection process.

Based on the aim of this scoping review, discretization methods, feature extraction, and feature groups are described. The articles are divided into the main imaging systems: CT, PET, MR, US, SPECT, and EPID.

A. Discretization methods

Discretization occurs after segmentation and before applying filters (if used), followed by feature extraction. All reviewed studies used lesion segmentation to focus on the tissue under investigation, mainly tumors or specific regions of phantoms simulating the human body. The most common discretization methods were the Fixed Bin Number (FBN) and the Fixed Bin Size (FBS). These methods consist of grouping the original intensity or gray values in a set number of range intervals (bins) according to the range of the original values. FBN, also called number of gray levels (Ng) or fixed bin count, sets the number of bins. Therefore, each bin size is automatically computed according to the range of original intensity values, as shown in equation (1).

$$\text{Bin Size} = \frac{\text{Set Number of bins}}{\text{Range of original gray values}} \quad (1)$$

Similarly, FBS, also called fixed bin width, sets the size of each bin. Then, the number of bins is then automatically computed, as shown in equation (2).

$$\text{Number of Bins} = \frac{\text{Set Bin Size}}{\text{Range of original gray values}} \quad (2)$$

FBN and FBS are uniform quantization methods, which means that for each image, all bins have the same size. A variation of the FBN discretization, shown in Wang et al [17], sets the minimum value of the range as -1000 (air HU) while maintaining the maximum value according to each sample. This method is here called FBN-0.

For imaging modalities where intensity or gray values have physical meanings, setting fixed minimum and maximum values guarantees that the same contrast will be used for the whole dataset while removing information not relevant to the specific medical task. For example, a lung cancer CT study might set a range of gray values that exclude air and bone since these tissues might not be present in the tumor. Besides, pre-defined range methods guarantee the same interval of intensity values included in each bin. For example, using the FBS method with a bin size 10 on an image with an intensity range of 1 to 300 results in the first bin having a 1 to 10 intensity value interval. However, if the original intensity values vary from 5 to 300, the first bin would group values from 5 to 14. The methods that use restricted ranges for the whole dataset are called group uniform quantization methods. We call FBSrange and FBNrange the corresponding FBS and FBN methods that use range limitation before discretization. The minimum and maximum values of the intensity interval might be determined priorly according to the medical task and image properties or according to the distribution of intensity values in the dataset [18].

FBN using equal probability (FBNequal) is an alternative FBN procedure where each bin has a different size set. For example, if you could choose a random gray value from the image, there would be an equal probability that it would be on each bin [19]. So predominant gray-value ranges end up in smaller bins, while less common ranges end up in bigger bins. So, each bin has roughly the same number of gray-value samples in the discretized image. The idea behind FBNequal is to discriminate high-density gray-value ranges that would end up homogeneous using FBN or FBS.

Another variation of FBN is the Lloyd-Max quantization (FBN-Lloyd), which tests multiple combinations of transitions and gray values through an iterative optimization procedure to determine the optimal quantization bins [17]. The best discretization procedure is determined by minimizing the mean-squared quantization error - where the error is computed, for each pixel or voxel, as the difference between the discretized values and the original ones.

B. Feature Extraction Software

There are multiple software and libraries available for feature extraction. In our review, we found that "In-house" developed or institutional software were used in 12 studies (28.6%) [20][21][22][23][24][25][26][27][28][29][30], followed by Pyradiomics, an open-source python package which was used by 10 studies (23.8%) [13][15][18][31][32][33][34][35][36][37]. The package provided MATLAB programming tools package for radiomics analysis were used by 5 studies (11.9%) [17][38][39][40][8], followed by LIFEx with 4 studies (9.5%) [41][42][43][44]. Imaging Biomarker Explorer (IBEX) was used in 2 studies (4.8%) [9][45] as well as Chang-Gung Image Texture Analysis (CGITA) [46][47]. GAN based extraction was used for one study (2.4%) [48]. Five studies (11.9%) do not clarify which features extraction procedure were used [49][50][51][52][53].

C. Feature Groups

Image features can be broadly divided into intensity-based features, morphological features, and texture-based features. Notably, intensity and texture-based features require prior discretization of intensities into gray level bins. We grouped the features into first-order, morphological, and texture features. Intensity features or first-order features describe the distribution of intensities within the image ROI using basic metrics as mean, variance, energy, entropy, beyond others [[54]]. Morphological features include

descriptors related to the size and shape of the ROI and its geometric properties; they are independent of the gray-level intensity distribution in the image. Texture features also called second-order features, are obtained by calculating the statistical inter-relationships between neighboring voxels or pixels [55].

All the 42 studies in this review extracted features from Gray Level Co-occurrence Matrix (GLCM). Other than that, 88.1% of the studies used texture features from Gray Level Run Length Matrix (GLRLM) group, 76.2% of the studies used Neighbouring Gray Tone Difference Matrix (NGTDM) features, also called Neighbouring Gray Level Difference Matrix (NGLDM), First Order features also appear in 76.2% of the studies, and 73.8% extracted features from the Gray Level Size Zone Matrix (GLSZM) group. Other feature groups that also appear in some of the articles are shape features in 38.1%. Texture-based features using Neighbouring Gray-Level Dependence Matrix (NGLDM) appears in 38.1% of the studies and Neighborhood Gray-Level Difference Statistics (NGLDS) in 4.8% of the studies. On the other hand, Voxel-Alignment Matrix (VAM), Normalized Gray Level Co-occurrence Matrix (NorGLCM) and Texture Spectrum were only used in one reviewed study (2.4%).

As can be seen, Neighbouring Gray Level Difference Matrix and Neighbouring Gray Level Dependence Matrix have the same acronym NGLDM. Researchers should be extra clear on which features group they are extracting. In this review, NGTDM refers to Neighbouring Gray Tone Difference Matrix, while NGLDM refers to Neighbouring Gray Level Dependence Matrix, as recommended by the Image Biomarker Standardisation Initiative (IBSI). All features are described in the IBSI reference manual [56].

D. Study review

We split the reviewed studies according to the imaging modality. For each modality, we summarized the image formation, and how gray or intensity values are computed. Each sub-section includes a table with all reviewed studies for the specific imaging modality.

1) *Computed tomography - CT*: In CT scans, the X-ray tube emits radiation as it revolves around the patient. The amount of x-rays passing through the body region is detected by multiple rows of detectors that are opposite the x-ray tube while the tube rotates around the body. This process generated a sinogram or collection of projections in all angles, which is used to reconstruct the image of each body slice [57] mathematically. The intensity value of the reconstructed images is measured in Hounsfield Units (HU), which varies from -1024 to 3071 (12 bits), and its primary reference (HU = 0) is the attenuation coefficient (μ) of water, as shown in equation 3.

$$HU \text{ Value} = 1000 \times \frac{\mu_{\text{tissue}} - \mu_{\text{water}}}{\mu_{\text{water}} - \mu_{\text{air}}} \quad (3)$$

Therefore, gray levels or intensity values in CT images have physical meaning, and specific values may be associated with specific tissues and organs [58]. In CT, HU = 100 indicates the same attenuation coefficient, even across different exams, patients, and scanners. This physical meaning needs to be considered when choosing a discretization method. Based on each CT scan's range of intensity values, relative discretization may not be ideal because feature extraction would occur with different contrasts for each sample, losing the physical information carried with HU values.

A total of ten studies that evaluated discretization methods in CT are included in this review (Table I).

Four studies investigated discretization methods for lung cancer. Bogowicz *et al.* [20] evaluated discretization in perfusion maps

and compared the following methods: FBN with three numbers of bins and FBS for blood flow, blood volume and mean transit time with three different bin sizes. It was found that the best discretization method was FBN with 64 bins. Another lung cancer study evaluated the impact of different values of FBN would impact in the resulting features [17]. It found that FBNrange (range was defined as the lowest and maximum gray values from the whole dataset) with 128 bins had the best performance. Reference [48] restricted the HU values to a range of -100 to 400, and then applied normalization from [0,1]. Then, low resolution and high resolution images were generated using a Generative Adversarial Network from which features were then extracted. FBNrange with different bin numbers ranging from 8 to 256 were investigated, and 64 bins showed the best performance. However, the value of 10 was the only evaluated size for FBS. The lung cancer study from Desseroit et al. [53] evaluated both FBS and FBN methods and found the best performance was FBS using a bin size of 10. However, the value of 10 was the only size evaluated for FBS.

Three CT oncological studies evaluated different types of carcinomas. The hepatocellular carcinoma study [50] study compared different FBN and FBNequal methods with four different bin numbers. The best performance was achieved with FBNequal with 8 bins. For renal cell carcinoma [33] three different FBN values were evaluated, and the best performance was achieved with 128 bins. Finally, the head-and-neck squamous-cell-carcinoma study [21] made a comparison between different FBN and FBS methods. The best performance was found with the FBNequal method, but they did not recommend a specific number of bins.

The remaining studies were related to vascular imaging on CT angiography (CTA), pancreatic neuroendocrine neoplasms (Pan-NENs), and a phantom study. The CTA study [32] compared FBN and FBS methods with different bin values, and the best result was found with a bin width of 25. The research regarding PanNENs study [49] used different numbers of bins in their analysis, and reached FBN with 64 bins as the best discretization method. Finally, in a phantom study [46] bin sizes (FBS) and HU ranges were compared. They recommended a bin size range between 15 and 25, and an HU range of 2000 or 1400 for higher feature reproducibility. Since the HU range needs to be applied according to the task, it may not be applicable to all organs.

Future CT radiomics studies from the organs already evaluated by this review should use our results as starting points for their discretization investigation. FBN with 64 and 128 bins and FBS with a bin size around 25 HU might be good starting values for other body parts. We also recommend that FBNrange be further investigated for CT radiomics studies, as IBSI also endorses range re-segmentation. This procedure might preserve the physical meaning information present in HU values.

2) Positron Emission Tomography - PET: PET is a molecular imaging technique that uses a combination of positron-emitting radioisotopes labeled with a molecule tracer [59]. The radioisotope emits positive beta particles emitted by radioactive nuclei, called positrons. A process called annihilation happens when the positron emitted by the radioisotope and the electron from the environment interact, and their masses turn into energy producing two 511 keV gamma photons emitted collinearly in opposite directions, with an angle of 180°. The PET scanner is designed to detect the paired photons generated in the annihilation event. A special advantage of PET is that the tissue radioactivity can be measured in absolute units (Bq/mL). Usually, the images are quantified using the standardized uptake values (SUV), a semiquantitative measure of the tracer uptake in a region of interest normalized by the tissue activity to the injected activity and the volume of distribution

(usually total body weight).

The SUV metrics can be used for tumor staging or treatment response assessment. However, SUV describes the overall tracer uptake in a tumor but does not contain information about its distribution within the tumor [60]. Therefore, it is possible to define radiomic features that describe tumor shape and first-order statistics and capture heterogeneity within the tumor.

In this review, a total of 18 PET studies were evaluated for the discretization methods in PET, which are presented in the Table II.

In total, seven studies used phantom images; four studies evaluated lung cancer, two studies were about cervical tumors, and one study was found for each of the following diseases: hepatocellular carcinoma, neuroendocrine tumor; brain tumor; solid cancer, and nasopharyngeal carcinoma.

The PET studies about hepatocellular carcinoma [53] and nasopharyngeal carcinoma [8] evaluated different number of bins. Reference [53] achieved 256 as the best bin number, for the purpose of the study while [8] found FBN equal to 64 as the best method. The study of neuroendocrine tumors [42] used two methods - FBSrange with bin size of 0.95 and FBN with 64 bin - and the best result occurred for the FBS method. The brain tumor study [36] used FBS with values between 0.05 and 0.2 SUV. The best method was the FBS with 0.1 - 0.15 SUV. Finally, the pediatric solid cancer study [9] compared FBN methods with 16 to 64 bins and FBS with 0.1 to 0.25 SUV. The best result was achieved with FBS with a bin size of 0.1 SUV.

For lung cancer, all four PET studies compared FBN and FBS discretization methods. Three of them found the FBS was the best method [41] [28] [26], but each one sets a different bin width. Reference [41] found the bin with of 0.25 using SUV values from 0-60 as the best method. Reference [28] found the bin size of 0.1, and [26] did not present a specific best bin size. The study [53] concluded FBN with 64 bins as the best method for discretization. We recommend further investigations of both FBS and FBN methods for PET lung cancer features extraction using similar bin sizes and number of bins, like the ones presented here.

The two PET studies about cervical tumors evaluated different FBN values. The first one [27] resulted in a combination between gray-level pairs 64-32 and 64-128, while the second one [43] found a bin number of 64 as the best value. Both extracted the same group of features: 1st Order; Shape; GLCM; GLSZM; GLRLM; NGTDM. These results show that, for cervical tumors, 64 bins had a high feature reproducibility, so we recommend further investigations with a similar number of bins.

In phantoms studies, three of them compared different FBN values [52] [24] [47], resulting in 64 as the best number of bins. However, it should be noted that [52] and [24] only evaluated GLCM features. The authors from [29] compared different FBS values tumor-to-background ratio (TTBR), resulting in FBS with a bin size of 0.01 TTBR as the best value for discretization. Reference [25] compared the two methodologies: FBN with 64 bins and FBS with 0.25 bin size, with the latter having the best performance. However, only one value (bin size and number of bins) was used for each method. Similarly, reference [45] compared FBN with 64 bins and FBS with 0.5 bin size. The best method of discretization was the FBS with 0.5 bin size. Opposed to the previous results, a study [47] compared FBS with 0.4 bin width and FBN with 64 bins and found that the latter was the best discretization method.

Overall, comparing the discretization methods in PET studies (FBN and FBS), 10 articles found FBS as the best method, while 8 articles found FBN as the best method. However, there is little agreement between the same organ/disease values.

TABLE I
LITERATURE REVIEW ON DISCRETIZATION IN RADIOMIC STUDIES FOR COMPUTED TOMOGRAPHY (CT) IMAGING

Ref.	Study 1st author	Year	Site/Organ	Filters	Feature groups	Discretization Methods	Discretization Parameters	Best Method
[50]	Chang	2021	Hepatocellular Carcinoma	Original; Wavelet	1st Order; Shape; GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBN; FBNequal	FBN*: 8; 16; 32; 64	FBNequal: 8
[48]	deFarias	2021	Lung Cancer	Original	GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBNrange	FBNrange: 8; 16; 32; 64; 128; 256	FBNrange: 64
[32]	Le	2021	CT Angiography	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBS; FBN	FBS: 10; 15; 20; 25; 30; 35 FBN: 8; 16; 32; 64; 128; 256	FBS: 25; 30
[21]	Chatterjee	2021	Head & neck Carcinoma	Original	GLCM; GLSZM; GLRLM; NGTDM; NGLDM; GLDZM	FBS; FBSequal; FBN; FBNequal	FBS*: 6.25; 12.5; 25; 50 FBN*: 8; 16; 32; 64	FBNequal
[33]	Nazari	2020	Renal Cell Carcinoma	Original; Wavelet; LoG	1st Order; Shape; GLCM; GLSZM; GLRLM; NGLDM	FBN	FBN: 32; 64; 128	FBN: 128
[49]	Li	2020	Phantom	Original	GLCM; GLSZM; GLRLM	FBSrange	FBSrange: 1; 5; 10; 15; 20; 25; 30; 35; 40; 45; 50	FBSrange: 15; 20; 25; range: 1400-2000 values
[46]	Loi	2020	Pancreatic Neuroendocrine Neoplasms	Original	GLCM; GLSZM; NGTDM; NGLDM; VAM; NorGLCM; Texture Spectrum	FBN	FBN: 32; 64; 128	FBN: 64
[17]	Wang	2019	Lung Cancer	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM	FBN; FBNrange; FBN-0; FBNLloyd;	FBN*: 4; 5; 6; 7;126; 127; 128	FBNrange: 128
[53]	Desseroit	2016	Lung Cancer	Original	1st Order; Shape; GLCM; GLSZM; NGTDM	FBS; FBN	FBS: 10 FBN: 8; 16; 32; 64; 128	FBS: 10
[20]	Bogowicz	2016	Lung Cancer Perfusion Maps	Original; Wavelet	1st Order; GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBN; FBS-BF; FBS-BV; FBS-MTT	FBN: 16; 32; 64 FBS-BF: 0.5%; 1%; 2% FBS-BV: 0.5%; 1%; 2% FBS-MTT: 5%; 10%; 20%	FBN: 64

Unless specified, all bin sizes for FBS and similar methods have their values shown in Hounsfield Units while FBN and similar methods are shown in number of bins. FBS* and FBN* in the Values column refers to, respectively, all FBS and FBN based methods used in the study. FBS-BF, FBS-BV and FBS-MTT refers to fixed bin size discretization in blood flow, blood volume and mean transit time, respectively.

In summary, for PET imaging, it is observed that the best discretization method is the FBN with 64 bins, so it is a good value to start the discretization evaluation. The FBS method with values between 0.01 to 1 SUV also showed a good performance and might lose less physical associated information. We recommend both methods be evaluated for PET studies with other organs or diseases not present in this review.

3) Magnetic Resonance - MR: Magnetic resonance imaging is a non-ionizing imaging technique where the primary origin of the signal comes from the hydrogen nuclei. Aside from the main magnet, other magnetic fields excite the hydrogens in the different tissues for signal production and localization [61]. As a result, the MR images have arbitrary image intensity values, influenced by the scanner magnetic field intensity, specific variations from each scanner, and tissue characteristics like proton density, T1/T2 relaxation times, and tissue susceptibility [62].

A total of 13 studies evaluating discretization methods for MR systems were included in this review, presented in Table III.

Six of them were related to brain tumors and, of those, four used several methods and values of FBN. Reference [39] used a number of bins between 8 and 128 with the FBNequal method. The best discretization method was 16 bins. The study of [51] used three methods (FBN, FBNequal and FBNLloyd) with a bin number range from 32 to 256. They found FBN with 128 bins was the best method but recommended multiparametric radiomic extraction since different features and feature groups performed better with other discretization values. References [40] and [23] used FBN with varying numbers of bins. The best result for [40] was with a bin

number of 32, while for [23] FBN with 16 and 32 bins had the best performance. Two other MR brain tumor studies [34] [18] compared FBN and FBS discretization methods. Reference [34] compared FBS with a bin width of 2 and FBN with a bin number of 64 and found that the best performance occurred with the FBS method. Reference [18] used different values for FBN and FBSrange (where the range was defined with the mean minimum and maximum gray values for the whole dataset), and the best result was achieved with FBN with 32 bins. Therefore, we suggest further investigation of the FBN method with 32 bins for brain tumor in MR, once at least three of the studies reached this as the best result.

Two MR prostate studies were found, both using different FBS values. Reference [63] could not find a significant difference among the values, so they used Pyradiomic standard bin width of 25. However, reference [15] found FBS with a bin width of 15 as the best discretization method. Therefore, further investigations are needed to find or confirm the best-suited discretization method for this type of tumor.

Two carcinoma MR studies were also found, one regarding nasopharyngeal Carcinoma [8] and the other about hepatocellular carcinoma [22]. Both of the studies used FBN with a number of bins between 8 and 128 in their analysis. Reference [8] achieved its best outcome with the bin number of 128. Reference [22] used different FBN based methods, that resulted in FBN as the best method with a number of gray values equal to 64. A study regarding lacrymal gland lesions and breast lesions [13] compared FBN and FBS methods with different values, and the best outcome was

TABLE II
LITERATURE REVIEW ON DISCRETIZATION IN RADIOMIC STUDIES FOR POSITRON EMISSION TOMOGRAPHY IMAGING (PET)

Ref.	Study 1st author	Year	Site/Organ	Filters	Feature groups	Discretiz. Methods	Discretization Parameters	Best Method
[24]	Tamal	2021	Phantom	Original	GLCM	FBN	FBN: 8; 16; 32; 64; 128; 256	FBN: 64
[44]	Mahmoud	2021	Hepatocellular Carcinoma	Original	1st Order; GLCM; GLRLM; NGTDM; GLZLM	FBN	FBN: 32; 64; 128; 256	FBN: 256
[43]	Crandall	2020	Cervical Tumor	Original	1st Order; Shape; GLCM; GLRLM; NGTDM; GLZLM	FBN	FBN: 32; 64; 128	FBN: 64
[42]	Liberini	2021	Neuroendocrine Tumor	Original	1st Order; Shape; GLCM; GLRLM; NGTDM; GLZLM	FBSrange; FBN	FBS: 0.95* FBN: 64*	FBS: 0.95
[36]	Barry	2021	Brain Tumor	Original; Square; SquareRoot; Exp; LoG; Wavelet; Logarithm	1st Order; Shape; GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBS	FBS: 0.05; 0.1; 0.15; 0.2	FBS: 0.1; 0.15
[41]	Kolinger	2021	Lung Cancer	Original	1st Order; Shape; GLCM; GLRLM; NGTDM; GLZLM	FBSrange; FBN	FBSrange: 0.25 FBN: 64	FBS: 0.25 Range: 0:60
[25]	Pfaehler	2019	Phantom	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM; NGLDM; GLDZM	FBS; FBN	FBS: 0.25 FBN: 64	FBS: 0.25
[8]	Yang	2020	Nasopharyngeal Carcinoma	Original; Wavelet	1st Order; GLCM; GLSZM; GLRLM; NGTDM	FBN	FBN: 8; 16; 32; 64; 128	FBN: 64
[52]	Tamal	2019	Phantom	Original	GLCM	FBN	FBN: 8; 16; 32; 64; 128; 256	FBN: 64
[29]	Papp	2018	Phantom	Original	Shape; GLCM; GLSZM; NGTDM	FBS	FBS: 0.01; 0.025; 0.05; 0.1 TTBR	FBS: 0.01 TTBR
[30]	Pfaehler	2018	Phantom	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM	FBS; FBN	FBS: 0.25 FBN: 64	FBS: 0.25 with 0.05 for low uptake data
[45]	Ger	2019	Phantom	Original	1st Order; GLCM; GLRLM; NGTDM	FBS; FBN	FBS: 0.5 FBN: 64	FBS: 0.5
[9]	Branchini	2019	Solid Cancer	Original	1st Order; Shape; GLCM; GLRLM	FBS; FBN	FBS: 0.1; 0.25 FBN: 16; 64	FBS: 0.1
[28]	Carles	2018	Lung Cancer	Original	1st Order; Shape; GLCM; GLRLM; NGTDM	FBS; FBN; FBNequal	FBS: 0.05; 0.1; 0.5 FBN: 16; 32; 64 FBNequal: 16; 32; 64	FBS: 0.1
[47]	Presotto	2018	Phantom	Original	GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBS; FBN	FBS: 0.4 FBN: 64	FBN: 64
[27]	Altazi	2017	Cervical Tumor	Original	1st Order; Shape; GLCM; GLSZM; GLRLM; NGTDM	FBN	FBN: 32; 64; 128; 256	FBN: 64 with 32 and 64 with 128
[53]	Desseroit	2016	Lung Cancer	Original	1st Order; Shape; GLCM; GLSZM; NGTDM	FBS; FBN	FBS: 0.5 FBN: 8; 16; 32; 64; 128	FBN: 64
[26]	Leijenaar	2015	Lung Cancer	Original	GLCM; GLSZM; GLRLM	FBS; FBN	FBS: 0.05; 0.1; 0.2; 0.5; 1 FBN: 8; 16; 32; 64; 128	FBS

Unless specified, all bin sizes for FBS and similar methods have their values shown in Standardized Uptake Values while FBN and similar methods are shown in number of bins. FBS* and FBN* in the Values column refers to, respectively, all FBS and FBN based methods used in the study.

reached with FBS, but the best bin sizes varied depending on the MR sequence. Another cervical tumor study [35] also compared FBN and FBS methods and the best discretization method found was the FBS with a bin size of 0.05. One of the studies used phantom images [19] with different methods and values of FBN, including FBNLloyd and FBNequal and bin numbers of 32 to 256. The FBNequal discretization with 64 bins resulted in the greatest number of features with the least variability.

It should be noted that FBNLloyd discretization was explored in three MR studies and didn't have the best performance in any of them. This might indicate that FBNLloyd is not a suitable quantization procedure in radiomics, at least for the MR imaging

modality.

Three of four MR studies that compared FBN and FBS methods resulted in FBS as the best discretization method. However, the IBSI recommends that FBS should not be used for imaging modalities with arbitrary intensity values, such as MR.

Future MR studies might explore the FBS method but should also include FBN for comparison, with bins between 16 and 128, as this bin range includes the best results for the reviewed MR studies. These recommendations apply to MR in general and other organs where discretization has no impact. We recommend exploring methods and similar values for diseases and organs shown in this review to those that showed the best performance.

TABLE III
LITERATURE REVIEW ON DISCRETIZATION IN RADIOMIC STUDIES FOR MAGNETIC RESONANCE IMAGING (MRI)

Ref.	Study 1st author	Year	Site/Organ	Filters	Feature groups	Discretization Methods	Discretization Parameters	Best Method
[22]	Dai	2021	Hepatocellular Carcinoma	Original	1st Order; Shape; GLCM; GLSZM; GLRLM; NGTDM; NGLDM; NGLDS	FBN; FBNequal; FBNLloyd	FBN*: 8; 16; 32; 64; 128	FBN: 64
[63]	Xue	2021	Prostate Tumor	Original; Wavelet; LoG	1st Order; GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBS	FBS: 10; 15; 20; 25; 30; 35; 40	No significant difference
[19]	Simpson	2020	Phantom	Original	GLCM; GLSZM; GLRLM; NGTDM	FBN; FBNequal; FBNLloyd	FBN*: 32; 64; 128; 256	FBNequal: 64
[18]	Carré	2020	Brain Tumor	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBSrange; FBN	FBSrange: (1/FBN) × mean range FBN: 8; 16; 32; 64; 128; 256; 512; 1024	FBN:32
[35]	Traverso	2020	Cervical Tumor	Original; Square; SquareRoot; Exp; LoG; Logarithm Gradient	1st Order; Shape; GLCM; GLSZM; GLRLM; NGTDM; NGLDM	FBS; FBN	FBS: 0.01; 0.05; 5; 15; 25 FBN: 64	FBS: 0.05
[8]	Yang	2020	Nasopharyngeal Carcinoma	Original; Wavelet	1st Order; GLCM; GLSZM; GLRLM; NGTDM	FBN	FBN: 8; 16; 32; 64; 128	FBN: 128
[13]	Duron	2019	Lacrimal gland & breast lesions	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM; NGLDM; NGLDS	FBS; FBN	FBS: 1; 5; 10; 20; 25; 50 FBN: 8; 16; 32; 64; 128; 256; 512; 1024	FBS: dif. values for dif. sequences
[15]	Schwier	2019	Prostate Tumor	Original; Square; SquareRoot; Wavelet; Exp; LoG; Logarithm	1st Order; Shape; GLCM; GLSZM; GLRLM	FBS	FBS: 10; 15; 20; 40	FBS: 15
[40]	Ortiz-Ramón	2018	Brain Metastases	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM	FBN	FBN: 8; 16; 32; 64; 128	FBN: 32
[34]	Goya-Outi	2018	Brain Tumor	Original	GLCM; GLSZM; GLRLM	FBS; FBN	FBS: 2 FBN: 64	FBS: 2
[39]	Ortiz-Ramón	2017	Brain Metastases	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM	FBNequal	FBNequal: 8; 16; 32; 64; 128	FBNequal: 16
[51]	Li	2017	Brain Tumor	Original	1st Order; GLCM; GLSZM; GLRLM; NGTDM	FBN; FBNequal; FBNLloyd	FBN*: 32; 64; 128; 256	FBN: 128
[23]	Molina	2016	Brain Tumor	Original	GLCM; GLRLM	FBN	FBN: 8; 16; 32; 64	FBN: 16; 32

Unless specified, all sizes for FBS and similar methods are shown in units of gray values while FBN and similar methods are shown in number of bins. FBS* and FBN* in the Values column refers to, respectively, all FBS and FBN based methods used in the study.

4) Ultrasound - US: Ultrasound (US) imaging delivers mechanical pulses of high-frequency to the tissues, and the fraction reflected as an echo goes back to the detector providing information about the tissues since different tissues reflect the pulses to a greater or lesser extent. B-mode is the electronic conversion of the information from the receiver into brightness-modulated dots. The brightness of each dot can be proportional to the echo signal amplitude, but the gain control can modify the range of brightness. In the US B-mode final image, the intensity value of each pixel carries information regarding the tissue that reflected it. These intensity values are distributed on the image according to the time it took between the pulse initiation and the reception of the echo [64].

There are many variables in the US image formation process. Aside from variations from different transducers and processing techniques, there are also physical limitations, like the amplitude of the original pulse that is reduced as parts of it are reflected. This results in progressively darkening at deep tissues in the body. Therefore, intensity values on US images are relative, meaning they have no physical significance by themselves.

Only one US study [37] investigating the impact of discretization in US imaging was found and it is presented in Table IV. The study investigated features repeatability in orbital lesions using B-mode US images. Four bin sizes and four number of bins were evaluated. FBS showed overall better performance, and the bin size 10 had the highest number of repeatable features. The study also confirmed that pre-processing has a significant impact on features extraction, as for other medical imaging modalities. However, the ideal discretization method found might not be the best for outcome prediction or diagnosis, which is the final goal of radiomics. Still, future US radiomics studies should continue to investigate the impact of discretization methods and should include FBS with bin size 10 as one of them, even for other regions.

5) Single-photon Emission Computed Tomography - SPECT: SPECT is a medical imaging technique based on conventional nuclear medicine and tomographic reconstruction methods. A labeled radioisotope is administered to a patient, and depending on the biodistribution properties, it is taken up by different organs and/or tissues. Different than PET, in SPECT images, there is not a standard method to quantify images, although there recent studies

TABLE IV
LITERATURE REVIEW ON DISCRETIZATION IN RADIOMIC STUDIES FOR ULTRASOUND IMAGING

Ref.	Study 1st author	Year	Site/Organ	Filters	Feature groups	Discretization Methods	Discretization Parameters	Best Method
[37]	Duron	2021	Orbital lesion	Original	1st Order; Shape; GLCM GLSZM; GLRLM; NGTDM	FBS; FBN	FBS: 2; 5; 10; 25 FBN: 16; 32; 64; 128 bins	FBS: 10

All sizes for FBS are shown in units of gray values

combining SPECT and CT to allow calibrated SUV images. SPECT images are usually presented in counts [65]. Only one study [66] investigated the impact of discretization in SPECT imaging was found and is presented in Table V.

The SPECT study analyzed the features reproducibility of cold (low) uptake in renal scan with 99mTc-Sestamibi. The SPECT images (presented in counts) were discretized using an FBN method with eight bin numbers to evaluate the most reproducible features. As a result, the quantization with 32 to 64 bins best captured tumors heterogeneity information of cold uptake regions of 99mTc-Sestamibi images [66].

Future radiomics studies regarding SPECT might explore other discretization methods since FBN was the only one investigated. Still, one should also include FBN with 32 and 64 bins as other similar values in analysis for comparison.

6) *Electronic Portal Imaging Device images - EPID*: Electronic Portal Imaging Devices (EPIDs) are digital x-ray imaging systems that acquire projection images of patients using high-energy photons, mainly in radiation therapy [67]. The image formation principle is similar to a conventional X-ray. EPIDs have mainly been used for accurate positioning of patients prior to radiation treatment but also have the potential for dosimetry verification [68].

EPID is not a traditional diagnostic modality. However, a study was included in this review. It explores the temporal stability and prognostic prediction of EPID and digitally reconstructed radiographs (DRRs) radiomic features in lung cancer patients with three number of bins [38]. FBN with 32, 64, and 128 number of bins were evaluated, and 64 bins were recommended for computation of radiomic features in EPID and DRR images, showing the highest intra-class correlation coefficient between features.

Future EPID radiomics studies might explore other discretization methods since FBN was the only one investigated but should also include FBN with 64 bins.

IV. APPENDIX

This section describes the search strategy used in this scoping review. Three different databases were used: PubMed, Scopus, and Science Direct. The following keywords were used in the search: medical, imaging, clinical, discretization, bin width, bin size, quantization, binning, radiomics. Reviews, abstracts and non-English articles were not considered. Next we describe the search query string for each database.

On PubMed: ("discretization" OR "bin width" OR "bin size" OR "quantization" OR "binning") AND ("radiomics") AND ("Medical" OR "Imaging" OR "Clinical")

On Scopus: TITLE-ABS-KEY (("discretization" OR "bin width" OR "bin size" OR "quantization" OR "binning") AND ("radiomics")) AND ("Imaging" OR "Medical" OR "Clinical")

On Science Direct: ALL ("Medical" OR "Imaging" OR "Clinical") AND Title, abstract and keywords ("discretization" OR "bin width" OR "bin size" OR "quantization" OR "binning") AND ("radiomics")

V. CONCLUSION

In this review, we analyzed multiple studies on different medical imaging modalities (CT, PET, MR, US, SPECT, and EPID), and mapped the impact of discretization methods in radiomics. Discretization methods recommendations were made for each modality in the corresponding section, according to each organ or disease.

Currently, the extraction of quantitative radiomic features has drawn increasing research interest as a tool to improve the diagnosis and prognosis of several diseases. There is, however, considerably less research dealing with the impact of different parameters on features extraction. In this review, significant variability was observed when searching for the best discretization method, even for the same imaging modality and organ, tissue or disease evaluated.

We highlight that a single discretization method is not ideal, and multiparametric extraction might be an option for some medical tasks [51], since different feature groups perform better with different discretization procedures. Future studies could find optimal parameters for each features group and imaging modality. For other medical tasks not included in this review, we recommend discretization methods for each modality. We reinforce the importance of investigating different discretization methods since they have a high impact on the resulting radiomic features.

REFERENCES

- [1] X. Tang, "The role of artificial intelligence in medical imaging research," *British Journal of Radiology*, vol. 2, no. 1, p. 20190031, Nov. 2020. [Online]. Available: <https://doi.org/10.1259/bjro.20190031>
- [2] R. J. Gillies, P. E. Kinahan, and H. Hricak, "Radiomics: Images are more than pictures, they are data," *Radiology*, vol. 278, no. 2, pp. 563–577, Feb. 2016. [Online]. Available: <https://doi.org/10.1148/radiol.2015151169>
- [3] B. Kocak, E. S. Durmaz, E. Ates, O. Kilickesmez, and and, "Radiomics with artificial intelligence: a practical guide for beginners," *Diagnostic and Interventional Radiology*, vol. 25, no. 6, pp. 485–495, Nov. 2019. [Online]. Available: <https://doi.org/10.5152/dir.2019.19321>
- [4] J. Feger and M. Idris, "Radiomics," Nov. 2017. [Online]. Available: <https://doi.org/10.53347/rid-56602>
- [5] Y. Yoon, T. Hwang, H. Choi, and H. Lee, "Classification of radiographic lung pattern based on texture analysis and machine learning," *Journal of Veterinary Science*, vol. 20, no. 4, 2019. [Online]. Available: <https://doi.org/10.4142/jvs.2019.20.e44>
- [6] S. S. F. Yip and H. J. W. L. Aerts, "Applications and limitations of radiomics," *Physics in medicine and biology*, vol. 61, no. 13, pp. R150–R166, Jul 2016, 27269645[pmid]. [Online]. Available: <https://pubmed.ncbi.nlm.nih.gov/27269645>
- [7] J. E. van Timmeren, D. Cester, S. Tanadini-Lang, H. Alkadhi, and B. Baessler, "Radiomics in medical imaging—'how-to' guide and critical reflection," *Insights into Imaging*, vol. 11, no. 1, Aug. 2020. [Online]. Available: <https://doi.org/10.1186/s13244-020-00887-2>
- [8] F. Yang, G. Simpson, L. Young, J. Ford, N. Dogan, and L. Wang, "Impact of contouring variability on oncological PET radiomics features in the lung," *Scientific Reports*, vol. 10, no. 1, Jan. 2020. [Online]. Available: <https://doi.org/10.1038/s41598-019-57171-7>

TABLE V

LITERATURE REVIEW ON DISCRETIZATION IN RADIOMIC STUDIES FOR SINGLE-PHOTON EMISSION COMPUTED TOMOGRAPHY IMAGING (SPECT)

Ref.	Study 1st author	Year	Site/Organ	Filters	Feature groups	Discretization Methods	Discretization Parameters	Best Method
[66]	Ashrafinia	2018	Renal Cell Carcinoma	Original	1st Order; Shape; GLCM; GLSZM; GLRLM; NGTDM NGLDM; GLDZM	FBN	FBN: 4; 8; 16; 32; 64; 128; 256; 512 bins	FBN: 32; 64 bins

TABLE VI

LITERATURE REVIEW ON DISCRETIZATION IN RADIOMIC STUDIES FOR ELECTRONIC PORTAL IMAGING DEVICE IMAGES (EPID)

Ref.	Study 1st author	Year	Site/Organ	Filters	Feature groups	Discretization Methods	Discretization Parameters	Best Method
[38]	Soufi	2018	Lung Cancer	Original	1st Order; GLCM; GLRLM	FBN	FBN: 32; 64; 128 bins	FBN: 64 bins

- [9] M. Branchini, A. Zorz, P. Zucchetta, A. Bettinelli, F. D. Monte, D. Cecchin, and M. Paiusco, "Impact of acquisition count statistics reduction and SUV discretization on PET radiomic features in pediatric 18f-FDG-PET/MRI examinations," *Physica Medica*, vol. 59, pp. 117–126, Mar. 2019. [Online]. Available: <https://doi.org/10.1016/j.ejmp.2019.03.005>
- [10] R. T. Leijenaar, G. Nalbantov, S. Carvalho, W. J. van Elmpt, E. G. Troost, R. Boellaard, H. J. Aerts, R. J. Gillies, and P. Lambin, "The effect of SUV discretization in quantitative FDG-PET radiomics: the need for standardized methodology in tumor texture analysis," *Scientific Reports*, vol. 5, no. 1, Aug. 2015. [Online]. Available: <https://doi.org/10.1038/srep11075>
- [11] F. H. P. van Velden, G. M. Kramer, V. Frings, I. A. Nissen, E. R. Mulder, A. J. de Langen, O. S. Hoekstra, E. F. Smit, and R. Boellaard, "Repeatability of radiomic features in non-small-cell lung cancer [18f]FDG-PET/CT studies: Impact of reconstruction and delineation," *Molecular Imaging and Biology*, vol. 18, no. 5, pp. 788–795, Feb. 2016. [Online]. Available: <https://doi.org/10.1007/s11307-016-0940-2>
- [12] S. S. Yip, C. Parmar, J. Kim, E. Huynh, R. H. Mak, and H. J. Aerts, "Impact of experimental design on PET radiomics in predicting somatic mutation status," *European Journal of Radiology*, vol. 97, pp. 8–15, Dec. 2017. [Online]. Available: <https://doi.org/10.1016/j.ejrad.2017.10.009>
- [13] L. Duron, D. Balvay, S. V. Perre, A. Bouchouicha, J. Savatovsky, J.-C. Sadik, I. Thomassin-Naggara, L. Fournier, and A. Lecler, "Gray-level discretization impacts reproducible MRI radiomics texture features," *PLOS ONE*, vol. 14, no. 3, p. e0213459, Mar. 2019. [Online]. Available: <https://doi.org/10.1371/journal.pone.0213459>
- [14] H. Um, F. Tixier, D. Bermudez, J. O. Deasy, R. J. Young, and H. Veeraraghavan, "Impact of image preprocessing on the scanner dependence of multi-parametric MRI radiomic features and covariate shift in multi-institutional glioblastoma datasets," *Physics in Medicine & Biology*, vol. 64, no. 16, p. 165011, Aug. 2019. [Online]. Available: <https://doi.org/10.1088/1361-6560/ab2f44>
- [15] M. Schwier, J. van Griethuysen, M. G. Vangel, S. Pieper, S. Peled, C. Tempny, H. J. W. L. Aerts, R. Kikinis, F. M. Fennessy, and A. Fedorov, "Repeatability of multiparametric prostate MRI radiomics features," *Scientific Reports*, vol. 9, no. 1, Jul. 2019. [Online]. Available: <https://doi.org/10.1038/s41598-019-45766-z>
- [16] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan, R. Chou, J. Glanville, J. M. Grimshaw, A. Hróbjartsson, M. M. Lalu, T. Li, E. W. Loder, E. Mayo-Wilson, S. McDonald, L. A. McGuinness, L. A. Stewart, J. Thomas, A. C. Tricco, V. A. Welch, P. Whiting, and D. Moher, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *PLOS Medicine*, vol. 18, no. 3, p. e1003583, Mar. 2021. [Online]. Available: <https://doi.org/10.1371/journal.pmed.1003583>
- [17] H. Y. C. Wang, E. M. Donovan, A. Nisbet, C. P. South, S. Alobaidli, V. Ezhil, I. Phillips, V. Prakash, M. Ferreira, P. Webster, and P. M. Evans, "The stability of imaging biomarkers in radiomics: a framework for evaluation," *Physics in Medicine & Biology*, vol. 64, no. 16, p. 165012, Aug. 2019. [Online]. Available: <https://doi.org/10.1088/1361-6560/ab23a7>
- [18] A. Carré, G. Klausner, M. Edjlali, M. Lerousseau, J. Briend-Diop, R. Sun, S. Ammari, S. Reuzé, E. A. Andres, T. Estienne, S. Niyoteka, E. Battistella, M. Vakalopoulou, F. Dhermain, N. Paragios, E. Deutsch, C. Oppenheim, J. Pallud, and C. Robert, "Standardization of brain MR images across machines and protocols: bridging the gap for MRI-based radiomics," *Scientific Reports*, vol. 10, no. 1, Jul. 2020. [Online]. Available: <https://doi.org/10.1038/s41598-020-69298-z>
- [19] G. Simpson, J. C. Ford, R. Lorente, L. Portelance, F. Yang, E. A. Mellon, and N. Dogan, "Impact of quantization algorithm and number of gray level intensities on variability and repeatability of low field strength magnetic resonance image-based radiomics texture features," *Physica Medica*, vol. 80, pp. 209–220, Dec. 2020. [Online]. Available: <https://doi.org/10.1016/j.ejmp.2020.10.029>
- [20] M. Bogowicz, O. Riesterer, R. A. Bundschuh, P. Veit-Haibach, M. Hüllner, G. Studer, S. Stieb, S. Glatz, M. Pruschy, M. Guckenberger, and S. Tanadini-Lang, "Stability of radiomic features in CT perfusion maps," *Physics in Medicine and Biology*, vol. 61, no. 24, pp. 8736–8749, Nov. 2016. [Online]. Available: <https://doi.org/10.1088/1361-6560/61/24/8736>
- [21] A. Chatterjee, M. Vallières, R. Forghani, and J. Seuntjens, "Investigating the impact of the CT hounsfield unit range on radiomic feature stability using dual energy CT data," *Physica Medica*, vol. 88, pp. 272–277, Aug. 2021. [Online]. Available: <https://doi.org/10.1016/j.ejmp.2021.07.023>
- [22] H. Dai, M. Lu, B. Huang, M. Tang, T. Pang, B. Liao, H. Cai, M. Huang, Y. Zhou, X. Chen, H. Ding, and S.-T. Feng, "Considerable effects of imaging sequences, feature extraction, feature selection, and classifiers on radiomics-based prediction of microvascular invasion in hepatocellular carcinoma using magnetic resonance imaging," *Quantitative Imaging in Medicine and Surgery*, vol. 11, no. 5, pp. 1836–1853, May 2021. [Online]. Available: <https://doi.org/10.21037/qims-20-218>
- [23] D. Molina, J. Pérez-Beteta, A. Martínez-González, J. Martino, C. Velásquez, E. Arana, and V. M. Pérez-García, "Influence of gray level and space discretization on brain tumor heterogeneity measures obtained from magnetic resonance images," *Computers in Biology and Medicine*, vol. 78, pp. 49–57, Nov. 2016. [Online]. Available: <https://doi.org/10.1016/j.compbiomed.2016.09.011>
- [24] M. Tamal, "A phantom study to investigate robustness and reproducibility of grey level co-occurrence matrix (GLCM)-based radiomics features for PET," *Applied Sciences (Switzerland)*, vol. 11, no. 2, p. 535, Jan. 2021. [Online]. Available: <https://doi.org/10.3390/app11020535>
- [25] E. Pfaehler, J. van Sluis, B. B. Merema, P. van Ooijen, R. C. Berendsen, F. H. van Velden, and R. Boellaard, "Experimental multicenter and multivendor evaluation of the performance of PET radiomic features using 3-dimensionally printed phantom inserts," *Journal of Nuclear Medicine*, vol. 61, no. 3, pp. 469–476, Aug. 2019. [Online]. Available: <https://doi.org/10.2967/jnumed.119.229724>
- [26] R. T. Leijenaar, G. Nalbantov, S. Carvalho, W. J. van Elmpt, E. G. Troost, R. Boellaard, H. J. Aerts, R. J. Gillies, and P. Lambin, "The effect of SUV discretization in quantitative FDG-PET radiomics:

- the need for standardized methodology in tumor texture analysis," *Scientific Reports*, vol. 5, no. 1, Aug. 2015. [Online]. Available: <https://doi.org/10.1038/srep11075>
- [27] B. A. Altazi, G. G. Zhang, D. C. Fernandez, M. E. Montejo, D. Hunt, J. Werner, M. C. Biagioli, and E. G. Moros, "Reproducibility of f18-FDG PET radiomic features for different cervical tumor segmentation methods, gray-level discretization, and reconstruction algorithms," *Journal of Applied Clinical Medical Physics*, vol. 18, no. 6, pp. 32–48, Sep. 2017. [Online]. Available: <https://doi.org/10.1002/acm2.12170>
- [28] M. Carles, T. Bach, I. Torres-Espallardo, D. Baltas, U. Nestle, and L. Martí-Bonmati, "Significance of the impact of motion compensation on the variability of PET image features," *Physics in Medicine & Biology*, vol. 63, no. 6, p. 065013, Mar. 2018. [Online]. Available: <https://doi.org/10.1088/1361-6560/aab180>
- [29] L. Papp, I. Rausch, M. Grahovac, M. Hacker, and T. Beyer, "Optimized feature extraction for radiomics analysis of 18f-FDG PET imaging," *Journal of Nuclear Medicine*, vol. 60, no. 6, pp. 864–872, Nov. 2018. [Online]. Available: <https://doi.org/10.2967/jnumed.118.217612>
- [30] E. Pfähler, R. J. Beukinga, J. R. Jong, R. H. J. A. Slart, C. H. Slump, R. A. J. O. Dierckx, and R. Boellaard, "Repeatability of 18 f- FDG PET radiomic features: A phantom study to explore sensitivity to image reconstruction settings, noise, and delineation method," *Medical Physics*, vol. 46, no. 2, pp. 665–678, Dec. 2018. [Online]. Available: <https://doi.org/10.1002/mp.13322>
- [31] X. Xie, Z. Zhong, W. Zhao, C. Zheng, F. Wang, and J. Liu, "Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: Relationship to negative RT-PCR testing," *Radiology*, vol. 296, no. 2, pp. E41–E45, Aug. 2020. [Online]. Available: <https://doi.org/10.1148/radiol.2020200343>
- [32] E. Le, L. Rundo, J. Tarkin, N. Evans, M. Chowdhury, P. Coughlin, H. Pavey, C. Wall, F. Zaccagna, F. Gallagher, Y. Huang, R. Sriranjani, A. Le, J. Weir-McCall, M. Roberts, F. Gilbert, E. Warburton, C.-B. Schönlieb, E. Sala, and J. Rudd, "Assessing robustness of carotid artery CT angiography radiomics in the identification of culprit lesions in cerebrovascular events," *Scientific Reports*, vol. 11, no. 1, 2021.
- [33] M. Nazari, I. Shiri, G. Hajianfar, N. Oveisi, H. Abdollahi, M. R. Deevband, M. Oveisi, and H. Zaidi, "Noninvasive fuhrman grading of clear cell renal cell carcinoma using computed tomography radiomic features and machine learning," *La radiologia medica*, vol. 125, no. 8, pp. 754–762, Mar. 2020. [Online]. Available: <https://doi.org/10.1007/s11547-020-01169-z>
- [34] J. Goya-Outi, F. Orhac, R. Calmon, A. Alentorn, C. Nioche, C. Philippe, S. Puget, N. Boddart, I. Buvat, J. Grill, V. Frouin, and F. Frouin, "Computation of reliable textural indices from multimodal brain MRI: suggestions based on a study of patients with diffuse intrinsic pontine glioma," *Physics in Medicine & Biology*, vol. 63, no. 10, p. 105003, May 2018. [Online]. Available: <https://doi.org/10.1088/1361-6560/aabd21>
- [35] A. Traverso, M. Kazmierski, M. L. Welch, J. Weiss, S. Fiset, W. D. Foltz, A. Gladwish, A. Dekker, D. Jaffray, L. Wee, and K. Han, "Sensitivity of radiomic features to inter-observer variability and image pre-processing in apparent diffusion coefficient (ADC) maps of cervix cancer patients," *Radiotherapy and Oncology*, vol. 143, pp. 88–94, Feb. 2020. [Online]. Available: <https://doi.org/10.1016/j.radonc.2019.08.008>
- [36] N. Barry, P. Rowshanfarzad, R. J. Francis, A. K. Nowak, and M. A. Ebert, "Repeatability of image features extracted from FET PET in application to post-surgical glioblastoma assessment," *Physical and Engineering Sciences in Medicine*, vol. 44, no. 4, pp. 1131–1140, Aug. 2021. [Online]. Available: <https://doi.org/10.1007/s13246-021-01049-4>
- [37] L. Duron, J. Savatovsky, L. Fournier, and A. Lecler, "Can we use radiomics in ultrasound imaging? impact of preprocessing on feature repeatability," *Diagnostic and Interventional Imaging*, vol. 102, no. 11, pp. 659–667, Nov. 2021. [Online]. Available: <https://doi.org/10.1016/j.diii.2021.10.004>
- [38] M. Soufi, H. Arimura, T. Nakamoto, T. aki Hirose, S. Ohga, Y. Umezū, H. Honda, and T. Sasaki, "Exploration of temporal stability and prognostic power of radiomic features based on electronic portal imaging device images," *Physica Medica*, vol. 46, pp. 32–44, Feb. 2018. [Online]. Available: <https://doi.org/10.1016/j.ejmp.2017.11.037>
- [39] R. Ortiz-Ramon, A. Larroza, E. Arana, and D. Moratal, "Identifying the primary site of origin of MRI brain metastases from lung and breast cancer following a 2d radiomics approach," in *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. IEEE, Apr. 2017. [Online]. Available: <https://doi.org/10.1109/isbi.2017.7950735>
- [40] R. Ortiz-Ramón, A. Larroza, S. Ruiz-España, E. Arana, and D. Moratal, "Classifying brain metastases by their primary site of origin using a radiomics approach based on texture analysis: a feasibility study," *European Radiology*, vol. 28, no. 11, pp. 4514–4523, May 2018. [Online]. Available: <https://doi.org/10.1007/s00330-018-5463-6>
- [41] G. D. Kolinger, D. V. García, G. M. Kramer, V. Frings, G. J. C. Zwezerijnen, E. F. Smit, A. J. D. Langen, I. Buvat, and R. Boellaard, "Effects of tracer uptake time in non-small cell lung cancer 18f-FDG PET radiomics," *Journal of Nuclear Medicine*, p. jnumed.121.262660, Dec. 2021. [Online]. Available: <https://doi.org/10.2967/jnumed.121.262660>
- [42] V. Liberini, B. D. Santi, O. Rampado, E. Gallio, B. Dionisi, F. Ceci, G. Polverari, P. Thuillier, F. Molinari, and D. Deandrei, "Impact of segmentation and discretization on radiomic features in 68ga-DOTA-TOC PET/CT images of neuroendocrine tumor," *European Journal of Nuclear Medicine and Molecular Imaging Physics*, vol. 8, no. 1, Feb. 2021. [Online]. Available: <https://doi.org/10.1186/s40658-021-00367-6>
- [43] J. P. Crandall, T. J. Fraum, M. Lee, L. Jiang, P. Grigsby, and R. L. Wahl, "Repeatability of 18f-FDG PET radiomic features in cervical cancer," *Journal of Nuclear Medicine*, vol. 62, no. 5, pp. 707–715, Oct. 2020. [Online]. Available: <https://doi.org/10.2967/jnumed.120.247999>
- [44] M. Mahmoud, M. Shihab, S. Saad, F. Elhussiny, and M. Houseni, "Effect of Standardized Uptake Value Discretization on Radiomics Features of Liver Tumors Using 18fdg-Pet/Ct Scan," *Russian Electronic Journal of Radiology*, vol. 11, no. 3, pp. 132–137, 2021.
- [45] R. B. Ger, J. G. Meier, R. B. Pahlka, S. Gay, R. Mumme, C. D. Fuller, H. Li, R. M. Howell, R. R. Layman, R. J. Stafford, S. Zhou, O. Mawlawi, and L. E. Court, "Effects of alterations in positron emission tomography imaging parameters on radiomics features," *PLOS ONE*, vol. 14, no. 9, p. e0221877, Sep. 2019. [Online]. Available: <https://doi.org/10.1371/journal.pone.0221877>
- [46] S. Loi, M. Mori, G. Benedetti, S. Partelli, S. Broggi, G. M. Cattaneo, D. Palumbo, F. Muffatti, M. Falconi, F. D. Cobelli, and C. Fiorino, "Robustness of CT radiomic features against image discretization and interpolation in characterizing pancreatic neuroendocrine neoplasms," *Physica Medica*, vol. 76, pp. 125–133, Aug. 2020. [Online]. Available: <https://doi.org/10.1016/j.ejmp.2020.06.025>
- [47] L. Presotto, V. Bettinardi, E. D. Bernardi, M. Belli, G. Cattaneo, S. Broggi, and C. Fiorino, "PET textural features stability and pattern discrimination power for radiomics analysis: An "ad-hoc" phantoms study," *Physica Medica*, vol. 50, pp. 66–74, Jun. 2018. [Online]. Available: <https://doi.org/10.1016/j.ejmp.2018.05.024>
- [48] E. C. de Farias, C. di Noia, C. Han, E. Sala, M. Castelli, and L. Rundo, "Impact of GAN-based lesion-focused medical image super-resolution on the robustness of radiomic features," *Scientific Reports*, vol. 11, no. 1, Nov. 2021. [Online]. Available: <https://doi.org/10.1038/s41598-021-00898-z>
- [49] Y. Li, G. Tan, M. Vangel, J. Hall, and W. Cai, "Influence of feature calculating parameters on the reproducibility of CT radiomic features: a thoracic phantom study," *Quantitative Imaging in Medicine and Surgery*, vol. 10, no. 9, pp. 1775–1785, Sep. 2020. [Online]. Available: <https://doi.org/10.21037/qims-19-921>
- [50] H.-C. Chang, Y.-H. Dai, P.-C. Shen, W.-C. Chang, C.-H. Lo, J.-F. Yang, C.-S. Lin, H.-L. Chao, S.-J. Tu, W.-Y. Huang, and J.-M. Hwang, "Using computed tomography-based radiomics to predict outcomes for hepatocellular carcinoma patients receiving stereotactic body radiotherapy," *Therapeutic Radiology and Oncology*, vol. 5, pp. 13–13, Sep. 2021. [Online]. Available: <https://doi.org/10.21037/tro-21-8>
- [51] Q. Li, H. Bai, Y. Chen, Q. Sun, L. Liu, S. Zhou, G. Wang, C. Liang, and Z.-C. Li, "A fully-automatic multiparametric radiomics model: Towards reproducible and prognostic imaging signature for prediction of overall survival in glioblastoma multiforme," *Scientific Reports*, vol. 7, no. 1, Oct. 2017. [Online]. Available: <https://doi.org/10.1038/s41598-017-14753-7>
- [52] M. Tamal, "Grey level co-occurrence matrix (GLCM) as a radiomics feature for artificial intelligence (AI) assisted positron emission tomography (PET) images analysis," vol. 646, no. 1. IOP Publishing, Oct. 2019, p. 012047. [Online]. Available: <https://doi.org/10.1088/1757-899x/646/1/012047>
- [53] M.-C. Desseroit, F. Tixier, W. A. Weber, B. A. Siegel, C. C. L. Rest, D. Visvikis, and M. Hatt, "Reliability of PET/CT shape and heterogeneity features in functional and morphologic components of non-small cell lung cancer tumors: A repeatability analysis in a prospective multicenter cohort," *Journal of Nuclear Medicine*, vol. 58, no. 3, pp. 406–411, Oct. 2016. [Online]. Available: <https://doi.org/10.2967/jnumed.116.180919>
- [54] J. J. van Griethuysen, A. Fedorov, C. Parmar, A. Hosny, N. Aucoin, V. Narayan, R. G. Beets-Tan, J.-C. Fillion-Robin, S. Pieper, and H. J. Aerts, "Computational radiomics system to decode the radiographic

- phenotype,” *Cancer Research*, vol. 77, no. 21, pp. e104–e107, Oct. 2017. [Online]. Available: <https://doi.org/10.1158/0008-5472.can-17-0339>
- [55] S. Rizzo, F. Botta, S. Raimondi, D. Origgi, C. Fanciullo, A. G. Morganti, and M. Bellomi, “Radiomics: the facts and the challenges of image analysis,” *European Radiology Experimental*, vol. 2, no. 1, Nov. 2018. [Online]. Available: <https://doi.org/10.1186/s41747-018-0068-z>
- [56] A. Zwanenburg, M. Vallières, M. A. Abdalah, H. J. W. L. Aerts, V. Andrearczyk, A. Apte, S. Ashrafinia, S. Bakas, R. J. Beukinga, R. Boellaard, M. Bogowicz, L. Boldrini, I. Buvat, G. J. R. Cook, C. Davatzikos, A. Depeursinge, M.-C. Desserot, N. Dinapoli, C. V. Dinh, S. Echegaray, I. E. Naqa, A. Y. Fedorov, R. Gatta, R. J. Gillies, V. Goh, M. Götz, M. Guckenberger, S. M. Ha, M. Hatt, F. Isensee, P. Lambin, S. Leger, R. T. Leijenaar, J. Lenkiewicz, F. Lippert, A. Losnegård, K. H. Maier-Hein, O. Morin, H. Müller, S. Napel, C. Nioche, F. Orlhac, S. Pati, E. A. Pfaehler, A. Rahmim, A. U. Rao, J. Scherer, M. M. Siddique, N. M. Sijtsema, J. S. Fernandez, E. Spezi, R. J. Steenbakkers, S. Tanadini-Lang, D. Thorwarth, E. G. Troost, T. Upadhaya, V. Valentini, L. V. van Dijk, J. van Griethuysen, F. H. van Velden, P. Whybra, C. Richter, and S. Löck, “The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping,” *Radiology*, vol. 295, no. 2, pp. 328–338, May 2020. [Online]. Available: <https://doi.org/10.1148/radiol.2020191145>
- [57] L. W. Goldman, “Principles of CT and CT technology,” *Journal of Nuclear Medicine Technology*, vol. 35, no. 3, pp. 115–128, Sep. 2007. [Online]. Available: <https://doi.org/10.2967/jnmt.107.042978>
- [58] B. P. Little, “Approach to chest computed tomography,” *Clinics in Chest Medicine*, vol. 36, no. 2, pp. 127–145, Jun. 2015. [Online]. Available: <https://doi.org/10.1016/j.ccm.2015.02.001>
- [59] C. C. Robilotta, “A tomografia por emissão de pósitrons: uma nova modalidade na medicina nuclear brasileira,” *Revista Panamericana de Salud Pública*, vol. 20, no. 2-3, Sep. 2006. [Online]. Available: <https://doi.org/10.1590/s1020-49892006000800010>
- [60] E. Pfaehler, “Methodological aspects and standardization of PET radiomics studies,” Ph.D. dissertation, University of Groningen, 2021. [Online]. Available: <https://doi.org/10.33612/diss.149306583>
- [61] G. Yves, H. Aline, G. Pierre, and V. Q. Lam, “Physics of magnetic resonance imaging: from spin to pixel,” *Journal of Physics D: Applied Physics*, vol. 43, no. 21, p. 213001, May 2010. [Online]. Available: <https://doi.org/10.1088/0022-3727/43/21/213001>
- [62] O. L. Wong, J. Yuan, Y. Zhou, S. K. Yu, and K. Y. Cheung, “Longitudinal acquisition repeatability of MRI radiomics features: An ACR MRI phantom study on two MRI scanners using a 3d t1w TSE sequence,” *Medical Physics*, vol. 48, no. 3, pp. 1239–1249, Feb. 2021. [Online]. Available: <https://doi.org/10.1002/mp.14686>
- [63] C. Xue, J. Yuan, D. M. Poon, Y. Zhou, B. Yang, S. K. Yu, and Y. K. Cheung, “Reliability of MRI radiomics features in MR-guided radiotherapy for prostate cancer: Repeatability, reproducibility, and within-subject agreement,” *Medical Physics*, vol. 48, no. 11, pp. 6976–6986, Sep. 2021. [Online]. Available: <https://doi.org/10.1002/mp.15232>
- [64] M. Mahesh, “The essential physics of medical imaging, third edition,” *Medical Physics*, vol. 40, no. 7, p. 077301, Jun. 2013. [Online]. Available: <https://doi.org/10.1118/1.4811156>
- [65] National Research Council (US) and Institute of Medicine (US) Committee on the Mathematics and Physics of Emerging Dynamic Biomedical Imaging, *Mathematics and Physics of Emerging Biomedical Imaging*. Washington (DC): National Academies Press (US), Feb. 1996. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK232492/>
- [66] S. Ashrafinia, K. Jones, M. A. Gorin, S. P. Rowe, M. S. Javadi, M. G. Pomper, M. E. Allaf, and A. Rahmim, “Reproducibility of cold uptake radiomics in 99m tc-sestamibi SPECT imaging of renal cell carcinoma,” in *2017 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)*. IEEE, Oct. 2017. [Online]. Available: <https://doi.org/10.1109/nssmic.2017.8533125>
- [67] A. L. Boyer, L. Antonuk, A. Fenster, M. Van Herk, H. Meertens, P. Munro, L. E. Reinstein, and J. Wong, “A review of electronic portal imaging devices (epids),” *Medical Physics*, vol. 19, no. 1, pp. 1–16, 1992. [Online]. Available: <https://aapm.onlinelibrary.wiley.com/doi/abs/10.1118/1.596878>
- [68] A. Grządziel, B. Smolińska, R. Rutkowski, and K. Ślosarek, “Epid dosimetry – configuration and pre-treatment imrt verification,” *Reports of Practical Oncology & Radiotherapy*, vol. 12, no. 6, pp. 307–312, 2007. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1507136710600697>